Steric aspects of the interaction of metal species with nucleic acid constituents

Thesis submitted for the degree

of

Doctor of Philosophy

in the

Department of Chemical Sciences

at the

Victoria University of Technology

by

Elizabeth Yuriev

March 1997

Publications, relevant to the scope of the thesis

1. Baker, A.T.; Crass, J.; Kok, G.B.; Orbell, J.D.; Yuriev, E. *Inorg. Chim. Acta.* **1993**, *214*, 169.

2. Watkins, P.; Yuriev, E. *AA Instruments at Work* **1995**, *118*, 1.

3. Yuriev, E.; Orbell, J.D. *Inorg. Chem.* **1996**, *35*, 7914.

4. Yuriev, E.; Orbell, J.D. *J. Comput.-Aided Mol. Design* **1996**, *10*, 589.

Conference presentations, relevant to the scope of the thesis

1. Baker, A.T.; Crass, J.K.; Kok, G.B.; Orbell, J.D.; Yuriev, E. *"The Design of Sterically Restrictive Metal Complexes and their Binding to DNA Constituents"*, **EUROBIC II**, Metal Ions in Biological systems, Societa' Chimica Italiana, Florence - Italy, August 30 - September 3, **1994**.

2. Orbell, J.D.; Yuriev, E. *"Modelling Steric Effects in DNA-binding Metal Complexes"*, **"Molecular Design Down Under"** Conference, Cairns, Australia, August-September, **1995**; the abstract is published in *Chemical Design Automation News*, **1995**, *10*, 38.

3. Orbell, J.D.; Yuriev, E. *"Towards Quantifying Steric Effects in Metal-Complex Nucleic Acid Systems"*, Second CSIRO Symposium **"Computational Challenges in Life Sciences"**, Melbourne, 5-7 February, **1996**.

4. Yuriev, E.; Orbell, J.D.; Watkins, P. *"Quantifying Steric Effects in Metal-Complex Nucleobase Systems"*, Fourth Congress of World Association of Theoretically Oriented Chemists (**WATOC-96**), Jerusalem, 7-12 July, **1996**.

5. Yuriev, E.; Orbell, J.D. *"Application of Repulsive Energy Strategy to Metal Binding Sites on Nucleobases: Rigid and Flexible Steric Probes"*, **Second Australian Molecular Modelling Workshop**, Canberra, 2-4 October, **1996**.

6. Yuriev, E.; Orbell, J.D. *"Modelling of Platinum Monofunctional Binding to Nucleobases. Steric and Conformational Aspects"*, **"Up and Coming Research in Australia"**, Royal Australian Chemical Institute, Medicinal and Agricultural Division 13th National Conference, Melbourne, 8-11 December, **1996**.

ABSTRACT

The delineation of steric effects in the interaction of metal species with biomolecules has been addressed. More specifically, a focus has been placed on the interaction of platinum complexes with nucleoconstituents. Steric effects thus characterised have been related to the biological properties of such complexes. Steric aspects of metal complex speciation have been investigated both from experimental and theoretical viewpoints. Manipulation of steric effects within model systems has been carried out by way of judicious ligand design. A sterically determined "switching" event under stoichiometric control has been proposed. An established repulsive energy methodology has been extended to bioinorganic models resulting in the quantification of steric effects and characterisation of a number of steric features potentially important for biological activity. Properties relating to the steric bulk of carrier ligand such as localisation, optimal size, flexibility and flatness have been demonstrated to exert specific influences. It has been shown that the relative steric demands of nucleobase binding sites can be quantified and the generated parameters can be used to rationalise the metal binding preference order. This is expected to be applicable to biomolecules in general. The conformational aspects of a metal complex adduct with a nucleobase have been shown to be related to steric demands exerted by both carrier ligand(s) and exocyclic functional groups in the vicinity of the target site.

ACKNOWLEDGMENTS

Abbreviations

AAS - atomic absorption spectroscopy

AMBER - Assisted Model Building and Energy Refinement

CRE - complex repulsive energy

CSD - Cambridge Structural Database

DA - dihedral angle

DDP - dichlorodiammineplatinum(II)

DNA - deoxyribonucleic acid

FF - force field

HMG - high mobility group

HOMO - highest occupied molecular orbital

HTH - head-to-head

HTT - head-to-tail

IR - infra-red

LFER - linear free energy relationship

LP - lone pair (of electrons)

LRE - ligand repulsive energy

LUMO - lowest unoccupied molecular orbital

MC - Monte Carlo

MD - molecular dynamics

MEP - molecular electrostatic potential

MLR - multiple linear regression

MM - molecular mechanics

NB - nucleobase

NMR - nuclear magnetic resonance

PCA - principal components analysis

PDB - Protein Databank

PLS - partial least squares

POS - point-on-a-sphere

QD - quenched dynamics

QM - quantum mechanics

PtN4 - platinum coordination plane in bis(N-bound nucleoconstituent) complexes with N-containing carrier ligands

RE - repulsive energy

RMS - root-mean-square

RNA - ribonucleic acid

RTS - room temperature simulation

QSAR - quantitative structure-activity relationship

QSPR - quantitative structure-property relationship

SA - simulated annealing

SAR - structure-activity relationship

SSRP - structure-specific recognition protein

UV-Vis - ultra-violet and visible

vdW - van der Waals

VFF - valence force field

<u>DNA constituents:</u>

Nucleobases: A - adenine, C - cytosine, G - guanine, H - hypoxanthine,

T - thymine, U - uracil

Nucleosides: Ade - adenosine, Cyt - cytidine, Guo - guanosine,

Ino - inosine, Thy - thymidine, Uri - uridine

Nucleotides: Amp - adenosine monophosphate, Cmp - cytidine mono-
phosphate, Gmp - guanosine monophosphate, Imp - inosine monophosphate,
Tmp - thymidine monophosphate, Ump - uridine monophosphate

<u>Carrier ligands:</u>

en - ethylenediamine

TMED - tetramethylethylenediamine

bpe - 1,2-bis(pyridin-2-yl)ethane

bmpe - 1,2-bis(6-methylpyridin-2-yl)ethane

dmp - 1,4-dimethylpiperazine

bispep - 1,2-di-(4-methylpiperazine)

<u>Other compounds:</u>

DMSO - dimethylsulfoxide

DSS - sodium 2,2-dimethyl-2-silapentane-5-sulfonate

TMS - tetramethylsilane

Abstract

Acknowledgments

Abbreviations

Chapter 2. Solution and solid state studies of metal complex – nucleic acid constituent interactions

Chapter 3. Structural studies of platinum complex – nucleic acid interactions using molecular modelling techniques

Chapter 5. Investigation of Quantitative Structure-Activity Relationships in platinum systems

Chapter 6. Summary

Appendices

Chapter 1

Introduction

1.1 Preamble

The involvement of metal species in nucleic acid biochemistry is well-recognised [1,2]. Although considerable information is now available on the interaction of various metal species with nucleic acids and their constituents, much work remains to be done in order to elucidate the details of such involvement. Indeed, it is not inappropriate to suggest that this crucial area of bioinorganic chemistry is still in its infancy.

As well as metal species being intrinsic components of nucleic acid biology [2] their interaction with nucleic acids and nucleoconstituents is also of interest with respect to metallopharmaceutics [3-5], metal complexes as tools in molecular biology [6] and in relation to environmental pollution by metal species [7].

With respect to metallopharmaceutics, probably the most famous example is the investigation of DNA-binding antitumour platinum drugs [8,9], the primary example being cisplatin (cis-dichlorodiammineplatinum(II)):

$$\begin{array}{ccc} Cl & & NH_3 \\ & \diagdown \diagup & \\ & Pt & \\ & \diagup \diagdown & \\ Cl & & NH_3 \end{array}$$
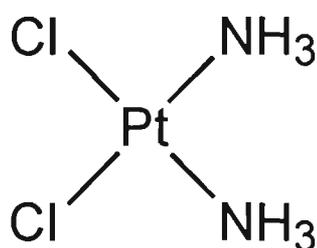
However, much still remains unanswered in regard to the mode of action of these compounds and some aspects, such as steric effects, have barely been addressed. It is anticipated that some of the findings of this thesis relating to the interaction of particular metal species with nucleoconstituents will be generally applicable. This is certainly expected to be the case for the characterisation of steric effects in such systems.

1.2 Mode of action of platinum drugs

1.2.1 Binding modes

Today, it is widely accepted that DNA is the major intracellular target of cisplatin [10-13]. A number of nucleic acid binding sites of cisplatin and its derivatives (analogues) have been established [14,15]; these are documented in Chapter 4 (Figs. 4.3 and 4.4). When cisplatin reacts with DNA there is a number of possible binding modes which are considered to be of importance. These include *interstrand*, *intrastrand* and *DNA-protein crosslinking*. The adducts which account for more than 90% of the bound platinum and supposedly lead to cell death [14,16] are shown in Fig.1.1:

1.   bifunctional binding to two adjacent guanine bases on the same strand

2.   bifunctional binding to adenine and guanine bases on the same strand

3.   bifunctional binding to two guanine bases on the same strand, separated by a third base

4.   bifunctional binding to two guanine bases of two separate strands

5.   monofunctional binding to a guanine base

6.   DNA-protein crosslinking

1.2.2 Mechanism hypotheses

In spite of DNA being almost universally acknowledged as the key target of the platinum drugs [13], and in spite of the enormous amount of information resulting from research in this area, it is still unclear as to the exact mechanism of action. However, a number of hypotheses of varying sophistication have been advanced.
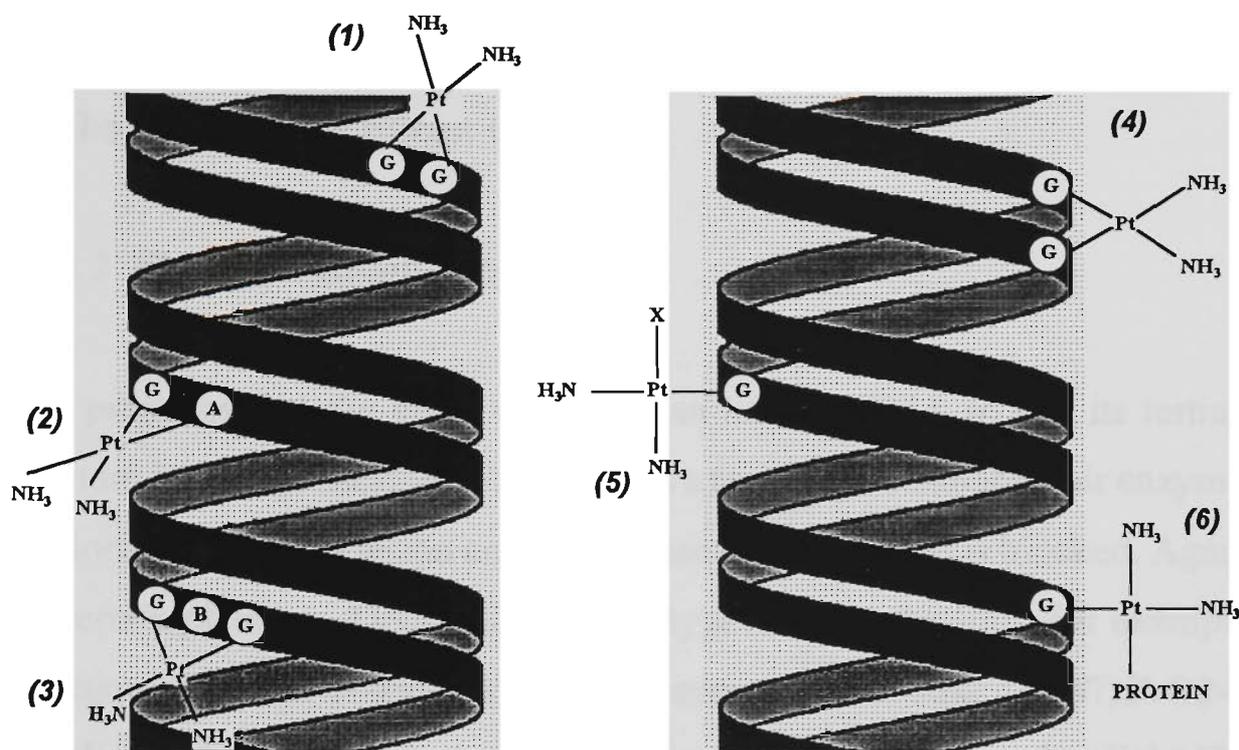
Figure 1.1. *The major cisplatin-DNA adducts. The adducts on the left (1-3) represent intrastrand crosslinks. A dduct (4) is an interstrand crosslink, adducts (5) and (6) represent monofunctional binding and DNA -protein crosslinking, respectively.*

1.2.2.1 The severe disruption hypothesis

This school of thought proposes that when the metal complex binds to DNA and forms an intrastrand crosslinkage, severe disruption to the DNA tertiary structure results. Such massive disruption, though potentially recognisable by the cellular DNA repair machinery, could be beyond the capability of the cell to address [17]. Subsequently, replication is arrested. This hypothesis is feasible for certain platinum–DNA adducts. For example, a potentially accessible C(N3)-Pt-C(N3) crosslink (where C(N3) is the N3 site of cytosine) may introduce significant rigid structural disruptions into the tertiary structure of DNA, as has been demonstrated on the basis of the conformational features of the *cis*-bis(1-

methylcytosine)-platinum(II) complex [18]. Thus, if one accepts that cancer cells have diminished repair capability [19], the lesions imposed by such crosslinks may be particularly difficult to excise.

1.2.2.2 The mild disruption hypothesis

This proposes that the metal complex can bind to DNA so that its tertiary structure is not significantly disrupted. As a result, the cellular repair enzymes overlook the induced lesion and, subsequently, replication is blocked. Again, for certain platinum–DNA adducts this hypothesis is feasible. For example, the degree of local DNA disruption caused in order to accommodate a G(N7)-Pt-G(N7) crosslink (where G(N7) is the N7 site of guanine) is known to be slight [20] in comparison with the potential C(N3)-Pt-C(N3) crosslink discussed above. This hypothesis is further supported by the comparative analysis of the binding features of *cis*- and *trans*-DDP. The latter also binds to DNA and blocks replication, but this complex is not antitumour active. Many studies have shown that the structural distortions introduced into the DNA tertiary structure by *trans*-DDP are more significant and, therefore, serve as stronger recognition signals for the cellular DNA repair machinery [21,22].

1.2.2.3 The specific structural motifs hypothesis

The recent publication of the crystal structure of a *cis*-[Pt(NH$_3$)$_2$\{d(GpG)\}] intrastrand crosslink within a double stranded DNA dodecamer [23] together with the report on the NMR structure of a *cis*-[Pt(NH$_3$)$_2$\{d(GpG)\}] interstrand crosslink within a double stranded DNA decamer [24] provide detailed information on unique distortions of duplex DNA which may be related to antitumour activity [8]. These studies have shown that when cisplatin binds to DNA it

induces very specific structural motifs, with associated unwinding and bending, in the DNA tertiary structure. Proteins, which recognise these motifs, bind quite specifically to such cisplatin-modified DNA. These proteins were found to contain the high mobility group (HMG) domain and have been designated as Structure-Specific Recognition Proteins (SSRPs) [25]. Although the exact mechanism is still unclear, several hypotheses for how SSRPs might explain the anticancer activity of cisplatin have been proposed [8,25]. Thus, binding of a HMG-related protein to a cisplatin-DNA adduct may prevent recognition of the damaged DNA by repair enzymes, leading to replication inhibition and eventually to cell death. Alternatively, the HMG-protein itself for which biological function is unclear could be required for transcription. Thus tying it up in cisplatin-DNA-protein complexes may interfere with the transcription process leading to cell death.

1.2.2.4 Hypotheses synergy

Is there more than one biochemical pathway by which cis-platinum complexes exert their antitumour activity? The actual mechanism may be a combination of two or more of the proposed hypotheses. However, it is clear that the antitumour activity of these compounds is related in some way to the nature of structural distortions introduced into DNA as the result of drug-binding. It is worth mentioning that the prevalence of a particular adduct does not necessarily implicate it as the critical lesion. It may well be that "minor adducts" which are feasible, albeit after the initial unwinding and disruption, are the important events.

1.2.3 Structural variation studies

Thus, even though significant progress is being made, more investigations are required into the mode of action of platinum complexes in order to further explore the various hypotheses and, indeed, to glean further fundamental information on the nature of metal-induced disruptions on DNA.

One approach to this problem focuses on varying the structural features of platinum complexes. Such structural variations are usually performed on the so-called "carrier-ligand(s)"[1] which is usually a nitrogen donor(s). This part of the complex is considered to be inert with respect to the Pt-N bond. The "leaving group(s)", which usually involve more labile oxygen or halogen donors, which are more associated with the kinetic aspects of the interaction, may also be varied [27-29]. Consequently, the structures, the DNA-binding and the biological activities of thus modified complexes are investigated.

1.2.3.1 Rational strategy in structural variation studies

Detailed structural studies, such as these mentioned previously [23,24], provide a basis for modifying the metal species so as to consolidate a given motif or to produce variants which may be related to specific biological outcomes [30,31].

1.2.3.2 Empirical strategy in structural variation studies

The more empirical studies relate to the derivation of structure-activity relationships

---

[1] These ligands have also been referred to as "transport ligands" by some authors [26].

(SARs) for platinum complexes. On the basis of such SARs new complexes may be designed which offer the possibility of improved characteristics including solubility [27,32,33], potency [5,28,29,34,35], tumour versatility [36-41], reduced side-effects [27,35,37,38,42-44] and tumour cell resistance [45,46].

1.3 Structure-activity relationships of platinum complexes

A body of work relating to the biological activity of miscellaneous platinum complexes has led to the formulation of several SARs resulting in the following requirements being proposed [34,47-49].

1.3.1 Structural features in the vicinity of the metal center

- *Cis-configuration around the metal center* is considered to lead to less severe structural distortions than *trans-* and, consequently, allows the adducts to escape the cellular repair mechanisms. A number of *trans*-complexes have been shown to have some antitumour activity but the mechanism of their action is likely to be different from cisplatin [50,51].

- *Square-planar geometry* of the platinum coordination sphere is considered to be essential for the antitumour activity.

- *Neutrality of the complex* is required in order for complexes to be able to transfer through cell membranes.

- *Pt(II) oxidation state.* Some octahedral Pt(IV) complexes also show certain activity [29,37,52,53], although it is thought that they undergo reduction to active Pt(II) species on route to cellular targets [47].

1.3.2 Nature of the leaving group(s)

Leaving groups are required to be of intermediate lability (e.g. halides, carboxylates). The nature of these ligands determine the kinetics and thermodynamics of the hydrolysis of complexes once they have entered the cell, an area of low chloride concentration. Therefore, the nature of these ligands determines the solubility of the complexes and, consequently, influences their biological activity [27-29].

1.3.3 Presence of at least one N-H group in the carrier ligand(s)

This arrangement is postulated to provide additional hydrogen-bonding stabilisation of a platinum–DNA adduct [20,54]. However, the platinum complexes with pyridine ligands [51,55], and more recently complexes with bis-imidazole based carrier ligands [56,57], which do not fulfil this requirement, have also been demonstrated to have antitumour activity. Although, the mechanism of action for such compounds could be partly or completely different from that of cisplatin and its analogues.

1.3.4 Inertness of the non-labile carrier ligand(s)

It is thought that the amine ligand(s) is preserved throughout, although in a few studies loss of the amine ligand during reaction with DNA has been demonstrated [58,59]. The inertness of carrier ligand(s) suggests that their structural features (rather than chemical properties) have an important role to play. In several studies it has been shown that the nature of the carrier ligand influences antitumour potency, cytotoxicity, solubility, distribution, and the tumour profiles of the complexes [28,55,60-62]. The analysis of these

SARs together with a review of quantitative structure-activity relationships (QSAR) studies of platinum complexes is covered in Chapter 5.

1.4 Significance of steric effects in metal complex interactions with nucleic acids

Ergo, the steric features of carrier ligands can influence interactions with nucleic acids and their constituents in a number of ways and may result in different biological outcomes both qualitatively (e.g. tumour profiles) and quantitatively (e.g. levels of antitumour potency, cytotoxicity and other indicators of biological activity). Steric variation of a carrier ligand can affect the kinetic [49,63-69] or thermodynamic [70,71] features of an interaction and can also determine the binding site [63]. Furthermore, steric properties of a carrier ligand can play an important role in determining the overall geometry of an adduct. This may be divided into local (in the immediate vicinity of the binding site) and remote geometric features.

1.4.1 The geometry/distortions in the immediate vicinity of the binding site

1.4.1.1 Dihedral angles

The dihedral angles (DA) between the coordinated nucleobases and the platinum coordination plane PtN_4, as well as those between the nucleobases themselves (Fig. 1.2) may be used to describe the geometry around the metal centre. These DA may be influenced by both intramolecular (e.g. between the coordinated nucleobases) and/or intermolecular (e.g. crystal packing forces) interactions. A *stereochemical convention* has been proposed [18,72] which allows a systematic comparison of the conformational features of platinum–DNA adducts and

an analysis of intra- and intercomplex steric effects in terms of these DA and also in the terms of the perpendicular displacement of the platinum atom from the plane of a nucleobase to which it is attached (ΔPt).
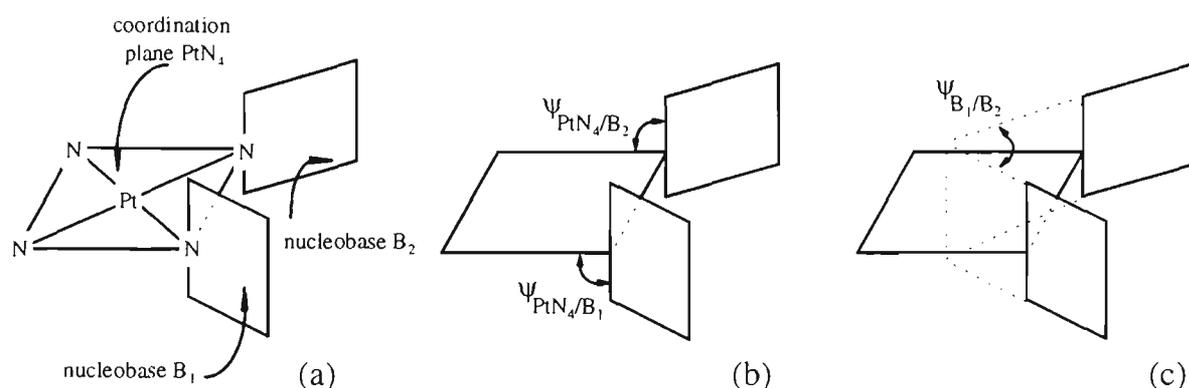
Figure 1.2. Dihedral angles. (a) Definition of complex planes; (b) Dihedral angles between the coordination plane and the nucleobase ligands; (c) Dihedral angle between the nucleobase ligands. In parts (a) and (b) arrows indicate the planes, whereas dihedral angle is calculated as an angle between the normals to these planes.

1.4.1.2 Relative orientation of cis-coordinated nucleobase ligands

Cis-coordinated nucleobase ligands may be oriented either "head-to-head" (HTH) or "head-to-tail" (HTT) [56,73-88] (Fig 1.3).
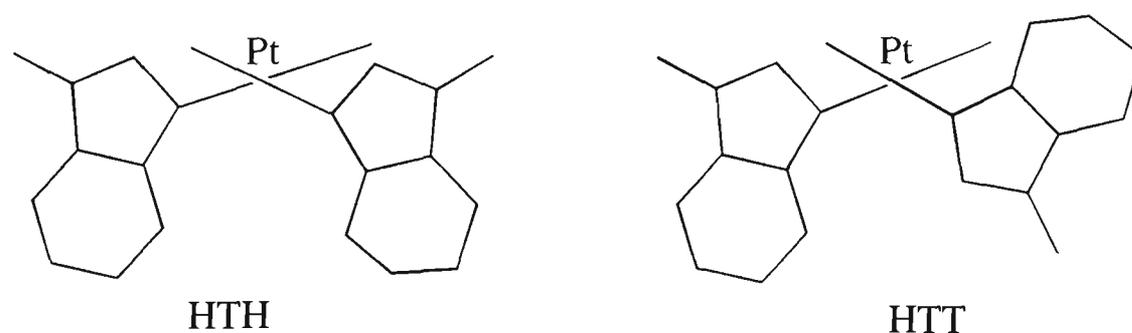
Figure 1.3. An illustration of rotational isomers in platinum complexes with two cis-coordinated nucleobase ligands for a bis(purine) system.

The existence of such rotamers is a consequence of a restricted rotation around the Pt-N bond which results from the steric constraints imposed on the adduct by the bulk of carrier ligand(s) and/or that of the exocyclic functional group(s) on the nucleoconstituents. Some authors introduce further classification to account for different types of intramolecular interaction. For example, "stepped head-to-head" and "head-to-side" have been used [31,89,90]. The number and nature of HTH and HTT isomers is sometimes further complicated by the chiral nature of the carrier ligand [86,87], of the coordinated nucleoconstituent [85] or of the metal center itself [85].

1.4.1.3 Nucleobase stacking

Distortions may also be reflected in departure from unstrained stacking patterns between adjacent nucleobases [1]. Nucleobase stacking-unstacking may be characterised by the base-overlap and also by values of the base/base dihedral angles.

1.4.1.4 Nucleotide conformation

The disruptions from "normality" in the nucleotide(side) conformation (that is, from the preferred conformation in the nucleic acid or nucleotide(side), unstrained by binding) are more pronounced in ternary complexes as compared to binary complexes [91]. This observation points to the significance of the steric influence exercised by the carrier ligand(s) on the nucleotide geometry, which is characterised by the following aspects (Fig. 1.4).
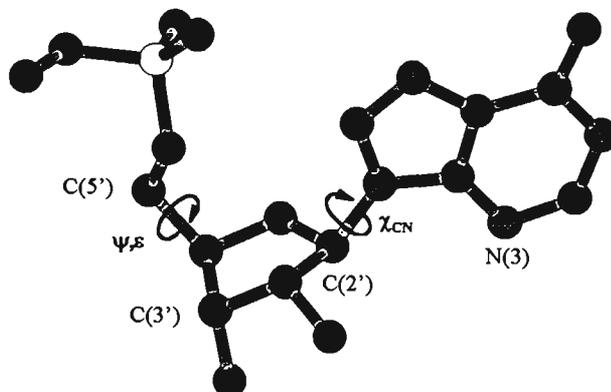
Figure 1.4. *An illustration of conformational features in the nucleotide unit for 5'-deoxyadenosine monophosphate. Hydrogen atoms are omitted for clarity.*

The orientation of the base moieties relative to the sugar ring is characterised by the value of χ_{CN} (glycosyl) torsion angle. Two broad conformational ranges may be defined depending upon whether the O(2) in pyrimidine bases or the N(3) in purine bases lies above the plane of the sugar or points away from it. Namely:

- χ_{CN} in the range of ~ 120° to ~ 300° - *syn*-conformation

- χ_{CN} in the range of ~ –60° to ~ 120° - *anti*-conformation

The *anti*-conformation is preferred in nucleotides [1].

The orientation of the phosphate moiety relative to the sugar ring is characterised by the value of ψ,ε-torsion angle, which may be divided into three categories:

- ψ,ε ~ 60° - gauche-gauche (*gg* or g^+) conformation

- ψ,ε ~ 180° - gauche-trans (*gt* or t^+) conformation

- ψ,ε ~ –60° - trans-gauche (*tg* or g^-) conformation

The gauche-gauche conformation is preferred in nucleotides [1].

The conformation of the sugar ring (sugar puckering) depends on the direction of the fifth atom "remaining" after the "most-planar" combination of four atoms of a 5-member sugar ring is selected. The displacement of this atom (usually C(3') or C(2')) relative to the above plane is considered with respect to the direction of C(5'):

- "+" displacement (same direction as C(5')) - *endo*-conformation
- "–" displacement (direction, opposite to C(5')) - *exo*-conformation

While RNA and ribonucleotides show preference for the C(3')-*endo* conformation, the C(2')-*endo* conformation appears to be preferred by DNA and *deoxy*ribonucleotides [1].

1.4.2 The overall geometric impact on the tertiary structure of DNA

The overall geometric impact of metal complex binding on the tertiary structure of DNA [23-25] is characterised by:

- the degree of unwinding of the duplex

- the extent of DNA bending or kinking

- shortening of the DNA chain.

It must be stressed at this point that the geometric features of an adduct, described above (in sections 1.4.1 and 1.4.2), may result not only from the steric requirements of the carrier ligand(s), but also from other aspects of the interaction. For example, the relative orientation of nucleobase ligands may depend significantly on the steric demands exerted by the nucleobase exocyclic groups [73,75,84], on the nature of counterions [78,81], on the presence of hydrogen bonding [73] and on the presence of other metal

centers [77]. Likewise, the actual location of the binding site(s) may be significantly affected by the steric demands of nucleobase exocyclic groups [72,92-94]. This aspect forms part of the present study [95][2] and is covered in Chapter 4.

Finally, overall variation of the structural speciation of the complex may influence complex–DNA interaction not only through the pathways described above, but may lead to a complete change of binding mode resulting in altered, rather than varied, biological outcome(s). Examples of such "extreme" structural variations include: incorporation of a second metal center into the complex [38,74,96-98], change of metal coordination geometry, e.g. octahedral geometry (usually associated with the Pt(IV) oxidation state) [29,37,52,53], *trans*-configuration of carrier ligands [50,51], and incorporation of π-systems in the carrier ligand(s) leading to DNA-intercalation [99,100]. Some authors even speculate that new clinically useful platinum compounds may not be direct analogues of cisplatin as defined by its basic structure and general structure-activity relationships [97].

1.5 Summary

Critical areas of investigation of steric effects which influence metal complex – nucleoconstituents interactions are as follows.

1.   Relationships between the steric features of the interacting moieties (e.g. nature of the carrier ligand) and the overall characteristic of the interaction itself (e.g. binding site, mode of coordination, stoichiometry, adduct stability).

---

[2] A reprint of this work is bound into the thesis

2.   Exploitation (by rational design) of the steric features of the interacting species in order to achieve a specific outcome such as site selectivity or a particular geometry.

3.   Quantification of steric effects in metal complex – nucleic acid interactions. Such quantification can be potentially used in order "to add Q to SPR or SAR". In other words, it is possible to quantitatively model the relationships between the structural (e.g. steric) features of the reacting species and the results of the interaction expressed in either chemical terms (e.g. equilibrium constant, reaction rate) or biological activity indicators (e.g. toxicity, potency, increased life span). The former of these two approaches models the quantitative structure-property relationships (QSPR) [101], the latter leads to the formulation of QSAR equations.

Thus, the above aspects virtually form a study loop between *the experiment*, producing information necessary for modelling, and *the modelling*, giving direction for further experimental work.

This thesis addresses all three aspects. Chapter 2 describes (i) design, synthesis and nucleoconstituent binding of a sterically restrictive metal complex and (ii) solution studies of the interactions between a series of metal complexes with varying degrees of carrier ligand steric bulk and nucleoconstituents. Chapters 3, 4 and 5 respectively report molecular modelling of Pt–nucleobase complexes, development of parameters to describe the steric effects in the interactions of platinum and other metal complexes with DNA constituents interactions, and investigation of QSARs of platinum complexes using the developed parameters.

1.6 References

1.   Saenger, W. *Principles of Nucleic Acids Structure;* Springer: New York, **1984**.

2.   Barton, J.K. In *Bioinorganic Chemistry;* Bertini, I.; Gray, H.B.; Lippard, S.J.; Valentine, J.S. (Eds.); University Science Books: Mill Valley, **1994**, pp. 455-504.

3.   Lippard, S.J. In *Bioinorganic Chemistry;* Bertini, I.; Gray, H.B.; Lippard, S.J.; Valentine, J.S. (Eds.); University Science Books: Mill Valley, **1994**, pp. 505-83.

4.   Anastassopoulou, J.; Theophanides, T. In *Bioinorganic Chemistry. An Inorganic Perspective of Life;* Kessissoglou, D.P. (Ed.); Kluwer: Dordrecht, **1995**, pp. 209-18.

5.   Otsuka, S. In *Metalloproteins. Chemical Properties and Biological Effects;* Otsuka, S.; Yamanaka, T. (Eds.); Elsevier: Amsterdam, **1988**, pp. 517-37.

6.   Pyle, A.M.; Barton, J.K. *Progr. Inorg. Chem.* **1990**, *38*, 414.

7.   Sakai, K.; Yamane, Y. In *Metalloproteins. Chemical Properties and Biological Effects;* Otsuka, S.; Yamanaka, T. (Eds.); Elsevier: Amsterdam, **1988**, pp. 509-16.

8.   Lippard, S.J. In *Bioinorganic Chemistry. An Inorganic Perspective of Life;* Kessissoglou, D.P. (Ed.); Kluwer: Dordrecht, **1995**, pp. 131-40.

9.   Rosenberg, B.; Van Camp, L.; Trosko, J.E.; Mansour, V.H. *Nature* **1969**, *222*, 385.

10.   Umapathy, P. *Coord. Chem. Rev.* **1989**, *95*, 129.

11.   Howe-Grant, M.E.; Lippard, S.J. In *Metal Complexes as Anticancer Agents;* Sigel, H. (Ed.); Marcel Dekker: New York, **1980**, pp. 63-125.

12.   Rosenberg, B.; Harder, H.C. *Int. J. Cancer* **1970**, *6*, 207.

13. Roberts, J.J.; Knox, R.J.; Friedlos, F.; Lydall, D.A. In *Biochemical Mechanisms of Platinum Antitumour Drugs;* McBrien, D.C.H.; Slater, T.F. (Eds.); IRL Press: Oxford, **1986**, 29.

14. Lippert, B. *Progr. Inorg. Chem.* **1989**, *37*, 1.

15. Fowless, A. In *Chemistry in an Australian Context Sourcebook;* Irvine, I. (Ed.); RACI: Parkville, **1989**, pp. 41-55.

16. Sherman, S.E.; Lippard, S.J. *Chem. Rev.* **1987**, *87*, 1153.

17. Roberts, J.J.; Pera, M.P.J. In *Platinum, Gold, and Other Metal Chemotherapeutic Agents;* Lippard, S.J. (Ed.); ACS: Washington D.C. **1983**, pp. 3-24.

18. Orbell, J.D.; Marzilli, L.G.; Kistenmacher, T.J. *J. Am. Chem. Soc.* **1981**, *103*, 5126.

19. Rosenberg, B. *Biochimie* **1978**, *60*, 859.

20. Sherman, S.E.; Gibson, D.; Wang, A.H.-J.; Lippard, S.J. *Science* **1985**, *230*, 412.

21. Bruhn, S.L.; Toney, J.H.; Lippard, S.J. *Progr. Inorg. Chem.* **1990**, *38*, 477.

22. Reedijk, J. *Pure Appl. Chem.* **1987**, *59*, 181.

23. Takahara, P.M.; Rosenzwieg, A.C.; Frederick, C.A.; Lippard, S.J. *Nature* **1995**, *377*, 649.

24. Huang, H.; Zhu, L.; Reid, B.R.; Drobny, G.P.; Hopkins, P.B. *Science* **1995**, *270*, 1842.

25. Pil, P.M.; Lippard, S.J. *Science* **1992**, *256*, 234.

26. Sen', V.D.; Golubev, V.A.; Volkova, L.M.; Konovalova, N.P. *J. Inorg. Biochem.* **1996**, *64*, 69.

27. Rose, W.C.; Schurig, J.E.; Huftalen, J.B.; Bradner, W.T. *Cancer Treat. Rep.* **1982**, *66*, 135.

28. Cleare, M.J.; Hoeschele, J.D. *Bioinorg. Chem.* **1973**, *2*, 187.

29. Cleare, M.J. *Coord. Chem. Rev.* **1974**, *12*, 349.

30. van Garderen, C.J.; Cornelis, A.; Reedijk, J. *Inorg. Chem.* **1990**, *29*, 1481.

31. Kline, T.P.; Marzilli, L.G.; Live, D. *J. Am. Chem. Soc.* **1989**, *111*, 7057.

32. Bednarski, P.J. *Biochem. Pharm.* **1992**, *43*, 2609.

33. Bitha, P.; Carvajal, S.G.; Citarella, R.V.; Delos Santos, E.F.; Durr, F.E.; Hlavka, J.J.; Lang, S.A.; Linsday, H.L.; Thomas, J.P.; Wallace, R.E.; Lin, Y. *J. Med. Chem.* **1989**, *32*, 2063.

34. Pasini, A.; Zinuno, F. *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 615.

35. Cleare, M.J. In *Structure-Activity Relationships of Anti-Tumour Agents;* Reinhoudt, D.N.; Connors, T.A.; Pinedo, H.M.; van de Poll, K.W. (Eds.); Martinus Nijhoff: The Hague, **1983**, pp. 59-92.

36. Farrell, N. *Cancer Invest.* **1993**, *11*, 578.

37. Kelland, L.R.; Murrer, B.A.; Abel, G.; Giandomenico, C.M.; Mistry, P.; Harrap, K.R. *Cancer Res.* **1992**, *52*, 822.

38. Farrell, N.; Qu, Y.; Hacker, M.P. *J. Med. Chem.* **1990**, *33*, 2179.

39. Twentyman, P.R.; Wright, K.A.; Mistry, P.; Kelland, L.R.; Murrer, B.A. *Cancer Res.* **1992**, *52*, 5674.

40. De Pooter, C.M.; Van Oosterman, A.T.; de Bruijn, E.A.; Scalliet, P.G.; Maes, R.A. *Biochem. Pharm.* **1996**, *51*, 629.

41. Dress, M.; Dengler, W.M.; Fiebig, H.H. *Eur. J. Cancer.* **1995**, *31*, 356.

42. Vickery, K.; Bonin, A.M.; Fenton, R.R.; O'Mara, S.; Russell, P.J.; Webster, L.K.; Hambley, T.W. *J. Med. Chem.* **1993**, *36*, 3663.

43. Khokhar, A.R.; Al-Baker, S.; Perez-Soler, R. *Anti-Cancer Drug Des.* **1988**, *3*, 177.

44. Colombo, A.; Giola, R.D.; Pasini, A.; Dasdia, T.; Zinuno, F. *Inorg. Chim. Acta.* **1986**, *125*, L1.

45. Gibbons, G.R.; Page, J.D.; Mauldin, S.K.; Husain, I.; Chaney, S.G. *Cancer Res.* **1990**, *50*, 6497.

46. Yoshida, M.; Khokhar, A.R.; Siddik, Z.H. *Anti-Cancer Drug Des.* **1994**, *9*, 425.

47. Lempers, E.L.M.; Reedijk, J. *Adv. Inorg. Chem.* **1991**, *37*, 175.

48. Rosenberg, B. In *Cisplatin: Current Status and New Developments;* Prestayko, A.W.; Crooke, S.T.; Carter, S.K. (Eds.); Academic Press: New York, **1980**, pp. 9-20.

49. Brown, D.B.; Khokhar, A.R.; Hacker, M.P.; McCormack, J.J.; Newman, R.A. In *Platinum, Gold, and Other Metal Chemotherapeutic Agents;* Lippard, S.J. (Ed.); ACS: Washington D.C. **1983**, pp. 3-24.

50. Weiss, R.B.; Christian, M.C. *Drugs* **1993**, *46*, 360.

51. Zou, Y.; Van Houten, B.; Farrell, N. *Biochemistry* **1993**, *32*, 9632.

52. Yoshida, M.; Khokhar, A.R.; Siddik, Z.H. *Cancer Res.* **1994**, *54*, 4691.

53. Anderson, W.K.; Kasliwal, R.; Houston, D.M.; Wang, Y.; Narayanan, V.L.; Haugwitz, R.D.; Plowman, J. *J. Med. Chem.* **1995**, *38*, 3789.

54. Reedijk, J. *Inorg. Chim. Acta.* **1992**, *198-200*, 873.

55. Meischen, S.J.; Gale, G.R.; Lake, L.M.; Frangakis, C.J.; Rosenblum, M.G.; Walker, E.M.; Atkins, L.M.; Smith, A.B. *JNCI* **1976**, *57*, 841.

56. Grehl, M.; Krebs, B. *Inorg. Chem.* **1994**, *33*, 3877.

57. Bloemink, M.J.; Engelking, H.; Karentzopoulos, S.; Krebs, B.; Reedijk, J. *Inorg. Chem.* **1996**, *35*, 619.

58. Wherland, S.; Deutsch, E.; Eliason, J.; Sigler, P.B. *Biochem. Biophys. Res. Comm.* **1973**, *54*, 662.

59.   Lippard, S.J. *Acc. Chem. Res.* **1978**, *11*, 211.

60.   Braddock, P.D.; Connors, T.A.; Jones, M.; Khokhar, A.R.; Melzack, D.H.; Tobe, M.L. *Chem.-Biol. Interactions* **1975**, *11*, 145.

61.   Das Sarma, B.; Daley, S.K.; Elespuru, R.K. *Chem.-Biol. Interactions* **1983**, *46*, 219.

62.   Beaumont, K.P.; McAuliffe, C.A.; Cleare, M.J. *Chem.-Biol. Interactions* **1976**, *14*, 179.

63.   Arpalahti, J.; Lehikoinen, P. *Inorg. Chim. Acta.* **1989**, *159*, 115.

64.   Shoukry, M.; Hohmann, H.; van Eldik, R. *Inorg. Chim. Acta.* **1992**, *198-200*, 187.

65.   Shih, K.-C.; Herber, R.H. *Inorg. Chem.* **1992**, *31*, 5444.

66.   Gust, R.; Schonenberger, H.; Kritzenberger, J.; Range, K.-J.; Klement, U.; Burgmeister, T. *Inorg. Chem.* **1994**, *32*, 5939.

67.   Gibson, D.; Arvanitis, G.M.; Berman, H.M. *Inorg. Chim. Acta.* **1994**, *218*, 11.

68.   Inagaki, K.; Dijt, F.J.; Lempers, E.L.M.; Reedijk, J. *Inorg. Chem.* **1988**, *27*, 382.

69.   Romeo, R.; Arena, G.; Scolaro, L.M.; Plutino, M.R.; Bruno, G.; Nicolo, F. *Inorg. Chem.* **1994**, *34*, 4029.

70.   Butour, J.-L.; Alvinerie, P.; Souchard, J.-P.; Colson, P.; Houssier, C.; Johnson, N.P. *Eur. J. Biochem.* **1991**, *202*, 975.

71.   Sigel, H. In *Mixed-Ligand Complexes;* Sigel, H. (Ed.); Marcel Dekker: New York, **1973**, pp. 63-125.

72.   Kistenmacher, T.J.; Orbell, J.D.; Marzilli, L.G. In *Platinum, Gold and other Metal Chemotherapeutic Agents;* Lippard, S.J. (Ed.); ACS: Washington D.C. **1983**, pp. 191-208.

73.   Hambley, T.W. *Inorg. Chem.* **1988**, *27*, 1073.

74. Grehl, M.; Krebs, B. *Inorg. Chem.* **1994**, *33*, 3877.

75. Schroder, G.; Kozelka, J.; Sabat, M.; Fouchet, M.-H.; Beyerle-Pfnur, R.; Lippert, B. *Inorg. Chem.* **1996**, *35*, 1647.

76. Preut, H.; Frommer, G.; Lippert, B. *Acta Cryst., C* **1991**, *47*, 852.

77. Micklitz, W.; Sheldrick, W.S.; Lippert, B. *Inorg. Chem.* **1990**, *29*, 211.

78. Micklitz, W.; Lippert, B.; Lianza, F.; Albinati, A. *Inorg. Chim. Acta.* **1994**, *227*, 5.

79. Baker, A.T.; Crass, J.; Kok, G.B.; Orbell, J.D.; Yuriev, E. *Inorg. Chim. Acta.* **1993**, *214*, 169.

80. Orbell, J.D.; Taylor, M.R.; Birch, S.L.; Lawton, S.E.; Vilkins, L.M.; Keefe, L.J. *Inorg. Chim. Acta.* **1988**, *152*, 125.

81. Schollhorn, H.; Randashl-Sieber, G.; Muller, G.; Thewalt, U.; Lippert, B. *J. Am. Chem. Soc.* **1985**, *107*, 5932.

82. Hambley, T.W.; Ling, E.C.H.; Messerle, B.A. *Inorg. Chem.* **1996**, *35*, 4663.

83. Pasini, A.; De Giacomo, L. *Inorg. Chim. Acta.* **1996**, *248*, 225.

84. Frommer, G.; Lippert, B. *Inorg. Chem.* **1990**, *29*, 3259.

85. Cramer, R.E.; Dahlstrom, P.L. *J. Am. Chem. Soc.* **1979**, *101*, 3679.

86. Xu, Y.; Natile, G.; Intini, F.P.; Marzilli, L.G. *J. Am. Chem. Soc.* **1990**, *112*, 8177.

87. Kiser, D.; Intini, F.P.; Xu, Y.; Natile, G.; Marzilli, L.G. *Inorg. Chem.* **1994**, *33*, 4149.

88. Marselis, A.T.M.; Korte, H.-J.; Krebs, B.; Reedijk, J. *Inorg. Chem.* **1982**, *21*, 4059.

89. Yohannes, P.G.; Zon, G.; Doetsch, P.W.; Marzilli, L.G. *J. Am. Chem. Soc.* **1993**, *115*, 5105.

90. Iwamoto, M.; Mukundan, S.; Marzilli, L.G. *J. Am. Chem. Soc.* **1994**, *116*, 6238.

91. Aoki, K. *Metal Ions in Biological Systems* **1996**, *32*, 91.

92. Kinjo, Y.; Tribolet, R.; Corfu, N.A.; Sigel, H. *Inorg. Chem.* **1989**, *28*, 1480.

93. Kinjo, Y.; Ji, L.-N.; Corfu, N.A.; Sigel, H. *Inorg. Chem.* **1992**, *31*, 5588.

94. Reily, M.D.; Marzilli, L.G. *J. Am. Chem. Soc.* **1986**, *108*, 6785.

95. Yuriev, E.; Orbell, J.D. *Inorg. Chem.* **1996**, *35*, 7914.

96. Frommer, G.; Lianza, F.; Albinati, A.; Lippert, B. *Inorg. Chem.* **1992**, *31*, 2434.

97. Qu, Y.; Bloemink, M.J.; Reedijk, J.; Hambley, T.W.; Farrell, N. *J. Am. Chem. Soc.* **1996**, *118*, 9307.

98. Hopp, M.; Erxleben, A.; Rombeck, I.; Lippert, B. *Inorg. Chem.* **1996**, *35*, 397.

99. Gibson, D.; Gean, K.; Najajareh, Y.; Mansur, N.; Gliko, I.; Binyamin, I.; Ramu, A.; Katzhendler, J. *EUROBIC II* **1994**, 142.

100. Romeo, R.; Arena, G.; Scolaro, L.M.; Pasternack, R.F. *Inorg. Chem.* **1995**, *34*, 2992.

101. Katritzky, A.R.; Lobanov, V.S.; Karelson, M. *Chem. Soc. Rev.* **1995**, *24*, 279.

Chapter 2

Solution and solid state studies of metal complex – nucleic acid constituent interactions

2.1 INTRODUCTION

2.1.1 Spectrophotometric studies

UV-visible spectroscopy (UV-Vis) has been extensively applied to the investigation of metal species interactions with nucleic acids [1,2]. This technique allows the detection of binding through the observation of spectral changes, and it is also valuable for determination of formation constants. More specifically, in the area of platinum – nucleic acid research, such applications as pH-dependent [3-11] and ratio-dependent [6,8,12-19] UV-Vis spectroscopy and difference spectroscopy techniques [3-9,13,14,16-23] (discussed below) have been employed in order to obtain information about the location of binding sites [4,5,9,24], stoichiometry of adducts [25], their acidity [10-24] and stability [13-16,19-21], reaction rates [12,20,26,27], hydrogen bonding [27], and nucleobase stacking [19]. These implementations also allow a study of the steric effects involved in metal complex – nucleoconstituent interactions including an investigation of the influence of steric effects on the thermodynamic and kinetic characteristics of such interactions [12].

Spectrophotometry is an important tool for studying chemical equilibria since the measurements are made without perturbing the equilibrium of the system under examination. One of the most widely used approaches within this method is *Difference UV-Vis Spectroscopy* [17,28]. This method is based on weighted subtraction of the absorbance of the original components from the absorbance of the reaction mixture resulting in a *corrected absorbance* [2,29]. The advantages of this technique are as follows:

- low concentration solutions may be used, reducing the probability of intermolecular interactions, aggregation, and precipitation;

- the relative ion concentration range can be extended by several orders of magnitude;

- difference spectra are more sensitive than simple (observed) UV-Vis spectra because common features in two (or more) spectra cancel, and subtle differences due to variation in concentration or pH are accentuated.

2.1.1.1 Obtaining the stoichiometry of complexes

In the *Continuous Variation* or *Job's* method [28-30] the absorbance is measured of a series of samples, in which ligand-to-metal concentration ratios (molar ratios) are varied and their combined total concentration is held constant. If all components absorb at the monitored wavelength, the *corrected absorbance* is used. If the only absorbing species is the product of the reaction then the observed absorbance is used directly for the stoichiometry determination. The absorbance, observed or corrected (whichever is appropriate), is plotted against the mole fraction F_m of one of the components. An extreme absorbance is reached when the mole fraction F_m in the sample equals the actual mole fraction of this component in the metal complex.

It is to be noted also that the Job's method is a technique for obtaining the stoichiometry of the *predominant* complex. That is, if several complexes coexist in solution in comparable proportions the extreme of the plot depends on the chosen wavelengths. One more disadvantage of the method

is that it may only be useful for complexes having small molar ratios of the components. While it is not difficult to distinguish between the values of ligand mole fraction for complexes **ML** (0.5), **ML$_2$** (0.67) and **ML$_3$** (0.75), it becomes increasingly hard for higher complexes, e.g. **ML$_4$** and **ML$_5$** (0.80 and 0.83 respectively).

In the ***Molar Ratio*** method [28,31] the absorbance is measured for a series of solutions with varied molar ratios and with constant total concentration for one of the components. The absorbance, observed or corrected (whichever is appropriate), is plotted against the molar ratio. The absorbance reaches a break point at the composition corresponding to the stoichiometry of the complex.

The ***Slope ratio*** method [28,32], used mainly for weak complexes, is based on the assumption that if one component is in large excess, then the formation reaction is forced to completion, the dissociation of the complex is negligible, and its concentration is defined by that of the limiting component. In this method the absorbance for two series of solutions is measured; firstly, with a constant large excess of the ligand and varying concentrations of the metal and, secondly, with a constant large excess of the metal and varying concentrations of the ligand. For both series, plots of absorbance vs concentration of the limiting component are constructed and the ratio of the slopes of these plots is equal to the molar ratio in the complex. A variation of the ***Slope Ratio*** method is the ***Mollard*** method [28], in which only two solutions are prepared, each with a large excess of either of the components. In this case the molar ratio of the complex is equal to $c_L A_M / c_M A_L$, where c_M and A_M are concentration and absorbance respectively of the solution with limiting concentration of the metal

component, and c_L and A_L are concentration and absorbance respectively of the solution with limiting concentration of the ligand.

Isosbestic points, which are the points of common absorbance for two spectra or a family of spectra, are commonly exploited in stoichiometry studies [32,33] and often applied to platinum – nucleic acid systems [5,7,9,13-15,21,23,24]. The exhibition of an isosbestic point(s) is often used as complimentary evidence in establishing the number of species or, rather, stoichiometric states in solution [33]. Usually if a family of spectra for solutions differing in pH or component(s) concentrations passes through a common point it means that only one reaction is taking place, typically leading to a conclusion that one-to-one complex formation is occurring.

2.1.1.2 Calculation of formation constants

Miscellaneous approaches for obtaining formation (stability) constants from spectrophotometric studies have been reviewed [33]. The *apparent*, or *conditional*, formation constants have been introduced by Ringbom [34]. The concept of conditional constants was introduced in order to simplify the calculations. Thus, the analytical target is the degree of completeness of the main reaction, not the forms and quantities of unreacted species. Generally, a complex ML_n is formed and the conditional constant is defined as $K = [ML_n]/[M][L]^n$ where [M] represents the concentration not only of a metal ion, but also of all metal-containing species that have not reacted with the ligand. [L] is defined in the same manner for the ligand itself. The term "conditional" means that such a "constant" is constant only under the particular experimental conditions (pH, concentration range, etc.). Henceforce the

term "constant" will be used instead of "conditional constant".

The spectrophotometric method for the determination of formation constants is based on the basic definition of the equilibrium constant K:

$$K = \frac{c}{(a - c)(b - nc)^n} \qquad (2.1)$$

and Beer's law [1]:

$$A_0 = \varepsilon_a a + \varepsilon_b b \qquad (2.2)$$

$$A = \varepsilon_a(a - c) + \varepsilon_b(b - nc) + \varepsilon_c c \qquad (2.3)$$

where a and b are the initial concentrations of components M and L respectively, and c is the concentration of the complex ML_n at equilibrium. A_0 is the absorbance of a system without any interaction (which corresponds to the sum of the absorbances of metal compound and ligand solutions at the same concentrations as in the mixture), A is the absorbance of a system where the complex ML_n is formed, ε_a and ε_b are the extinction coefficients of the components and ε_c is the extinction coefficient of the complex ML_n.

From eqns 2.2 and 2.3 the expression for the corrected absorbance is:

$$\Delta A = A - A_0 = c(\varepsilon_c - \varepsilon_a - n\varepsilon_b) = \varepsilon^* c \qquad (2.4)$$

---

[1] Pathlength of the cell l, present in Beers's law, may be omitted if it is equal to 1 cm.

For 1:1 complexation (n = 1), by elimination of the complex concentration c between eqns 2.1 and 2.4, several expressions for the formation constant K may be derived, one of the *general* forms being [35]:

$$\frac{1}{K} = \frac{\Delta A}{\varepsilon^{\star}} - a - b + \frac{ab\varepsilon^{\star}}{\Delta A} \tag{2.5}$$

These equations are rather laborious to solve analytically. Various approaches which have been developed for their treatment are given below.

1. Calculation of formation constants for **weak complex(es)**. In this case, the equilibrium concentration of a complex is much smaller than that of components **M** and **L** and the formation constant for complex **ML** reduces to [36]:

$$\frac{1}{K\varepsilon^{\star}} = \frac{ab}{\Delta A} - \frac{a + b}{\varepsilon^{\star}} \tag{2.6}$$

This expression does not yield the values of the formation constant directly and requires further treatment, such as iteration or the use of favourable experimental conditions.

2. The **use of favourable experimental conditions** may be achieved by carrying out an experiment with a large excess of one of the components so that the concentration of the other component and that of the complex become insignificant. For instance, for 1:1 complexation (n = 1) and $a \gg b$, the general eqn 2.5 is reduced to the Benesi-Hildebrand equation [37]:

$$\frac{1}{K\varepsilon^{\star}} = \frac{ab}{\Delta A} - \frac{a}{\varepsilon^{\star}} \tag{2.7}$$

Eqn 2.7 yields the value of the formation constant directly and gives the

extinction coefficient of a complex from the values of the slope and the intercept of the *double-reciprocal plots*. The Benesi-Hildebrand method has been used for the investigation of platinum-am(m)ine complexes interactions with nucleic acids and their constituents [19].

3. The *iterative techniques* have been developed to treat the absorbance data without any assumptions and omissions. These include: the thorough but very laborious method of *Drago and Rose* [35], simple and accurate approaches of the *improved Benesi-Hildebrand* method [36] and the *improved Rossotti-Rossotti* method [38]. The *improved Rossotti-Rossotti* method deals with the 1:1 case and uses the following form of the general eqn **2.5**:

$$\frac{ab}{\Delta A} = \frac{1}{\varepsilon^* K} + \frac{a + b - c}{\varepsilon^*} \tag{2.8}$$

The values of c are unknown but may be approximated iteratively. The equilibrium constant and ε^* may then be calculated from the slope and the intercept of the final regression line.

2.1.2 Nuclear Magnetic Resonance (NMR) studies

Nuclear Magnetic Resonance (NMR) spectroscopy is one of the most extensively used methods in the study of metal ion/complex interactions with nucleic acids and their constituents in solution [39]. While ^{1}H NMR remains the most widely used single technique, ^{195}Pt [40-42], ^{13}C [40-45], ^{15}N [43,46-48] and ^{31}P [44,48-51] are also potentially suitable for such studies. In the field of platinum – nucleic acid research, NMR has been used for structure elucidation [52,53], for investigation of conformational

changes in single- and double-stranded oligonucleotides upon platination [49,50], and in kinetic [43,44] and thermodynamic [54-56] studies. In particular, monitoring chemical shifts of specific nuclei is capable of giving information about: the location of a binding site [47,50,54,56], the stoichiometry of an adduct [57,58], the hydrogen-bonding [27,51], and the nucleobase stacking [49,50]. Thus, correlating the data obtained from NMR studies with structural information (e.g. for a series of platinum complexes with varied carrier ligands) provides a better understanding of how various structural features influence the interaction.

^{1}H NMR evidence for platinum–nucleobase binding includes a downfield chemical shift of ~ 0.2–1 ppm for proton(s) adjacent to the binding site, an absence of shift due to deprotonation at the binding site, and a characteristic [3]J ([195]Pt–[1]H) coupling in the form of H8 signal "satellites" at low field [59].

The calculation of formation constants from NMR data is based on the same basic principles described for spectrophotometry. For example, the difference in chemical shift between free and bound ligands, ΔH_p, may be related to a component mole ratio and, consequently, to the equilibrium constant, thus:

$$\Delta H_p = (1 + P_m^{-1}) \times \Delta H_a$$

where ΔH_a is the chemical shift difference between a ligand solution and a ligand/metal mixture and P_m is the mole ratio of bound and unbound ligands [60].

Such a method has been applied to the investigation of equilibria between

metal salts and imidazole, purine and pyrimidine derivatives [61]. These workers used the following expression for the observed chemical shift ν:

$$\nu = \frac{B_0 - x(MB_x)}{B_0}\nu_f + \frac{x(MB_x)}{B_0}\nu_c = \nu_f + \frac{x(MB_x)}{B_0}(\nu_c - \nu_f)$$

where ν_f and ν_c are characteristic frequencies of free and complexed ligands, respectively, B_0 is the initial concentration of a ligand, and (MB_x) is the equilibrium concentration of a complex MB_x.

Miscellaneous approaches, similar to those described for UV-Vis, and involving approximations, favourable experimental conditions or iterative treatment of data, have been developed for the quantification of formation constants from NMR data (on basis of equations analogous to the above) [45,54,56,60-64].

2.1.3 Other solution studies

Further spectroscopic techniques useful for the elucidation of structural information of metal – nucleic acid complexes are *Infra-Red* (IR) and *Atomic Absorption* (AA) spectroscopy. In the field of IR-spectroscopy such developments as Fourier Transform methods [65], RAMAN spectroscopy [45,48,57] and extension into the far-infrared region [66] have useful application to coordination compounds in general, and to metal – nucleic acid complexes in particular [67-70]. In the area of platinum research, AA spectroscopy has been used to measure the distribution of platinum in biological fluids [71,72] as well as being applied to study of platinum coordination reactions [73,74].

Apart from spectroscopic techniques, electrochemical methods are also valuable. One very powerful method for the determination of stability constants is *potentiometry*. This technique has been extensively used by Sigel's group to compile an impressive database on the stabilities of a broad range of metal/nucleoconstituent complexes [1].

Conductivity methods have been effectively used for studies of complex formation [75]. When a metal ion forms a complex with a ligand(s) the number of charged species usually decreases and the mobility characteristics of the equilibrium mixture change. Therefore conductivity may be used as a sensor for such changes. One frequently used approach is based on *difference conductivity*, i.e. the weighted subtraction of the conductivity of the original components, κ_{sum}, from the conductivity of the reaction mixture, κ_{mix}. Thus for a two component system:

$$\Delta\kappa = \kappa_{mix} - \kappa_{sum} = \kappa_{mix} - (a \times \kappa_1 - (1 - a) \times \kappa_2)$$

where a is the mole fraction of one of the components, and κ_1 and κ_2 are the specific conductivities of two components respectively. Similar to UV-Vis spectroscopy (section 2.1.1), $\Delta\kappa$ is called *corrected conductivity*, because the conductivity of the reaction mixture is "corrected" for the conductivities of the reaction components. Usually, the experimental data is presented in one of two ways. It may be presented as an *absolute* difference $\Delta\kappa$ (as expressed above) [76,77] or as a *percentage* difference $\%\Delta\kappa$ [78-81]:

$$\%\Delta\kappa = \Delta\kappa \times 100 / \kappa_{mix}$$

This could be either percentage decrease [80] or percentage increase [81] depending on the system under study.

To establish stoichiometry of complex formation by the above technique the method of continuous variation, similar to that described for UV-Vis spectroscopy (section 2.1.1.1) is frequently used. It has been applied to such systems as cupric amino nitrates and sulfates [78,81], chloromercuric acids [79], amminecobalt sulfates [76], the reactions of anhydrous formic acid with amides [82], the interactions of Hg^{2+} and Ag^+ ions with the $K(CN)_6^{5-}$ species [83], and charge-transfer complexes [84,85]. The only references found in the literature with respect to the use of conductivity for the study of heavy metal – nucleic acid interactions are for Cu^{2+} – guanosine derivatives systems [86] and a Ni – ATP system [87].

Miscellaneous other methods have been applied to monitoring of metal binding to nucleic acids. Namely, mass-spectroscopy [88], circular dichroism [19,48,49], calorimetry [89], electrophoresis [90,91], melting behaviour [8], electron microscopy [91], and enzymatic techniques [11].

2.1.4 Solid state studies

X-Ray crystallography is complementary to the methods described above. The limitations of crystallography are well known. For example, its results usually represent only one product, crystallised from solution. Solid structures may differ significantly from solution structures. Although this problem is particularly relevant for biological systems with their diversity of equilibrium states, these systems are exceptional in that the interactions on the surface of or within macromolecules could be considered as intermediate between solution and solid phase. In such circumstances, the information obtained from crystallographic studies is not necessarily less relevant, than that from solution experiments. Numerous crystal structures

have been solved for Pt-nucleoconstituent adducts [92]. X-Ray studies of these systems offer a wealth of information regarding binding sites and binding patterns, conformational distributions, hydrogen bonding interactions, etc.

The above mentioned limitations characteristic of crystallography can be ameliorated by use in conjunction with other techniques for obtaining structural information, such as those described previously. In addition, besides the experimental methods, theoretical calculations are being increasingly applied to metal – nucleic acid interactions in general and platinum systems in particular. These are discussed in Chapter 3.

2.2 METHODS

2.2.1 Materials and equipment

Nucleoconstituents and K_2MCl_4 salts (M = Pt,Pd) were purchased from Sigma. These were of the highest grade and were used without further purification. D_2O and DMSO-d_6 were obtained from Cambridge Isotope Laboratories, NaOD and DCl were obtained from Sigma. AAS standards were purchased from Aldrich and BDH. Common chemicals were obtained from other supply houses (Table 2.1). Ligands 1,2-bis-(pyridin-2-yl)ethane (bpe) and 1,2-bis(6-methylpyridin-2-yl)ethane (bmpe) were synthesised and supplied by collaborators[2].

Table 2.1. Chemicals.

Compound	Supplier (grade)
potassium iodide	Ajax (lab)
celite	Aldrich (anal)
tetramethylsilane (TMS)	Aldrich (NMR)
sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS)	Aldrich (NMR)
ammonia (28% solution)	Ajax (anal)
ethylenediamine (en)	May&Baker (lab)
tetramethylethylenediamine (TMED)	Sigma (anal)
1,2-di-(4-methylpiperazine) (bispep)	IDT
silver nitrate	BDH (anal)
sodium perchlorate	BDH (lab)
1,4-dimethylpiperazine (dmp)	Aldrich (anal)

---

[2] Dr. G.B. Kok, Monash University, Australia.

UV-Vis spectra were recorded at 25.0 ± 0.1 °C on a Cary1 UV-Vis Spectrophotometer equipped with the Cary Temperature Controller. NMR spectra were recorded on a 60 MHz HITACHI R-1200 NMR Spectrometer, a Brucker DPX300 NMR Spectrometer and a Brucker AMX300 NMR Spectrometer in 507-PP and MG 5mm NMR tubes from Wilmad Glass Co. Platinum and palladium concentration was measured using a Varian SpectrAA-400 Atomic Absorption Spectrometer equipped with a GTA-96 Graphite Tube Atomizer. pH and Cl^- concentration were monitored by electrodes supplied by Orion Research Ltd attached to an Expandable IonAnalyzer EA 940 also supplied by Orion Research Ltd. To achieve the most precise sample handling an EDP PLUS Motorized Microliter electronic pipette (model EP-250) from Rainin Instr. Co. was used. Conductivity was measured with the use of a digital T.P.S. direct reading conductivity meter (model 2100) featuring RAVTC (precision automatic variable temperature compensation). Elemental analyses were carried out by National Analytical Laboratories Pty Ltd.

2.2.2 Synthesis

Synthesis of the complex of general formula cis-[Pd(carrier ligand)Cl$_2$].
General procedure: a solution of carrier ligand in an appropriate solvent is added to a saturated solution of K_2PdCl_4 in deionised water with stirring at room temperature (1:1 molar ratio). A yellow or orange precipitate usually forms immediately. The resulting mixture is then stirred overnight, the precipitate is filtered, washed with hot water, ice-cold ethanol and ether, and dried in a desiccator over silica gel. Yields, appropriate solvents and comments are shown in Table 2.2a.

Table 2.2a. Synthesis of the complex of general formula *cis*-[Pd(carrier ligand)Cl₂].

carrier ligand	solvent	yield	comments	microanalysis (%)	
				found	calc.
en	H_2O	83%	synthesised by the alternative method [93]	–	
TMED	H_2O	91%	wash with *cold* water	C: 24.8 H: 5.5 N: 9.0	C: 24.5 H: 5.5 N: 9.5
bpe	H_2O	87%	–	C: 36.5 H: 3.3 N: 6.9	C: 39.9 H: 3.3 N: 7.7
bmpe	acetone	92%	an alternative method is available [94]	C: 43.0 H: 4.3 N: 7.3	C: 43.2 H: 4.14 N: 7.19
bispep	H_2O	91%	centrifuge after the precipitation, wash with water and acetone [95]	C: 24.6 H: 4.8 N: 9.4	C: 24.8 H: 4.5 N: 9.6
dmp	H_2O^a	89%	wash with *cold* water	C: 24.7 H: 4.8 N: 9.6	C: 24.8 H: 4.8 N: 9.2

ᵃLigand is a liquid at room temperature, but adding it neat decreases the yield.

Synthesis of the complex of general formula cis-[Pt(carrier ligand)I₂].
Attempts have been carried out to synthesise platinum complexes, analogous to the above palladium complexes. Although they are not as useful for equilibria studies (see section 2.3.2.1), such complexes are required for the investigation of biological activity, [195]Pt NMR and solid state structural studies. The general procedure has been based on that of Dhara [96] for the synthesis of cisplatin, which is also generally useful for preparing cisplatin analogues [97]. It involves reacting K_2PtI_4 (formed *in*

situ by adding KI to a solution of K$_2$PtCl$_4$) with the ligand solution. For the carrier ligands NH$_3$, en and TMED the procedure is straightforward. The synthesis of *cis*-[Pt(bpe)I$_2$] has been carried out with slight variations according to Ref. [98]. The synthesis of *cis*-[Pt(bmpe)I$_2$] and *cis*-[Pt$_2$(bispep)I$_4$] requires very slow addition *of* the complex solution *to* the ligand solution with vigorous stirring to prevent polymerisation (by ensuring an excess of the ligand). The complex *cis*-[Pt(dmp)Cl$_2$] has been quantitatively synthesised by the method of Watson and Mann [95] as described for the palladium complex above and identified by t_{melt}. Yields and microanalysis results are presented in Table 2.2b.

Table 2.2b. Synthesis of the complex of general formula *cis*-[Pt(carrier ligand)I$_2$].

carrier ligand	yield	microanalysis (%)	
		found	calc.
NH$_3$	94%	–	–
en	95%	C: 5.0	C: 4.7
		H: 1.5	H: 1.6
		N: 4.0	N: 5.5
TMED	95%	C: 12.9	C: 8.9
		H: 2.9	H: 3.0
		N: 4.8	N: 5.2
bpe	89%	C: 22.4	C: 22.8
		H: 1.9	H: 1.9
		N: 4.1	N: 4.4
bmpe	31%	C: 24.0	C:25.4
		H: 2.6	H: 2.4
		N: 3.8	N: 4.2
bispep	75%	C:10.7	C: 12.8
		H: 2.2	H: 2.3
		N: 3.7	N: 5.0

Preparation of solvated species cis-[M(carrier ligand)(solvent)$_2$](NO$_3$)$_2$, where the "solvent" is H_2O in preparative and spectrophotometric studies, D_2O and DMSO-d_6 in NMR studies and M is platinum or palladium. The general procedure involves addition of a $AgNO_3$ solution to a stirred suspension of *cis*-[M(carrier ligand)Hal$_2$] (Hal = Cl or I) in a dark bottle (~ 1:2 complex-to-silver molar ratio; slightly less than the required two equivalents of Ag^+ are added to avoid Ag^+ contamination). The mixture is then stirred at ~ 60°C for 2h. The AgCl is removed by filtration through celite and a sintered glass filter producing a pale yellow filtrate which is assumed to contain the solvated species. The concentration of a metal complex in the filtrate is determined using atomic absorption spectroscopy as described below, and the yield is calculated based on the initial weight of $AgNO_3$ as a limiting reagent:

	NH$_3$	en	TMED	bpe	bmpe	dmp	bispep
Pt	86%	100%	53%	–	–	–	–
Pd	–	100%	100%	76%	100%	100%	44%

Since Cl^- greatly affects the equilibrium, solutions were also tested for the presence of chloride ion after filtering through celite.

2.2.3 AAS determination of metal concentration in solvated samples

Although the removal of chlorine ions from a complex using $AgNO_3$ is often quantitative (see above), for equilibrium studies it is important to know the precise concentration of the metal in solution. Therefore AAS methods which are suitable for the determination of platinum and palladium concentrations for the system under study, i.e. in the nitrate-am(m)ine

matrix, were developed [99][3].

Platinum and Palladium SpectrAA hollow cathode lamps were used. Samples were dispensed automatically into pyrolytically coated graphite tubes without a modifier, normal gas type flow was employed. For platinum determination the following lamp settings were used: Pt resonance line at 265.0 nm, 0.2 nm slit width, and 5 mA lamp current. Standard solutions (BDH) were used as follows: stock – 1000 mg/L, intermediate – 10 mg/L in 0.1% v/v HNO_3, working solutions were prepared by serial dilution in 0.1% v/v HNO_3. For palladium determination the following lamp settings were used: Pd resonance line at 244.8 nm, 0.3 nm slit width, and 5 mA lamp current. Standard solutions (Aldrich) were used as follows: stock – 1000 mg/L, intermediate and working solutions were prepared as for platinum.

The graphite tube was purged with Argon in no-interrupt mode and cleaned and burnt-out at maximum temperature to minimise the carry-over of a metal from sample to sample. Furnace parameters were optimised by altering ashing and atomisation temperatures and are presented in Table 2.3. The automixing facility was chosen and checked for reproducibility, autosampler parameters are shown in Table 2.4. The precision of the determination was checked and found to be very good for both metals (95% – 99.4%).

---

[3] A reprint of this work is bound in the thesis.

Table 2.3. Furnace parameters.

Step number	Platinum			Palladium		
	T (°C)	Time (sec)	Gas flow (L/min)	T (°C)	Time (sec)	Gas flow (L/min)
1	85	5.0	3.00	85	5.0	3.00
2	95	40.0	3.00	95	40.0	3.00
3	120	10.0	3.00	120	10.0	3.00
4	700	5.0	3.00	800	5.0	3.00
5	700	2.0	3.00	800	1.0	3.00
6	800	5.0	3.00	800	2.0	0.00
7	800	2.0	3.00	2600	1.0	0.00
8	800	2.0	0.00	2600	2.0	0.00
9	2700	1.3	0.00	2600	2.0	3.00
10	2700	2.0	0.00			
11	2700	2.0	3.00			

Table 2.4. Autosampler parameters.

	Platinum		Palladium	
	Solution (μL)	Blank (μL)	Solution (μL)	Blank (μL)
Blank	–	20	–	20
Standard 1	5	15	2	18
Standard 2	10	10	4	16
Standard 3	15	5	6	14
Sample	10	10	10	10

2.2.4 Difference UV-Vis spectrophotometry

To prepare samples, 0.1 M $NaClO_4$ was used to maintain a constant ionic strength. Use of buffers was avoided because of their potential ligating ability towards the metal ion [4]. All experiments were carried out under constant pH ~ 6. Stock solutions of nucleoconstituent (nucleosides only were used for these investigations) and metal complex were mixed and diluted to a constant volume so as to give solutions with the following characteristics: a constant total concentration of components whilst varying their ratios (continuous variation experiments) or a constant concentration of metal complex (molar ratio experiments).

An in-house computer program was developed for quick-and-easy calculation of volume and analytical concentration, Appendix I(a). Solutions were incubated at room temperature for 1–2 weeks before recording the spectra.

Difference spectra were created by one of the following methods. Spectra of the metal complex alone, the nucleoside alone and the reaction mixture are measured sequentially, and the difference spectrum is then simulated by means of a program developed in the Department of Environmental Management, VUT, Australia [100]. In another method spectra of the components and the reaction mixture are transferred into the ASCII format, Appendix I(b), and the difference spectra are then simulated by the external spreadsheet software [101].

2.2.5 NMR spectroscopy

For experiments in D_2O, the complexes $Pd(en)Cl_2$ and $Pd(bmpe)Cl_2$ were treated with $AgNO_3$ in D_2O as described above. For experiments in DMSO-d_6 several ways of transferring a complex into solution were employed: (i) direct dissolution in DMSO-d_6 resulting in a maximum concentration of ~ 0.02 M; (ii) $AgNO_3$ treatment in DMSO-d_6 yielding a filtrate with ~ 0.1 M concentration, but since in this procedure it is necessary to expose a complex to $AgNO_3$ for several days there is a risk of DMSO-d_6 absorbing water; (iii) $AgNO_3$ treatment in H_2O with consequent complete H_2O vaporisation, drying the precipitate, washing with acetone or ether and eventual dissolving in DMSO-d_6. The last method gave an ~ 0.1 M solution, the molecular formula of the intermediate precipitate was established as $[Pd(carrier\ ligand)(H_2O)_2](NO_3)_2$ by microanalysis. Weighted amounts of nucleoconstituents (nucleotides only were used for these investigations) were dissolved in dry solvent to provide definite concentrations and then mixtures of particular component molar ratios were prepared at room temperature (sample volumes were 0.4 – 0.5 ml). For ratio-dependent experiments in D_2O, pH values were adjusted to ~ 8.5 (values uncorrected for pD) by trace quantities of concentrated solutions of NaOD and DCl [102]. Most experiments were performed at 0.03 – 0.18 M for Pd and 0.03 – 0.1 M for nucleotide. NMR spectra were recorded after 4 – 5 days of storage in the dark.

2.2.6 Other methods

2.2.6.1 Conductivity

All measurements were carried out at 20 ± 1 °C and normalised to 25 °C. 0.005 M potassium chloride was used as a standard (conductivity 0.714 mS/cm).

Samples for batch titrations were prepared by mixing aliquots of 0.001 M guanosine solution with aliquots of metal complex solution (Pd(en), 0.006 M; Pd(bmpe), 0.001 M) so as to give solutions with a 0.001 M total concentration of components, whilst varying their ratios (continuous variation experiments). This concentration value was chosen in order to be as high as possible but not too high to cause the precipitation of product(s) and to distort the conductivity readings. Mixtures were prepared in deionised water with pH ~ 6.0. No background electrolyte was used. Thus all titrations were carried out at low ionic strength, so that, the concentration and hence the conductivity of background electrolyte should not effect conductivity change due to complex formation. Reference solutions of the corresponding metal complex and guanosine were prepared at appropriate concentrations.

2.2.6.2 Potentiometry

Stock solutions of inosine (0.033 M) and metal complex (Pd(en), 0.0008 M; Pd(bmpe), 0.0011 M) were prepared in 0.1 M $NaClO_4$ which was used to maintain a constant ionic strength. Metal solutions were titrated with inosine and pH readings monitored.

An aforementioned in-house computer program was used for samples preparation and continuous titration calculations in NMR, conductivity and potentiometry studies, Appendix I(a).

2.3 RESULTS AND DISCUSSION

2.3.1 Case study. The design, synthesis and nucleoconstituent binding properties of a sterically restrictive metal complex

2.3.1.1 Steric aspects in the judicious design of sterically restrictive metal complexes

As discussed in the Introduction the nature of the binding of a metal complex to a nucleic acid may be influenced to varying degrees by steric factors which may even be determinative of the molecular conformation of the adduct [103]. Such steric demands may be dictated by the bulk on the carrier ligand and by the bulk on the nucleic acid in the vicinity of the binding site, such as nucleobase exocyclic groups [104]. Therefore, there is scope for influencing the molecular conformation of an adduct by judicious design of the metal complex via the carrier ligand.

In order to expedite the rational design of metal complexes a systematisation of geometrical parameters in model systems is necessary. Thus, for the interaction of platinum-am(m)ine complexes with nucleoconstituents, a previously mentioned *stereochemical convention* [103] (section 1.4.1.1) has been employed. It allows to ascertain the flexibility of adducts under a variety of influences and to compare the conformational aspects of model systems, from which steric influences may be inferred. Such comparisons have been mainly restricted to complexes with primary amines as the carrier ligands and *cis*-coordinated nucleobases in the head-to-tail (HTT) configuration. There is still a paucity of information on the steric influence of bulkier carrier ligands and on systems with nucleobases coordinated in

the head-to-head (HTH) configuration [105].

The following examples help to illustrate the influence of steric interactions on the adduct geometry. For instance, in the cis-[Pt(NH$_3$)$_2$(1-MeC)$_2$] [103] (Fig. 2.1a), the nonbulky amine carrier ligands do not contribute to the geometry of the complex (Base/PtN$_4$ dihedral angles are 101° and 102°). On the other hand, replacing the NH$_3$ groups with the bulkier tetramethylethylenediamine ligand (TMED) as in the structure cis-[Pt(TMED)(1-MeC)$_2$] [106] (Fig. 2.1b) introduces more steric features into the complex including contacts between the carrier ligand and the nucleobases *and* between the nucleobases themselves (Base/PtN$_4$ = 94.5°).

2.3.1.2 Steric restriction of the adduct geometry

One possible way of exerting specific steric influence for controlling the geometry of *cis* bis-coordinated nucleobases, nucleosides and nucleotides is to use carrier ligands and metal complexes which have the potential to enforce a HTH conformation. This is an important goal since in the conventional nucleic acid tertiary structures bases are oriented in a HTH conformation (albeit, prior to any disruptive influence). The crystal structure analyses of systems modelling platinum – nucleic acid interactions show that in the majority of model cases the nucleobases are arranged in a HTT conformation. Only in cases of a fortuitous nature [107] and in several complexes containing oligonucleotides (i.e. the conformation is dictated by the polymeric character of the ligand) [108,109] have HTH conformations been found. Thus, the prevalence of the HTT orientation in simple model systems for which crystal structures have been obtained has permitted the properties and the flexibility of this kind of adduct to be systematically explored [110].
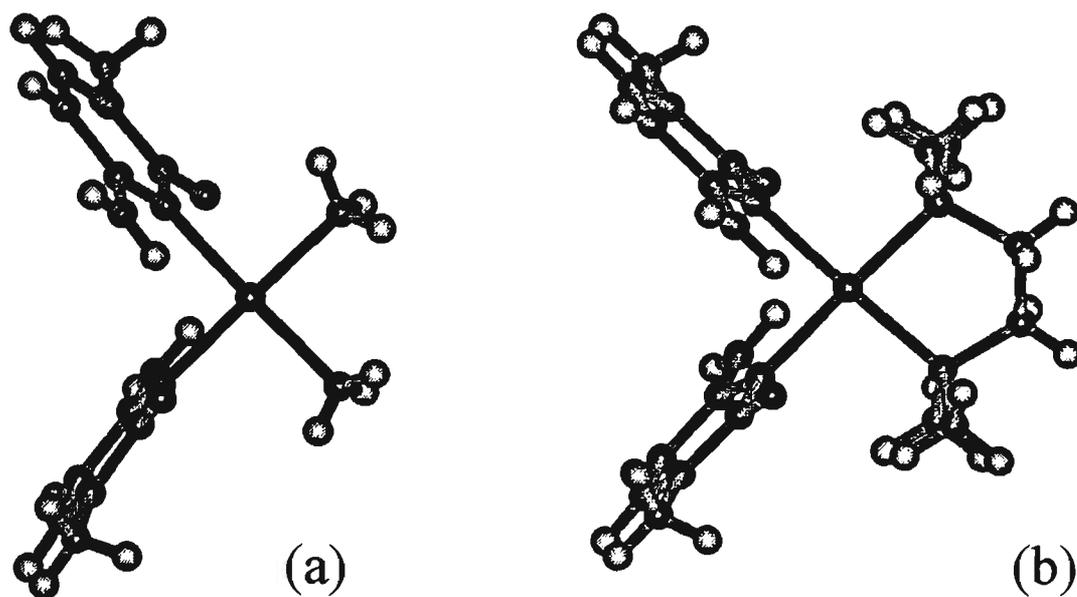
Figure 2.1. Crystal structures of platinum complexes of cis bis-coordinated nucleobases. (a) cis-[Pt(NH₃)₂(1-MeC)₂ [103]. (b) cis-[Pt(TMED)(1-MeC)₂ [106].

Accordingly, enforcing a HTH conformation would allow a similar systematic exploration of such adducts and might assist in determining the extent to which a nucleic acid would be disrupted. As has already been mentioned, unique structural motifs introduced into DNA upon metal binding have been implicated in the molecular mechanism of action of cisplatinum complexes [111]. Thus, the rational design of nucleic acid-binding complexes, inducing specific distortions in nucleic acids, may also provide a means for the control of various biological functions related to nucleic acid topology.

It has been recognised for some time now that a carrier ligand which specifically blocks off one side of the coordination square plane, PtN_4, has the potential to enforce a HTH geometry of nucleobases. Subsequent structural modifications of the carrier ligand could result in a wide range of conformations of varying rigidity which could be reflected in the tertiary structure of a coordinated nucleic acid. Reedijk et al. [98] were the first to test this approach using the ligand **bpe** (bpe = 1,2-bis(pyridin-2-yl)ethane, Fig. 2.2a) which provides different environments above and below the platinum coordination plane (Fig. 2.3). However, it was revealed by the X-Ray analysis of $Pt(bpe)(9\text{-}MeH)_2(NO_3)_2$ (H = hypoxanthine) that the carrier ligand **bpe** does not influence nucleobase ligands enough to enforce the HTH orientation.

To further investigate the possibility of enforcing a HTH arrangement by a **bpe**-like ligand several series of bidentate ligands, where each ligand may have subtly different steric properties in terms of interactions with nucleobases, may be envisaged. An example of one such series of ligands which has been designed for these investigations is shown in Fig. 2.2. For a range of ligands such as this the possibility exists that there may be a fine distinction between enforcing a HTH arrangement and hindering nucleobase binding completely.
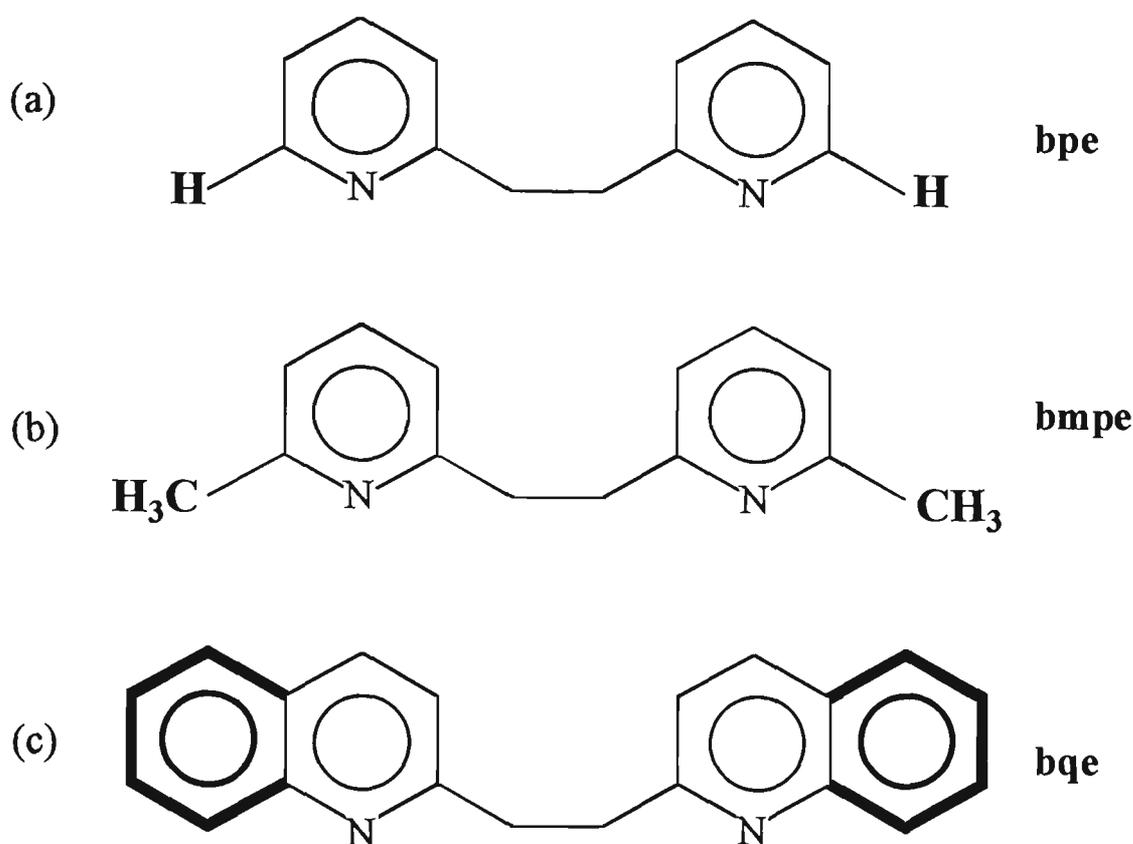
Figure 2.2. *A series of potential sterically restrictive ligands. (a) 1,2-bis(pyridin-2-yl)ethane (bpe); (b) 1,2-bis(6-methylpyridin-2-yl)ethane (bmpe); (c) 1,2-bis(quinolin-2-yl)ethane (bqe).*

Of particular interest is the possibility of a confluence of physical and chemical conditions dictated by steric requirements, where "crowding out" may occur with respect to *cis*-biscoordination, resulting in a stoichiometrically controlled destabilisation of a complex. This hypothesis will be elaborated subsequently.

Upon inspection of the series of ligands in Fig. 2.2 it could be expected that, when coordinated to a square-planar metal centre, the presence in **bmpe** of the bulky methyl groups adjacent to the pyridinyl donor nitrogens, compared to just hydrogen atoms in **bpe**, will result in an enhanced hindrance of one side of the coordination plane and may provide the steric conditions for the enforcement of a HTH *cis*-coordination of two nucleobase ligands.
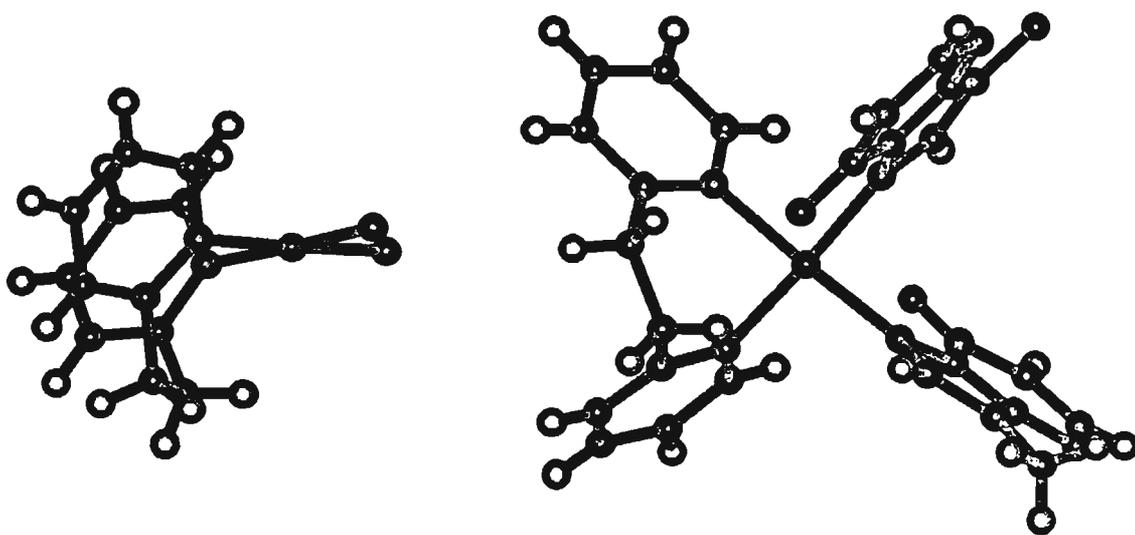
Figure 2.3. Crystal structure of *Pt(bpe)(9-MeH)₂* *[98].* *(left)* A side view, 9-MeH ligands are omitted for clarity. *(right)* A view perpendicular to the coordination plane.

Thus, the ligand 1,2-bis(6-methylpyridin-2-yl)ethane (**bmpe**, Fig. 2.2b) was synthesised and its 1:1 complex with palladium(II) prepared and crystallised as the malonate trihydrate. The structure was subsequently characterised by X-Ray analysis [94][4].

2.3.1.3 The crystal and molecular structure of complex Pd(bmpe)(malonato)·3H₂O

The crystal and molecular structure of the title compound were determined as part of a collaborative venture, full details of the structure may be found in the corresponding publication [94]. In the context of this section selected structural features are presented as follows. Three views the square-planar Pd(bmpe)(malonato) moiety are shown in Fig. 2.4. Both methyl substituents can be seen to be projecting over one side of the coordination plane successfully blocking it off (Fig. 2.4b, side view). The ethylene bridge also projects to some extent over the other side of the plane. Thus, success has been achieved with this ligand in building up the bulk preferentially on one side of the plane as compared to the **bpe** system.

The coordination geometry around Pd(II) is typically square-planar. The **bmpe** ligand forms a boat conformation and pyridine moieties are twisted from the coordination plane by 66.2° and 73°; these angles compare well with the 62.5° and 69.4° for the pyridine moieties in the complex Pt(bpe)(9-MeH)$_2$(NO$_3$)$_2$ [98]. Ethylene bridge protons show distinct chemical shifts in solution (as in the **bpe** system) which, together with the similarity in the angles between the coordinated **bpe** and **bmpe** systems, reflects the rigidity of this type of ligand.

---

[4] A reprint of this work is bound in the thesis.
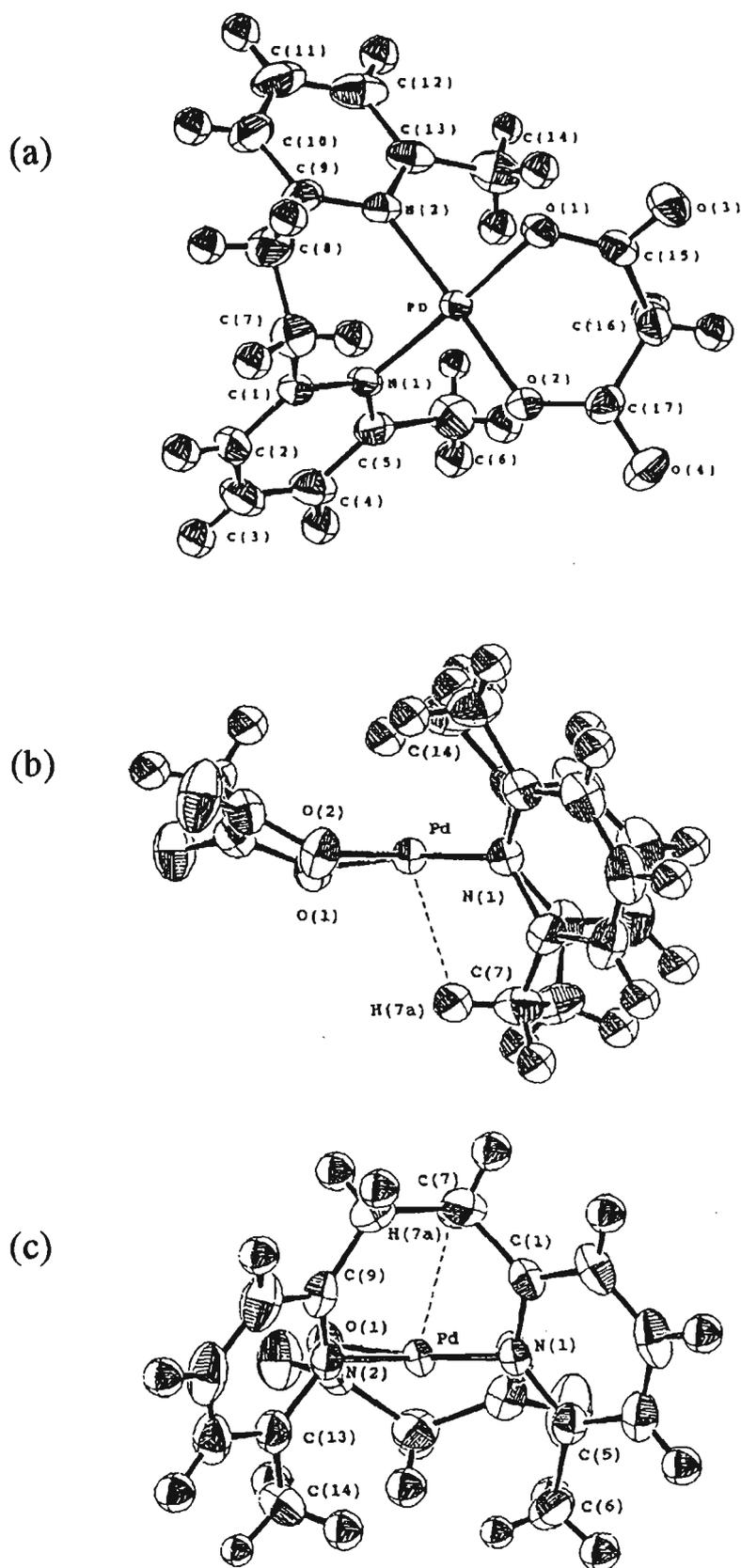
Figure 2.4. Crystal structure of the Pd(bmpe)(mal) molecule (a) A view perpendicular to the coordination plane together with atom numbering scheme. (b) A side view. (c) An end view.

An interesting feature of this structure is that one of the ethylene bridge protons, H(7a), exhibits a very short approach to the palladium atom, 2.436(5) Å (Fig. 2.4(b,c)). If this represents a weak attractive interaction it could contribute to the difference in dihedral angles between the pyridine moieties and the coordination plane. Selected bond length and angles associated with the coordination plane are typical of those reported in related systems. However, in this sterically crowded molecule several angles, in particular those around the pyridine nitrogen atoms, and the distortion of the square-plane itself reflect a "pushing-down" influence of the methyl substituents.

2.3.1.4 Nucleoconstituent binding to the Pd(bmpe)$^{2+}$ moiety

2.3.1.4.1 ^1H NMR investigations

Coordination of Pt(II) or Pd(II) to the N7 of purines and the N3 of pyrimidines is expected to result in a downfield shift of at least 0.4 ppm for the H8 (purines) and at least of 0.2 ppm for the H5 and H6 (pyrimidines) proton resonances compared to those of the free ligand [59]. For 2:1 stoichiometric mixtures of nucleoconstituents with the solvated Pd(bmpe)$^{2+}$ complex the observed differences in chemical shifts of monitored protons in nucleoconstituents in the absence and presence of metal complex were considerably smaller in magnitude than what would be expected if binding had occurred (Tables 2.5 and 2.6). This is consistent with an inability to isolate *any* products preparatively. This is surprising, since on the basis of a comparison of the metrical parameters associated with the projection of the ethylene bridge across the coordination plane in the complexes Pd(bmpe)(malonato)·3H$_2$O and Pt(bpe)(9-MeHX)$_2$(NO$_3$)$_2$ (Table 2.7), it is deemed possible for *at least* one ligand to coordinate to the metal centre.

Table 2.5. H8 proton chemical shifts (δ, ppm) for guanine derivatives in the absence and presence of Pd(bmpe)$^{2+}$.

	9-EtG[a]	Guo	5'-GmpNa$_2$
Guanine derivative in the absence of metal complex	7.70	7.98	8.20
Guanine derivative in the presence of metal complex	7.70	8.03	8.23
Difference	0.00	+0.05	+0.03

[a]Downfield from TMS in DMSO-d_6. Other values are from DSS in D$_2$O.

Table 2.6. H5 and H6 proton chemical shifts (δ, ppm)[a] for 1-methylcytosine in the absence and presence of Pd(bmpe)$^{2+}$.

	H(5)[b]	H(6)[b]
1-methylcytosine in the absence of metal complex	5.94	7.55
1-methylcytosine in the presence of metal complex	5.87	7.51
Difference	-0.07	-0.04

[a]Downfield from DSS in D$_2$O. [d]Values are averages of the doublet signals for H(5) and H(6).

Table 2.7. A comparison of the metrical parameters in the complexes Pd(bmpe)(malonato)·3H$_2$O and Pt(bpe)(9-MeHX)$_2$(NO$_3$)$_2$.

Contacts[a]		Distance[a] (Å)		Difference between the two structures (Å)
C8-Pd	(C11-Pt)	2.906	(3.028)	0.122
C7-Pd	(C31-Pt)	3.270	(3.327)	0.057
C8-O1	(C11-N27)	3.893	(4.150)	0.257
C7-O2	(C31-N17)	4.204	(4.420)	0.216

[a]Contacts and distances in parentheses are for the bpe complex.

Although complete hindering of nucleobase binding could be attributed to the influence of the steric bulk of the exocyclic methyl substituents, it is unlikely that the subtle change in coordination geometry between **bpe** and **bmpe** could produce such a dramatic effect on the relative coordination abilities of these systems. The above results have prompted an investigation into the stoichiometric dependence of coordination for this system.

All ratio-dependence experiments for the Pd(bmpe)$^{2+}$ system were accompanied by simultaneous control experiments using a non-sterically restrictive system, namely Pd(en)$^{2+}$. Ratio-dependence NMR experiments (Fig. 2.5) suggest that at low nucleoconstituent-to-complex molar ratios (R) coordination of Pd(bmpe)$^{2+}$ to the nucleotide and the nucleoside does indeed occur. Comparison of plots (a) and (b) shows a remarkable difference between the binding behaviour of Pd(bmpe)$^{2+}$ and Pd(en)$^{2+}$ (which is capable of coordinating two nucleoside/nucleotide ligands [12]). In the experiment with 5'-GmpNa$_2$ in D$_2$O (Fig. 2.5b) Pd(en)$^{2+}$ produces a smooth δ(H8(G)) vs R curve which is consistent with the formation of a 2:1 complex. The smoothness of this curve is a result of various equilibria contributing to the average chemical shift value; in other words the H8(G) resonance for the free ligand, the 1:1 bound ligand and the 2:1 bound ligand are not resolved on the NMR time scale. On the contrary, in the experiment with Pd(bmpe)$^{2+}$ (Fig. 2.5a) the data shows a sharp change in chemical shift value from that of a complexed to that of uncomplexed species. Thus, these results suggest that the 1:1 adduct becomes destabilised when the nucleotide-to-complex ratio exceeds a value of $\sim$ 0.75.
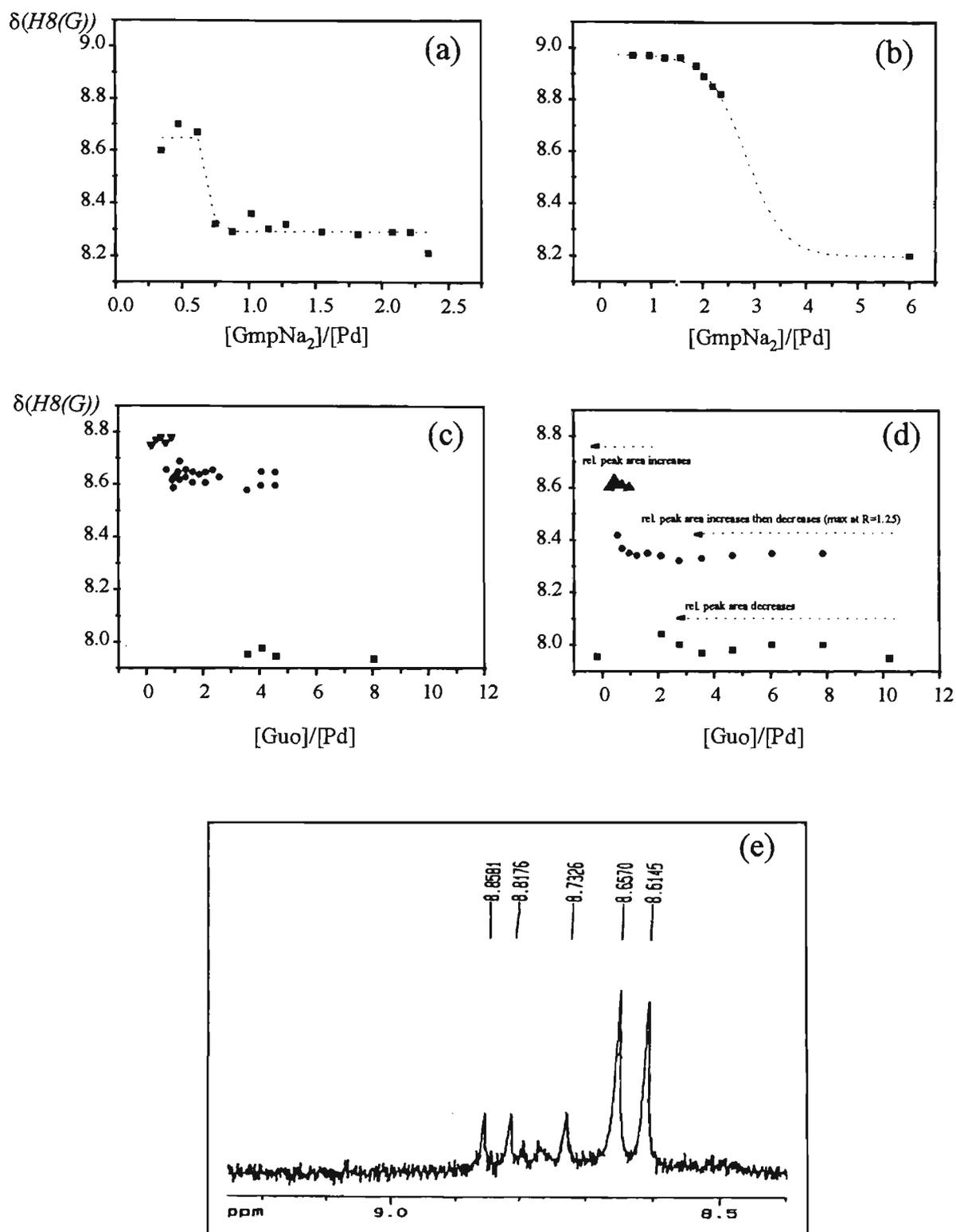
Figure 2.5. NMR titrations: H8(G) chemical shift as a function of nucleoconstituent-to-complex ratio. (a) Pd(bmpe)/5'-GmpNa₂ in D₂O; (b) Pd(en)/5'-GmpNa₂ in D₂O; (c) Pd(bmpe)/guanosine in DMSO-d₆; (d) Pd(en)/guanosine in DMSO-d₆; (e) Pd(bmpe)/guanosine in DMSO-d₆ on 300 MHz, R ~ 1.

Due to the low aqueous solubility of most nucleosides, the titrations with guanosine were carried out in DMSO-d_6 (Fig. 2.5(c,d)). Controversy exists with respect to using DMSO in binding studies. Since DMSO itself acts as a ligand towards platinum and palladium the results of such studies may be ambiguous [97]. Pd(bmpe)$^{2+}$ and Pd(en)$^{2+}$ titrations with guanosine produce multiple peaks in the H8(G) frequency range for samples with R: ~ 0.5 to ~ 4.5. The following explanations may be advanced for this data. It is possible for DMSO to be taking part in the reaction. However, perhaps a more plausible interpretation is that DMSO slows down the proton exchange on the NMR time scale and allows the H8(G) proton signals from different species to be resolved.

While the above experiments were carried out on a low field NMR (due to limited resources at the time), selected samples of the Pd(bmpe)$^{2+}$/Guo system in DMSO at R ~ 1 were scanned on a high field NMR spectrophotometer which allowed a better resolution of the H8(G) frequency. The splitting pattern in the range 8.5 – 9.0 ppm (Fig. 2.5e) indicates the existence of two different H8(G) protons, resolved on the NMR time scale. In such a sterically restricted environment where both *cis*-coordination sites are being attacked by nucleoside ligands competing for space it is not unreasonable to propose a "quasi-binding" to the ligand. Fig. 2.6 depicts possible coordination modes for this system, involving both full-coordination and quasi-binding. While mode (A) does not explain the splitting of signals and modes (B) and (C) are not very probable for a reaction mixture with ≤ 1:1 molar ratio of components, mode (D) is feasible. In this mode ligands jostling for position may result in complex destabilisation.

Pd   Pd   Pd   Pd

G   G   G   G   G   G   G

HTH or HTT   HTH or HTT   HTH or HTT

(A)   (B)   (C)   (D)
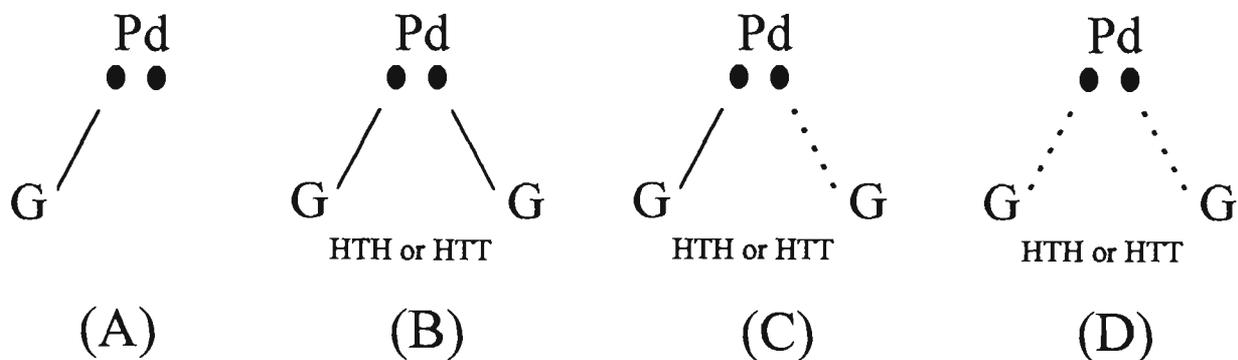
Figure 2.6. *Possible coordination modes in Pd(bmpe)(guanosine) system. Black dots symbolise coordination sites, dotted lines represent quasi-binding, G stands for guanosine ligand.*

2.3.1.4.2 UV-Vis, conductivity and potentiometry investigations

Further evidence for $Pd(bmpe)^{2+}$ coordinating at least one nucleoconstituent is provided by UV-Visible difference studies. The Job's plot from a continuous variation experiment with guanosine (Fig. 2.14a) supports the notion of formation of a 1:1 complex and suggests that this adduct is a predominant species in solution. Further analysis of UV-Vis data from the continuous variation experiments is carried out later (section 2.3.2) together with data for other complexes.

The UV-Vis molar ratio titrations were carried out for $Pd(en)^{2+}$ and $Pd(bmpe)^{2+}$ with guanosine. To establish stoichiometry, molar ratio plots have been constructed at λ_{min} and λ_{max}, where the absorbance difference between samples is most pronounced, as well as at several selected wavelengths to check for wavelength-dependence of the determined stoichiometry. Fig. 2.7 shows representative molar ratio plots at selected wavelengths for each system.
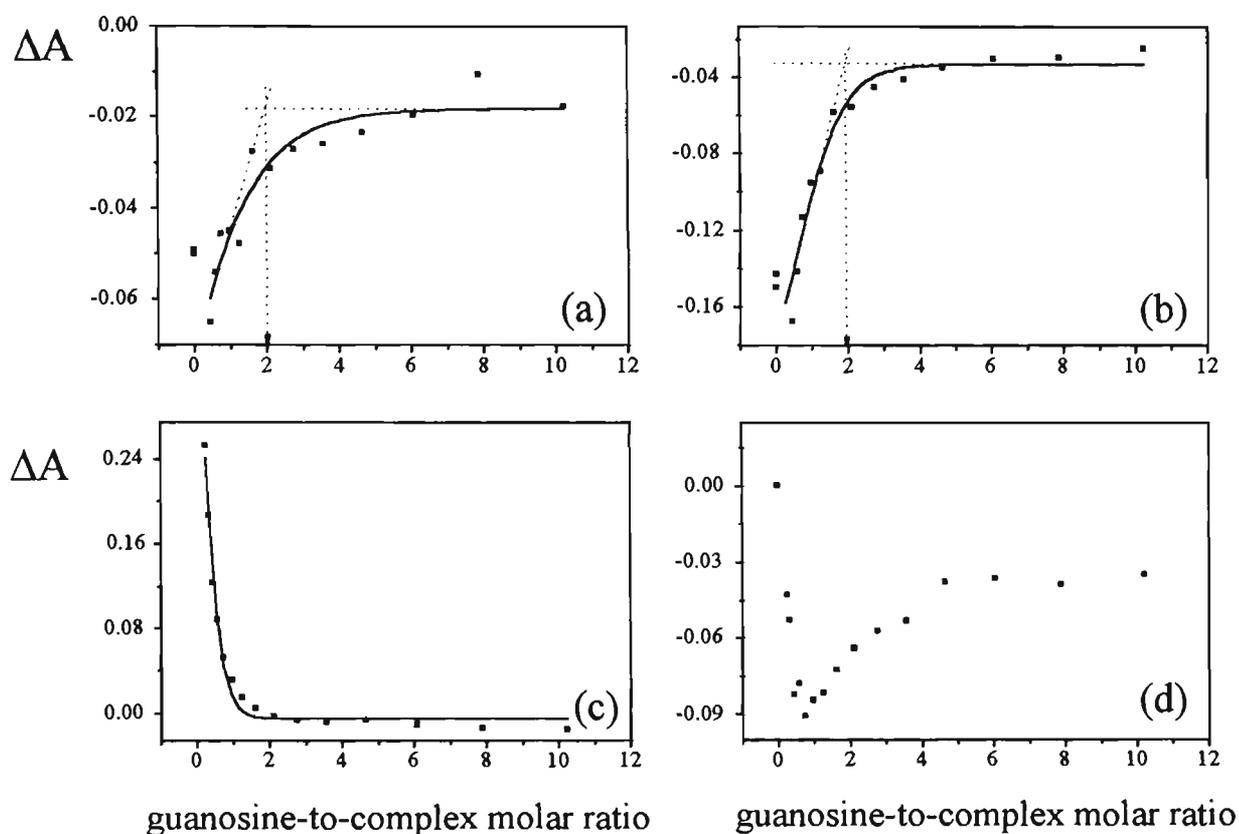
Figure 2.7. *UV-Vis molar ratio titrations of palladium complexes with guanosine. (a) Pd(en)$^{2+}$, $\lambda = 277$ nm; (b) Pd(en)$^{2+}$, $\lambda = 250$ nm. (c) Pd(bmpe)$^{2+}$, $\lambda = 277$ nm; (d) Pd(bmpe)$^{2+}$, $\lambda = 250$ nm.*

While the Pd(en)$^{2+}$ system behaves according to the formation of 2:1 complex, displaying a break point around R ~ 2 at all selected wavelengths (e.g. Fig. 2.7(a,b), $\lambda = 277$ and 250 nm respectively), the coordinating repertoire of Pd(bmpe)$^{2+}$ is more diverse. Whereas at some wavelengths this system demonstrates a break point at R ~ 1 (e.g. Fig. 2.7c, $\lambda = 277$ nm), at other wavelengths it actually displays an extremum point around R ~ 1 (e.g. Fig. 2.7d, $\lambda = 250$ nm). These observations correspond to the formation of a 1:1 Pd(bmpe)(guanosine) complex and the latter in particular suggests that beyond an approximately 1:1 molar ratio the complex dissociates.

The wavelength-dependence of the molar ratio plots of Pd(bmpe)$^{2+}$ may be

Figure 2.7. *UV-Vis molar ratio titrations of palladium complexes with guanosine. (a) Pd(en)$^{2+}$, $\lambda = 277$ nm; (b) Pd(en)$^{2+}$, $\lambda = 250$ nm. (c) Pd(bmpe)$^{2+}$, $\lambda = 277$ nm; (d) Pd(bmpe)$^{2+}$, $\lambda = 250$ nm.*

While the Pd(en)$^{2+}$ system behaves according to the formation of 2:1 complex, displaying a break point around R ~ 2 at all selected wavelengths (e.g. Fig. 2.7(a,b), λ = 277 and 250 nm respectively), the coordinating repertoire of Pd(bmpe)$^{2+}$ is more diverse. Whereas at some wavelengths this system demonstrates a break point at R ~ 1 (e.g. Fig. 2.7c, λ = 277 nm), at other wavelengths it actually displays an extremum point around R ~ 1 (e.g. Fig. 2.7d, λ = 250 nm). These observations correspond to the formation of a 1:1 Pd(bmpe)(guanosine) complex and the latter in particular suggests that beyond an approximately 1:1 molar ratio the complex dissociates.

The wavelength-dependence of the molar ratio plots of Pd(bmpe)$^{2+}$ may be

due to the different nature of optical transitions at different wavelengths. While absorbance in the region of 250 – 260 nm is due to $\pi \rightarrow \pi^*$ transitions in a nucleobase ring [112], the absorbance around 270 – 280 nm (d-d transition zone) is usually assigned to a charge-transfer from a ligand to a metal ion upon binding [113,114]. Therefore, both the disappearance of the absorbance difference between the sum of the components and the Pd(bmpe)/guanosine mixture (ΔA = 0) at λ = 277 nm and the reversal of absorbance difference at λ = 250 nm beyond a 1:1 molar ratio corroborate the conclusion about dissociation of the 1:1 complex.

Conductivity and potentiometry titrations (Fig. 2.8(a,b)) also support the results described above, with respect to both the 1:1 coordinating ability of Pd(bmpe)$^{2+}$ and the differences between the binding behaviours of Pd(bmpe)$^{2+}$ and Pd(en)$^{2+}$. These experiments have been carried out for Pd(bmpe)$^{2+}$ and Pd(en)$^{2+}$ with nucleosides. Nucleosides are preferred over nucleobases for such titrations because of their greater solubility, and over nucleotides because of their neutrality.

In conductivity titrations (Fig. 2.8a), the difference between a sample reading and a reference solution conductivity has been used. Out of the two aforementioned approaches to presenting difference conductivity results (section 2.1.3), the *absolute* difference $\Delta \kappa$ was opted for in order to maintain consistency with the UV-Vis experiments. Thus, $\Delta \kappa$ was plotted against the mole fraction of guanosine, resulting in continuous variation plots, analogous to Job's plots. In case of the Pd(en)$^{2+}$ system (Fig. 2.8a, inset), the data asymptotes rather than displaying a minimum. This profile is consistent with those observed for other metal/nucleoside systems where 2:1 formation is indicated (for example, that of Pb^{2+}/Cytidine [115]).
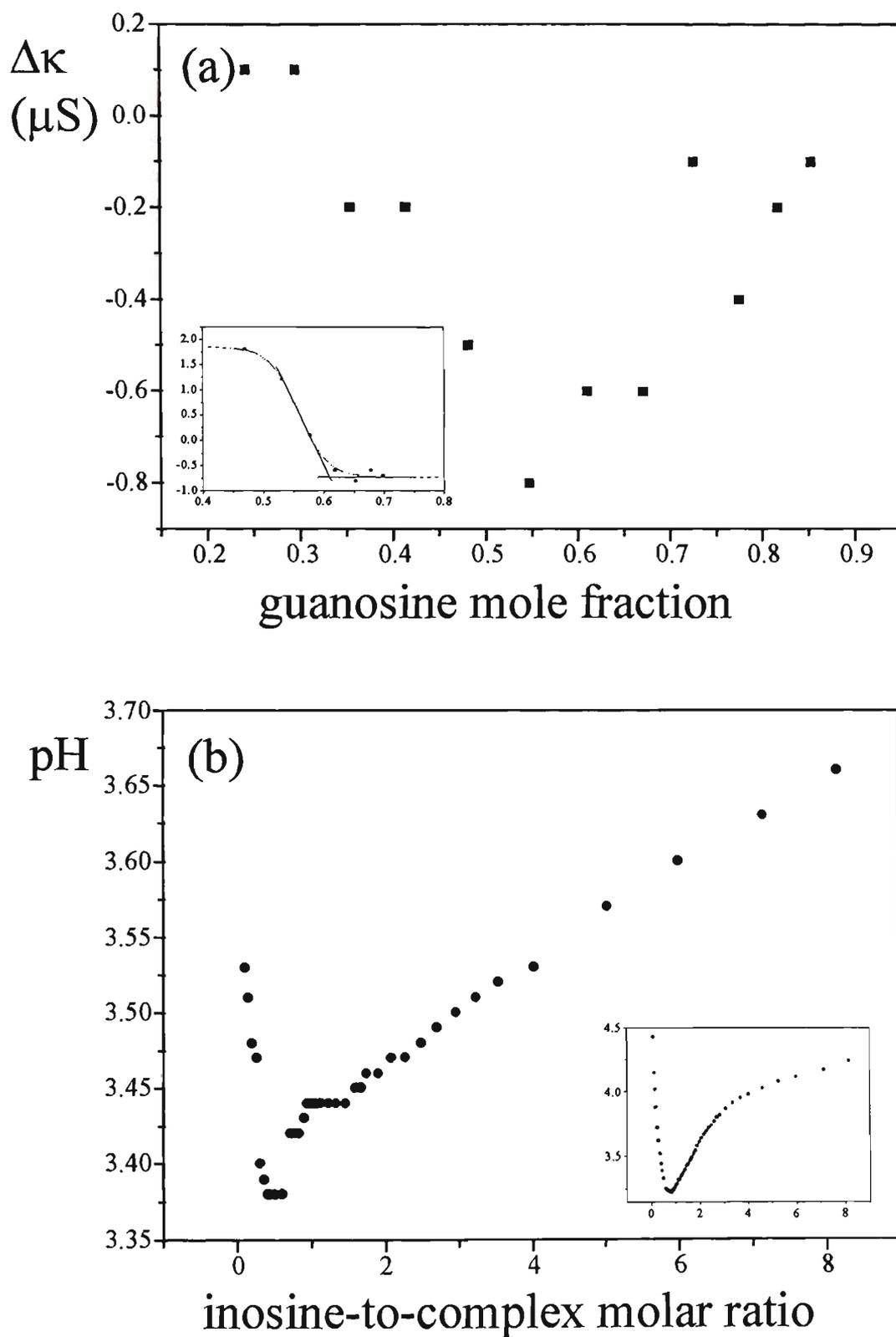
Figure 2.8. Conductivity *(a) and potentiometry (b) titrations of Pd(bmpe)²⁺ with nucleosides. Insets - control experiments with Pd(en)²⁺.*

The asymptotic behaviour of $\Delta\kappa$ for the $Pd(en)^{2+}$/Guo system after a 2:1 ratio has been achieved indicates that no new conducting species are formed (by either association or dissociation). However, in the $Pd(bmpe)^{2+}$ case, $\Delta\kappa$ (absolute values) decreases after a 1:1 ratio has been achieved, indicating a destabilisation process.

A clear difference in behaviour between the two systems may also be observed in the potentiometric titration curves, where the pH is plotted against inosine-to-complex molar ratio (Fig. 2.8b)[5]. For the $Pd(en)^{2+}$ system (Fig. 2.8b, inset), the minimum is observed corresponding to a 1:1 equilibrium, and the break point corresponding to a 2:1 equilibrium. Thus, as expected, a two step 2:1 complex formation is suggested. For the $Pd(bmpe)^{2+}$ system the curve displays some unusual features. The minimum is achieved at a lower inosine-to-complex molar ratio of ~ 0.5. The break point is undetectable in this case due to the linear rather than sigmoidal shape of the titration curve after R ~ 1. Such a shape indicates the absence of any equilibria at a molar ratio R > 1. Additionally, an interesting phenomenon is the observation of a plateau (8 data points, ratio range 0.94...1.47), which could be related to an event occurring, such as complex destabilisation and dissociation.

2.3.1.4.3 A switching event under stoichiometric control?

From the investigations above an anomalous binding behaviour of the sterically restrictive $Pd(bmpe)^{2+}$ system towards nucleoconstituents is suggested. For nucleoconstituent binding to a non-sterically restrictive system $Pd(en)^{2+}$, the ^1H NMR, UV-Vis, conductivity and potentiometry data are consistent with

---

[5] The abscissa units which are different from those used in the conductivity experiments above are standard in potentiometry studies.

a stepwise formation of a 2:1 complex, as expected. Namely,

$$Pd(en)^{2+} + L \rightarrow Pd(en)L^{2+}$$
$$Pd(en)L^{2+} + L \rightarrow Pd(en)L_2^{2+}$$
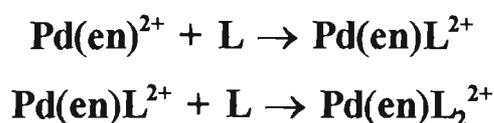
For the sterically restrictive $Pd(bmpe)^{2+}$ at low nucleoconstituent-to-complex ratios (up to $R \sim 1$), the above-presented data is consistent with a 1:1 complexation. However, when the molar ratio exceeds a critical value ($R \sim 1$), the 1:1 complex appears to become unstable and dissociate. After this point the $Pd(bmpe)^{2+}$ complex appears to remain inert to coordination.

It may be hypothesised that the steric crowding associated with a nucleobase ligand, competing against itself for the remaining coordination position on the metal complex may result in a catastrophic destabilisation of the 1:1 complex and lead to its full dissociation. The bmpe system may represent a fortuitous confluence of physical and chemical features for such an "on-off" 1:1 coordination to occur. The following scheme is suggested:

$$Pd(bmpe)^{2+} + L \rightarrow Pd(bmpe)L^{2+}$$
$$Pd(bmpe)L^{2+} + L \rightarrow Pd(bmpe)^{2+} + 2L$$
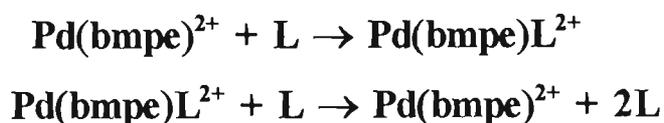
Thus, once a *critical ratio* has been exceeded a dissociative cascade eventuates. The exact value of the *critical ratio* appears to be dependent on the conditions of the experiment. What is being proposed is essentially a sterically induced switching event under stoichiometric control. If there is substance in this hypothesis and if it represents a physical phenomenon which can occur in other related systems, it may have important biological significance. That is, since metal species play varied and complex roles in the metabolism of

living systems, such sterically restrictive systems which are sensitive to varying concentration of ligands within the cell (under metabolic control) could play an important regulatory role. The molecular basis for such events is an area of active research [116].

2.3.2 Spectrophotometric studies of carrier ligand steric effects on nucleoconstituent binding

The previous studies on the ratio-dependent binding of Pd(bmpe)$^{2+}$ to nucleoconstituents (section 2.3.1.4) suggest atypical behaviour associated with the designed sterically restrictive bmpe carrier ligand. This has prompted further investigations, utilising spectrophotometric techniques, into the influence of carrier ligand steric demands on metal complex – nucleoconstituent binding. Thus, product stoichiometry is investigated as well as the effect on coordination of nucleoconstituent-to-metal ratio and reagent concentration level.

2.3.2.1 Experimental design considerations

When designing and carrying out spectroscopic experiments for studying complex equilibria one attempts to create from the data the best model for the system in terms of estimating the number of species in solution, their stoichiometry and formation constants. The obtained characteristics of the model should be wavelength independent, hence analysing the data at several different wavelengths is a good diagnostic tool for checking the model. It may also be advantageous to carry out experiments at different concentration levels to check if the determined characteristics are concentration dependent. Their independence may provide evidence of original assumption being correct, but it is also possible that at different concentration levels the system behaves differently. The absence of a sharp isosbestic point or the presence of more than one isosbestic point is an indication of system complexity, but this is not always the case (see section 2.3.2.2 for more details).

The spectrophotometric equilibria studies presented here have been carried

out on a series of palladium complexes with increasing steric bulk on the carrier ligand. Palladium instead of platinum has been chosen for these studies. Both metal ions form diamagnetic, planar complexes, possess similar radii, and strongly prefer nitrogen donors [117]. However, the kinetic behaviour of Pt(II) complexes is up to five orders of magnitude slower than that of Pd(II) [117]. The slow reactions of platinum may result in polymerisation reactions and complicate the equilibria studies [118]. The comparable stereochemistry of Pt(II) and Pd(II), together with more favourable kinetic characteristics of palladium, make it a preferable vehicle for relative equilibria studies.

2.3.2.2 Continuous variation experiments, description of spectral features

Spectra were generated in continuous variation experiments with four palladium complexes of the carrier ligands **en**, **bpe**, **bmpe** and **bispep**, interacting with
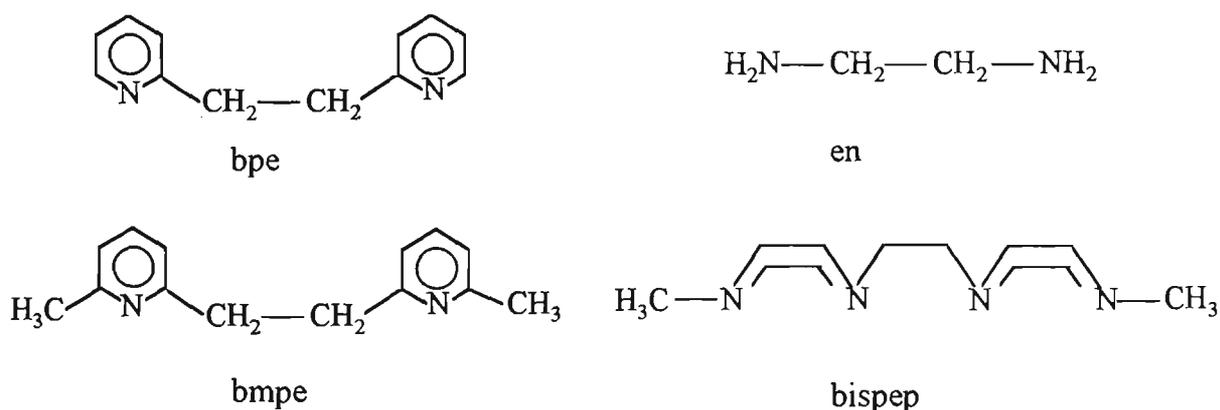
the nucleosides guanosine and inosine (Figs. 2.9 – 2.13). The wavelength values corresponding to maximum and minimum absorbance and to isosbestic points in observed and difference spectra are summarised in Table 2.8 together with the established nucleoside-to-complex stoichiometries of the predominant products.
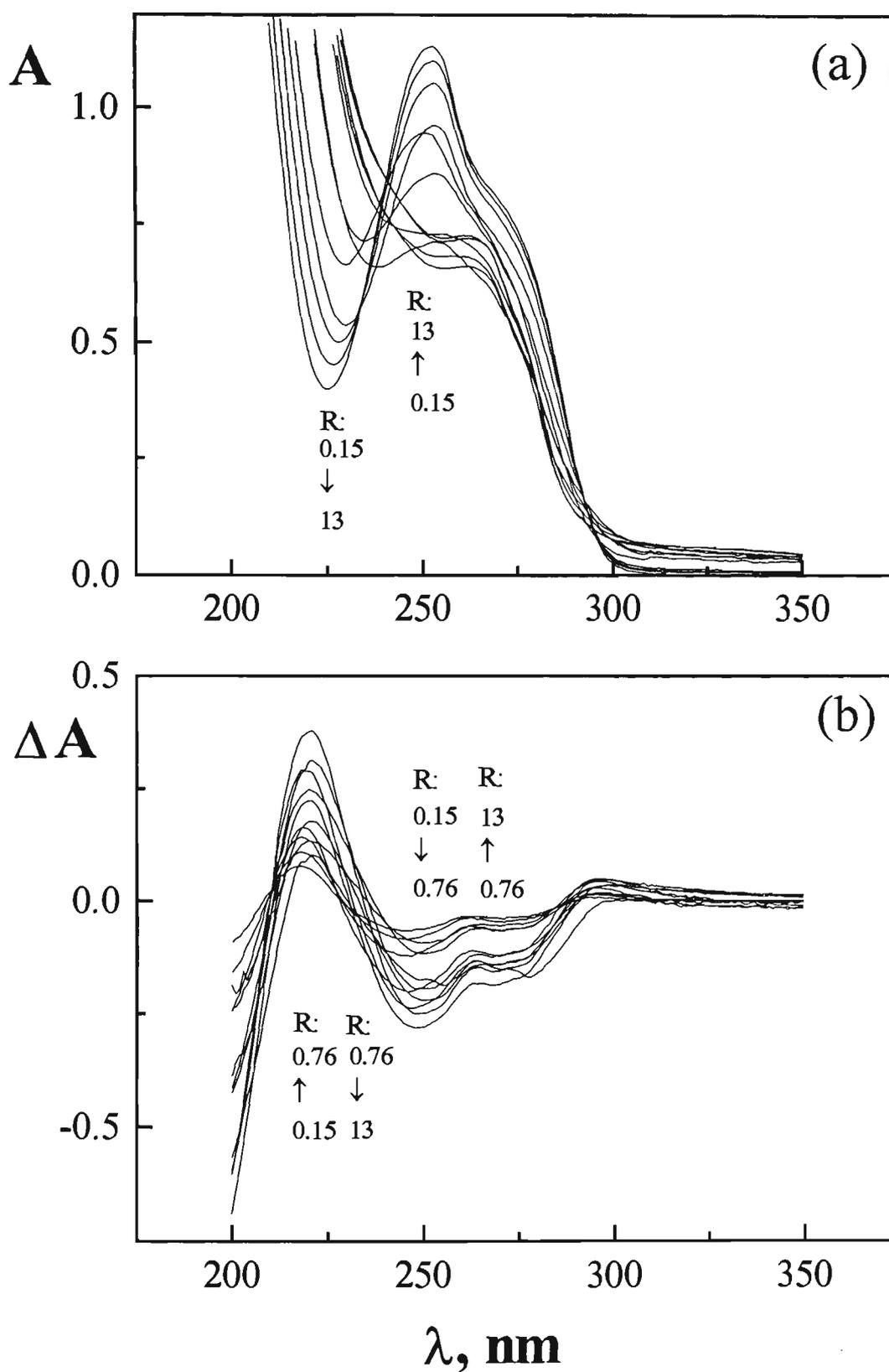
Figure 2.9. Continuous variation experiment. System: Pd(bmpe)/guanosine. Total concentration - 10^{-4} M. Molar ratio R = [Guo]/[Pd] range: 0.15...13.
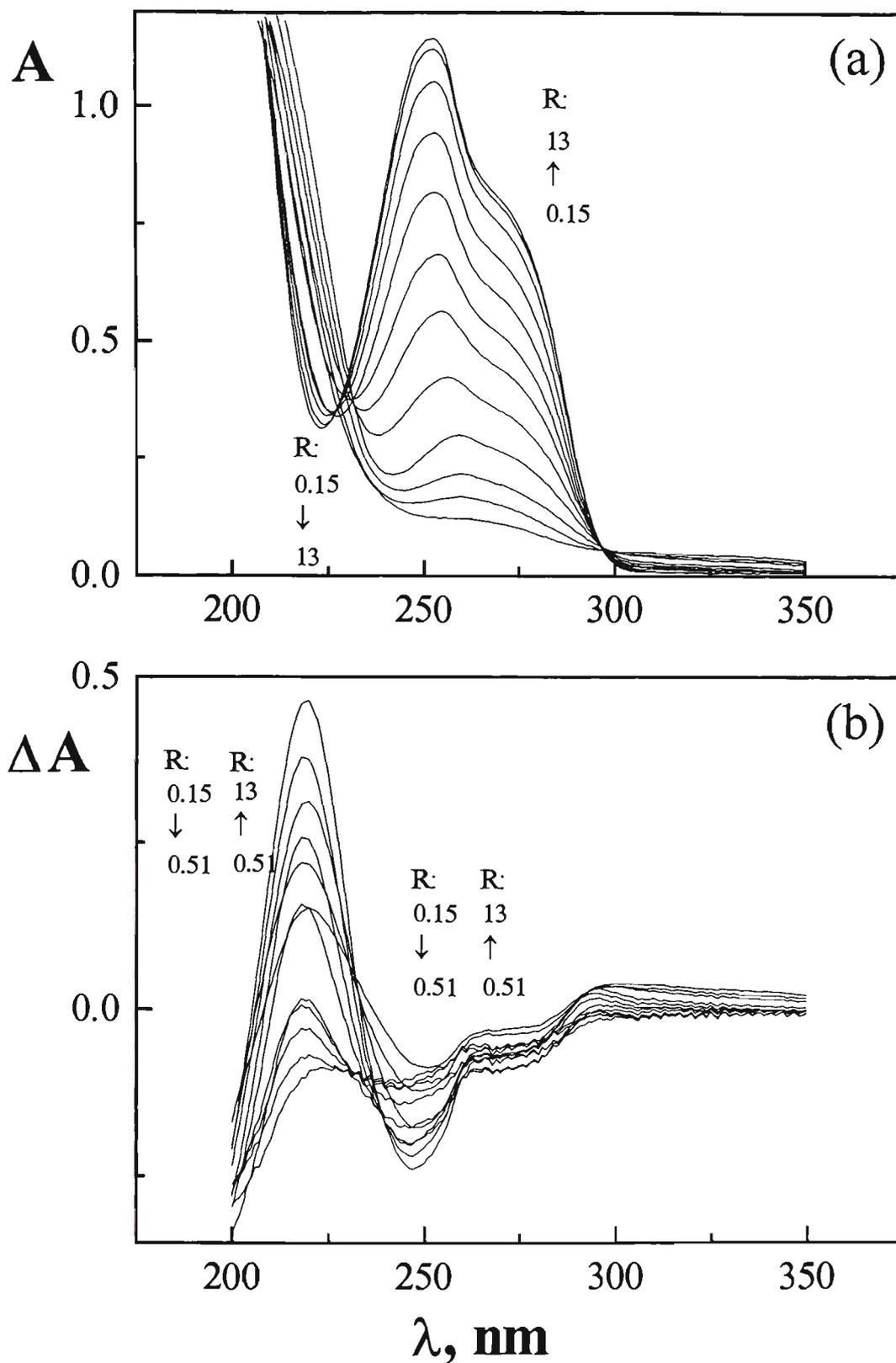
Figure 2.10. *Continuous variation experiment. System: Pd(en)/guanosine. Total concentration - 10^{-4} M. Molar ratio R = [Guo]/[Pd] range: 0.15...13.*
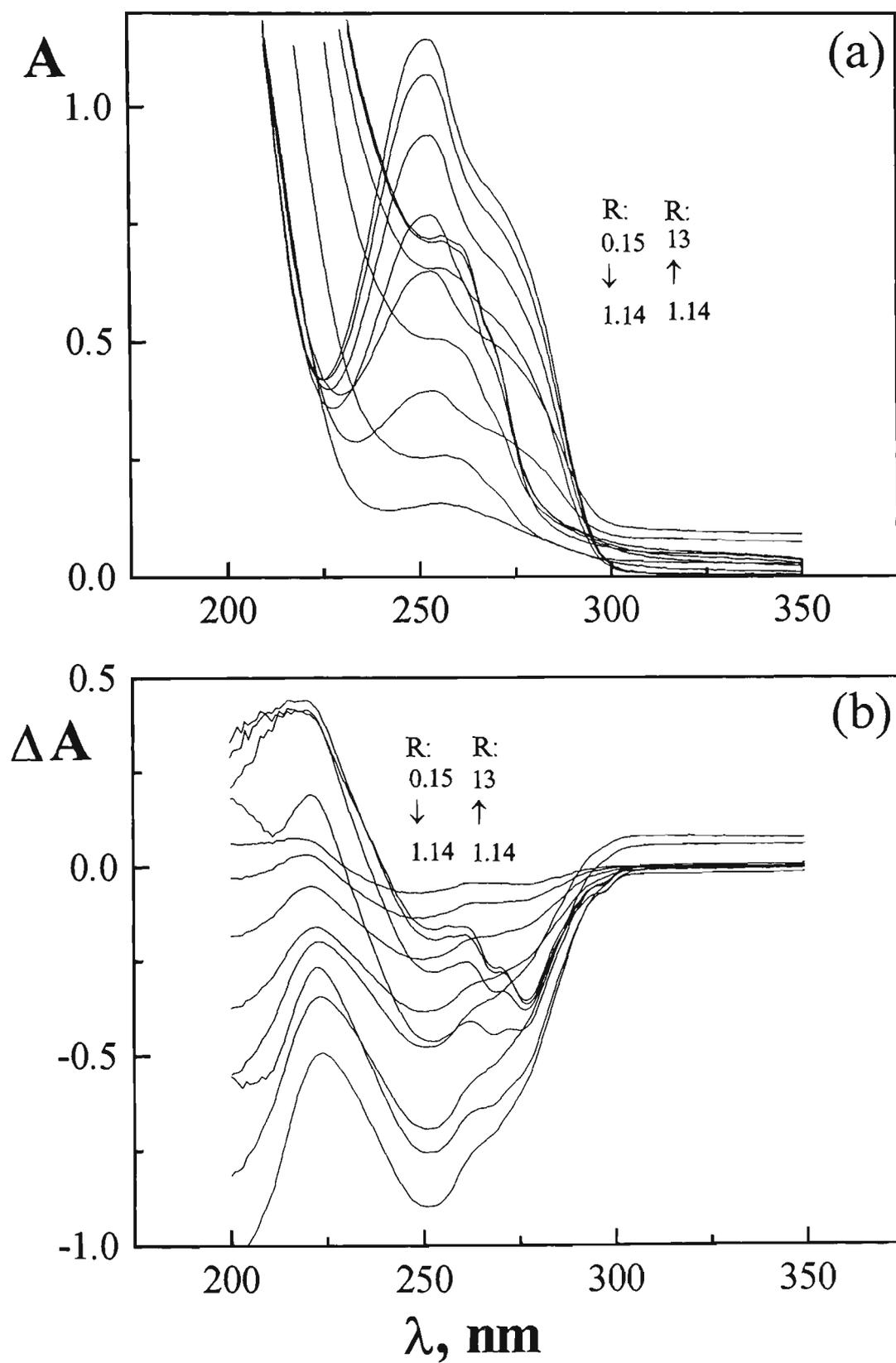
Figure 2.11. Continuous variation experiment. System: Pd(bpe)/guanosine. Total concentration - $10^{-4}M$. Molar ratio R = [Guo]/[Pd] range: 0.15...13.
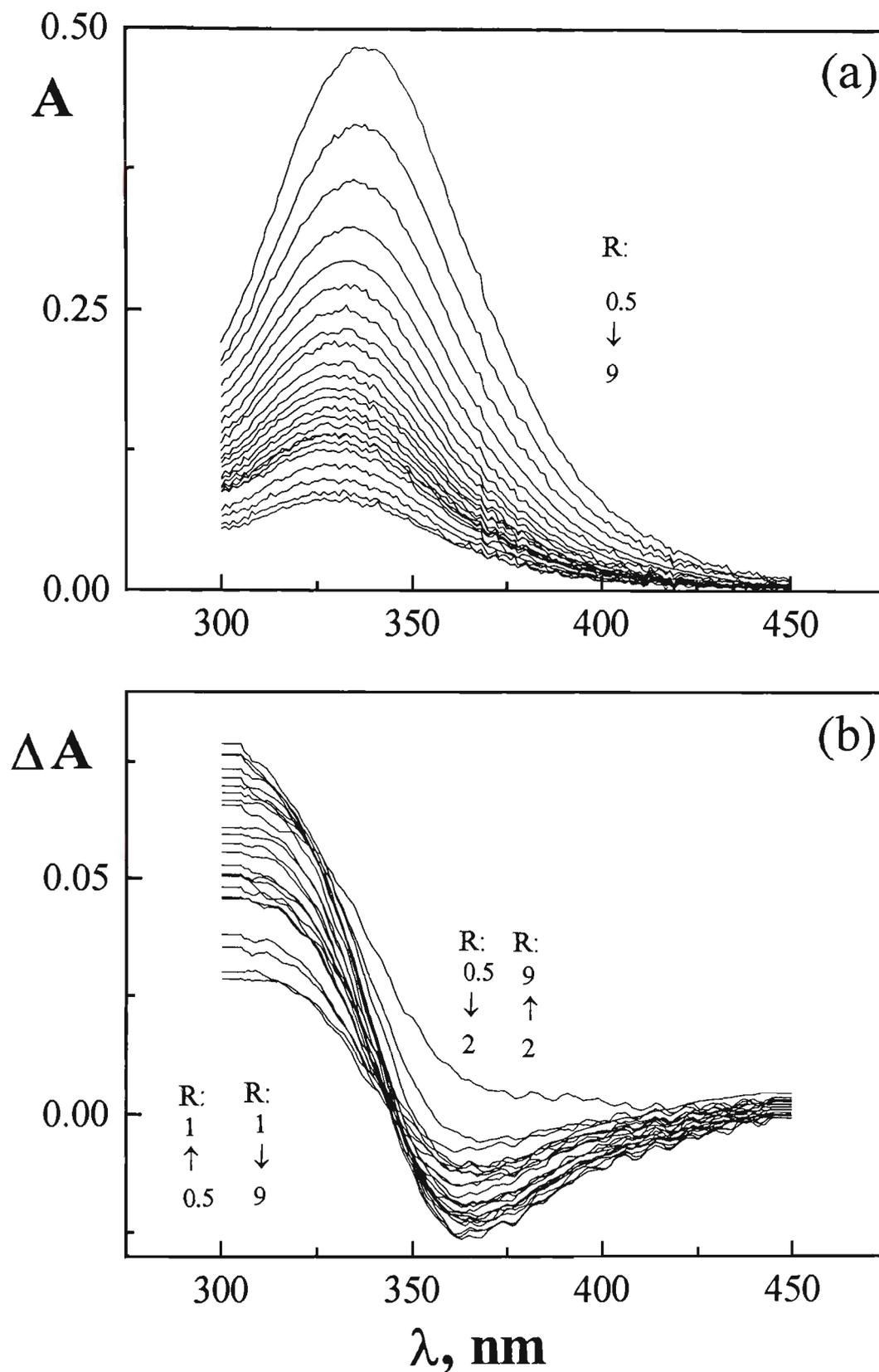
Figure 2.12. Continuous variation experiment. System: $Pd_2(bispep)/inosine$. Total concentration $2.5 \times 10^{-4}M$. Molar ratio $R = [Ino]/[Pd]$ range: $0.15...9$.
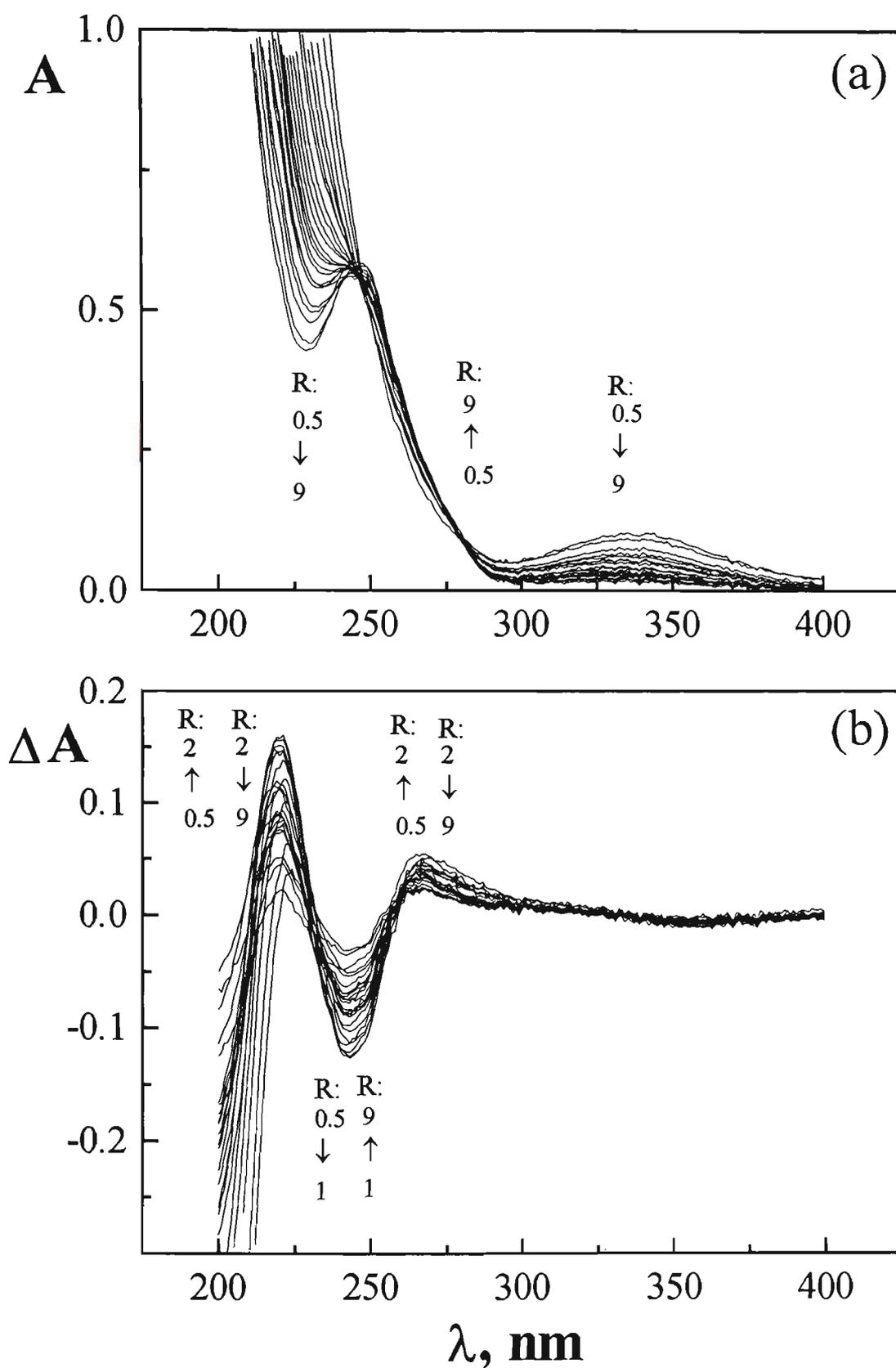
Figure 2.13. *Continuous variation experiment. System: Pd(bispep)/inosine. Total concentration - 5 × 10⁻⁵ M. Molar ratio R = [Ino]/[Pd] range: 0.15...9.*

Table 2.8. Continuous variations experiments.

(a) Observed spectra characteristics

System	λ_{max}	λ_{min}	λ_{iso}
Pd(bmpe)/Guo	252 (>1.7)[a] 265[sh][b] (.15-1.14)	230→225[c] (>1.7)	241(.3-3.8) 234(>3.8) 293(>3.8)
Pd(en)/Guo	260→252 271[sh]	247→223	232(.5-~2) 227(>~2) 296
Pd(bpe)/Guo	258→251 269[sh]	250→225	225 (>~2)
Pd₂(bispep)/Ino[d]	337→327	–	–
	242[sh]→246 (2-9) 335	240[sh]→229 (2-9) 293	251(.5-1.25) 280(>1)

(b) Difference spectra characteristics and established stoichiometries

System	λ_{max}	λ_{min}	λ_{iso}		Stoich.
Pd(bmpe)/Guo	220 263[sh] 297[sh]	251→245 272[sh]	209(.15-.75) 234(.15-.75) 287(.15-.75)	210(>1.7) 227(>1.7) 290(>1.7)	1 : 1
Pd(en)/Guo	218 263[sh] 294[sh]	252→244 276[sh]	232(.15-.5) 239(.5-~2) 231(>~2)		2 : 1
Pd(bpe)/Guo	220 262[sh] 272[sh]	252 269[sh]	–		1 : 1
Pd₂(bispep)/Ino[d]	–	363	325(.5-2) 350(2-4)		2 : 1
	~220 266	242	232[e] 266[e]		2 : 1

[a]Henceforth the value in parentheses is the nucleoside-to-complex molar ratio range, for which the observed wavelength is valid. The ratio range is specified if λ_{max}, λ_{min} or λ_{iso} is not constant over the entire range. [b]Henceforth [sh] stands for shoulder. [c]An arrow is showed if λ_{max} or λ_{min} shifts steadily over the entire or specified ratio range; λ is shown from smaller to larger values of molar ratio. [d]Top row corresponds to the experiment with the total concentration of components 2.5×10^{-4}, bottom row corresponds to the experiment with the total concentration of components 5×10^{-5}.

For the **en, bpe** and **bmpe** systems with guanosine the observed spectra represent families with maximum and minimum absorbance at ~ 255 and ~ 230 nm, respectively, and shoulders at ~ 265 nm (Figs. 2.9 – 2.11, parts (a)). With increasing nucleoside-to-complex ratio, the values of λ_{min} and λ_{max} undergo hypsochromic shifts. The observed spectra exhibit one or more isosbestic points, these vary with ratio range.

For the **en, bpe** and **bmpe** systems with guanosine the difference spectra have absorbance maxima at ~ 220 nm, minima at ~ 250 nm, and shoulders at ~ 260, 270 and 295 nm. With increasing nucleoside-to-complex ratio, the values of λ_{min} undergo hypsochromic shifts in **en** and **bmpe** systems (Figs. 2.9 – 2.11, parts (b)). Isosbestic points are not detected in the **bpe** experiment. In the experiments for the **en** and **bmpe** systems, difference spectra are divided into three and two groups respectively, each having their own isosbestic points. It is interesting to point out that, for the **bmpe** carrier ligand, isosbestic points observed for two ratio ranges have different absorbance values but are located at very close wavelength values. Combined with the inspection of the observed spectra, the appearance of isosbestic points for the **en** and **bmpe** systems suggests that different reactions are taking place at different ratio ranges, consistent with stepwise complex processes.

Regarding the use of isosbestic points a certain degree of caution needs to be exercised. It is possible, for example, that a system in which one-to-one complexation is occurring may exhibit more than one isosbestic point or, if the spectra of initial component(s) and product(s) do not superimpose, no isosbestic points. If there is an intersection of any two members of a spectral family and if only one reaction does indeed occur, then all spectra should go through this point. The corollary is that if two or more, but not all, spectra

pass through a common point then more than one reaction takes place; that is, more than one complex is formed or side reactions are significant. Thus, in the case of stepwise complex formation, when simultaneous existence of successive complexes is assumed, the presence or absence of an isosbestic point gives valuable information about the stoichiometry of a product [33]. However, in some circumstances an isosbestic point may be present even when the system contains more than one complex. Therefore, the presence or absence of an isosbestic point is not definitive and can only provide supporting evidence for a particular complex formation. Thus, isosbestic points should be used only to complement other methods and observations.

The $Pd_2(bispep)^{4+}$ system, being binuclear, represents a different type of molecular system, as compared to those described above. Continuous variation experiments for this system with inosine have been carried out at two concentration levels (Table 2.8). Carrying out experiments at different concentration levels allows an analysis of the system in different spectral regions since the components have significantly different extinction coefficients in different regions. In this system, inosine absorbs most strongly in the 200 to 300 nm range, whilst $Pd_2(bispep)^{4+}$ absorbs most strongly in the range 300 to 400 nm. The **bispep** experiment, carried out with the total concentration of components equal to 2.5×10^{-4} M, allows an analysis of the system in the range 300 – 450 nm. The observed spectra represent a family with an absorbance maximum at ~ 325 nm, undergoing a hypsochromic shift (Fig. 2.12a). There are no isosbestic points, because inosine does not absorb in this region. Difference spectra have an absorbance minimum at ~ 363 nm and two isosbestic points, characteristic for two ratio ranges (Fig. 2.12b). The **bispep** experiment, carried out with the total concentration of components equal to 5×10^{-5} M, allows an analysis of the system in the range 200 – 300

nm. The observed spectra (Fig. 2.13a) have an atypical appearance compared to spectra described above for the **en, bpe** and **bmpe** systems (Figs. 2.9 – 2.9, parts (a)). Thus they exhibit maximum and minimum absorbance at ~ 245 and ~ 235 nm respectively. These begin to emerge for samples with R > 2, first as shoulders, and then as clearly defined maxima and minima, undergoing bathochromic and hypsochromic shifts respectively as R continues to increase. Additional extremum features are observed at ~ 335 and ~ 293 nm for all samples. There are two quite sharp isosbestic points, characteristic for two ratio ranges. Difference spectra produce absorbance maxima at ~ 220 and 265 nm and an absorbance minimum at ~ 240 nm with not very sharp isosbestic points at ~ 232 and ~ 266 nm (Fig. 2.13b).

2.3.2.3 Continuous variation experiments, Job's plots

From the above experiments Job's plots have been constructed for all systems under study. To establish stoichiometry, plots have been constructed at λ_{min} and λ_{max}, where the absorbance difference between samples is most pronounced, as well as at several selected wavelengths to check for wavelength-dependence of the determined stoichiometry. Fig. 2.14 shows representative Job's plots at selected wavelengths for each system.

Data from the **en** experiment represented as a Job's plot corresponds to a 2:1 adduct (Fig. 2.14b), as expected [12]. A plateau appears at nucleoside mole fraction 0.4 – 0.5, probably indicating the intermediate formation of a 1:1 adduct, anticipated for the stepwise process.
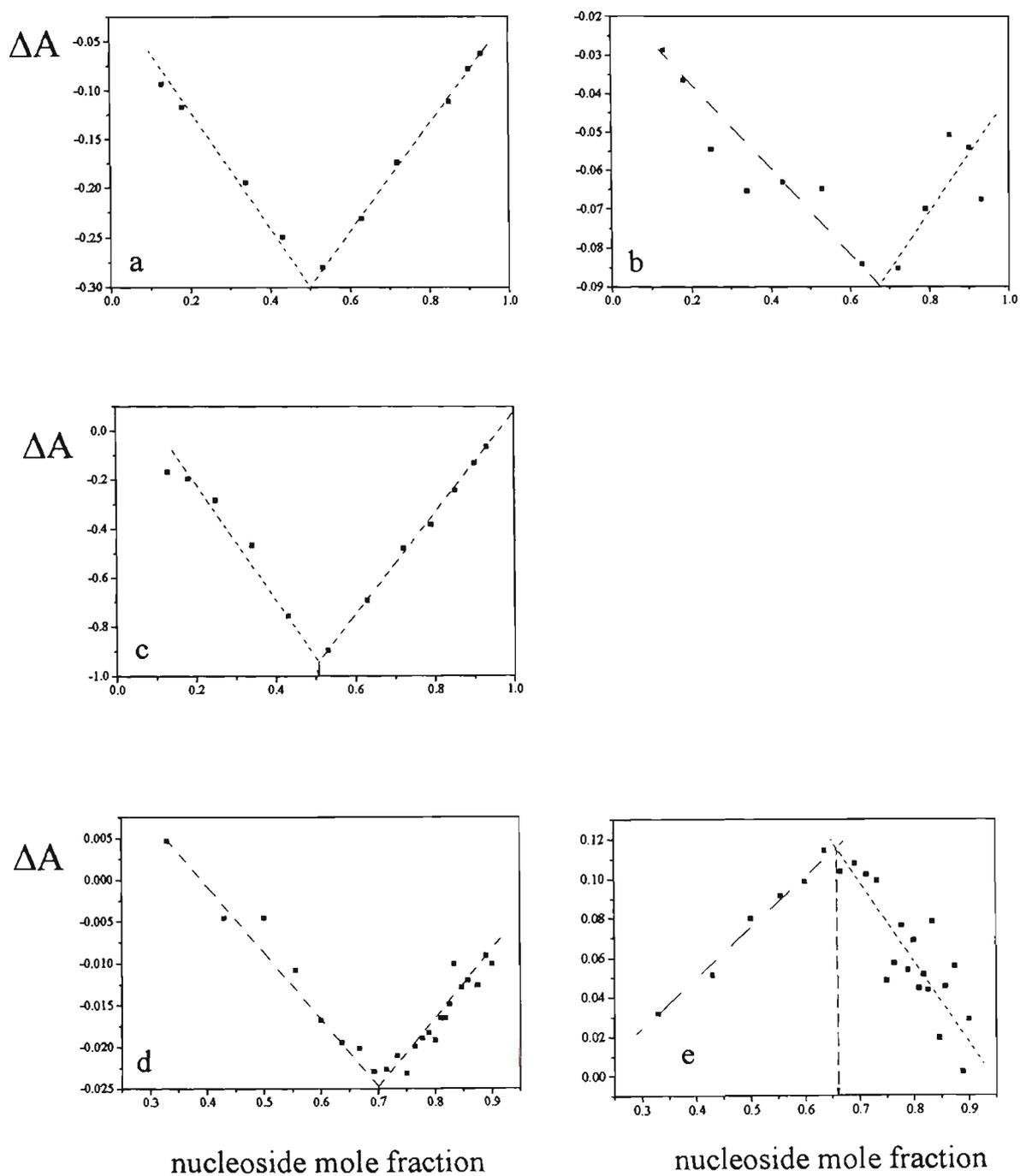
Figure 2.14. Continuous variation experiments, Job's plots.

a: Pd(bmpe)/guanosine, λ = 250 nm. b: Pd(en)/guanosine λ = 275 nm.

c: Pd(bpe)/guanosine, λ = 252 nm.

Total concentration - $10^{-4} M$. Molar ratio R = [Guo]/[Pd] range: 0.15...13.

d: Pd_2(bispep)/inosine, λ = 375nm. Total concentration - 2.5 × $10^{-4} M$.
Molar ratio R = [Ino]/[Pd] range: 0.5...9.

e: Pd_2(bispep)/inosine, λ = 225 nm. Total concentration - 5 × $10^{-5} M$.
Molar ratio R = [Ino]/[Pd] range: 0.5...9.

Contrary to the stoichiometry determined from the crystal structure of the Pd(bpe)-nucleobase system [98], data from the **bpe** experiment represented as a Job's plot produces the ligand mole fraction value corresponding to a 1:1 adduct (Fig. 2.14c). This could be an example of a system where the predominant species in solution differs from the crystallisable form, which may be less soluble. Thus, results obtained in this investigation may be complementary to those from X-ray analysis.

For the **bmpe** system the Job's plot indicates 1:1 stoichiometry (Fig. 2.14a). Together with the observation of isosbestic points at molar ratio range 0.15 – 0.75, this is indicative of coordination at the low end of molar ratio scale, as has been discussed in the previous section (2.3.1.4).

With respect to the intersection of the linear parts of Job's plots at the mole fraction of nucleoside equal to 0.5, it must be noted that this may also correspond to the formation of a 2:2 adduct. Such a possibility arises if ligands undergo base stacking [119]. In these experiments it has been ensured that at the concentrations used and under the experimental conditions employed (pH, T, etc) nucleobases alone do not undergo base stacking, i.e. the linearity between a monitored parameter, absorbance or chemical shift, and concentration is not compromised. However, it is possible for nucleobases to undergo stacking if exposed to metal species, since metal ions are good initiators of base stacking [119].

Both experiments for the **bispep** system yield a 2:1 nucleoside-to-complex ratio, corresponding to a 1:1 nucleoside-to-metal ratio (Fig. 2.14(d,e)). This suggests that the steric bulk of the piperazine ligand may possibly exercise steric control similar to that of **bmpe** and described in section 2.3.1.

2.3.2.4 Formation constants

Attempts have been made to calculate formation constants of Pd(carrier ligand)(nucleoside) systems from the molar ratio experiments. Unfortunately, the obtained results are contradictory and require further investigations. Hence, these are not presented here. However, prior to these experiments an analytical treatment of stepwise complex formation was carried out by the author which, in principle, allows the determination of formation constants without any of the omissions or assumptions described in section 2.1.1.2. This derivation is considered to be of sufficient novelty to be included in this work (Appendix II).

2.3.3 Overview

Experiments described in this chapter have been designed and carried out in order to investigate the relationship between the carrier ligand steric bulk and the binding of metal complexes to nucleic acid constituents. An attempt has been made to determine the extent to which carrier ligand steric bulk may control adduct stoichiometry, up to the complete preclusion of coordination, or the extent to which steric bulk may destabilise a system under certain conditions, e.g. stoichiometry. Thus it has been demonstrated that with judicious ligand design experiments can be constructed which have the potential for exploring such steric demands. However, a larger set of complexes with a more varied carrier ligand profile is required to further pursue this goal. Another possible future direction for this research is to proceed to combine such experimental results with the theoretical studies described in the following chapters. Though beyond the scope of this project, this may include relating theoretically quantified steric parameters to thermodynamic and kinetic data.

2.4 REFERENCES

1. Sigel, H.; Song, B. In *Interactions of Metal Ions with Nucleotides, Nucleic Acids, and Their Constituents;* Sigel, H. (Ed.); Marcel Dekker: New York, **1996**, 135-205.

2. Sorokin, V.A.; Valeev, V.A.; Gladchenko, G.O.; Sysa, I.V.; Blagoi, Y.P.; Volchok, I.V. *J. Inorg. Biochem.* **1996**, *63*, 79.

3. den Hartog, J.H.J.; van den Elst, H.; Reedijk, J. *J. Inorg. Biochem.* **1984**, *21*, 83.

4. Mansy, S.; Rosenberg, B.; Thomson, A.J. *J. Am. Chem. Soc.* **1973**, *95*, 1633.

5. Inagaki, K.; Kidani, Y. *J. Inorg. Biochem.* **1979**, *11*, 39.

6. Umapathy, P.; Harnesswala, R.A.; Dorai, C.S. *Polyhedron* **1985**, *4*, 1595.

7. Inagaki, K.; Kidani, Y. *Inorg. Chim. Acta* **1983**, *80*, 171.

8. Umapathy, P.; Harnesswala, R.A. *Polyhedron* **1983**, *2*, 129.

9. Inagaki, K.; Kuwayama, M.; Kidani, Y. *J. Inorg. Biochem.* **1982**, *16*, 59.

10. Lippert, B.; Schollhorn, H.; Thewalt, U. *Inorg. Chim. Acta* **1992**, *198-200*, 723.

11. Inagaki, K.; Kidani, Y. *Inorg. Chim. Acta* **1984**, *92*, L9.

12. Hohmann, H.; Hellquist, B.; van Eldik, R. *Inorg. Chem.* **1992**, *31*, 345.

13. Scovell, W.M.; O'Connor, T. *J. Am. Chem. Soc.* **1977**, *99*, 120.

14. Mital, R.; Ray, K.S.; Srivastava, T.S.; Bhattacharya, R.K. *J. Inorg. Biochem.* **1986**, *27*, 133.

15. Mansuri-Torshizi, H.; Srivastava, T.S.; Chavan, S.J.; Chitnis, M.P. *J. Inorg. Biochem.* **1992**, *48*, 63.

16. Mital, R.; Srivastava, T.S.; Parekh, H.K.; Chitnis, M.P. *J. Inorg. Biochem.* **1991**, *41*, 93.

17. Inagaki, K.; Kidani, Y. *Inorg. Chem.* **1980**, *46*, 35.

18. Paul, A.K.; Mansuri-Torshizi, H.; Srivastava, T.S.; Chavan, S.J.; Chitnis, M.P. *J. Inorg. Biochem.* **1993**, *50*, 9.

19. Odani, A.; Shimata, R.; Masuda, H.; Yamauchi, O. *Inorg. Chem.* **1991**, *30*, 2133.

20. Kleinwachter, V.; Zaludova, R. *Chem.-Biol. Interactions* **1977**, *16*, 207.

21. O'Connor, T.; Scovell, W.M. *Chem.-Biol. Interactions* **1979**, *26*, 227.

22. Kidani, Y. *Yakugaku Zasshi* **1985**, *105*, 909.

23. Mansuri-Torshizi, H.; Mital, R.; Srivastava, T.S.; Parekh, H.K.; Chitnis, M.P. *J. Inorg. Biochem.* **1991**, *44*, 239.

24. Inagaki, K.; Kidani, Y. *Bioinorg. Chem.* **1978**, *9*, 157.

25. Pillai, C.K.S.; Nandi, U.S. *Biopolymers* **1978**, *17*, 709.

26. Arpalahti, J.; Lehikoinen, P. *Inorg. Chim. Acta.* **1989**, *159*, 115.

27. Struik, A.F.; Zuiderwijk, C.T.M.; van Boom, J.H.; Elding, L.I.; Reedijk, J. *J. Inorg. Biochem.* **1991**, *44*, 249.

28. Ingle, J.D.; Crouch, S.R. *Spectrochemical Analysis;* Prentice-Hall: Englewood Cliffs, **1988**, 386-396.

29. Harris, D.C. *Quantitative Chemical Analysis;* W.H.Freeman & Co: New York, **1991**, 512-521.

30. Job, P. *Anal. Chim.* **1928**, *9*, 113.

31. Hargis, L.G. *Analytical Chemistry. Principles and Techniques;* Prentice-Hall: Englewood Cliffs, **1988**, 424-427.

32. Polster, J.; Lachmann, H. *Spectrometric Titrations;* VCH: Weinheim, **1989**, 11-21.

33. Connors, K.A. *Binding Constants;* Wiley: New York, **1987**, 141-187.

34. Ringbom, A.; Wanninen, E. In *Treatise in Analitical Chemistry, Vol.2;* Koltoff, E. (Ed.); Interscience: New York, **1962**, 461-501.

35. Rose, N.J.; Drago, R.S. *J. Am. Chem. Soc.* **1959**, *81*, 6138.

36. Moriguchi, I.; Kaneniwa, N. *Chem. Pharm. Bull.* **1969**, *17*, 2173.

37. Benesi, H.A.; Hildebrand, G.P. *J. Am. Chem. Soc.* **1949**, *71*, 2703.

38. Kakemi, K.; Sezaki, H.; Suzuki, E.; Nakano, M. *Chem. Pharm. Bull.* **1969**, *17*, 242.

39. Theophanides, T.; Polissiou, M. In *Spectroscopy of Biological Molecules. Theory and Applications - Chemistry, Physics, Biology, and Medicine;* Sandorfy, C.; Theophanides, T. (Eds.); D. Reidel: Dordrecht, **1984**, 291-301.

40. Rochon, F.D.; Melanson, R.; Doyon, M.; Butler, I.S. *Inorg. Chem.* **1994**, *33*, 4485.

41. Uemura, T.; Shimura, T.; Nakanishi, H.; Tomohiro, T.; Nagawa, Y.; Okuno, H. *Inorg. Chim. Acta.* **1991**, *181*, 11.

42. Reily, M.D.; Marzilli, L.G. *J. Am. Chem. Soc.* **1986**, *108*, 6785.

43. Menard, R.; Phan Viet, M.T.; Zador, M. *Inorg. Chim. Acta.* **1987**, *136*, 25.

44. Bose, R.N.; Cornelis, R.D.; Viola, R.E. *J. Am. Chem. Soc.* **1986**, *108*, 4403.

45. Marzilli, L.G.; de Castro, B.; Caradonna, J.P.; Stewart, R.C.; Van Vuuren, C.P. *J. Am. Chem. Soc.* **1980**, *102*, 916.

46. Berners-Price, S.J.; Kuchel, P.W. *J. Inorg. Biochem.* **1990**, *38*, 305.

47. Schenetti, L.; Mucci, A.; Longato, B. *J. Chem. Soc., Dalton Trans.* **1996**, 299.

48. Uchida, K.; Toyama, A.; Tamura, Y.; Sugimura, M.; Mitsumori, F.; Furukawa, Y.; Takeuchi, H.; Harada, I. *Inorg. Chem.* **1989**, *28*, 2067.

49. van Garderen, C.J.; Cornelis, A.; Reedijk, J. *Inorg. Chem.* **1990**, *29*, 1481.

50. Admiraal, G.; Alink, M.; Altona, C.; Dijt, F.J.; van Garderen, C.J.; de Graaf, R.A.G.; Reedijk, J. *J. Am. Chem. Soc.* **1992**, *114*, 930.

51. Bloemink, M.J.; Heetebrij, R.J.; Inagaki, K.; Kidani, Y.; Reedijk, J. *Inorg. Chem.* **1992**, *31*, 4656.

52. Frommer, G.; Mutikainen, I.; Pesch, F.J.; Hillgeris, E.C.; Lippert, B. *Inorg. Chem.* **1992**, *31*, 2429.

53. Frommer, G.; Schollhorn, H.; Thewalt, U.; Lippert, B. *Inorg. Chem.* **1990**, *29*, 1417.

54. Vestues, P.I.; Martin, R.B. *J. Am. Chem. Soc.* **1981**, *103*, 806.

55. Schollhorn, H.; Thewalt, U.; Lippert, B. *Inorg. Chim. Acta* **1985**, *108*, 77.

56. Inagaki, K.; Kidani, Y. *Bioinorg. Chem.* **1978**, *9*, 333.

57. Chu, G.Y.H.; Mansy, S.; Duncan, R.E.; Tobias, R.S. *J. Am. Chem. Soc.* **1978**, *100*, 593.

58. Mansy, S.; Chu, G.Y.H.; Duncan, R.E.; Tobias, R.S. *J. Am. Chem. Soc.* **1978**, *100*, 607.

59. Reedijk, J. In *NMR Spectroscopy in Drug Research*; Jaroszewski, J.W.; Schaumburg, K.; Kofod, H. (Eds.); Munksgaard: Copenhagen, **1988**, 341-54.

60. Granot, J.; Fiat, D. *J. Am. Chem. Soc.* **1977**, *99*, 70.

61. Wang, S.M.; Li, N.C. *J. Am. Chem. Soc.* **1966**, *88*, 4592.

62. Jordan, F.; McFarquhar, B.Y. *J. Am. Chem. Soc.* **1972**, *94*, 6557.

63. Kan, L.S.; Li, N.C. *J. Am. Chem. Soc.* **1970**, *92*, 4823.

64. Shimokawa, S.; Fukui, H.; Sohma, J.; Hotta, K. *J. Am. Chem. Soc.* **1973**, *95*, 1777.

65. Theophanides, T. (Ed.) *Fourier Transform Infrared Spectroscopy;* Kluwer: Dordrecht, **1984**.

66. Buijs, H. In *Fourier Transform Infrared Spectroscopy;* Theophanides, T. (Ed.); Kluwer: Dordrecht, **1984**, 43-54.

67. Theophanides, T. In *Fourier Transform Infrared Spectroscopy;* Theophanides, T. (Ed.); Kluwer: Dordrecht, **1984**, 105-24.

68. Manfait, M.; Theophanides, T. In *Spectroscopy of Biological Molecules;* Clark, R.J.H.; Hester, R.E. (Eds.); Wiley: Chichester, **1986**, 311-46.

69. Manfait, M.; Theophanides, T.; Alix, A.J.P.; Jeannesson, P. In *Spectroscopy of Biological Molecules;* Sandorfy, C.; Theophanides, T. (Eds.); D.Reidel: Dordrecht, **1984**, 113-36.

70. Theophanides, T.; Tajmir-Riahi, H.A. In *Spectroscopy of Biological Molecules;* Sandorfy, C.; Theophanides, T. (Eds.); D.Reidel: Dordrecht, **1984**, 137-52.

71. LeRoy, A.F.; Wehling, M.L.; Sponseller, H.L.; Friauf, W.S.; Solomon, R.E.; Dedrick, R.L. *Biochem. Med.* **1977**, *18*, 184.

72. McGahan, M.C.; Tyczkowska, K. *Spectrochim. Acta* **1987**, *42B*, 665.

73. Macquet, J.-P.; Theophanides, T. *Biochim. Biophys. Acta* **1976**, *442*, 142.

74. Macquet, J.-P.; Theophanides, T. *AA Newsletter* **1975**, *14*, 23.

75. Wilson, C.L.; Wilson, D.W.; Strouts, C.R.N. (Eds.) *Comprehensive Analytical Chemistry, XXI;* Elsevier: Amsterdam, **1964**, 154-163.

76. Katayama, S.; Tamamushi, R. *Bull. Chem. Soc. Japan.* **1970**, *43*, 2354.

77. Warren, R.M.L.; Lappin, A.G.; Tatehata, A. *Inorg. Chem.* **1992**, *31*, 1566.

78. Bhattacharya, B.; Dey, A.K. *Curr. Sci.* **1945**, *14*, 69.

79. Dey, A.K. *Curr. Sci.* **1946**, *15*, 24.

80. Gelb, R.I.; Schwartz, L.M.; Murray, C.T.; Laufer, D.A. *J. Am. Chem. Soc.* **1978**, *100*, 3553.

81. Dey, A.K. *Nature* **1946**, *158*, 95.

82. Nawrot, C.F.; Veis, A. *J. Am. Chem. Soc.* **1970**, *92*, 3903.

83. Zmikic, A.; Cvrtila, D.; Pavlovic, D.; Murati, I.; Reynolds, W.; Asperger, S. *J. Chem. Soc., Dalton Trans.* **1973**, 1284.

84. Gutmann, F. *Electrochim. Acta* **1966**, *11*, 1163.

85. Gutmann, F.; Keyzer, H. *Electrochim. Acta* **1967**, *12*, 1255.

86. Tu, A.T.; Friederich, C.G. *Biochemistry* **1968**, *7*, 4367.

87. Csaszar, J.; Kiss, L. *Acta Phys. et Chem.* **1980**, *26*, 151.

88. Roos, I.A.G.; Thomson, A.F.; Eagles, J. *Chem.-Biol. Interactions* **1974**, *8*, 421.

89. Kagemoto, A.; Takagi, H.; Naruse, K. *Thermochim. Acta* **1991**, *190*, 191.

90. Huang, H.; Zhu, L.; Reid, B.R.; Drobny, G.P.; Hopkins, P.B. *Science* **1995**, *270*, 1842.

91. Tullius, T.D.; Ushay, H.M.; Merkel, C.M.; Caradonna, P.; Lippard, S.J. In *Platinum, Gold, and Other Metal Chemotherapeutic Agents*; Lippard, S.J. (Ed.); ACS: Washington D.C. **1983**, 51-74.

92. Sherman, S.E.; Lippard, S.J. *Chem. Rev.* **1987**, *87*, 1153.

93. McCormick, B.J.; Jaynes, E.N.; Kaplan, R.I. In *Inorganic Syntheses, V.XIII*; Cotton, F.A. (Ed.); McGraw-Hill: New York, **1972**, 216-8.

94. Baker, A.T.; Crass, J.; Kok, G.B.; Orbell, J.D.; Yuriev, E. *Inorg. Chim. Acta.* **1993**, *214*, 169.

95. Mann, F.G.; Watson, H.R. *J. Chem. Soc.* **1958**, 2772.

96. Dhara, S.C. *Indian J. Chem.* **1970**, *8*, 193.

97. Abrams, M.J. In *The Chemistry of Antitumour Agents;* Wilman, D.E.V. (Ed.); Blackie: Glasgow, **1990**, 331-41.

98. Marcelis, A.T.M.; Korte, H.-J.; Krebs, B.; Reedijk, J. *Inorg. Chem.* **1982**, *21*, 4059.

99. Watkins, P.; Yuriev, E. *AA Instruments at Work* **1995**, *118*, 1.

100. Watkins, P. *ADL News* **1993**, *78*.

101. ORIGIN, Technical Graphics and Data Analysis in Windows™. Yang, C.P.; Version 2. Northhampton, MA: MicroCal Software, Inc. **1992**.

102. den Hartog, J.H.J.; Salm, M.L.; Reedijk, J. *Inorg. Chem.* **1984**, *23*, 2001.

103. Orbell, J.D.; Marzilli, L.G.; Kistenmacher, T.J. *J. Am. Chem. Soc.* **1981**, *103*, 5126.

104. Kistenmacher, T.J.; Orbell, J.D.; Marzilli, L.G. In *Platinum, Gold and other Metal Chemotherapeutic Agents;* Lippard, S.J. (Ed.); ACS: Washington D.C. **1983**, 191-208.

105. Lippert, B. *Progr. Inorg. Chem.* **1989**, *37*, 1.

106. Preut, H.; Frommer, G.; Lippert, B. *Acta Cryst., C* **1991**, *47*, 852.

107. Schollhorn, H.; Randashl-Sieber, G.; Muller, G.; Thewalt, U.; Lippert, B. *J. Am. Chem. Soc.* **1985**, *107*, 5932.

108. Sherman, S.E.; Gibson, D.; Wang, A.H.-J.; Lippard, S.J. *J. Am. Chem. Soc.* **1988**, *110*, 7368.

109. Admiraal, G.; van der Veer, J.L.; de Graaf, R.A.G.; den Hartog, J.H.J.; Reedijk, J. *J. Am. Chem. Soc.* **1987**, *109*, 592.

110. Orbell, J.D.; Wilkowski, K.; de Castro, B.; Marzilli, L.G.; Kistenmacher, T.J. *Inorg. Chem.* **1982**, *21*, 813.

111. Lippard, S.J. In *Bioinorganic Chemistry. An Inorganic Perspective of Life;* Kessissoglou, D.P. (Ed.); Kluwer: Dordrecht, **1995**, 131-40.

112. Bloomfield, V.A.; Crothers, D.M.; Tinoco, I. *Physical Chemistry of Nucleic Acids;* Harper & Row: New York, **1974**.

113. Williams, P.A.M.; Baran, E.J. *J. Inorg. Biochem.* **1992**, *48*, 15.

114. Onori, G. *Biophys. Chem.* **1987**, *28*, 183.

115. Fiskin, A.M.; Beer, M. *Biochemistry* **1965**, *4*, 1289.

116. Frantz, B.; O'Halloran, T. In *Metal-DNA Chemistry;* Tullius, T.D. (Ed.); ACS: Washington, DC, **1989**, 97-105.

117. Hartley, F.R. *The Chemistry of Pt and Pd;* Applied Science: London, **1973**.

118. Haring, U.K.; Martin, B.R. *Inorg. Chim. Acta.* **1983**, *78*, 259.

119. Rifkind, J.M.; Eichhorn, G.L. *Inorg. Chem.* **1972**, *94*, 6526.

Chapter 3

Structural studies of platinum complex – nucleic acid interactions using molecular modelling techniques

3.1 INTRODUCTION

3.1.1 Preamble

Drug–biomolecular interactions and associated structural features represent an important area of research. Anti-tumour agents such as certain platinum complexes [1] are considered to exert their effect by targeting DNA [2,3] rather than proteins. Hence, the characteristic features of cancer cells (uncontrolled cell growth, reduced repair resources) have been successfully exploited over the past 40 years in the development of such drugs which inhibit cell growth through binding to DNA [4].

As a result of the perceived structural regularity of DNA compared to proteins, it is easy to assume that "DNA-active" drugs are less specific than those which interact with proteins. However, a growing body of evidence [5] suggests that the tertiary structure of DNA may also be quite directive of an interaction. This warrants the necessity to explore such interactions in more detail with a view to their application to rational drug design. In this regard, molecular modelling of drug–DNA interactions and, more specifically, metallochemotherapeutics–DNA interactions, represents an emerging field of study which primarily aims to delineate the mechanistic aspects of DNA-affinity and binding [6]. Due to its ability to handle macromolecular biological systems, the most widely used modelling method in this area is molecular mechanics. Thus, the following sections briefly outline the concept of molecular mechanics and its general application to coordination and bioinorganic chemistry with regard to platinum complex–DNA interactions in particular.

3.1.2 Concept of molecular mechanics (MM)

Reviews of *molecular mechanics* (MM) in the literature are numerous, e.g. Ref. [7]. MM approaches chemical problems from the "nuclear structure" point of view. From this standpoint a molecule is regarded as a collection of masses, interacting with each other through harmonic forces (the ball-and-spring model), described by potential energy functions. Any departure of the model from an unstrained geometry will result in an increase in the energy of a molecule.

The fundamental paradigm of MM is that the total energy of the molecular system, E_{tot}, may be represented as a sum of energy terms attributed to bond stretching, E_b, angle bending, E_a, and non-bonded, torsional and electrostatic interactions, E_{vdw}, E_t and E_e respectively:

$$E_{tot} = E_b + E_a + E_t + E_{vdw} + E_e \qquad (3.1)$$

Other energy terms, such as hydrogen-bonding and cross-terms (e.g. stretch-bend) can be included in the equation to account for specific phenomena. The purpose of the MM calculation is to minimise the value of the total energy by the modification of trial coordinates and therefore to determine an optimum structure for a molecule. It is generally agreed that the value of the total energy has no absolute meaning, but provides a valid basis for comparison between configurational isomers and conformations [8].

Bond stretching and *angle bending* are basically described by Hooke's law, which approximates the Morse curve over an appropriate range of values for bond lengths and angles.

Thus:

$$E_b = \frac{1}{2}k_r(r - r_0)^2 \qquad (3.2)$$

$$E_a = \frac{1}{2}k_a(\theta - \theta_0)^2 \qquad (3.3)$$

where k_r and k_a are the force constants, r_0 and θ_0 are unstrained geometric parameters and r and θ are their actual values (bond length and angle respectively).

The *internal rotation* energy E_t is a three-term Fourier series expansion:

$$E_t = \frac{V_1}{2}(1 + \cos\phi) + \frac{V_2}{2}(1 - \cos2\phi) + \frac{V_3}{2}(1 + \cos3\phi) \qquad (3.4)$$

where V_1 are torsional force constants and ϕ is a torsional angle.

Out of a variety of potential functions used to describe *van der Waals interactions*, the Lennard-Jones 6-12 potential is most widely used:

$$E_{vdw} = \varepsilon \times \left[\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^6\right] \qquad (3.5)$$

where ε is a potential well depth, r_0 is the sum of the van der Waals radii of the interacting atoms and r is the distance between them.

Electrostatic interactions may also be represented by various methods. One approach is based on bond dipole moments μ:

$$E_e = \frac{\mu_i \mu_j}{Dr^3} (\cos\chi - 3\cos\alpha_i\cos\alpha_j) \qquad (3.6)$$

where D is the effective dielectric constant of the solvent and parameters in brackets are correcting terms for the angles: χ being the angle made by the bond dipoles and α_i and α_j being the angles made by each of the bond dipoles with the r vector. The alternative approach is to use point charges q according to Coulomb's law:

$$E_e = \frac{q_i q_j}{Dr} \qquad (3.7)$$

Taken together, equations such as 3.1 to 3.7 and parameters such as force constants and unstrained geometrical values define the energy surface of a molecule and are referred to as a **force field** (FF). Force field parameters are assigned to **atom types**, not atoms. Atom type defines the chemical environment of an atom; for example, its hybridisation. The reliability of a force field depends on the choice of both the equations and parameters. Force field parameters are either derived from experimental data (structural or spectroscopic) or from theoretical simulations (*ab initio*). Two principal approaches in obtaining force field parameters are: (i) the development of specialised parameters for the system of interest [9-11] and (ii) the development of universal force fields, in which force constants are based only on the element, its hybridisation, and its connectivity [12]. The former is advantageous in that the results are more accurate, but it is obviously limited in use, because a force field is tailored for a specific system. The later force field is easier to parameterise and is more generally applicable, but the trade-off is a loss of accuracy.

3.1.3 Major problem areas in MM

3.1.3.1 Supplementation of missing parameters

MM calculations require all parameters associated with an input molecule to be available. Three scenarios could be envisaged where a program encounters a missing parameter.

(i) It stops and requests the user to supply the missing parameter. This may require parameters derivation from *ab initio* calculations or a parameter may be fitted to an experimental value. The latter parameters will obviously be more accurate and reliable if obtained from sufficient experimental data. However, frequently a user selects a parameter "by analogy" between a system under study and a system for which a parameter is already available. A program has been written to automate such selection [13].

(ii) Many commercial molecular modelling packages supplement the missing parameters using a built-in algorithm and alert the user to the "default" parameters. The "wild-card" approach to the problem of missing parameters, as it is termed in the MM+ force field of HyperChem [14], is described in greater detail in the Methods section of this chapter. An analogous approach of obtaining such generic parameters for implementation in MM2 and MM3 force fields[1] is described in Ref. [15]. For example, with respect to bond lengths, this approach employs covalent radii of bonded atoms, their electronegativities and Schomaker-Stevenson correction terms in the bond length expression. For the bonding force

---

[1] MM2 force constants have been derived to reproduce structures and energies, MM3 force constants have been derived to reproduce vibrational spectra as well as structures and energies.

constant, the expression is based on the nature of the bonded atoms. Similar rules and expressions are derived for calculating parameters for angle bending, torsional rotation, and van der Waals interactions.

(iii) The third choice is where the user agrees to ignore the missing parameter and the program proceeds with the calculation. More often than not this leads to unrealistic structures and meaningless values of strain energy and should be avoided where it affects the significant parameters. However, it could be used for trial structure optimisation.

3.1.3.2 Finding the global energy minimum

Currently, there is no universal algorithm which allows the global minimum in the multidimensional conformational space to be found without testing all possible conformations. Therefore, several methods have been developed for global minimum energy conformation searching without such testing [16]. It should be borne in mind that these methods, described below, are not limited to application within the framework of molecular mechanics alone.

Databases, such the as Cambridge Structural Database (CSD) [17] and the Brookhaven Protein Databank (PDB) [18] may be exploited in order to obtain conformational distributions for structures of interest.

Systematic search methods. The major structural differences between conformations are reflected in torsion angles. Hence, systematic variation of torsion angles provides a means for conformational sampling. The major disadvantage of this approach is in its combinatorial complexity and the search may become intractable as a system size grows. In some circumstances,

however, constraints may be applied which make such methods more manageable.

Molecular Dynamics (MD) may be utilised for conformational searching. The time-dependent motional behaviour of a molecule is simulated (usually at room temperature) and conformations, chosen at regular time periods, are optimised to the corresponding minimum energy structures. In some cases selected conformations are used for further simulations. The use of elevated temperature [19] or Quenched Dynamics (QD), which is a combination of high temperature MD with energy minimisation [20], are common approaches when high potential energy barriers need to be overcome. However, a molecular system can be trapped in a local minimum if it is optimised directly after QD, so it is advisable to follow a QD simulation with Simulated Annealing (SA) (slow cooling) and Room Temperature Simulation (RTS) [20].

MonteCarlo (MC) methods are based on a random variation of torsion angles. The conformations thus created are minimised and an algorithm, comparing them with other conformations and validating them against a set of predefined options, is used to determine whether they should be accepted. "Improved" MC methods were shown to be superior to MD methods in generating conformations in regard to not getting trapped in local minima [21].

Distance geometry methods which generate conformations by varying all pairwise atomic distances in a molecule are particularly suited for macromolecular problems [22].

3.1.3.3 Solvent contribution

This is one of the most important issues in computational chemistry and molecular modelling in general and their applications to biomolecular systems in particular [23]. It is a problem common to the majority of MM methods. In general, solvents have rarely been explicitly included in drug–DNA modelling studies [4]. The most straightforward way to simulate solvent media is to apply the Coulombic electrostatic term $E = c \times r$, where c is a distance-dependent dielectric constant.

3.1.4 Bioinorganic molecular mechanics

A recent review [24] gives an excellent account of the history, achievements and frustrations of the MM of transition metal bioinorganic systems, a subset of MM studies of coordination compounds. Inorganic and bioinorganic FFs are usually developed by modification of organic FFs [24]. Such development is based on the assumption that the geometry of an organic backbone (ligand) is not significantly altered upon coordination to a metal. Thus, it is not considered necessary to develop an entirely new FF.

3.1.4.1 Extending existing FFs to the treatment of metal complexes

Several methods are available for extending existing FFs to the treatment of metal complexes.

Valence force field (VFF) method. The metal atom is treated in a similar way to all other atoms, i.e. parameters are added for all M-L bonds, M-L-X and L-M-L angles, M-L-X-X and L-M-L-X torsions and M...X non-bonded

contacts (M, metal; L, ligand donor atom; X, any other atom). Provision for different atom types for donor atoms depending on atom geometry (e.g. *cis* and *trans*, radial and axial) has to be made. Several approaches are in use to treat angles around the metal center [25]. The rotational barriers around M-L bonds are usually low, allowing L-M-L-X torsion parameters to be omitted, as well as those describing non-bonded interactions involving the metal. The latter have been shown not to significantly affect the final conformations.

Point-on-a-sphere (POS) method is based on an assumption that the donor atoms tend to locate themselves in space (on a sphere) so as to minimise repulsions between them. Therefore, all L-M-L angle terms are replaced by L...L van der Waals terms. This approach is better than VFF in requiring fewer parameters and usually yields the minimal energy conformations which correspond to appropriate geometries. Exceptional deviation (from expected geometries) occurs in cases of square-planar geometry complexes, where minimum repulsion of four donor atoms would result in a tetrahedral arrangement. Platinum(II) complexes are one of the most important examples in this regard. For such systems measures are required to enforce a planar arrangement [26]. These include: tethering metal and donor atoms to a plane, bonding the metal atom to two dummy atoms (hydrogens or electron lone pairs) above and below the coordination plane, or using an out-of-plane bending restraint for the metal atom.

Ionic method is based solely upon using M...L non-bonded terms and is extremely sensitive to charges and vdW parameters. Such FFs are easy to parameterise, but their accuracy is questionable [27].

3.1.4.2 Parameterisation of bioinorganic FFs

The parameters required for the description of the coordination sphere are derived in the same way as parameters for organic FFs. However, bioinorganic MM calculations are complicated because parameters are required for a range of different metals in all their different states (oxidation, spin, coordination geometry). Parameterisation of bioinorganic force fields is further restricted by a paucity of high-resolution crystal structures, low metal-ligand vibrational frequencies (entangling the process of deriving force constants from infra-red spectra), complexity of *ab initio* calculations of metal centers and strong electronic effects complicating their analysis.

In *parameter derivation from crystal structures* the initial metal parameters are commonly added to a well-proven force field (i.e. they are initially approximated using available parameters for similar systems). Consequently, they are refined by systematically altering them and fitting the computed geometries to the crystal structures until an agreement is achieved. Such a fit between calculated and experimental structures is usually estimated by establishing the root-mean-squared standard deviation (RMS) between the coordinates of the two structures, by comparing the internal coordinates or some other geometric properties of interest. Force fields so derived have been shown to be a good tool for predicting solid-state geometries [24].

3.1.4.3 Electrostatic interactions

The major problem in simulating electrostatic interactions in coordination complexes is related to the calculation of partial charges around the metal center(s). Three main methods to derive such charges are as follows.

- *Ab initio* calculations. Though very accurate these methods have tremendous computer appetites.

- Using well-established point charges for the ligand and optimising/fitting the metal charge(s) [28].

- Charge equilibration [29,30] which allows charges as functions of geometry to be obtained.

The above methods, when applied to coordination complexes, are not as accurate or reliable as for organic compounds [31-33]. Given these difficulties, it should be emphasised that even if electrostatic terms are not explicitly included into the FF, it does not mean that electrostatic interactions are neglected; they are simply intermingled into other functions [24]. Therefore, the question of whether charges are actually required for bioinorganic force fields is currently being actively debated in the molecular mechanics community [24].

3.1.4.4 Conformational searching

Up until recently it has not been considered necessary to carry out conformational searches in the area of bioinorganic modelling [24]. This was for two reasons. Firstly, the metal reduces the number of conformations by constraining the geometry of ligand. Secondly, many metal-ligand systems are cyclic which complicates methods using internal coordinate frames (torsion angles). The former reason has led to the common practice of individually drawing and minimising all possible conformations instead of proper conformational searching. The latter problem could be solved by employing recently developed methods for the conformational searching of cyclic molecules, e.g. torsional flexing [34].

However, it is desirable for conformational searching to be included in

bioinorganic molecular modelling experiments, to simplify preparation of starting geometries for minimisation and, more importantly, to ensure that the conformational space is more thoroughly covered so that the global minimum has a higher chance of being located.

3.1.5 Development of MM force fields for platinum compounds

Molecular mechanics has been extensively employed in platinum – nucleic acid studies [26,35-54]. This is a result of the desire to understand the mode of action of platinum anti-tumour drugs in a situation where experimental structural information for these compounds is insufficient. Most of platinum – nucleic acid molecular modelling studies have used AMBER (a force field developed for proteins and nucleic acids) as their starting point [10,11,40]. In these studies the following issues were investigated:

- structural effects of cisplatin on DNA [35,36,41,42,44,45,47,55,56]
- the existence of monomeric macrochelate complexes between platinum, the nucleobase, and the phosphate [57]
- hydrogen-bonding schemes in cisplatin [47,53,56]
- isomer preference and ligand rotation barriers [51,58]
- stereochemical factors affecting monofunctional and bifunctional binding [43,59]
- solution structures [54,60]
- solvent and counter-ion effects [40]
- relative stabilities of HTH and HTT conformations [40,51]

The MMX force field has also been used for the study of reactants and products of the hydrolysis of cisplatin and substituted $Pt^{II}(en)Cl_2$ complexes, but details of parameterisation were not provided [37].

3.2 METHODS

All modelling studies were carried out using HyperChem modelling package (version 4.0) [14] on IBM compatible computers: AST 486 (33 MHz) and COMPAQ Pentium-100.

3.2.1 HyperChem MM+ functional form and parameters

The MM+ force field of HyperChem is an extension of Allinger's MM2 force field [61]. The main objective of Allinger et al. was to develop a specific method for organic systems that would give accurate results, rather than a generic force field with its associated loss of accuracy. When using the MM2 force field, there is often a problem with missing parameters requiring their development for a specific system. The HyperChem advancement of MM2 introduces a "wild-card approach" which automates decisions on missing parameters, employing a generic technique that considers hybridisation, bond type and standard covalent radii [20]. Thus, "Default" MM+ force field retains Allinger's 1977 MM2 functional form [9], the 1991 MM2 parameters[2] and uses the above scheme to obtain missing parameters. In HyperChem the MM+ functional form is modified for nonbonded cutoffs, periodic boundary conditions and the bond-stretch term (switched from cubic to quadratic form at long range) [20]. Similar to MM2, MM+ internally uses ergs as energy units and millidyne-based units for force constants. But all energetic results are reported in kcal/mol (more traditional units in MM). Therefore, the appropriate conversion factors are required. These are included into the equations below.

---

[2] These parameters are provided directly by Dr. Allinger to HyperCube, Inc.

The standard expression for the potential energy of the molecular system in the MM+ force field is:

$$E_{tot} = E_{Stretch} + E_{Bend} + E_{Tor} + E_{vdW} + E_E + E_{SB} + E_{OOPB} \qquad (3.8)$$

For *bond stretching* MM+ uses the cubic term:

$$E_{Stretch} = 143.88 \sum \frac{1}{2} K_r (r - r_0)^2 \times S \qquad (3.9)$$

$$S = 1 + switch(r - r_0, -\frac{1}{3}CS, -\frac{4}{3}CS)CS(r - r_0) \qquad (3.10)$$

where K_r is the stretch force constant (mdyn/Å), r is the actual and r_0 is the unstrained distance. The default value for the cubic-stretch factor CS in MM2 and MM+ is -2.0. For distances significantly larger than unstrained ones, MM+ applies the switching function, turning the cubic term off.

The *angle bending* term of MM+ is sextic:

$$E_{Bend} = 0.043828 \sum \frac{1}{2} K_\theta (\theta - \theta_0)^2 \times [1 + SF(\theta - \theta_0)^4] \qquad (3.11)$$

where K_θ is the bending force constant (mdyn×Å/rad²), θ is the actual and θ_0 is the unstrained angle. The default value for the scale factor SF in MM+ is 7.0×10^{-8}. MM+ has special sets of bending parameters for 3- and 4-member rings.

The *torsion term* of MM+ explicitly includes 1-, 2- and 3-fold symmetries:

$$E_{Tor} = \sum \frac{V_1}{2}(1 + \cos\phi) + \frac{V_2}{2}(1 - \cos2\phi) + \frac{V_3}{2}(1 + \cos3\phi) \quad (3.12)$$

where V_n (n = 1, 2, 3) are energies of torsional barriers (kcal/mol) and ϕ is a torsion angle. MM+ has special sets of torsional barriers for 4-member rings.

For *van der Waals interactions* MM+, as well as MM2 and MM3, does not use the Lennard-Jones potential, employed in many molecular mechanics calculations. It replaces the twelfth power term for repulsion with the exponential Buckingham potential and combines it with the sixth power term for attraction, forming an "exponential-6" expression:

$$E_{vdW} = \sum \varepsilon_{ij} (2.9 \times 10^5 \exp(-12.5\rho_{ij}) - 2.25\rho_{ij}) \quad (3.13)$$

where $\rho = R_{ij}/r_{ij}^*$, R_{ij} is the distance between atoms of types i and j and r_{ij}^* is the minimum energy separation distance between them:

$$r_{ij}^* = r_i^* + r_j^* \quad (3.14)$$

ε_{ij} is the well depth for two atoms of types i and j:

$$\varepsilon_{ij} = \sqrt{\varepsilon_i + \varepsilon_j} \quad (3.15)$$

r_i^* and ε_i are the vdW radius (Å) and hardness parameter (kcal/mol) respectively of atom i. The MM+ force field uses different equation at short distances ($\leq$ 3.311 Å) and special parameters for C-H interactions.

The *electrostatic interactions* are accounted for in MM+ by defining a set of bond dipole moments:

$$E_E = 14.39418 \; \varepsilon \sum \mu_i \mu_j [\frac{\cos\chi - 3\cos\alpha_i\cos\alpha_j}{r}] \qquad (3.16)$$

where ε is the effective dielectric constant of the media (default value for the gas phase is 1.5), χ is the angle made by the bond dipoles and α_i and α_j are the angles made by each of the bond dipoles with the r vector.

MM+ contains the *cross-term between bond stretching and angle bending* for the angle including atoms i, j and k:

$$E_{SB} = 2.51118 \sum K_{SB}(\theta - \theta_0)_{ikj}[(r - r_0)_{ik} + (r - r_0)_{jk}] \qquad (3.17)$$

where K_{SB} is the stretch-bend constant and other parameters are as defined above.

The *out-of-plane-bending interactions*, E_{OOPB}, account for the tendency of sp^2-hybridised atoms to be coplanar with atoms they are attached to. If such a central atom X is connected to atoms A, B and C, then the angle bending energy calculation uses the parameters for the AYB, AYC and BYC angles, where Y is the projection of atom X on the ABC plane. In addition, the out-of-plane angles XAY, XBY and XCY are computed as well, using the standard angle bending equation with $\theta = 0$ and a special set of K_θ out-of-plane bending constants.

3.2.2 HyperChem AMBER functional form and parameters

The AMBER force field of HyperChem is based on that of Kollman et al. AMBER (Assisted Model Building and Energy Refinement) force field [10,11,62] was developed for simulation of nucleic acids and proteins, with few atom types and parameters for small organic compounds and inorganic systems.

The standard expression for the potential energy of the molecular system in the AMBER force field is:

$$E_{tot} = E_{Stretch} + E_{Bend} + E_{Tor} + E_{vdW} + E_E + E_{IMP} + E_{HB} \qquad (3.16)$$

The AMBER functional forms for *bond stretching*, $E_{Stretch}$, and *angle bending*, E_{Bend}, are quadratic only and are identical to those described in the Introduction to this chapter (eqns **3.2** and **3.3**). The units for bond stretching and angle bending force constants are $kcal/(mol \times A^2)$ and $kcal/(mol \times rad^2)$ respectively.

The energy of *internal rotation*, E_{Tor}, is associated with the tendency of torsional angles to have a certain n-fold symmetry $(n$ – periodicity of the Fourier term) and minima, corresponding to *cis-*, *gauche-* or *trans-*conformations:

$$E_{Tor} = \sum \frac{V_n}{2} [1 + \cos(n\phi - \phi_0)] \qquad (3.17)$$

where V_n are energies of torsional barriers (kcal/mol) and ϕ and ϕ_0 are the torsional and the phase offset angles, respectively.

Out-of-plane-bending is accounted for by the improper torsional energy, E_{Imp}, the functional form of which is identical to the above (eqn **3.19**) but with a separate parameter set.

The AMBER functional form for *non-bonded interactions*, E_{vdW}, is identical to that described in the Introduction (eqn **3.5**). The units for the potential well depth ε (minimum energy separation) are kcal/mol.

The *electrostatic interactions*, E_E, in AMBER are characterised on the basis of classical Coulomb's law, as described in the Introduction (eqn **3.7**). A set scaling of the dielectric constant is used when solvent molecules are included explicitly into the simulation, and a distance dependent scaling is used for the emulation of solvent effects without explicitly adding solvent molecules. Standard atomic charges required for these calculations could be obtained from template or database sources, calculated by quantum mechanical methods or manually added to the molecular structure.

The AMBER force field employs a special term E_{HB} for an explicit description of *hydrogen bonding* interactions as a recognition of the importance of these interactions in biological molecular systems.

$$E_{HB} = \sum \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right] \tag{3.20}$$

C_{ij} and D_{ij} are the coefficients for appropriate donor–acceptor pairs. If a non-bonded atom pair has hydrogen bond parameters then the interaction between them is calculated using the expression for hydrogen bonding, rather than the vdW term (except for 1,4-interactions).

3.2.3 Computational details

Geometry optimisation. The molecular structures were optimised using the conjugate gradient algorithm (Polak-Ribiere) *in vacuo* with the termination condition being RMS gradient < 0.1 kcal/(mol×Å) for an initial optimisation and during a conformational analysis. For a final structure optimisation a RMS gradient < 0.001 kcal/(mol×Å) was employed as termination condition unless otherwise specified. No cutoffs were used, scale factors of 0.5 were set in AMBER for 1,4-nonbonded interactions.

Conformational analysis. The MD and MC procedures have been used to search for the lowest energy conformation and usually produced similar lowest energy structures. The MC search not only allows the lowest energy structure to be found, but also allows the conformational space in the low energy region to be more thoroughly explored than does the MD protocol used here. Hence, the majority of structures have been simulated using the MC search, which was also shown to be computationally more effective.

Input molecular systems were either taken from crystal structures (where available from CSD [17]) or model-built by modifying the crystal structures of related molecular systems. Prior to conformational searching, molecular systems were initially energy minimised to remove residual strain and bad contacts resulting from modified experimental structures and model building. This way MD simulation could be started with atoms having representative velocities. When MD is carried out primarily for purposes of a global minimum search there is no need for regular data collection during simulation at high temperature, as the SA procedure takes the system to the lowest energy. The technique of simulated annealing should, in principle,

converge to the global energy minimum with a "slow" decrease in temperature, but since there is no definite way to establish how "slow" the cooling should proceed, it is often necessary to repeat QD-SA-RTS steps several times to increase chances of arriving at a global energy minimum.

Molecular Dynamics. Ten repeated simulations were run for each molecule, each starting with a structure resulting from the previous run of quenched dynamics simulation. Thus, an overall run of 8 picoseconds of quenched dynamics was carried out for each structure. Molecular dynamics options are summarised in Table 3.1(a,b).

MonteCarlo. MonteCarlo analysis was carried out using the Conformational Search module of ChemPlus (a set of extension modules for HyperChem) [63]. Conformational search options are summarised in Table 3.2.

Table 3.1a. Quenched Dynamics and Simulated Annealing options.

Times (ps)		Temperature (K)	
Heat time	0.2	Starting temperature	0
Run time	0.8	Simulation temperature	600
Cool time	0.5	Final temperature	300
Step size	0.0005	Temperature step	30

Table 3.1b. Room Temperature Simulation options.

Times (ps)		Temperature (K)	
Heat time	0	Starting temperature	–
Run time	0.5	Simulation temperature	300
Cool time	0	Final temperature	–
Step size	0.0005	Temperature step	–

Table 3.2. Conformational search options.

geometric parameters	1) all relevant heavy atom torsion angles varied 2) for ring structures endocyclic torsion angles *torsional flexing* employed
ranges for variation	1) 1 to 8 simultaneous variations 2) ± 60° to 180° for acyclic torsion variation 3) ± 30° to 120° for torsion flexing
search method	1) initial conformations to vary chosen by the *Usage Directed* method 2) acceptance energy criterion for new conformations set to 3 kcal/mol above best
pre-optimisation tests	1) skip if atoms are closer than 0.5 Å 2) skip if torsions within 15° of previous
post-optimisation and duplication tests	1) accept conformation if chiral centers changed 2) energy within 0.05 kcal/mol – duplicated 2) RMS error ≤ 0.25 Å (heavy atoms are matched; ignore hydrogens in RMS fit) – duplicated
limits	1) 200 iterations and/or optimisations 2) 100 lowest energy conformations kept

3.3 RESULTS AND DISCUSSION

3.3.1 Force field options

For the molecular mechanics modelling of platinum complexes of nucleic acid constituents two force field options were considered and tested. The MM+ force field of HyperChem [14] is an advanced implementation of the Allinger MM2 force field [61], enhanced by the missing-parameter supplementation algorithm. This force field was originally developed for organic compounds and is therefore a favourable choice for metal complexes containing organic carrier ligands. However, this choice, when applied to platinum complexes of nucleic acids leads to a large number of missing parameters associated with the nucleoconstituent ligands. The AMBER force field [10,11,62] was designed *specifically* for the molecular mechanics simulation of nucleic acids and proteins and, therefore, it is an alternative method of choice for the systems of interest. The performance of both force fields has been tested in order to develop a suitable method for the systems under investigation.

The force field development was carried out in three stages. Initially, the complex cis-[Pt(NH$_3$)$_2$(9-EtG)$_2$] was modelled, since, for platinum–guanine bis-adducts, a large amount of experimental [64-77] and theoretical [24] structural information is available. This provides a sound basis for parameterisation. The acquired force field was then further developed to enable the simulation of the mono-adduct precursor, cis-[Pt(NH$_3$)Cl(9-EtG)$_2$], carrying one unsubstituted leaving group. Finally, the force field was extended to include the parameters required for the simulation of the monodentate moieties, Pt(A)$_3$, where (A)$_3$ is a triam(m)ine carrier ligand,

coordinated to a range of common nucleobases. The justification for choosing a given system in particular is explained in Chapter 4.

3.3.2 Conserving planarity in platinum complexes

In simulating platinum(II) complexes, maintaining planarity requires special attention [26,37,40,78,79] since the PtD_4 group (where D is a ligand donor atom) acquires tetrahedral geometry if no restraints are used. Several approaches, available for ensuring square-planar geometry of metal complexes, are compared here using the unmodified MM+ force field and cis-[Pt(NH$_3$)$_2$Cl$_2$] as a test molecule. These approaches are: restraining four or more atoms to a plane (option 1), tethering of atoms to particular positions using a function applied to their cartesian coordinates (option 2), and modelling the electronic density of platinum above and below the coordination plane by binding it to two dummy atoms (option 3), e.g. electron lone pairs LP (LP-Pt-LP angle constrained to 180°).

The results presented in Table 3.3 are absolute values of the improper torsion angles depicted in the diagram below.
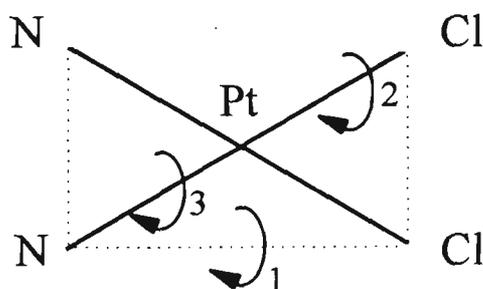
Since the two chlorine ligands are not necessarily equivalent, two different torsion angles N-N-Pt-Cl generally result. A similar situation applies to the amine ligands (torsion angles N-Pt-Cl-Cl). These results are compared in

Table 3.3 with the crystal structure of cis-[Pt(NH₃)₂Cl₂] (refcode CUKRAB from CSD[3]). These investigations tested only the computational feasibility of the approaches for maintaining planarity; their ability to reproduce bond lengths and angles was not tested at this stage.

Table 3.3. Comparison of options to conserve planarity in platinum complexes.

	torsion	CUKRAB (deg)	option 1 (deg)	option 2 (deg)	option 3 (deg)
1	N-N-Cl-Cl	2.4	0.1	0	0.25
2	N-N-Pt-Cl	0.6 1.8	0.0 0.2	2.6 2.8	0.0 2.0
3	N-Pt-Cl-Cl	0.6 1.7	0.1 0.15	2.5 3.7	0.1 2.0

Option 1 (with the restraining force constant set at 105 kcal/(mol×degree²)) allows the retention of a planar platinum atom, but the D-Pt-D angles significantly deviate from the values of 90° expected for a square geometry, necessitating force field parameters describing platinum-centered angles. Option 2 retains planarity *and* maintains square geometry but, since all three cartesian coordinates have to be tethered, too strong a restraint is imposed on the optimisation and the resulting geometry significantly depends on the starting geometry. Furthermore, this method may not be appropriate for some systems because experimental evidence exists for some deviation of the platinum atom from the coordination plane [80]. Hence, this method would apply unrealistic restraints to the calculated structures.

[3]All Cambridge Structural Database [17] entries referred to henceforce are compiled in the Appendix III. They are tabulated in the alphabetical order according to their refcodes. Molecular structures and bibliographic references are provided.

Like options 1 and 2, option 3 gives satisfactory reproduction of experimental planarity and square geometry but, again, difficulty arises with respect to modelling of sterically crowded structures, where the clash between LPs and carrier ligand(s) or nucleobase exocyclic group(s) is highly probable and could lead to structural distortions.

With regard to the above considerations, the development of appropriate force field parameters appears to be the most acceptable option in order to maintain planarity. Two groups of parameters are required in this regard. Firstly, an out-of-plane bending term with parameters for improper torsions, containing platinum as central atom. Secondly, explicit angle parameters associated with the coordination plane are required. Hence, new atom types associated with different donor atoms need to be specified.

3.3.3 Force field selection and parameter development for the modelling of the cis-[Pt(NH$_3$)$_2$(9-EtG)$_2$] complex

For cis-[Pt(NH$_3$)$_2$(9-EtG)$_2$] the results of force field bench-marking against the experimental data (refcode DEGXES [65] from CSD) and the alternative theoretical study [58] are presented in Table 3.4. It would be desirable to compare the results of this study with results from other theoretical investigations of related systems. Unfortunately, in few studies, where bis-nucleobase adducts were simulated [40,51], the atomic coordinates or the values of selected geometric parameters of the models have not been reported. Also, the majority of studies have been focused on Pt-oligonucleotide structures [35,36,38,39,41,48-50,52,56,57]. In some of these studies the values of geometric parameters around the platinum atom have been reported [42,47] and these are summarised in Table 3.5 for comparison.

Table 3.4. Comparison of the crystal structure and geometry-optimised structures for cis-[Pt(NH₃)₂(9-EtG)₂].

atoms	crystal structure [65]	structure from Ref. [58]	unmodified MM+	modified MM+	modified AMBER
Pt-N(11)	2.046(7)	2.037	2.003	2.029	2.029
Pt-N(10)	2.044(6)	2.037	2.004	2.031	2.030
Pt-N(71)	2.022(7)	2.022	1.975	2.019	2.010
Pt-N(72)	2.002(6)	2.020	1.973	2.016	2.010
N(11)-Pt-N(10)	87.8(3)	94.9	87.3	88.0	89.7
N(71)-Pt-N(72)	89.2(3)	82.3	91.9	90.7	89.4
N(11)-Pt-N(71)	91.8(3)	91.3	91.5	91.1	90.6
N(72)-Pt-N(10)	91.3(3)	91.3	88.7	89.2	90.3
Pt-N(7)-C(8)	124.6(6)	126.6	127.8	126.0	125.2
	126.0(6)	126.3	128.0	127.6	125.5
Pt-N(7)-C(5)	128.8(5)	127.7	130.0	130.1	129.7
	128.2(5)	127.6	129.2	128.5	129.4
C(6)-N(1)-C(2)	123.8(7)	124.9	120.8	120.8	124.4
	123.4(7)	124.8	120.7	120.8	124.4
N(2)-C(2)-N(3)	122.9(8)	119.3	118.6	118.5	120.0
	119.9(7)	119.3	118.6	118.6	120.0
N(2)-C(2)-N(1)	114.2(7)	118.1	118.7	118.6	116.8
	115.8(7)	118.1	118.5	118.4	116.8
C(2)-N(3)-C(4)	114.3(7)	113.3	119.7	119.3	113.3
	112.6(7)	113.3	119.6	119.2	113.3
N(9)-C(4)-N(3)	126.9(7)	126.2	126.3	127.4	126.7
	124.8(7)	126.1	126.7	127.8	126.5
N(9)-C(4)-C(5)	105.5(7)	106.3	109.5	108.0	106.7
	107.6(7)	106.4	109.2	107.7	106.4
N(3)-C(4)-C(5)	127.6(7)	127.4	124.2	124.6	127.2
	127.6(8)	127.4	124.02	124.3	127.1

N(7)-C(5)-C(4)	108.6(7)	109.6	110.7	110.4	109.5
	108.9(7)	109.6	110.8	110.4	109.6
N(7)-C(5)-C(6)	133.7(7)	130.3	131.9	132.2	130.8
	131.0(7)	130.3	131.5	131.8	130.7
C(4)-C(5)-C(6)	117.7(7)	120.9	117.4	117.4	119.7
	119.6(7)	119.9	117.7	117.8	119.7
N(1)-C(2)-N(3)	122.8(8)	122.5	122.7	122.9	123.1
	124.2(7)	122.5	122.9	123.0	123.1
O(6)-C(6)-C(5)	127.9(7)	127.7	126.2	126.2	128.3
	129.2(7)	127.8	126.0	126.1	128.2
O(6)-C(6)-N(1)	119.4(7)	120.4	118.7	118.8	119.3
	119.0(7)	120.4	119.0	119.0	119.5
C(5)-C(6)-N(1)	113.6(7)	111.9	115.1	115.0	112.3
	111.8(7)	111.8	115.0	114.8	112.3
N(7)-C(8)-N(9)	110.2(7)	112.6	114.7	112.1	113.1
	112.0(8)	112.7	114.4	111.9	113.1
C(9)-N(9)-C(8)	127.2(7)	127.7	130.0	128.5	129.6
	126.1(8)	128.0	127.3	126.3	128.2
C(9)-N(9)-C(4)	123.7(7)	126.0	127.2	125.9	124.9
	127.3(7)	125.8	128.4	127.4	126.1
C(8)-N(9)-C(4)	109.1(7)	105.9	102.8	105.5	105.5
	105.9(7)	105.9	103.1	105.5	105.7
C(9')-C(9)-N(9)	113.1(8)	109.9	114.5	114.7	112.7
	107.4(9)	110.4	110.9	110.9	110.1
C(8)-N(7)-C(5)	106.6	–	102.3	103.8	105.2
	105.7	–	102.4	103.9	105.1

Units: bond lengths - angstroms, angles - degrees.

Table 3.5. Geometric parameters around the platinum atom in cis-Pt(NH$_3$)$_2$$^{2+}$ complexes with oligonucleotides.

geometric parameter	values from Ref. [42]		values from Ref. [47]	
Pt-N	–		1.99 ± 0.01 2.00 ± 0.01	
N(7)-Pt-NH$_3$	91.8 90.4 90.1	90.7 90.7 89.0	89 89 91	91 91 91
	171.1 176.5 176.4	177.1 176.0 179.1	179 179 176	178 179 175
N(7)-Pt-N(7)	85.3 85.3 89.1		89 89 89	
H$_3$N-Pt-NH$_3$	92.1 93.1 91.8		91 91 90	

Several values for each parameter correspond to different oligonucleotide conformations.
Units: bond lengths - angstroms, angles - degrees.

The structure of the complex cis-[Pt(NH$_3$)$_2$(9-EtG)$_2$] is shown in Fig. 3.1. The atomic coordinates of the optimised structures are compiled in Appendix IV (Tables IV.1). Overlays of the optimised geometries with the experimental structure are shown in Fig. 3.2. The overlays were created based on the best fit between the coordination planes of two structures.

Unmodified MM+. Geometry optimisation was carried out by attaching two electron lone pairs to the platinum atom as described above (option 3 in section 3.3.2). The default parameters generated for this optimisation are summarised in Table 3.6.
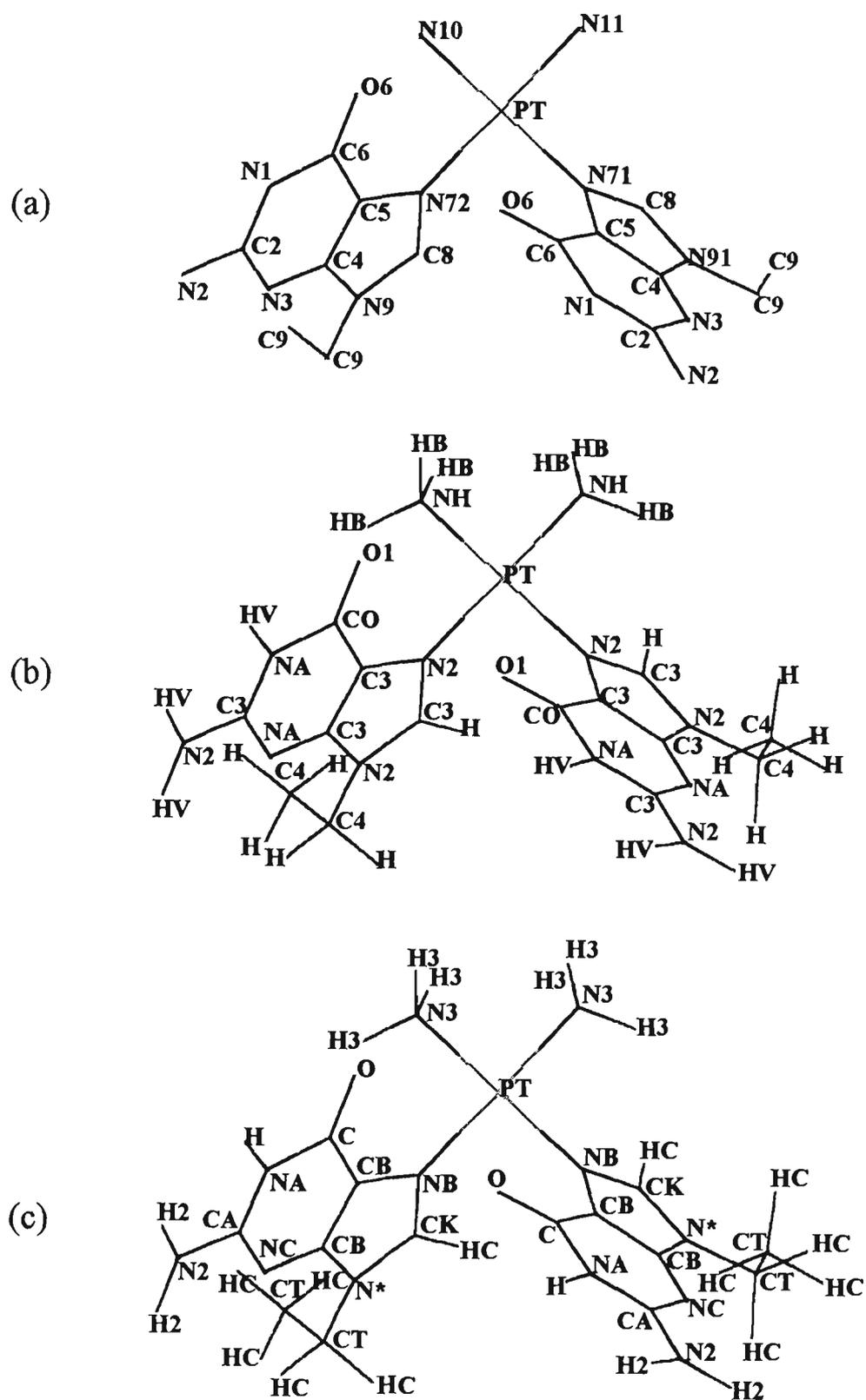
Figure 3.1. *Structure of cis-[Pt(NH₃)₂(9-EtG)₂] complex [65]. (a) Numbering scheme, hydrogen atoms removed for clarity. (b) MM+ force field atom types. (c) AMBER force field atom types.*
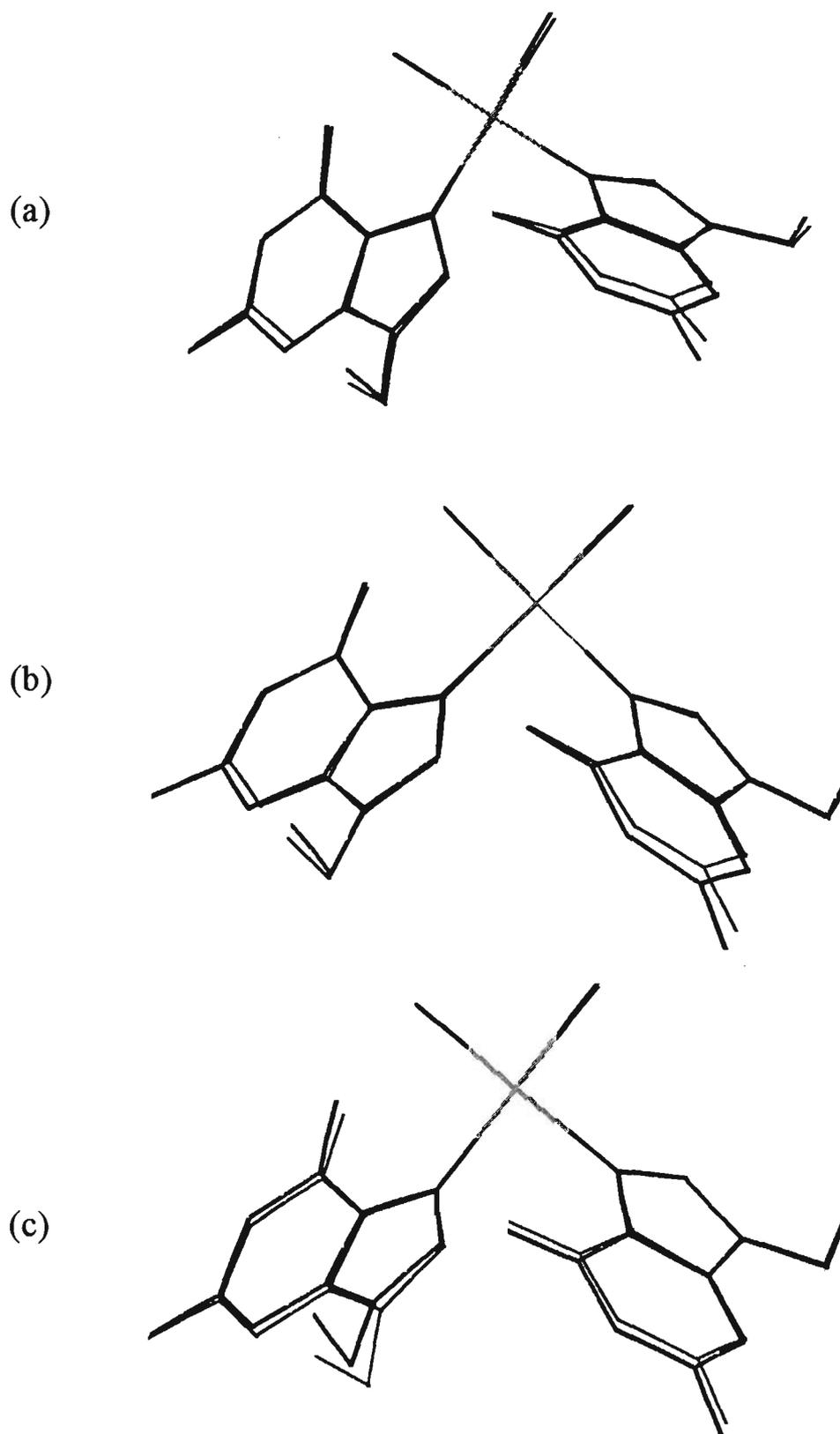
Figure 3.2. *Overlays of the optimised geometries with the experimental structures for cis-[Pt(NH₃)₂(9-EtG)₂] complex [65], experimental structure is shown in bold. (a) Unmodified MM+. (b) Modified MM+. (c) Modified AMBER.*

Table 3.6. Default MM+ parameters for the geometry optimisation of the cis-[Pt(NH₃)₂(9-EtG)₂] complex.

(i) Bond Stretch parameters

bond	r₀(Å)	K_r(mdyne/Å)
PT-NH	2.002	5.0
PT-N2	1.975	5.0
PT-LP	1.620	5.0

(ii) Angle Bend parameters

angle	θ₀(deg)	K_θ(mdyne×Å/rad²)
PT-NH-HB	106.7	0.695
PT-N2-C3	120.0	0.695
N2-C3-NA	120.0	0.695
N2-C3-CO	120.0	0.695

(iii) Torsional barriers

torsion angle	V₁(kcal/mol)	V₂(kcal/mol)	V₃(kcal/mol)
PT-N2-C3-C3	0.00	1.25	0.00
PT-N2-C3-CO	0.00	1.25	0.00
PT-N2-C3-H	0.00	1.25	0.00
PT-N2-C3-N2	0.00	1.25	0.00
N2-C3-NA-C3	0.00	2.50	0.00
N2-CO-C3-N2	0.00	1.25	0.00
C3-N2-CO-C3	0.00	1.25	0.00

N2-C3-N2-CO	0.00	1.25	0.00
NA-C3-N2-CO	0.00	1.25	0.00
NA-C3-N2-HV	0.00	1.25	0.00
NA-C3-C3-CO	0.00	1.25	0.00
NA-C3-C3-N2	0.00	1.25	0.00
NA-C3-N2-C3	0.00	1.25	0.00
NA-C3-N2-C4	0.00	1.25	0.00
C3-N2-C3-N2	0.00	1.25	0.00
C3-N2-C3-H	0.00	1.25	0.00
C3-CO-N2-HV	0.00	1.25	0.00
CO-C3-C3-N2	0.00	1.25	0.00
CO-C3-N2-C3	0.00	1.25	0.00
O1-CO-C3-N2	0.00	1.25	0.00
N2-C3-C3-N2	0.00	1.25	0.00
C4-N2-C3-H	0.00	1.25	0.00
C3-N2-CO-C3	0.00	1.25	0.00

(iv) Nonbonded parameters

atom	$r^*(\text{Å})$	$\epsilon(\text{kcal/mol})$
PT	1.75	0.1

(v) Out-of-plane-bending parameters

central atom	$K_\theta(\text{mdyne} \times \text{Å/rad}^2)$
PT	0.05

For MM+ atom types in cis-[Pt(NH$_3$)$_2$(9-EtG)$_2$] see Fig. 3.1b. In this calculation platinum was assigned the "wild-card" (unknown) atom type.

It may be seen from Table 3.4 and Fig. 3.2a that this force field, even with the default parameters describing the platinum coordination sphere, is capable of reproducing experimental geometry of the studied system. It generates a good reproduction of N-Pt-N angles and of selected angles associated with nucleobases. However, the overall performance of this force field is not satisfactory; large deviations from expected values are observed in many cases, e.g. for platinum-nitrogen bonds and nucleobase ring angles, such as C(6)-N(1)-C(2), C(2)-N(3)-C(4) and C(8)-N(9)-C(4). Therefore, this force field requires the missing parameters associated with the platinum atom and nucleobases to be specified.

Modified MM+. Geometry optimisation was carried out by attaching two electron lone pairs to the platinum atom as described above (option 3 in section 3.3.2). The parameters, added to the MM+ force field in order to model the *cis*-[Pt(NH$_3$)$_2$(9-EtG)$_2$] complex, are summarised in Table 3.7. When two dummy atoms are bound to the platinum atom on both sides of the coordination plane, platinum acquires "octahedral" geometry. In such a case the POS-approach [27] (see section 3.1.4 for a description) is invoked and the MM+ force field includes 1,3-contacts into the calculation of nonbonded interactions. For this reason, parameters describing the N-Pt-N angles do not need to be explicitly added to the set. For other parameters, required for the optimisation of *cis*-[Pt(NH$_3$)$_2$(9-EtG)$_2$] and not specified in Table 3.7, the default approach was used (Table 3.6).

Table 3.7. Parameters added to the MM+ force field for the geometry optimisation of the *cis*-[Pt(NH$_3$)$_2$(9-EtG)$_2$] complex.

(i) New atom types

atom type	description
PT	platinum atom

(ii) Bond Stretch parameters[a]

bond	r$_0$(Å)	K$_r$(mdyn/Å)
PT-NH	2.030	2.6
PT-N2	2.010	2.6

(iii) Angle Bend parameters[a]

angle	θ$_0$(deg)	K$_\theta$(mdyn×Å/rad^2)
PT-N2-C3	127.3	0.308
C3-C3-N2	107.0[b]	0.970
C3-N2-C3	107.0[c]	0.970
N2-C3-N2	112.3	0.970

(iv) Nonbonded parameters[d]

atom type	r*(Å)	ε(kcal/mol)
PT	1.75	0.02

(v) Out-of-plane-bending parameters[a]

central atom	$K_\theta(\text{mdyn} \times \text{Å/rad}^2)$
PT	1.023

For MM+ atom types of atoms in cis-[Pt(NH$_3$)$_2$(9-EtG)$_2$] see Fig. 3.1b.
[a]These parameters were taken from Ref. [58]. The force field used in that study is somewhat different to MM+; however, the functional forms for bond stretching and angle bending are similar to these in MM+. [b]This value is an average of two values reported in Ref. [58] for C(4)-C(5)-N(7) and C(5)-C(4)-N(9) angles. Without creating new different atom types for N(7) and N(9) atoms in guanine it is impossible to assign unique parameters to these angles. [c]Unmodified MM+ value. [d]These parameters were set as suggested in Ref. [20].

A comparative inspection of the data in Table 3.4 reveals that defining parameters associated with platinum improves the reproduction of values of bonds and angles in the platinum coordination sphere. That is, the modified MM+ force field generates a better reproduction of Pt-N bond lengths and N-Pt-N angles and in some cases it surpasses the results presented in Ref. [58]. However, its ability to reproduce the geometry of coordinated nucleobases is still deficient. Whilst the supplementation of missing parameters for the imidazole ring of guanine (angle bend parameters for C3-C3-N2, C3-N2-C3 and N2-C3-N2) brings the bond length and angle values closer to the experimental structure, those in the pyrimidine ring remain inadequate. Therefore, it is clear that the force field required for the simulation of platinum–nucleobase complexes has to contain specific atom types and associated parameters for atoms in nucleobase rings. The AMBER force field meets such a requirement.

Modified AMBER. Geometry optimisation was carried out by attaching two electron lone pairs to the platinum atom as described above. No angle restraining for LP-PT-LP was employed, being replaced by the explicit angle parameter in the force field. The parameters, added to the AMBER force field in order to model the cis-[Pt(NH$_3$)$_2$(9-EtG)$_2$] complex, are summarised in Table 3.8.

Table 3.8. AMBER parameters added for the geometry optimisation of cis-[Pt(NH₃)₂(9-EtG)₂].

(i) New atom types

atom type	description
PT	platinum atom

(ii) Bond Stretch parameters

bond	r_0(Å)	K_r(kcal/(mol×Å²))
PT-N3[a]	2.030	366
PT-NB[a]	2.010	366
PT-LP[b]	1.615	300

(iii) Angle Bend parameters

angle	θ_0(deg)	K_θ(kcal/(mol×rad²))
N3-PT-N3[a]	90	42
NB-PT-NB[a]	90	42
PT-NB-CB[a]	127.95	20
PT-NB-CK[a]	127.95	20
N*-CK-NB[c]	112.3	70
CB-NB-CK[d]	106	70
LP-PT-N3	90	20[e]
LP-PT-NB	90	20[e]
LP-PT-LP	180	20[e]

(iv) Torsional parameters[a]

torsion angle	ϕ_0(deg)	n	V/2(kcal/mol)
N3-PT-NB-CK	90	2	0.25
N3-PT-NB-CB	90	2	0.25
NB-PT-NB-CK	90	2	0.25
NB-PT-NB-CB	90	2	0.25
CK-NB-CB-PT[f]	180	2	5

(v) Nonbonded parameters[g]

atom type	r*(Å)	ε(kcal/mol)
PT	1.75	0.1

For AMBER atom types of atoms in cis-[Pt(NH₃)₂(9-EtG)₂] see Fig. 3.1c.
[a]These parameters were taken from Ref. [40]. [b]This bond distance was taken to be the sum of the covalent radii of platinum and hydrogen atoms. The corresponding stretching force constant was set to be slightly smaller than the force constant for PT-N bond. This ensures equal distribution on both sides of the coordination plane, but not too strong to ease avoiding steric clashing. [c]These parameters were taken from Ref. [58]. [d]This angle value was found to give a better match to crystal structure data than the angle value developed in Ref. [40]. [e]This force constant was set by analogy to those developed in Ref. [40]. [f]Improper torsion. [g]These parameters were set as suggested in Ref. [20].

A final inspection of Table 3.4 shows the improvement in prediction of geometric parameters associated with nucleobase rings, e.g. angle values involving atoms C(2) and C(5). Perhaps it comes as no surprise that the AMBER force field, containing parameters for nucleobases, together with the introduced parameters describing the platinum environment (Table 3.8) is superior to the MM+ in its ability to model platinum–nucleobase complexes. Therefore, it has been finally decided to employ this force field in order to model related systems.

3.3.4 Modelling of cis-[Pt(NH₂R)₂Cl(9-EtG)] complexes

A series of complexes of general formula cis-[Pt(NH₂R)₂Cl(9-EtG)] were modelled using the modified AMBER force field described above. To choose the equilibrium values associated with the chlorine atom the crystal structures of relevant platinum complexes containing chlorine ligands were analysed with respect to the values of the geometric features of interest. These are summarised in Appendix V (Table V.1). The AMBER force field was modified by adding parameters developed for the modelling of monoadducts (Table 3.9), in addition to those derived initially for the modelling of bisadducts (Table 3.8).

To bench-mark the force field thus developed, the structure of the complex cis-[Pt(NH₃)₂(N2-dimethyl-9-MeG)Cl] was optimised and values of selected bond lengths and angles are compared with the experimental data (refcode BOHDAD [69] from CSD) (Table 3.10). The structure of the complex cis-[Pt(NH₃)₂(N2-dimethyl-9-MeG)Cl] is shown in Fig. 3.3. The atomic coordinates of optimised structures are compiled in Appendix IV (Tables IV.2). Overlays of the optimised geometries with the experimental structure are shown in Fig. 3.4. The overlays were created based on the best fit between coordination planes of two structures.

Optimisation with dummy atoms. Initially, geometry optimisation was carried out by attaching two electron lone pairs to the platinum atom, as described above. In this optimisation, amine donor nitrogens were assigned atom type N3 (Fig. 3.3b) and the corresponding parameters were utilised (Table 3.8).

Table 3.9. AMBER parameters for the geometry optimisation of complexes with general formula cis-[Pt(NH$_2$R)$_2$Cl(9-EtG)].

(i) New atom types[a]

atom type	description
N3C	Pt-bound ligand nitrogen, cis to Cl
N3T	Pt-bound ligand nitrogen, trans to Cl

(ii) Bond Stretch parameters

bond	r$_0$(Å)	K$_r$(kcal/(mol×Å^2))
PT-CL	2.305[b]	366[c]

(iii) Angle Bend parameters

angle	θ$_0$(deg)	K$_\theta$(kcal/(mol×rad^2))
NB-PT-N3T[d]	90	42
NB-PT-N3C[d]	180	42
NB-PT-CL[e]	90	42
N3C-PT-CL[e]	90	42
N3T-PT-CL[e]	180	42
PT-N3*[f]-CT	109.47[g]	20[c]
PT-N3*[f]-H3	109.47[g]	20[c]
LP-PT-CL	90	20[c]

(iv) Torsional parameters

torsion angle	ϕ_0(deg)	n	V/2(kcal/mol)
CK-NB-PT-CL^e	90	2	0.25
CB-NB-PT-CL^e	90	2	0.25

ᵃAll relevant parameters, involving the N3 atom types, available in the HyperChem AMBER force field and added to it previously (Table 3.8), were duplicated for N3C and N3T atom types. ᵇThis value was chosen upon analysis of available crystal structures. ᶜThis force constant was set by analogy to those developed in Ref. [40]. ᵈThese parameters were taken from Ref. [40]. ᵉThese parameters were set by analogy to those developed in Ref. [40]. ᶠN3*: N3C and N3T. ᵍThe equilibrium value for tetrahedral geometry.

Optimisation without dummy atoms. Two new atom types were created to describe different ligand donor atoms (*cis* and *trans* with respect to the chlorine atom), see the diagram below and also Fig. 3.3c.
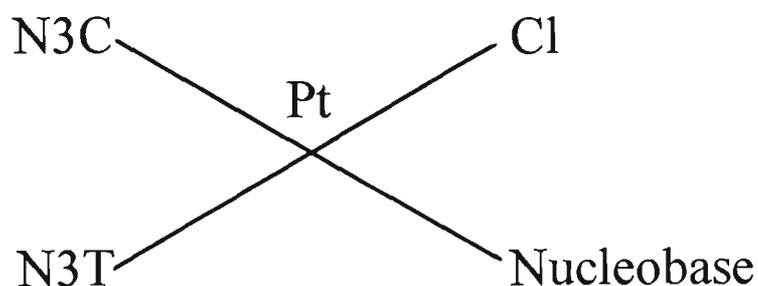
For optimisation without dummy atoms, new angle bend parameters for the coordination plane associated with these new atom types were introduced (Table 3.9). Thus, the dummy atom approach employed above in order to conserve planarity is rendered unnecessary by this force field. This force field now permits the risk of a steric clashing between the dummy atoms and atoms of a carrier ligand(s) or a nucleobase exocyclic group(s) in sterically crowded systems to be avoided. It also improves the reproduction of the adduct geometric features in the immediate environment of the platinum atom (angles Pt-N(7)-C(5)/C(8) in the nucleobase plane and Cl-Pt-N(7)/N(11) in the coordination plane).
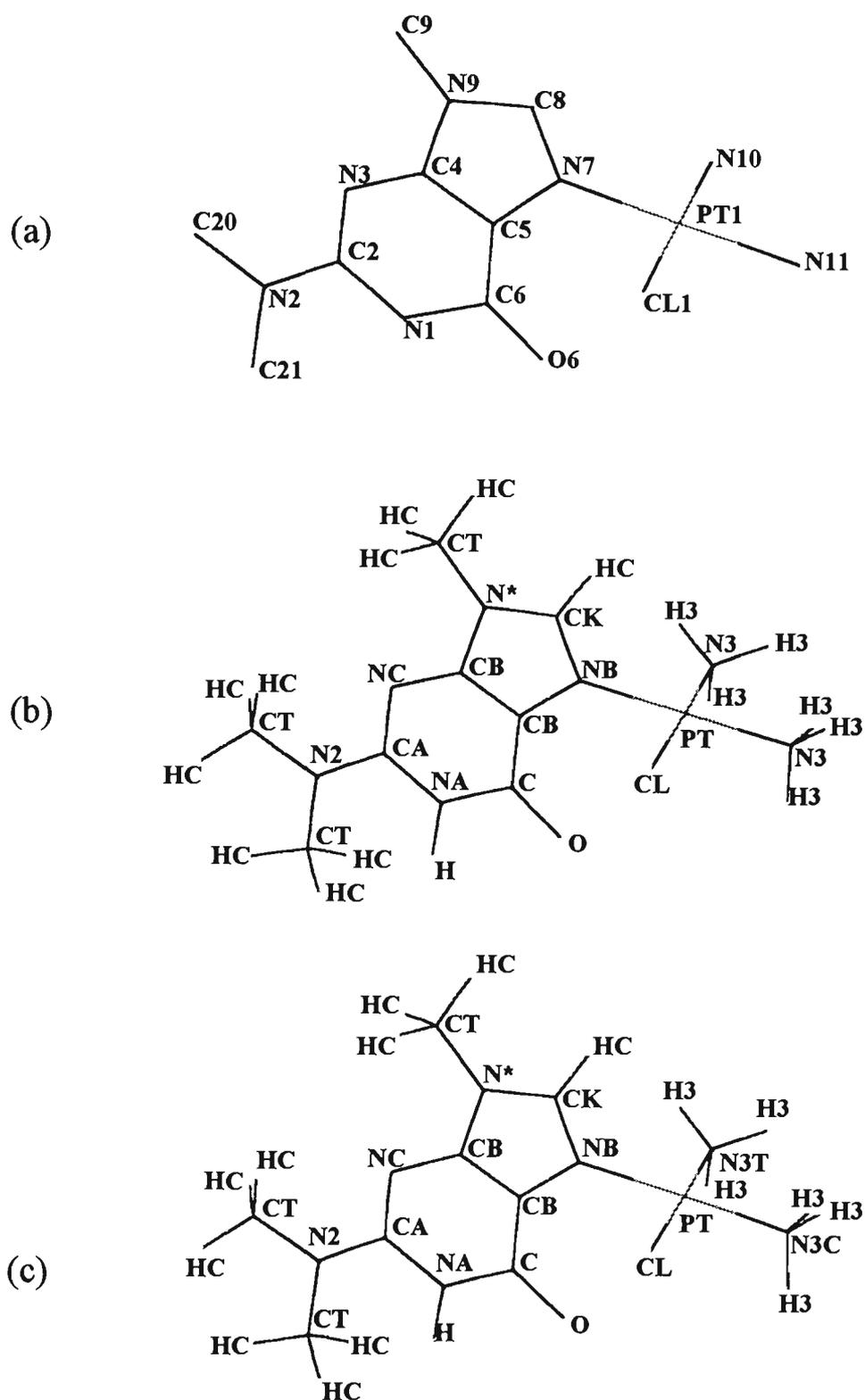
Figure 3.3. Structure of cis-[Pt(NH₃)₂(N2-dimethyl-9-MeG)Cl] complex [69]. (a) Numbering scheme, H atoms removed for clarity. (b) AMBER force field atom types in optimisation with dummy atoms. (c) AMBER force field atom types with new atoms for am(m)ine nitrogen atoms.

(a)

(b)
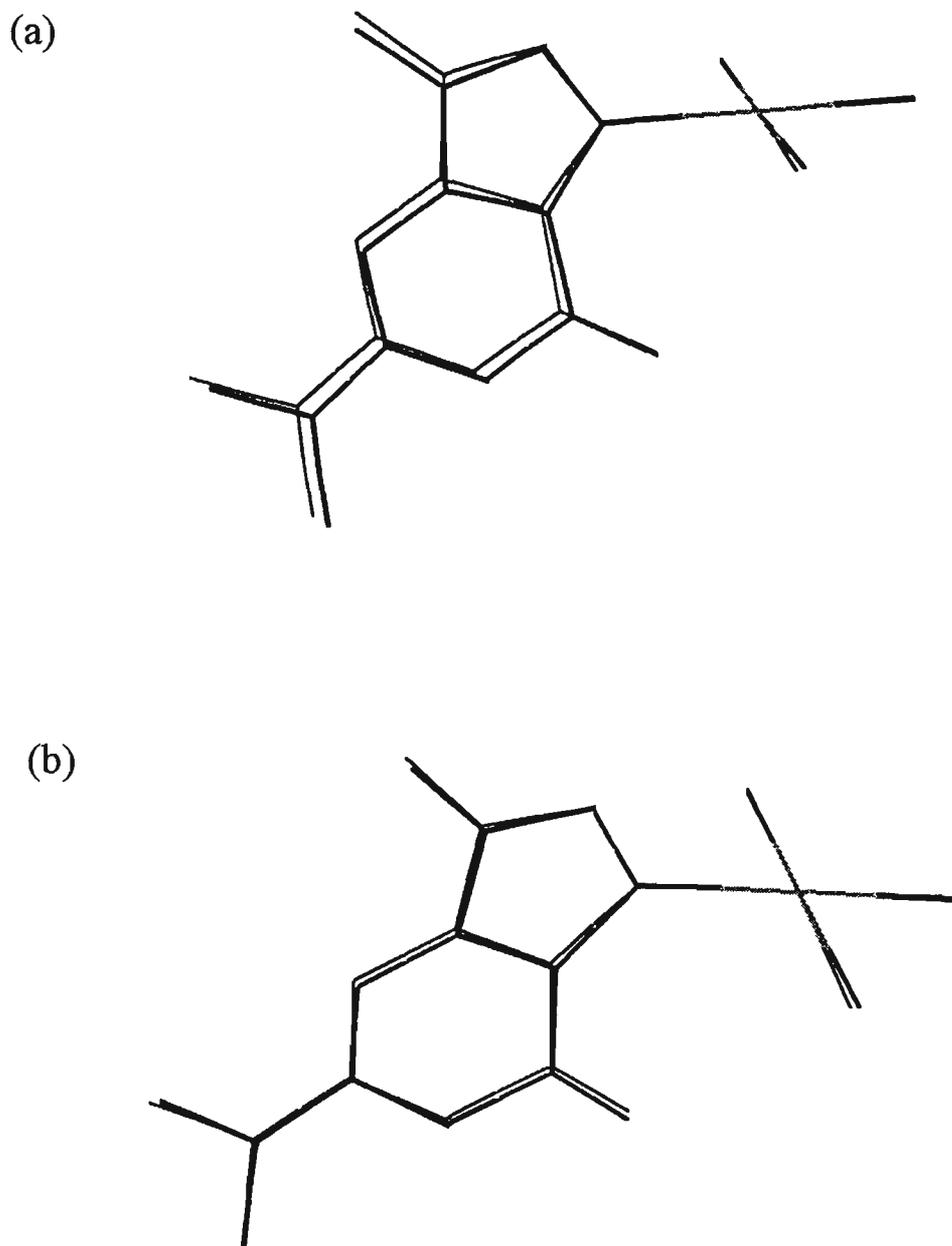
Figure 3.4. Overlays of the optimised geometries with the experimental structures for cis-[Pt(NH₃)₂(N2-dimethyl-9-MeG)Cl] complex [69], experimental structure is shown in bold. (a) Optimisation carried out with dummy atoms. (b) Optimisation carried out without dummy atoms.

Table 3.10. Comparison of the crystal structure and AMBER geometry-optimised structures for cis-[Pt(NH₃)₂(N2-dimethyl-9-MeG)Cl].

atoms	crystal structure [69]	structure calculated with LP-dummies	structure calculated without LP-dummies
Pt-N(10)	2.059	2.030	2.030
Pt-N(11)	2.027	2.030	2.030
Pt-N(7)	2.033	2.012	2.009
Pt-Cl	2.300	2.305	2.305
N(10)-Pt-N(11)	91.7	90.5	90.2
N(10)-Pt-N(7)	88.7	89.5	89.8
Cl-Pt-N(11)	87.6	90.2	89.9
Cl-Pt-N(7)	92.0	89.8	90.1
Pt-N(7)-C(5)	126.2	130.7	128.3
Pt-N(7)-C(8)	126.5	124.3	126.6
C(6)-N(1)-C(2)	125.2	124.9	124.8
N(1)-C(2)-N(2)	116.4	117.7	117.8
N(1)-C(2)-N(3)	123.4	122.2	122.2
N(2)-C(2)-N(3)	120.1	120.1	120.0
C(2)-N(3)-C(4)	113.4	113.9	113.9
N(3)-C(4)-N(9)	126.4	126.2	126.4
N(9)-C(4)-C(5)	104.7	106.6	106.4
N(3)-C(4)-C(5)	128.8	127.3	127.1
N(7)-C(5)-C(4)	108.6	109.6	109.7
N(7)-C(5)-C(6)	132.9	131.1	130.6
C(4)-C(5)-C(6)	118.6	119.4	119.7
O(6)-C(6)-C(5)	129.5	128.4	128.0
O(6)-C(6)-N(1)	120.0	119.1	119.7

C(5)-C(6)-N(1)	110.5	112.4	112.3
N(7)-C(8)-N(9)	110.9	113.1	113.0
C(8)-N(9)-C(4)	108.5	105.7	105.7
C(8)-N(7)-C(5)	107.3	105.0	105.1

Units: bond lengths - angstroms, angles - degrees.

3.3.5 Modelling of [Pt(A)₃(Nucleobase)] complexes

Three series of complexes of general formula [Pt(A)$_3$(Nucleobase)] were modelled using the modified AMBER force field, A $\equiv$ NH$_3$ and (A)$_3$ $\equiv$ di-ethylenetriamine (dien) and 1,1,4,7,7-pentamethyldiethylenetriamine (pmdien):

dien   pmdien
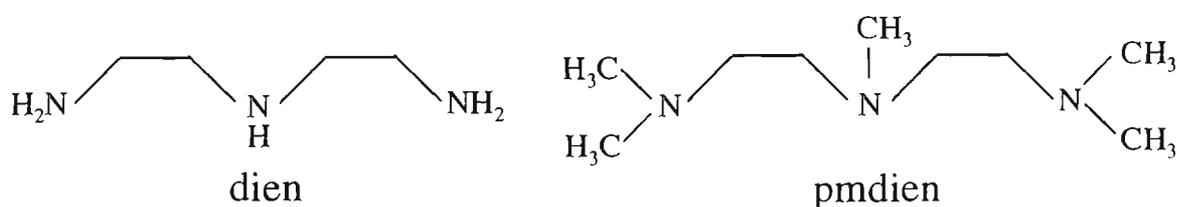
Six common nucleobases with a variety of potential platinum-binding sites were selected for this study (Fig. 3.5).

To choose the equilibrium values associated with a variety of nucleobase binding sites the crystal structures of relevant platinum complexes were analysed with respect to the values of geometric features of interest. These are summarised in Appendix V (Table V.2). The AMBER force field was modified by adding parameters developed for the modelling of such monofunctional adducts (Table 3.11), in addition to those derived previously (Tables 3.8 and 3.9).
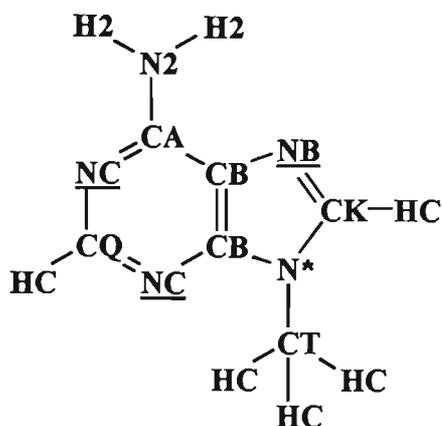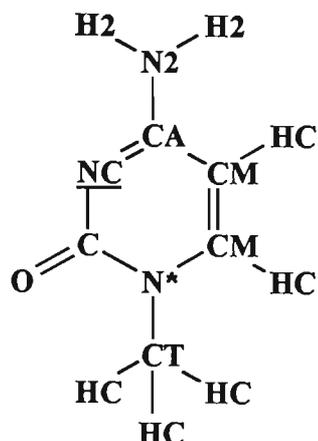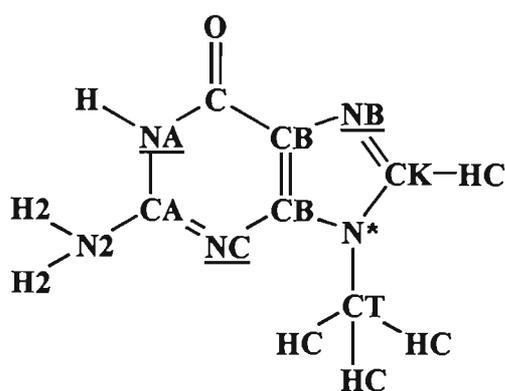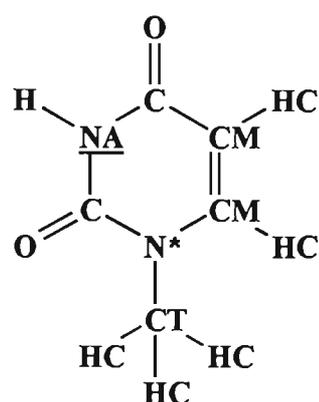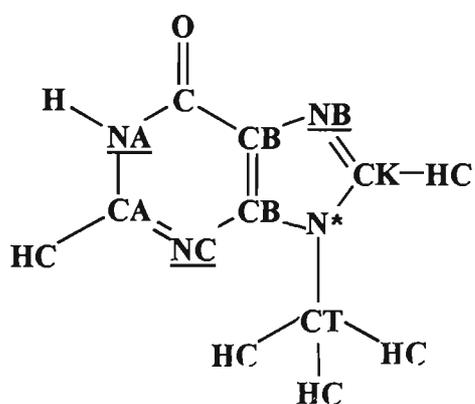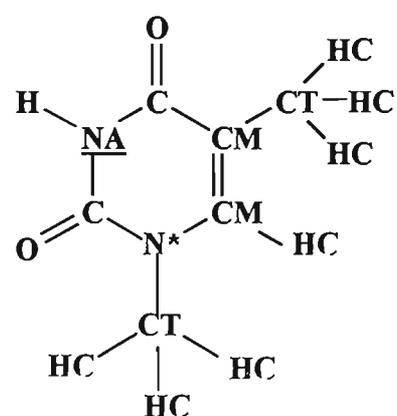
Figure 3.5. AMBER atom types in common nucleobases. Endocyclic nitrogen binding sites are underlined.

Table 3.11. AMBER parameters for the geometry optimisation of complexes with general formula [Pt(A)$_3$(Nucleobase)].

(i) Bond Stretch parameters

bond	$r_0(\text{Å})^a$	$K_r(\text{kcal/(mol} \times \text{Å}^2))^b$
PT-NA^c	2.035	366
PT-NC^d	2.038	366

(ii) Angle Bend parameters

angle	$\theta_0(\text{deg})$	$K_\theta(\text{kcal/(mol} \times \text{rad}^2))^e$
PT-NC-CA^f	122.35	20
PT-NC-CQ^g	116.00	20
PT-NC-CB^h	126.15	20
PT-NC-C^i	116.75	20
PT-NA-CA^j	119.90	20
PT-NA-C^k	119.15	20
HC-CA-NA^l	115.00	35
HC-CA-NC^l	119.10	35
N*-PT-N3T^m	90.00	42
N*-PT-N3C^m	180.00	42
N3T-PT-N3T	180.00	42

(iii) Torsional parameters[n]

torsional angle	ϕ_0(deg)	n	V/2(kcal/mol)
N3*[o]-PT-NC-C*[p]	90	2	0.25
N3*[o]-PT-NA-C*[r]	90	2	0.25

For AMBER atom types of atoms in nucleobases see Fig. 3.5.

[*]These bond lengths were derived by averaging Pt-N bond length values from available crystal structures of platinum complexes (Appendix III, Table III.2). [b]The stretching force constant was set analogous to Pt-N(7) of guanine [40]. [c]NA: N(1) position of guanine and hypoxanthine, N(3) position of thymine and uracil. [d]NC: N(1) position of adenine, N(3) position of cytosine, N(3) position of purines. [e]The bending force constants for angles involving platinum were set by analogy to the corresponding angles Pt-N(7)-* and *-Pt-* for guanine complexes [40]. Bending constants for angles around C(2) of hypoxanthine were set analogous to adenine. [f]Pt-N(3)-C(2) angles of guanine and hypoxanthine, Pt-N(1)-C(6) angle of adenine and Pt-N(3)-C(4) angle of cytosine. [g]Pt-N(1)-C(2) and Pt-N(3)-C(2) angles of adenine. [h]Pt-N(3)-C(4) angles of guanine and hypoxanthine and Pt-N(3)-C(4) angle of adenine. [i]Pt-N(3)-C(2) angle of cytosine. [j]Pt-N(1)-C(2) angles of guanine and hypoxanthine. [k]Pt-N(1)-C(6) angles of guanine and hypoxanthine. [l]H(2)-C(2)-N(1) and H(2)-C(2)-N(3) angles of hypoxanthine respectively. [m]N*: NA and NC atom types. [n]Torsional parameters involving platinum were set analogous to Ref. [40]. [o]N3*: N3C and N3T. [p]C*: CQ, CA, CB and C atom types. [r]C*: CA and C atom types.

Carrier ligand nitrogen atoms were assigned atom types as shown below:
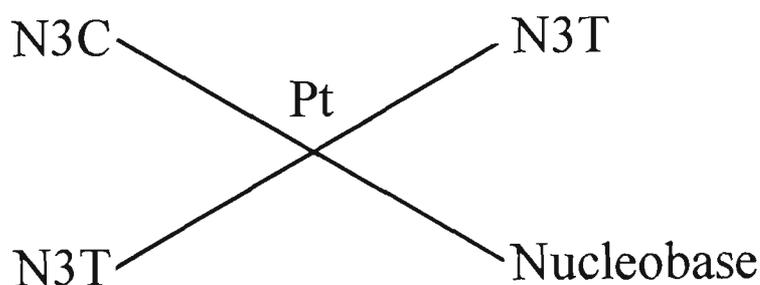
To bench-mark the force field thus developed, the values of selected bond lengths and angles for several complexes were compared with the experimental data for related systems (Table 3.12). The atomic coordinates of optimised structures are compiled in Appendix V (Tables V.3). Overlays of the optimised geometries with the experimental structures are shown in Fig. 3.6. The overlays were created based on the best fit between coordination planes of two structures.

Table 3.12. Comparison of crystal structures and geometry-optimised structures for [Pt(A)₃(Nucleobase)] complexes.

(a) [Pt(NH₃)₃(pyrimidine)] complexes, N3-binding.

atoms[a]	1-MeT	DEYXUA[b]	1-MeC	VEPROX[b]	1-MeU	SEBVIE[b]
Pt-N(amine)	2.030 2.030 2.029	2.052 2.001	2.030 2.030 2.030	2.053 2.061 2.056	2.030 2.030 2.030	2.031 2.023
Pt-N3	2.035	1.974	2.039	2.044	2.035	2.048
N(amine)-Pt-N3	89.86 89.91 179.99	88.15 178.33	89.99 89.99 179.97	91.71 89.40 178.90	89.98 89.96 179.96	89.84 176.49
Pt-N3-C2	117.16	117.21	116.53	116.96	117.21	118.04
Pt-N3-C4	117.33	121.84	122.27	121.39	117.41	118.30
C2-N3-C4	125.50	120.73	121.21	121.61	125.44	123.38
N1-C6-C5	121.71	121.53	121.39	121.77	121.55	120.43
O2-C2-N3	121.08	118.70	121.00	121.54	121.07	119.28
DA[c]	76.15	77.12	71.91	80.75	75.81[d]	65.67[d]
RMS fit[e]	0.07 Å		0.12 Å		0.45 Å	

(b) [Pt(NH₃)₃(purine)] complexes, N1-binding.

atoms[a]	9-MeG	BAHNUT[b]	9-MeA	SISCIG[b]	9-MeH	ENXPTA10[b]
Pt-N(amine)	2.031 2.020 2.030	2.060 2.020 2.035	2.030 2.020 2.031	2.032 2.020 2.043	2.030 2.020 2.030	2.050 2.040 1.997
Pt-N1	2.035	2.043	2.037	2.043	2.034	2.052
N(amine)-Pt-N1	91.88 91.79 179.99	97.11 94.27 178.54	91.82 91.80 179.97	95.79 94.20 177.76	91.82 91.76 179.99	97.97 93.33 177.15
Pt-N1-C2	119.09	124.21	116.99	114.36	117.79	118.90
Pt-N1-C6	116.82	115.27	123.75	123.62	118.37	119.17
C2-N1-C6	124.09	120.42	119.25	122.01	124.17	121.55
N3-C4-C5	127.19	127.25	126.58	127.62	127.19	128.09
N6-C6-N1	119.67	120.91	118.91	119.58	119.70	121.95
DA[c]	72.91	61.59	66.22	66.14	66.53[f]	45.02[f]
RMS fit[e]	0.18 Å		0.12 Å		0.24 Å	

(c) [Pt(NH₃)₃(purine)] complexes, N3-binding.

atoms[a]	9-MeG	DASCUV[b]
Pt-N(amine)	2.030 2.029 2.030	2.038 2.050 2.048
Pt-N3	2.038	2.045
N(amine)-Pt-N3	89.87 89.95 179.99	90.334 88.27 176.75
Pt-N3-C2	119.25	117.45
Pt-N3-C4	127.57	126.15
C2-N3-C4	113.18	115.96
N1-C6-C5	111.81	112.32
N2-C2-N3	120.08	117.96
DA[c]	84.94	81.13
RMS fit[e]	0.22 Å	

Units: bond lengths - angstroms, angles - degrees. [a]To produce a representative depiction of the optimised structures the bonds and angles are selected which describe the immediate environment of the platinum atom and the geometries of nucleobase rings; the latter are characterised by three angles with respect to the platinum atom: close to it, removed from it and involving an exocyclic functional group adjacent to it. [b]CSD refcode. [c]PtN₄/nucleobase dihedral angle; the least-square-plane equations for the platinum coordination plane and nucleobase plane in all optimised structures and in the experimental structures mentioned in this table are summarised in Appendix VII. [d]Experimental structure is a bis-adduct; hence, a larger difference between dihedral angles in experimental and calculated structures results. [e]Based on the best overlay of nucleobase ring atoms and atoms of the platinum coordination plane, only positions of the heavy atoms are compared. [f]Experimental and calculated structures present mirror images of each other, which gives rise to the difference in dihedral angles; the overlays in Fig. 3.6(e,f) are carried out for direct and inverted optimised structures, respectively.

It may be seen from Table 3.12 and Fig. 3.6 that the designed force field simulates experimental structures with good accuracy. The largest deviation occurs for platinum-related geometric parameters. The accuracy of reproduction of the nucleobase orientation with respect to the coordination plane is one particularly impressive outcome of the simulations presented here.
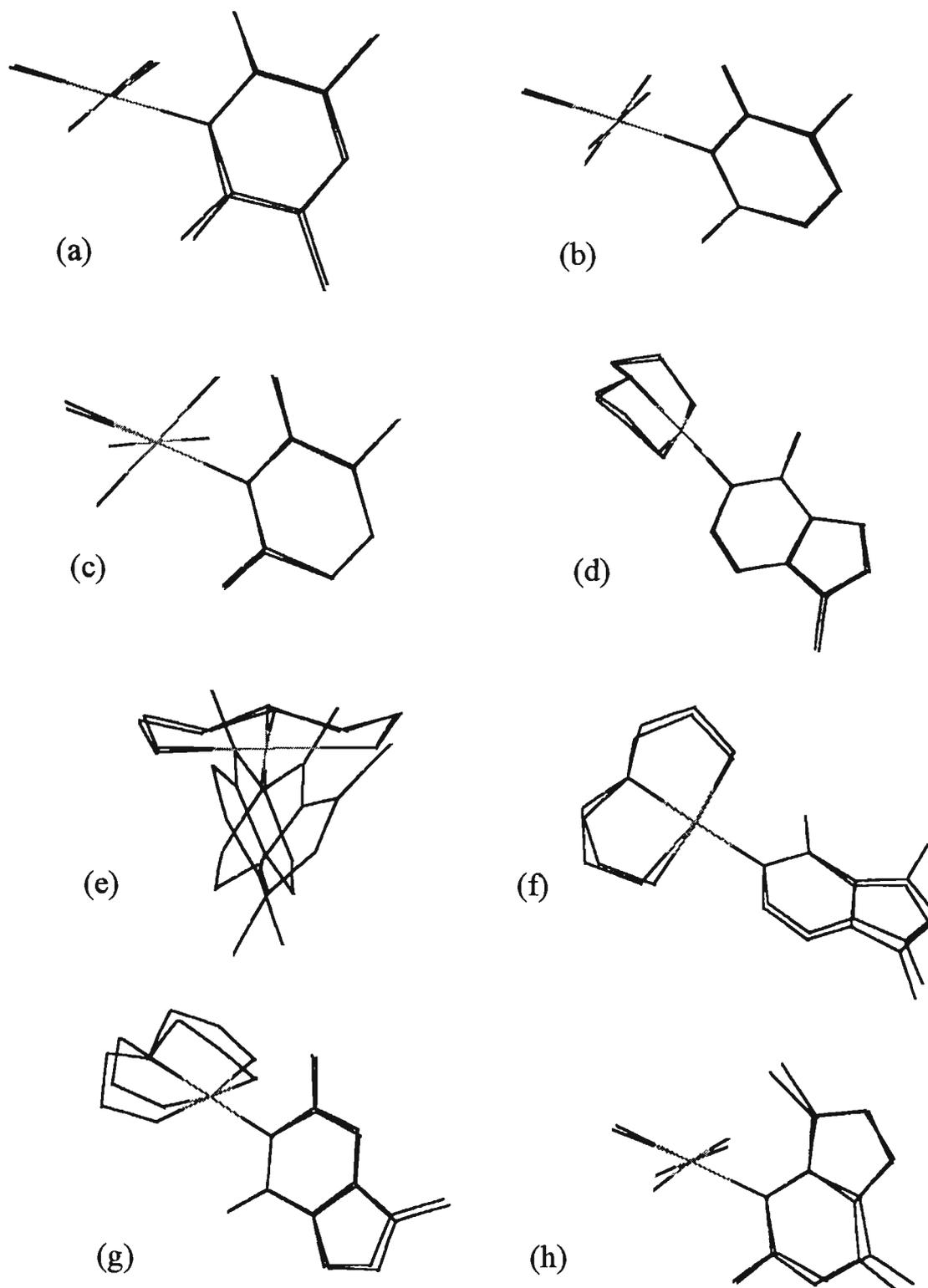
Figure 3.6. Overlays of the optimised geometries with the experimental structures for [Pt(A)₃(Nucleobase)] complexes. In the experimental structures only atoms equivalent to those in the optimised geometries are depicted. (a) [Pt(NH₃)₃(1-MeT-N3)] and DEYXUA. (b) [Pt(NH₃)₃(1-MeC-N3)] and VEPROX. (c) [Pt(NH₃)₃(1-MeU-N3)] and SEBVIE. (d) [Pt(dien)(9-MeA-N1)] and SISCIG. (e) [Pt(dien)(9-MeH-N1)] and ENXPTA10 (optimised structure prior to inversion). (f) [Pt(dien)(9-MeH-N1)] and ENXPTA10 (optimised structure after inversion). (g) [Pt(dien)(9-MeG-N1)] and BAHNUT. (h) [Pt(NH₃)₃(9-MeG-N3)] and DASCUV.

3.3.6 Electrostatic interactions

All AMBER-based force fields described above were implemented without the incorporation of the electrostatic energy term into the total molecular energy expression (eqn **3.18**). It has been suggested that while electrostatic interactions significantly influence molecular energetics, they do not affect the structures to the same extent [81].

In order to enhance the force fields described above, an attempt has been made to cautiously include electrostatic interactions. The charge distribution scheme, based on the approach developed by Yao et al. [40] for the platinum–guanine system, has been applied to the complexes under investigation in this work. Charge distribution diagrams are compiled in Appendix VIII. The simulation results obtained with the use of these point charges have shown that they do not affect final structures significantly, maintaining similar findings gained for other coordination systems [79]. However, in some cases assignment of atomic point charges according to the Yao-Marzilli scheme [40] has led to "unreasonable" structural distortions, such as the deviation of exocyclic oxo substituents in pyrimidine nucleobases from the base plane and "incorrect prediction" of base plane – coordination plane dihedral angles as compared to the experimental structures. Therefore, it may be concluded that this charge distribution scheme requires further development. On the other hand, it is clear that the "neglect" of electrostatic interactions does not distort the modelled geometry. This is a widely accepted to be the case with respect to electrostatics in an inorganic force field (see section 3.1.4 for more background).

3.3.7 Caveat

As an adjunct to the narrative of method development for modelling of Pt – nucleoconstituent adducts, it must be emphasised that the parameters developed in this work are devised empirically to provide agreement with experimental data (crystal structures), or based on analogy with related systems. Therefore, it should be recognised that such parameterisation may introduce a bias towards the available structural data. Accordingly, care should be taken in the application of the force field described here to various platinum – nucleic acid systems. However, the results presented in this chapter (Tables 3.4, 3.10 and 3.12) bolster confidence in its application to the systems under study in this work.

3.4 REFERENCES

1. Kelland, T.L.R.; McKeage, M.J. *Drugs and Aging* **1994**, *5*, 85.

2. *Molecular Aspects of Anti-Cancer Drug Action;* Neidle, S.; Waring, M.J. (Eds.); McMillan Press: London, **1983**.

3. *Metal-Based Anti-tumour Drugs;* Gielen, M.F. (Ed.); Freund: London, **1988**.

4. Neidle, S. In *The Design of Drugs to Macromolecular Targets;* Beddell, C.R. (Ed.); Wiley: Chichester, **1992**, 173-210.

5. Saenger, W. *Principles of Nucleic Acids Structure;* Springer: New York, **1984**.

6. *Nucleic Acid Targeted Drug Design;* Propst, C.L.; Perun, T.J. (Eds.); Marcel Dekker: New York, **1992**.

7. Bowen, J.P.; Allinger, N.L. In *Reviews in Computational Chemistry, Vol. 2;* Lipkowitz, K.B.; Boyd, D.B. (Eds.); VCH: New York, **1991**, 81-97.

8. Hambley, T.W.; Hawkins, C.J.; Palmer, J.A.; Snow, M.R. *Aust. J. Chem.* **1981**, *34*, 45.

9. Allinger, N.L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

10. Weiner, S.J.; Kollman, P.A.; Nguen, D.T.; Case, D.A. *J. Comput. Chem.* **1986**, *7*, 230.

11. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

12. Rappe, A.K.; Casewit, C.J.; Colwell, K.S.; Goddard, W.A.I.; Skiff, W.M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

13. Schnur, D.M.; Grieshaber, M.V.; Bowen, J.P. *J. Comput. Chem.* **1991**, *12*, 844.

14. HyperChem™ Version 4. Canada: Hypercube, Inc. **1994.**

15. Allinger, N.L.; Zhou, X.; Bergsma, J. *J. Mol. Struct. (Theochem)* **1994**, *312*, 69.

16. Leach, A.R. In *Reviews in Computational Chemistry, Vol. 2*; Lipkowitz, K.B.; Boyd, D.B. (Eds.); VCH: New York, **1991**, 1-55.

17. Allen, F.H.; Kennard, O. *CDA News* **1993**, *8*, 31.

18. Lauher, J.W. *J. Am. Chem. Soc.* **1986**, *108*, 1521.

19. Brooks, B.R. In *Supercomputer Research in Chemistry and Chemical Engineering*; Jensen, K.F.; Truhlar, D.G. (Eds.); ACS: Washington, DC, **1987**, 123-45.

20. *HyperChem*^{TM}, *Computational Chemistry*; Hypercube, Inc. Canada, **1994**.

21. Howard, A.E.; Kollman, P.A. *J. Med. Chem.* **1988**, *31*, 1669.

22. Crippen, G.M.; Havel, T.F. *Distance Geometry and Molecular Conformation, Chemometrics Research Studies Series 15.* Wiley: New York, **1988**.

23. Cramer, C.J.; Truhlar, D.G. In *Reviews in Computational Chemistry, Vol.6*; Lipkowitz, K.B.; Boyd, D.B. (Eds.); VCH: New York, **1995**, 1-72.

24. Zimmer, M. *Chem. Rev.* **1995**, *95*, 2629.

25. Kiss, F.; Tvaruzek, P.; Hladka, E. In *Fifth International Symposium "Platinum and Other Metal Coordination Compounds in Cancer Chemotherapy"*; Nicolini, M.; Bandoli, G. (Eds.); Cleup Padua: Padua, **1988,** 369.

26. Hambley, T.W.; Hawkins, C.J.; Palmer, J.A.; Snow, M.R. *Aust. J. Chem.* **1981**, *34*, 2525.

27. Hay, B.P. *Coord. Chem. Rev.* **1993**, *126*, 177.

28. Lin, W.; Welsh, W.J.; Harris, W.R. *Inorg. Chem.* **1994**, *33*, 884.

29. Rappe, A.K.; Goddard, W.A.I. *J. Phys. Chem.* **1991**, *95*, 3358.

30. Medforth, C.J.; Senge, M.O.; Forsyth, T.P.; Hobbs, J.D.; Shelnutt, J.A.; Smith, K.M. *Inorg. Chem.* **1994**, *33*, 3865.

31. Bosnich, B. *Chem. Soc. Rev.* **1994**, *94*, 387.

32. Comba, P. *Coord. Chem. Rev.* **1993**, *123*, 1.

33. Landis, C.R.; Root, D.M.; Cleveland, T. In *Reviews in Computational Chemistry*, *Vol.6*; Lipkowitz, K.B.; Boyd, D.B. (Eds.); VCH: New York, **1995**, 73-148.

34. Kolossvary, I.; Guida, W.C. *J. Comput. Chem.* **1993**, *14*, 691.

35. Chottard, J.-C. *Rec. Trav. Chim. Pays-Bas* **1987**, *106*, 192.

36. Herman, F.; Kozelka, J.; Stoven, V.; Guitet, E.; Girault, J.-P.; Huynh-Dinh, T.; Igolen, J.; Lallemand, J.-Y.; Chottard, J.-C. *Eur. J. Biochem.* **1990**, *194*, 119.

37. Nokolov, G.S.; Trendafilova, N.; Schonenberger, H.; Gust, R.; Kritzenberger, J.; Yersin, H. *Inorg. Chim. Acta* **1994**, *217*, 159.

38. Ling, E.C.H.; Allen, G.W.; Hambley, T.W. *J. Chem. Soc., Dalton Trans.* **1993**, 3705.

39. Hambley, T.W. *Inorg. Chim. Acta* **1987**, *137*, 15.

40. Yao, S.; Plastaras, J.P.; Marzilli, L.G. *Inorg. Chem.* **1994**, *33*, 6061.

41. Mazeau, K.; Vovelle, F.; Rahmouni, A.; Leng, M.; Ptak, M. *Anti-Cancer Drug Des.* **1989**, *4*, 63.

42. Kozelka, J.; Archer, S.; Petsko, G.A.; Lippard, S.J. *Biopolymers* **1987**, *26*, 1245.

43. Hambley, T.W. *J. Chem. Soc., Chem. Comm.* **1988**, 221.

44. Hambley, T.W. *Drug Design and Delivery* **1988**, *3*, 153.

45. Hambley, T.W. *Comments Inorg. Chem.* **1992**, *14*, 1.

46. Kozelka, J.; Savinelli, R.; Berthier, G.; Flament, J.-P.; Lavery, R. *J. Comput. Chem.* **1993**, *14*, 45.

47. McCarthy, S.L.; Hinde, R.J.; Miller, K.J.; Anderson, J.S.; Basch, H.; Krauss, M. *Biopolymers* **1990**, *29*, 823.

48. Sip, M.; Schwartz, A.; Vovelle, F.; Ptak, M.; Leng, M. *Biochemistry* **1992**, *31*, 2508.

49. Miller, K.J.; McCarthy, S.L.; Krauss, M. *J. Med. Chem.* **1990**, *33*, 1043.

50. Dunham, S.U.; Lippard, S.J. *J. Am. Chem. Soc.* **1995**, *117*, 10702.

51. Schroder, G.; Kozelka, J.; Sabat, M.; Fouchet, M.-H.; Beyerle-Pfnur, R.; Lippert, B. *Inorg. Chem.* **1996**, *35*, 1647.

52. Laoui, A.; Kozelka, J.; Chottard, J.-C. *Inorg. Chem.* **1988**, *27*, 2751.

53. Vickery, K.; Bonin, A.M.; Fenton, R.R.; O'Mara, S.; Russell, P.J.; Webster, L.K.; Hambley, T.W. *J. Med. Chem.* **1993**, *36*, 3663.

54. van Garderen, C.J.; Van Houte, L.P.A. *Eur. J. Biochem.* **1994**, *225*, 1169.

55. Kozelka, J.; Petsko, G.A.; Lippard, S.J.; Quigley, G.J. *J. Am. Chem. Soc.* **1985**, *107*, 4079.

56. McCarthy, S.L.; Hinde, R.J.; Miller, K.J.; Anderson, J.S.; Basch, H.; Krauss, M. *Biopolymers* **1990**, *29*, 785.

57. Reily, M.D.; Hambley, T.W.; Marzilli, L.G. *J. Am. Chem. Soc.* **1988**, *110*, 2999.

58. Hambley, T.W. *Inorg. Chem.* **1988**, *27*, 1073.

59. Hambley, T.W. *Inorg. Chem.* **1991**, *30*, 937.

60. Jankowski, K.; Turkkan, N.; Brostow, W. *J. Mol. Struct. (Theochem)* **1986**, *137*, 299.

61. Burkert, U.; Allinger, N.L. *Molecular Mechanics*; ACS: Washington, D.C. **1982**.

62. Kollman, P.A. *Chem. Rev.* **1993**, *93*, 2395.

63. ChemPlus™: Extensions for HyperChem^R Version 1. Canada: Hypercube, Inc. **1993.**

64. Sherman, S.E.; Lippard, S.J. *Chem. Rev.* **1987**, *87*, 1153.

65. Schollhorn, H.; Randashl-Sieber, G.; Muller, G.; Thewalt, U.; Lippert, B. *J. Am. Chem. Soc.* **1985**, *107*, 5932.

66. Sherman, S.E.; Gibson, D.; Wang, A.H.-J.; Lippard, S.J. *Science* **1985**, *230*, 412.

67. Orbell, J.D.; Wilkowski, K.; Marzilli, L.G.; Kistenmacher, T.J. *Inorg. Chem.* **1982**, *21*, 3478.

68. Orbell, J.D.; Wilkowski, K.; de Castro, B.; Marzilli, L.G.; Kistenmacher, T.J. *Inorg. Chem.* **1982**, *21*, 813.

69. Orbell, J.D.; Solorzano, C.; Marzilli, L.G.; Kistenmacher, T.J. *Inorg. Chem.* **1982**, *21*, 3806.

70. Marzilli, L.G.; Chalilpoyil, P.; Chiang, C.C.; Kistenmacher, T.J. *J. Am. Chem. Soc.* **1980**, *102*, 2480.

71. Kistenmacher, T.J.; Castro, B.; Wilkowski, K.; Marzilli, L.G. *J. Inorg. Biochem.* **1982**, *16*, 33.

72. Frommer, G.; Mutikainen, I.; Pesch, F.J.; Hillgeris, E.C.; Lippert, B. *Inorg. Chem.* **1992**, *31*, 2429.

73. Heyl, B.L.; Shinozuka, K.; Miller, S.K.; van der Veer, D.G.; Marzilli, L.G. *Inorg. Chem.* **1985**, *24*, 661.

74. Lippert, B.; Raudaschl, G.; Lock, C.J.L.; Pilon, P. *Inorg. Chim. Acta.* **1984**, *93*, 43.

75. Sherman, S.E.; Gibson, D.; Wang, A.H.-J.; Lippard, S.J. *J. Am. Chem. Soc.* **1988**, *110*, 7368.

76. Admiraal, G.; van der Veer, J.L.; de Graaf, R.A.G.; den Hartog, J.H.J.; Reedijk, J. *J. Am. Chem. Soc.* **1987**, *109*, 592.

77. Frommer, G.; Schollhorn, H.; Thewalt, U.; Lippert, B. *Inorg. Chem.* **1990**, *29*, 1417.

78. Norrby, P.-O.; Akermark, B.; Haeffner, F.; Hansson, S.; Blomberg, M. *J. Am. Chem. Soc.* **1993**, *115*, 4859.

79. Allured, V.S.; Kelly, C.M.; Landis, C.R. *J. Am. Chem. Soc.* **1991**, *113*, 1.

80. Orbell, J.D.; Marzilli, L.G.; Kistenmacher, T.J. *J. Am. Chem. Soc.* **1981**, *103*, 5126.

81. Pettit, B.M.; Karplus, M. *J. Am. Chem. Soc.* **1985**, *107*, 1166.

Chapter 4

Quantification of steric effects in metal complex – nucleic acid systems

4.1 INTRODUCTION

4.1.1 Potential steric effects in metal complex – nucleic acid systems

The multiple binding sites on nucleic acids present the possibility of a diversity of significant steric interactions [1]. With respect to many metal species, nucleobase binding sites are favoured for coordination [2]. Therefore, in order to develop a systematic method for the quantitative assessment of steric effects in platinum complex – nucleic acid systems, platinum(II)-am(m)ine complex/nucleobase adducts should be investigated first. For square-planar complexes which may coordinate to nucleobases via a bidentate intrastrand crosslinkage [3], potential intramolecular steric interactions in the immediate vicinity of the binding site may be characterised schematically as in Fig. 4.1. These include interactions between nucleobases and carrier ligand(s) and between nucleobases themselves, both below and above the coordination plane. They may operate independently or cooperatively to restrict the geometry of the adduct.

Prior work in this area has been concentrated on the systematisation of geometrical parameters in model systems. Although these approaches do not allow quantification of steric effects *per se*, they do provide some assessment of these effects, albeit qualitatively. For example, the **stereochemical convention** referred to in the Introduction (section 1.4.1.1) [1] employs Base/Base and Base/PtN$_4$ coordination plane dihedral angles and ΔPt (perpendicular displacement of the platinum atom from the base plane) to ascertain the flexibility of adducts under a variety of influences. In spite of its suggested drawbacks [4] this convention has been widely used [5,6] and allows a detailed comparison of the conformational aspects of model systems, such as *cis*-bis(nucleobase-bound)platinum(II) complexes from which steric influences may be inferred.
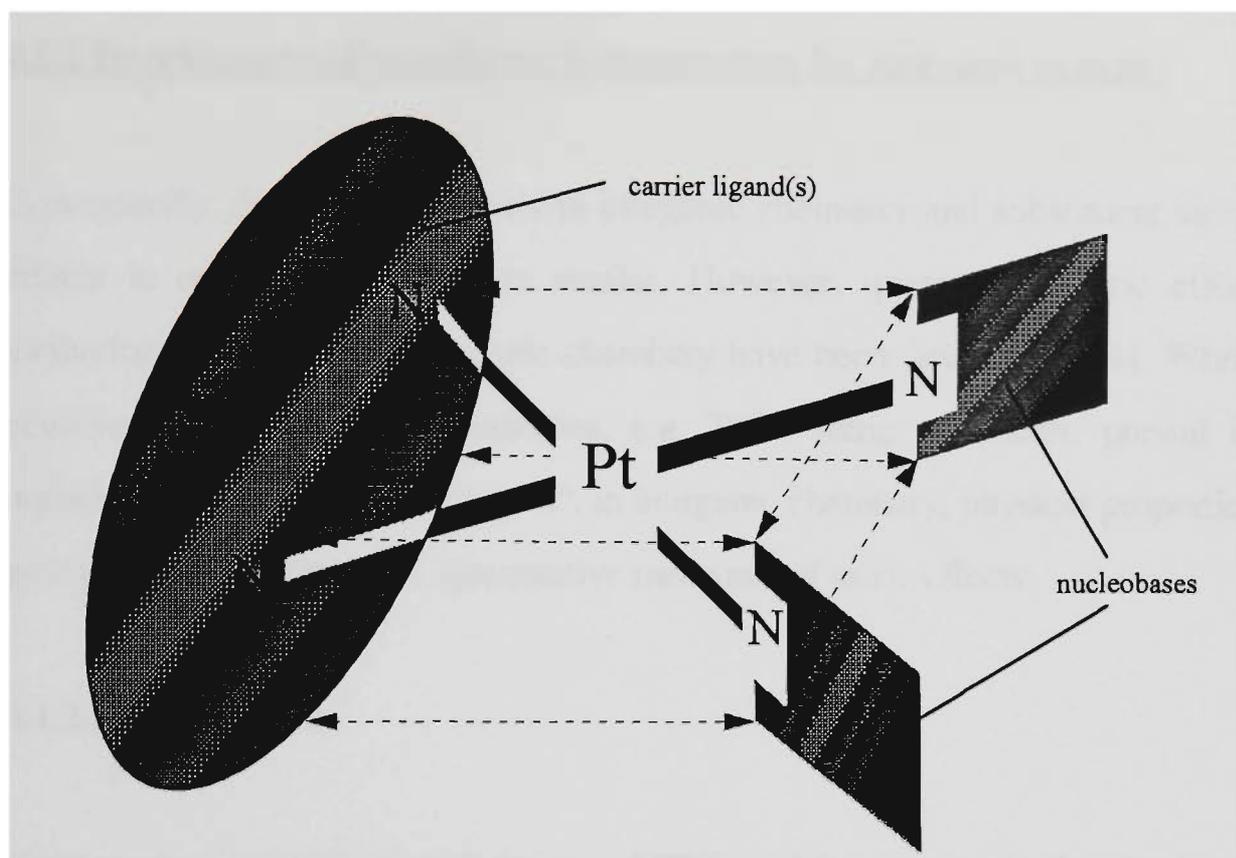
Figure 4.1. Schematic representation of platinum square-planar complexes coordinated to nucleobases via a bidentate intrastrand crosslinkage. Potential intramolecular steric interactions in the immediate vicinity of the binding site depicted as arrows.

Steric considerations prior to, during and after coordination [7] are dictated by the nature of the metal complex (usually by the carrier ligand) and by the features of the binding site itself, such as the neighbouring exocyclic substituent(s) in the case of the nucleobases [8-10]. Therefore, for the quantitative description of steric effects in these interactions, a thorough analysis of steric demands of both carrier ligands and nucleoconstituent binding sites is required. The former problem is related to the quantification of ligand steric effects in inorganic systems [11]. The latter is associated with quantifying the *accessibility* of binding sites and the *steric fit* of intermediate and/or final adducts. The following sections overview these issues.

4.1.2 Development of specific steric parameters for inorganic systems

Conceptually, ligand steric effects in inorganic chemistry and substituent steric effects in organic chemistry are similar. However, quantitative steric effect methodologies unique to inorganic chemistry have been developed [11]. While consideration of chemical properties, e.g. Taft's steric parameter, prevail in organic chemistry (see next chapter), in inorganic chemistry, physical properties govern the development of quantitative measures of steric effects.

4.1.2.1 Cone angles

Cone angles Θ, calculated with the use of CPK models, were originally introduced by Tolman [12,13] for phosphorus ligands in complexes with nickel. A cone angle Θ is the apex angle of a cylindrical cone, centred 2.28 A from the center of ligand donor atom, which touches the van der Waals radii of the outermost atoms of the model (Fig. 4.2a). The distance 2.28 A was chosen to represent the Ni-P bond length. Because cone angles were originally generated for a specific system (specific metal radii and metal–ligand bond length), they provide a *relative order* of ligand steric effects. However, it is assumed [12] that this relative order is generally valid for other systems. Tolman's concept is attractive because of its simplicity but it has certain limitations [11]; namely:

- It does not account for the conformational flexibility of ligands. The chosen conformation does not necessarily represent the energy minimum. Besides, a cone angle quantifies the steric bulk of a *free* ligand in a "folded-back" conformation [12], which often differs from a conformation of a metal-bound ligand.

- The ligands are assumed to have cylindrical symmetry. This limits the

generality of the method, although a correction technique for unsymmetrical ligands has been proposed [13].

- The effects of the molecular environment are not considered.
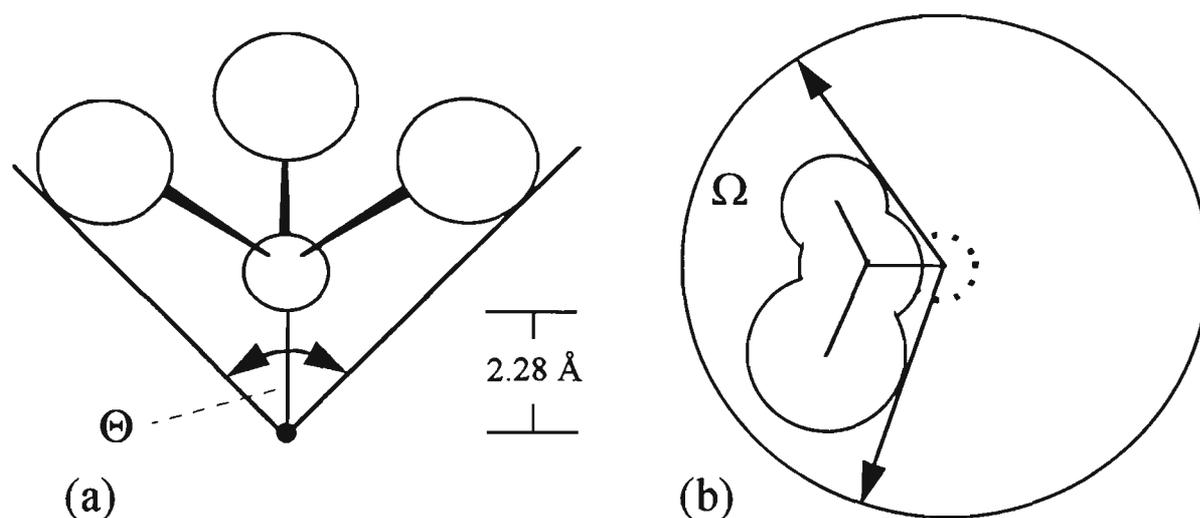
Figure 4.2. *Steric parameters specific for inorganic systems. (a) Cone angle illustration. (b) Solid angle illustration.*

4.1.2.2 Solid angles

Solid angles were recently applied by White et al. [14] in an attempt to refine the cone angle approach and to overcome its limitations. The solid angle approach was devised to satisfy the following criteria: generality; physical meaningfulness; ability to be calculated from atomic radii, bond lengths and angles only; accountability for group meshing and conformer possibilities. The solid angle Ω at a point of a surface can be represented by the integral

$$\Omega = \int_s \frac{R \times dS}{r^3} \tag{4.1}$$

where R is the position vector of an element of the surface with respect to

O and r is the magnitude of R. The solid angle concept is illustrated in Fig. 4.2b. The solid angle of a ligand can be envisaged as the area of shadow of the solid body (solid line) projected onto the inside of a sphere, with the light source centered at the metal (dotted line). The solid angle is measured in steradians (sr).

4.1.2.3 Ligand repulsive energy

The *Ligand Repulsive Energy* E_R has been developed and utilised by Brown et al. as a quantitative measure of ligand steric effects [15-19]. Originally they used ΔE (the change of the strain energy upon complexation) to estimate steric effects [20]. It was found to be an inadequate parameter, because it contains both electronic (attractive vdW forces) and steric (repulsive vdW forces) components. Consequently, Brown et al. defined steric effects in terms of purely *repulsive forces* between ligands and their molecular environment; more specifically, in terms of repulsive vdW (nonbonded) interactions. In their studies a prototype metal center $Cr(CO)_5$, which henceforth will be called the "steric probe", was chosen. Also $Cr(CO)_5$ moiety is a good representative of a transition metal species in terms of the degree of crowding about the metal centre. The rigidity, rotational symmetry and possibility to thoroughly parameterise $Cr(CO)_5$-complexes for molecular mechanics calculations promoted this choice. Brown et al. also investigated how much the variation in E_R values is caused by varying the geometrical character of the metal centre to which the ligands are bound [17]. They received a generally good correlation between E_R values of phosphites and phosphines computed for $CpRh(CO)$ and $Cr(CO)_5$ steric probes. It was concluded that the *relative* E_R values calculated for $Cr(CO)_5$-complex can be quite confidently applied to other systems as a measure of ligand steric effects. The procedure for E_R calculation (described in the Methods

section in more detail) involves the computation of the repulsive component of vdW energy $E_{vdW}(rep)$ (eqn **4.2**) for the energy minimised structure of a $Cr(CO)_5$(ligand-D) complex (D, ligand donor atom). Then the gradient of this energy with respect to the Cr-D distance r_{Cr-D} is multiplied by the Cr-D bond length in the energy minimised structure r_e (eqn **4.3**).

$$E_{vdW}(rep) = \Sigma D_0 \exp\left[\gamma\frac{r_0-r}{r_0}\right] \tag{4.2}$$

$$E_R = -r_e \times \frac{\partial E_{vdW}(rep)}{\partial r_{Cr-D}} \tag{4.3}$$

where D_0 is a potential well depth, γ is a scaling factor (=12.5), r are interatomic distances in the energy minimised structure and r_0 are unstrained interatomic distances. Multiplication of the gradient by r_{Cr-D} is done to distinguish between steric effects of ligands which give similar slope values (for plots of $E_{vdW}(rep)$ vs r_{Cr-D}) but have different r_{Cr-D} values.

Brown's concept has certain advantages in comparison to Tolman's. It considers a ligand in a most appropriate conformation. Also, it can be applied to asymmetrical ligands. Ligand repulsive energy values are versatile, subject only to adequate parameterisation of a system under study.

The limitations of Brown's concept are well recognised [15]. Thus, ligand repulsive energy is calculated for a specific system and the results depend on the force field used. Calculation is carried out in a gas-phase so solvent interactions are not considered. Finally, ligand repulsive energy values are of questionable interpretation when different conformations of similar strain energy exist.

4.1.2.4 Modified vdW energy

Modified vdW energy $E_{vdW}(mod)$ was formulated by Woo and Ziegler [21] in one more effort to find a MM-derived characteristic of steric effects. This energy has the same analytical form as $E_{vdW}(rep)$ (eqn **4.2**) but with modified D_0, γ and r_0 parameters, fitted so that $E_{vdW}(mod)$ reproduces the repulsive stretch of E_{vdW}. The advantages of this method are the simplicity of the calculation procedure and the applicability to any kind of ligand, multidentate as well as monodentate. Its drawbacks should also be noted. Namely, it requires a separate scaling factor for each atom pair, which significantly complicates the force field parameterisation. The arrival at a correct minimum energy structure is doubtful, since $E_{vdW}(mod)$ overlooks attractive nonbonded interactions. Rejection of Brown's gradient approach leads to losing the conformational aspect. The gradient approach allows for "larger" ligands (in scaler terms, e.g. MV) to have lower ligand repulsive energy values than those for "smaller" ligands in cases where larger ligands minimise their steric requirements via conformational changes. In contrast to the gradient approach, $E_{vdW}(mod)$ seems to be clearly proportional to the scalar size of a ligand.

4.1.3 Development of steric parameters for nucleobase metal-binding sites

Nucleic acids exhibit a large variety of potential metal binding sites. These may be grouped as follows.

- endocyclic nitrogens on nucleobases (Fig. 4.3): N1, N3 and N7 sites of purines and N3 sites of pyrimidines

- exocyclic groups: O6(G), N4(C), N6(A), O6(T) in the major groove and N2(G), O2(T), O2(C) in the minor groove

- sugar and phosphate oxygens
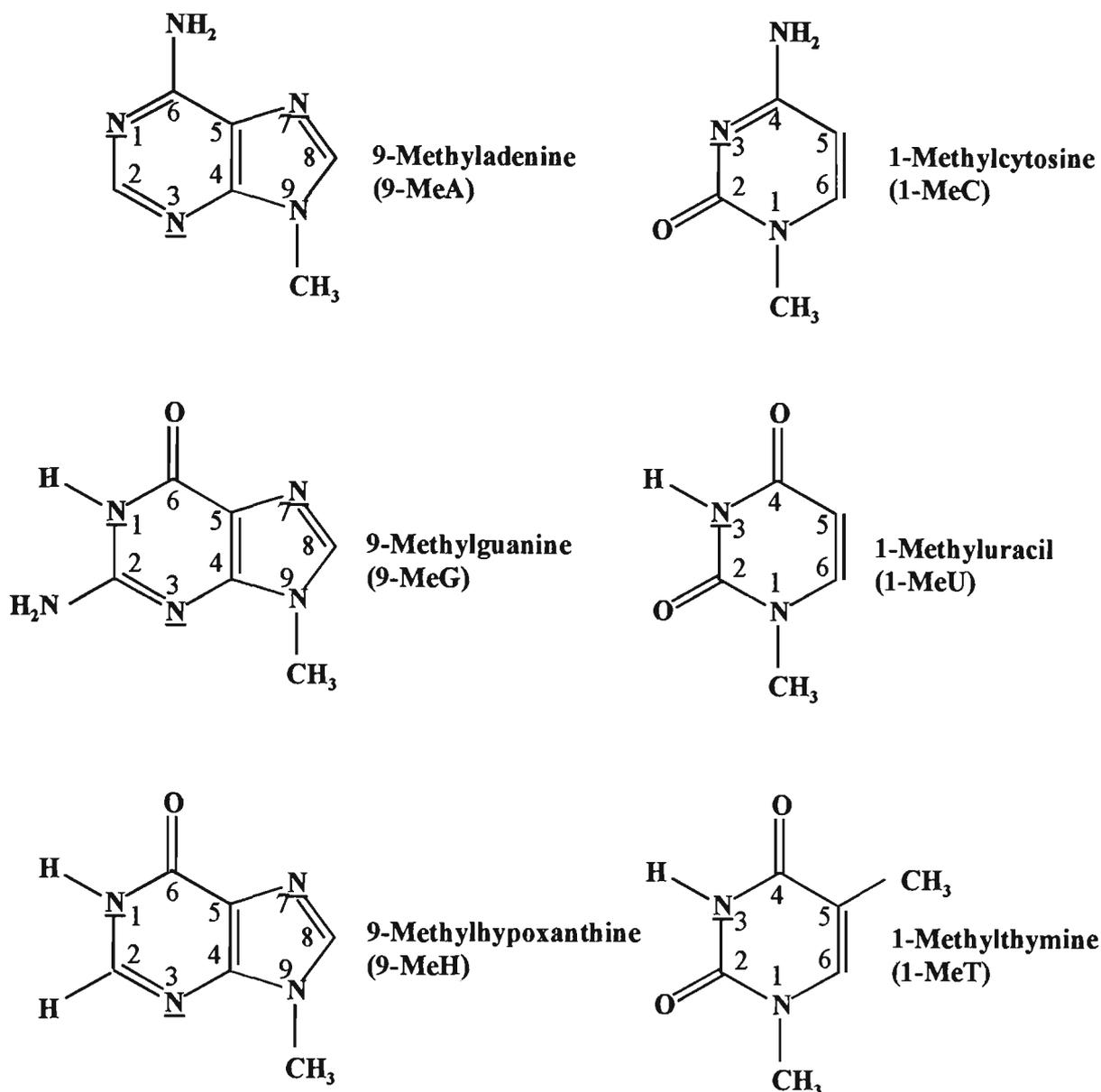
Figure 4.3. *Common nucleobases with atom numbering scheme. Endocyclic nitrogen metal-binding sites are underlined. N1 site in guanine and hypoxanthine and N3 site in uracil and thymine require deprotonation prior to metal binding.*

The endocyclic nitrogens of the common nucleobases (Fig. 4.3) are of particular interest with respect to metallochemotherapeutics, such as the antitumour drug cisplatin, since these sites are considered to be key molecular targets for such compounds [22,23]. Of the potential binding sites listed above, several were observed in platinum binding; the majority of these sites have been

unambiguously confirmed by X-ray crystallography, some also verified spectrophotometrically (NMR, Raman, UV, IR) [24]. The long established sites include: N7(purines), N1(purines), N3(pyrimidines), N7N1(purines). The novel recently verified patterns incorporate: N1N3N7(G), N3N4(C) bridging and chelating, N4(C), N3O4(T,U), C5(U) and others in heteronuclear complexes [24]. The above are summarised in Fig. 4.4.
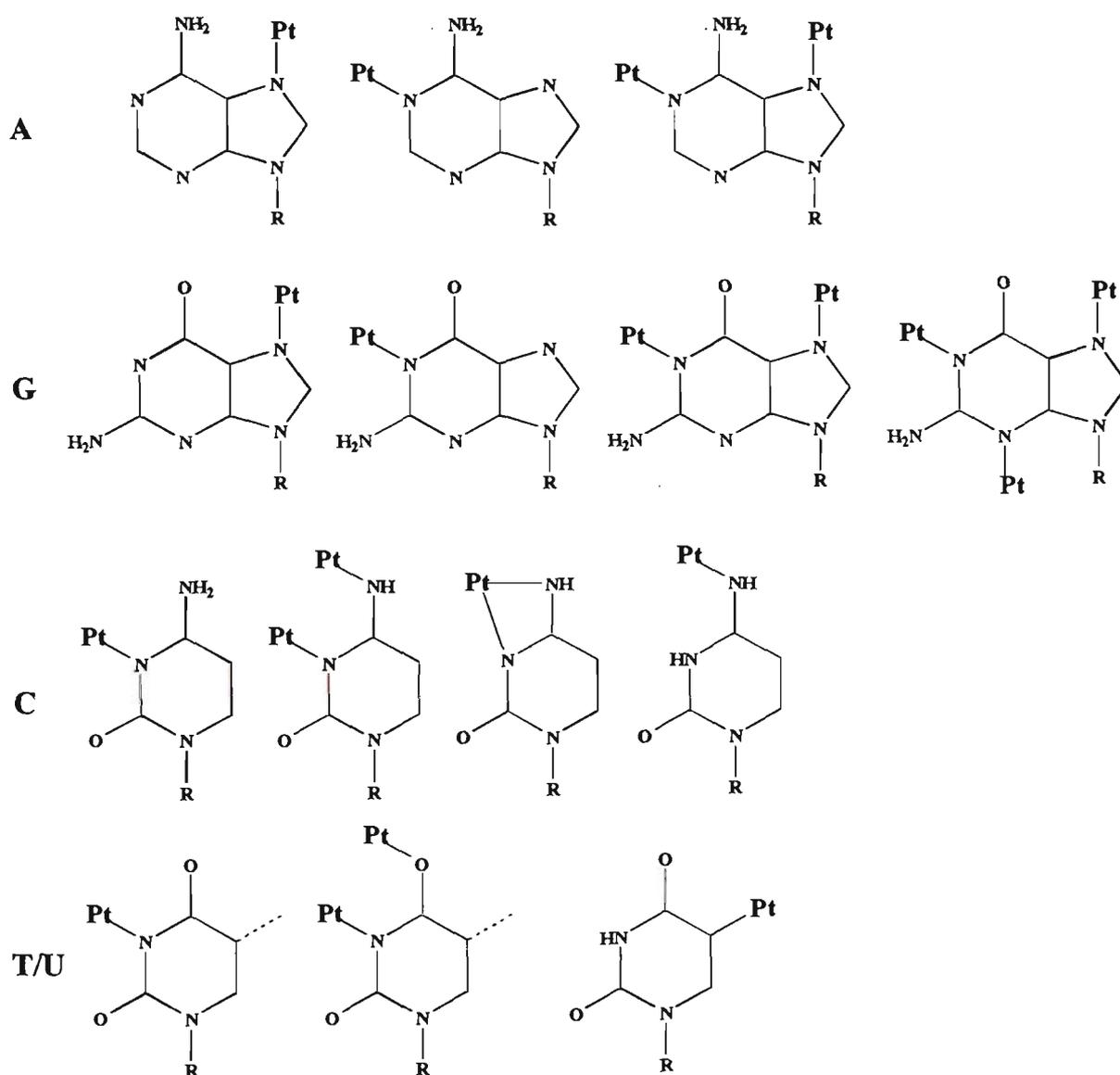
Figure 4.4. Experimentally observed Pt binding sites and modes of coordination for the four major bases A, G, C and T/U. Dotted line correspond to methyl group in thymine.

Factors which influence a particular mode of coordination are varied, complex and interrelated and include the following:

- electronic factors, including basicity [25];
- steric factors, including fitting and accessibility;
- metal speciation (oxidation state, hydrolysed form);
- interligand interactions.

With respect to electronic factors, the preferential binding (in particular to the N7 position of guanine) has been quantitatively rationalised through pK_a values [25]. Relative coordination strengths of different binding sites have also been rationalised on the basis of *electrostatic potential energy distributions* [8,26,27] and *ab initio SCF bonding energies* [28], for a range of endocyclic and exocyclic nucleobase sites.

The relative steric demands of such binding sites have been considered by some workers to be difficult, if not impossible, to quantify [29]. An attempt to quantify steric influences of nucleobase binding sites, among other sites within a B-DNA macromolecule, was made by Lavery et al. [30]. Using a "steric probe"-like approach, they calculated static steric accessibilities of atoms within B-DNA, presented in the form of *accessible area AA* and *intrinsic accessible area IAA*, the latter associated with atoms within macromolecule subunits (bases, sugars and phosphates). This concept is a steric equivalent of the molecular electrostatic potential technique in that it is characteristic of what a reactant (steric probe) "feels" when approaching a binding site. This methodology enabled the comparison of reactional potentialities for a range of nucleobase endocyclic, as well as exocyclic, sites, and resulted in a good agreement with experimental data. However, it suffers from several drawbacks, limiting its otherwise very promising applicability. It does not

consider structural changes in the reacting species, leading to the "static" nature of obtained parameters. Although it takes into consideration what authors term the configurational flexibility of the attacking molecule, the corresponding procedure is not rigorous enough. In principle, both limitations can be easily improved by the use of geometry optimisation and conformational search procedures.

4.2 METHODS

4.2.1 Repulsive Energy (RE) Methodology

Structural effects at the binding interface between a small molecule and a biological macromolecular target (e.g. protein, DNA) are determinative of the thermodynamic and kinetic nature of the interaction and its biological consequences. Quantitative relationships between various structural parameters and indicators of biological activity for platinum complexes are discussed in the following chapter. In this chapter the attention is focused specifically on the quantification of parameters suitable for the description of steric effects related to the binding of metal complexes to nucleoconstituents. Fig. 4.5 illustrates an interaction between a metal complex and a biomolecular target. Steric effects arising from both carrier ligand and binding site could be operative in this interaction.

In the present study a Ligand Repulsive Energy (LRE) strategy [16], has been employed and extended for the quantification of steric effects involved in the above interactions. Henceforth it will be referred to as Repulsive Energy (RE) strategy, since it has been applied to the characterisation of steric demands of metal complex moieties as well as those of isolated ligands. Generally, the RE method (Fig. 4.6) involves calculation of the repulsive component of the vdW energy between a ligand and a metal-based steric probe (eqn 1 in Fig. 4.6). Subsequently, the gradient of this energy with respect to varied ligand–metal distance is calculated for the linear segment the curve in order to obtain a resulting steric parameter (eqn 2 in Fig 4.6).
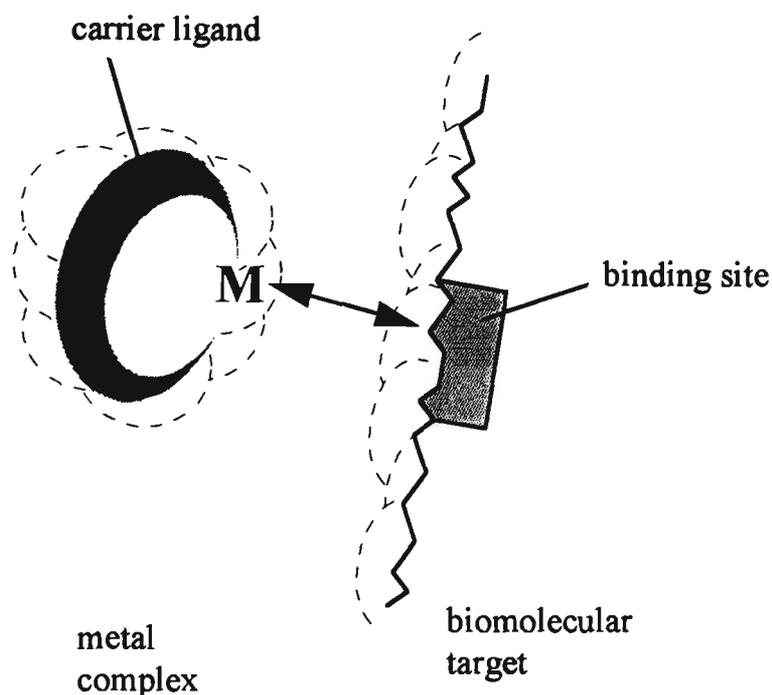
Figure 4.5. *Schematic representation of the interaction between a metal complex and a biomolecular target. In the context of this study, "binding site" refers to a particular donor atom on the target and its immediate surroundings, such as exocyclic substituents on a nucleobase, one or two atoms removed from an endocyclic nitrogen. In subsequent schematic diagrams the dotted lines representing the molecular surface are omitted for clarity.*
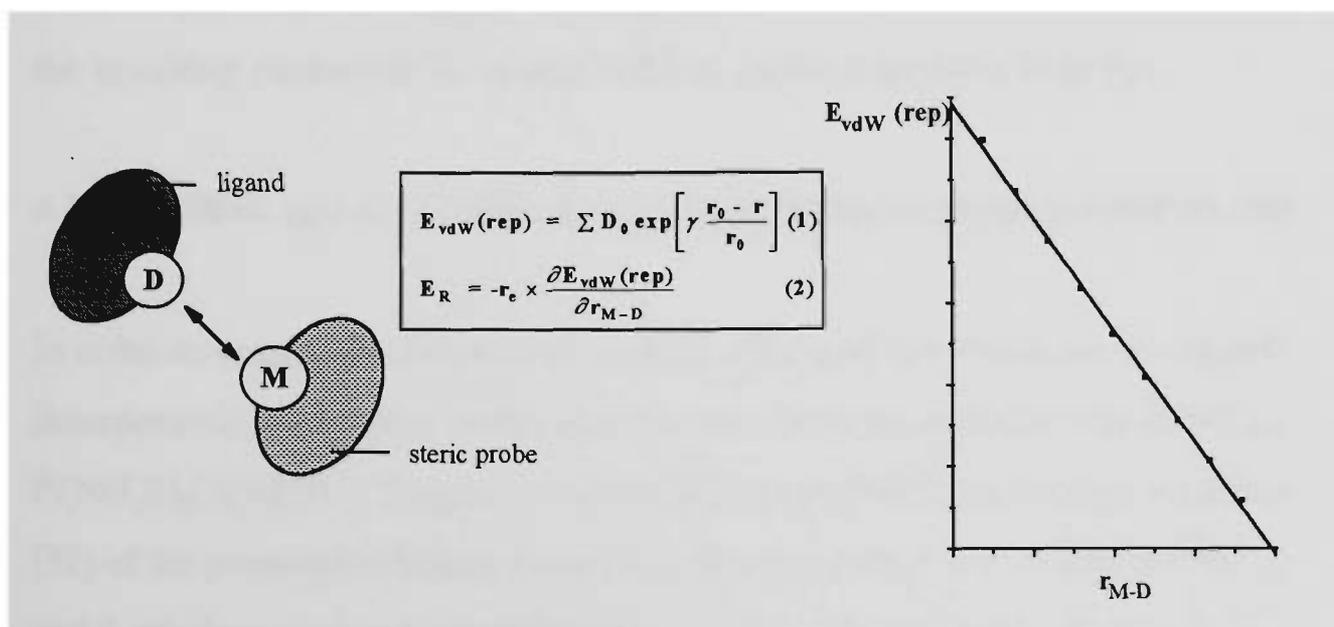
Figure 4.6. *Schematic representation of the RE method. D is the ligand donor atom, M is the metal atom of the steric probe. The symbols in equations (1) and (2) are described in section 4.2.2.*

Steric effects in metal complex interactions with nucleic acids and their constituents are diverse. The goal of this study is to develop a systematic method for the quantitative assessment of steric effects in Pt(II)-am(m)ine complex – nucleobase systems. The variety of approaches used to pursue this goal (combined under the term *Repulsive Energy Methodology*) are outlined below and depicted in Fig. 4.7.

4.2.1.1. Application of the RE strategy to steric effects of carrier ligands

4.2.1.1.1 Steric effects of isolated carrier ligands

The original repulsive energy strategy [16] has been utilised to calculate steric requirements of four series of am(m)ine carrier ligands NH_2R (Fig. 4.8). These compounds were selected because they represent a range of potential steric requirements and the biological profiles of their platinum complexes cis-[Pt(NH$_2$R)$_2$Cl$_2$] are well characterised [31]. Therefore they may be used in consequent QSAR studies. This approach is illustrated in Fig.4.7(A) and the resulting parameter is termed *LRE* (Ligand Repulsive Energy).

4.2.1.1.2 Steric effects of carrier ligands incorporated in complex metal species

In order to quantitatively assess the steric effects of the above carrier ligands incorporated in complex metal species, the steric repulsion is calculated for $Pt(NH_2R)_2Cl(9-EtG-N7)$ species. This models the potential monodentate precursor [32] of the proposed bidentate interaction between complexes cis-[Pt(NH$_2$R)$_2$Cl$_2$] and 9-ethylguanine, a representative model of a binding target on DNA. This technique is represented by the approach of B_1 to B_2 (Fig. 4.9) and illustrated in Fig.4.7(B).

Ligand	Steric probe	Steric parameter
NH₂R *varied*	Cr(CO)₅ *constant*	*LRE*
9-EtG *constant*	Pt(NH₂R)₂Cl *varied*	*CRE*
nucleobase binding site *varied*	Cr(CO)₅ *constant*	*Eᴿ(Cr)*
nucleobase binding site *varied*	Pt(A)₃ *constant*	*Eᴿ(Pt)*
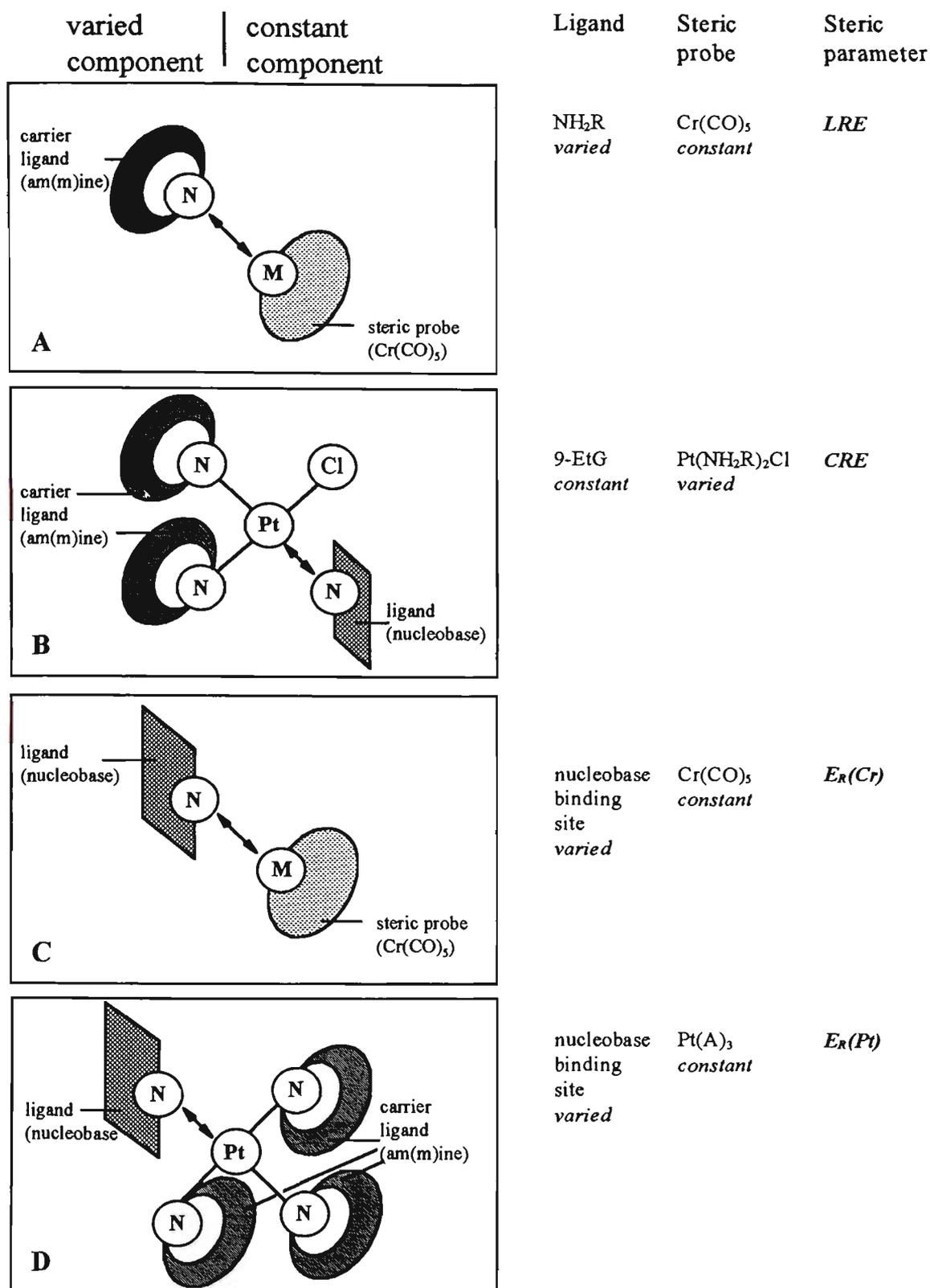
Figure 4.7. Different approaches within the Repulsive Energy Methodology. Parts A and B illustrate the quantification of carrier ligand steric effects, isolated and incorporated in platinum complexes respectively. Parts C and D illustrate the quantification of nucleobase steric effects, using chromium and platinum steric probes respectively.

STRAIGHT CHAINS:

$H_3C\!-\!NH_2$ *(20)*

$C_2H_5\!-\!NH_2$ *(21)*

$n\text{-}C_3H_7\!-\!NH_2$ *(22)*

$n\text{-}C_4H_9\!-\!NH_2$ *(23)*

$n\text{-}C_5H_{11}\!-\!NH_2$ *(24)*

$n\text{-}C_6H_{13}\!-\!NH_2$ *(25)*

$n\text{-}C_7H_{15}\!-\!NH_2$ *(26)*

$n\text{-}C_8H_{17}\!-\!NH_2$ *(27)*

CYCLES:

▷$-NH_2$ *(2)*

◇$-NH_2$ *(3)*

⬠$-NH_2$ *(4)*

⬡$-NH_2$ *(5)*

⬭$-NH_2$ *(6)*

◯$-NH_2$ *(7)*

$HO\!-\!\bigcirc\!-\!NH_2$ *(18)*

$H_3C\!-\!\bigcirc\!-\!NH_2$ *(19)*

BRANCHED CHAINS:

$\rangle\!-\!NH_2$ *(12)*

$-NH_2$ *(13)*

$-NH_2$ *(14)*

$-NH_2$ *(15)*

$-NH_2$ *(16)*

NH_2 *(17)*

POLYCYCLES:

$-NH_2$ *(8)*

$-NH_2$ *(9)*

$-NH_2$ *(10)*

NH_2 *(11)*
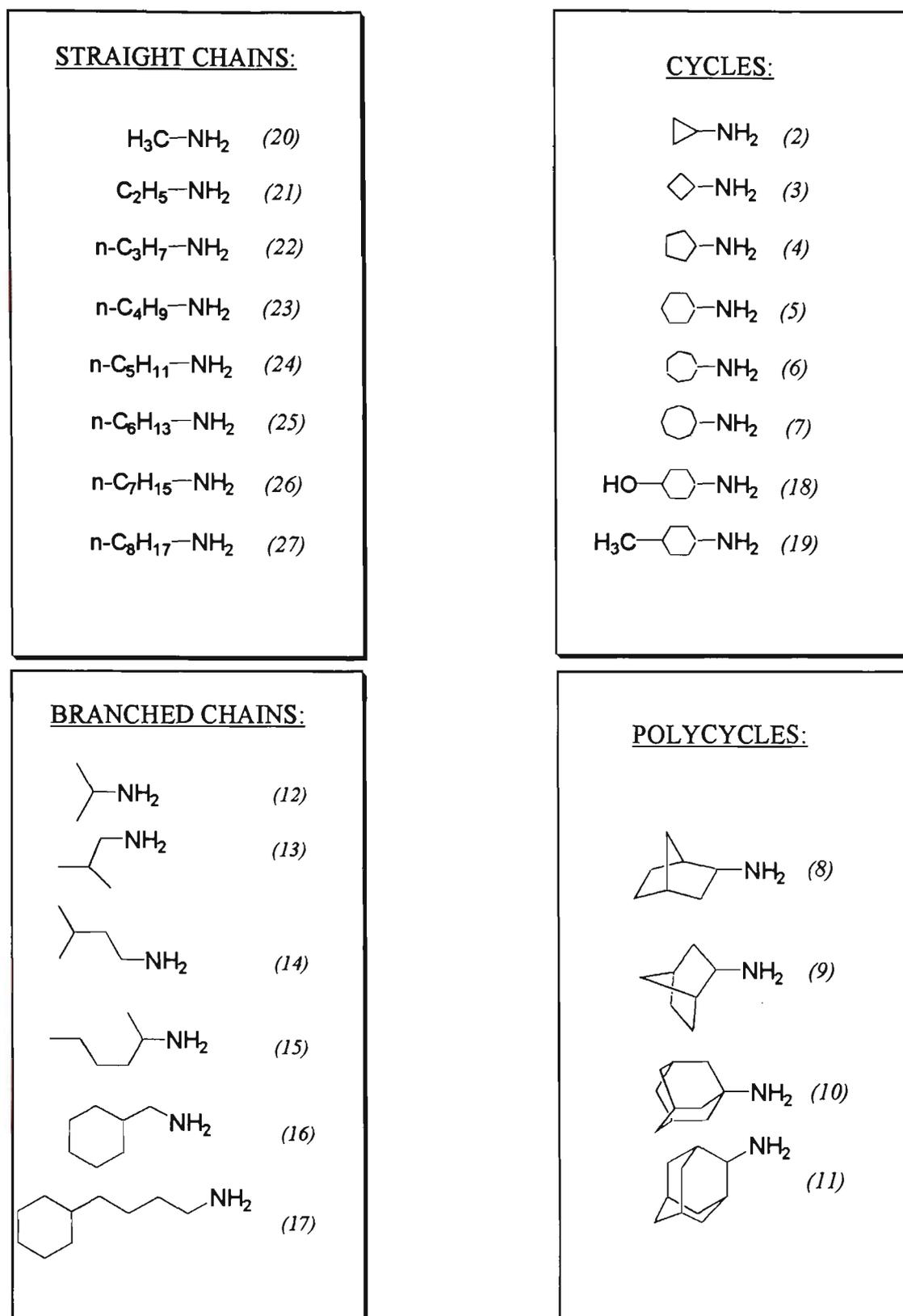
Figure 4.8. Carrier ligand series. Numbers in brackets correspond to numbering of complexes in Table 5.1.
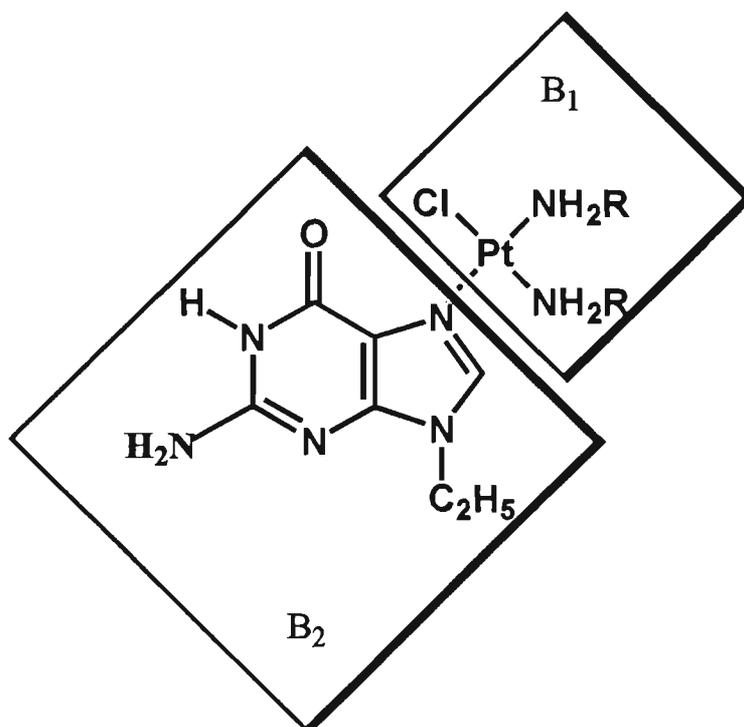
Figure 4.9. A schematic representation of a monoadduct between cis-platinum complexes (B_1) and 9-ethylguanine (B_2).

The resulting parameter is called **CRE** (Complex Repulsive Energy), since the metal containing steric probe $Pt(NH_2R)_2Cl$ is the varied species. The **CRE** parameter is characteristic of the steric features of a metal complex rather than these of a nucleobase. This adduct is chosen on the basis of an assumption that an initial monodentate attack occurs on the N7 position of the guanine moiety without a replacement of the remaining chloro ligand [33]. It is possible, of course, that the steric demands of both initial and consolidated interactions are relevant to the biological outcome.

The first of these two approaches, **LRE**, simplifies the calculation of parameters and in its nature parallels the use of traditional substituent constants (steric, electronic and hydrophobic) in classical QSAR studies of organic compounds [34]. The second approach, **CRE**, is more reflective of ligand – receptor modelling

in drug – protein studies and may also provide insight into the mechanism of action of the platinum drugs.

4.2.1.2. Application of the RE strategy to steric effects associated with nucleobase binding sites

4.2.1.2.1 Steric parameters of nucleobase binding sites, calculated using the Cr steric probe

The original repulsive energy strategy is utilised in order to calculate steric requirements of twelve typical binding sites on common nucleobases (Fig. 4.3). This is illustrated in Fig.4.7(C) and the resulting parameter is called $E_R(Cr)$.

4.2.1.2.2 Steric parameters of nucleobase binding sites, calculated using Pt steric probes

In order to calculate steric requirements of the above binding sites with respect to their interactions with platinum complexes, monodentate platinum species of the type $Pt^{II}(A)_3$ are employed as steric probes ($A \equiv NH_3$, $(A)_3 \equiv tria(m)mine$, see diagram in section 3.3.5). This technique does not differ methodologically from the previous approach. However, the steric and conformational outcomes are expected to be different due to the more "planar" nature of these probes. This, therefore, warrants separate classification. This approach is illustrated in Fig.4.7(D) and the resulting parameter is called $E_R(Pt)$.

4.2.2 The procedure of repulsive energy calculation

To accommodate alternative software [35], the procedure as described in Ref. [16] was slightly modified. Using Refs. [16,36] as controls, identical outcomes were established for *LRE* values calculated for selected ligands. The general strategy for the calculation of parameters described in sections 4.2.1.1 and 4.2.1.2 is similar and may be described as follows.

1) Obtain the lowest energy structure for the appropriate complex. The parameters added to the MM+ force field of HyperChem [35] to model the chromium complexes of amine ligands NH_2R (Fig. 4.8) and nucleobases (Fig. 4.3) are compiled in Tables 4.1 and 4.2. MM+ atom types in nucleobases are compiled in Appendix VI. The platinum complexes were geometry optimised using the AMBER force field of HyperChem as described in Chapter 3. The conformational searches were carried out for all complexes as described in the Methods section of Chapter 3.

2) For a given complex, vary $r_{Me-N} \pm 0.08$ Å, with all other internal coordinates frozen, to create a set of structures[1].

3) Using the nonbonded parameters of a modified MM+ force field for both chromium and platinum complexes[2], compute the repulsive portion of E_{vdw} for each structure in the above set, according to the eqn **4.4**.

---

[1] In order to do this in HyperChem the metal-nitrogen bond is name-selected, a steric probe is selected, and a molecular system is aligned with LINE. HyperChem shortkeys then allow to translate the selection (a translation-step and selection-translation are specified in the Tool Preferences Dialog Box).

[2] In order to use MM+ nonbonded parameters for platinum complexes, the atom types are converted from AMBER to MM+ force field. The use of MM+ non-bonded parameters for repulsive energy calculation for platinum complexes secures the internal consistency between the calculation of repulsive energies for Pt and Cr systems. The MM+ nonbonded parameters associated with Pt atom are provided in Table 3.7.

Table 4.1. Parameters added to the MM+ force field for geometry optimisation of $Cr(CO)_5(NH_2R)$ complexes.

(i) New atom types

atom type	description
CM	chromium
CR1[a]	sp-carbon radial
CR2[a]	sp-carbon radial
C2A	sp-carbon axial
OR1[a]	carbonyl oxygen radial
OR2[a]	carbonyl oxygen radial
O1A	carbonyl oxygen axial

(ii) Nonbonded parameters[b]

atom type	r*(Å)	ε(kcal/mol)
CM[c]	2.220	0.416

(iii) Bond Stretch parameters[c]

bond	r_0(Å)	K_r(mdyn/Å)
CM-N3	2.14	2.625[d]
CM-C(sp, radial)	1.88	2.100
CM-C(sp, axial)	1.85	2.100
C(sp)-O(carbonyl)	1.12	17.040
O(carbonyl)-LP	0.6	4.600

(iv) Angle Bend parameters[c]

angle	θ_0(deg)	K_θ(mdyn×Å/rad^2)
C(sp, radial)-CM-N3	90.00	0.500
C(sp, axial)-CM-N3	180.00	0.000
CM-N3-C4	115.00	0.210
CM-NH-HN	105.25	0.210
CM-C(sp)-O(carbonyl)	180.00	0.500
C(sp)-O(carbonyl)-LP	180.00	0.521
CR1-CM-CR1	180.00	0.000
CR2-CM-CR2	180.00	0.000
CR1-CM-CR2	90.00	0.550
C(sp, radial)-CM-C2A	90.00	0.550

(v) Torsional parameters

All torsional barriers involving chromium were set to 0[c].

[a]Two pairs of radial sp-carbon and carbonyl oxygen atom types are required to correctly describe angles in the radial plane. [b]Nonbonded parameters associated with the new atom types for C[rp] and carbonyl O are reproduced from standard MM2 atom types C2 and O1. [c]Values acquired from Brown et al. {415,424}. [d]The force constant was set larger than that in Brown et al. {415} to compensate for the inclusion of 1,3-metal centered contacts into the vdW interactions in the MM+ force field. The force constant value was fitted to reproduce control structures.

Table 4.2. Parameters added to the MM+ force field for geometry optimisation of $Cr(CO)_5$(Nucleobase) complexes.

(i) Bond Stretch parameters

bond	r_0(Å)	K_r(mdyn/Å)
CM-NA[a]	2.14	2.625
CM-N2[a]	2.14	2.625
NA-CO[b]	1.358	3.176

(ii) Angle Bend parameters[c]

angle	θ_0(deg)	K_θ(mdyn×Å/rad^2)
Cr-N(sp^2)-C(sp^2)[d]	120.00	0.695
N2-CA-NA[e]	120.00	0.486
N2-C3-CO[f]	130.00	0.486
N2-C3-NA[g]	122.90	0.486
C3-NA-CO[h]	120.50	0.486
N2-CO-NA[i]	118.60	0.486
NA-CO-O1[j]	122.50	0.556

For standard numbering in nucleobases see Fig. 4.3. [a]Parameters for this bond were set to the same values as for Cr-N(sp^3) as justified Brown et al. {415}. [b]Cytosine C2-N3 bond. [c]Nucleobase angle bend parameters are taken directly from AMBER since, as characterised by X-Ray crystallography, the effect of metal complexation on the base geometry is only marginal. Furthermore, for the purpose of this particular study it is desirable to keep nucleobase geometry as native as possible. [d]To accommodate the diversity of Cr-N-C angles the average value of 120° was chosen; the HyperChem "wild-card" approach assigns this value with a force constant of 0.695 mdyn×Å/rad^2 to these angles. HyperChem "wild-card" approach was also used to assign default values to torsional barriers. [e]Adenine N(3)-C(4)-N(9) (126.20) and N(6)-C(6)-N(1) (119.30) angles are averaged. [f]Guanine N(7)-C(5)-C(6) angle. [g]Guanine N(1)-C(2)-N(3) (123.30), N(2)-C(2)-N(3) (119.30) and N(9)-C(4)-N(3) (126.20) angles are averaged. [h]Cytosine C(4)-N(3)-C(2) angle. [i]Cytosine N(1)-C(2)-N(3) angle. [j]Cytosine N(3)-C(2)-O(2) angle.

$$E_{vdW}(rep) = \sum D_0 \exp\left[\gamma\frac{r_0 - r}{r_0}\right] \tag{4.4}$$

where D_0 represents the potential well depth, γ is a scaling factor (typically 12.5), r are the interatomic distances in the energy minimised structure and r_0 are the unstrained interatomic distances. Since HyperChem does not allow force field functionality to be changed [37], it was necessary to develop a procedure allowing the recalculation of the repulsive portion of E_{vdW} with the use of external software. For this purpose the Single Point calculations are carried out for each structure in the set, created as described above, and output information containing vdW and other interactions is saved.

4) Calculate the repulsive energy according to:

$$E_R = -r_e \times \left[\frac{\partial E_{vdW}(rep)}{\partial r_{Me-N}}\right] \tag{4.5}$$

where r_{Me-N} represents a varied metal-to-nitrogen distance and r_e is the metal-to-nitrogen distance in the energy minimised structure. This calculation was carried out with the help of the spreadsheet package Origin [38]. A particular modification to the original protocol [16] was necessary for LRE and $E_R(Cr)$ calculation due to a POS treatment of octahedral complexes in MM+ force field (see section 3.1.4 for explanation of the POS method). Since in MM+ force field metal-centred 1,3-interactions are included into the vdW energy calculation [37], it was necessary to exclude the correspondent lines, containing "bad" contacts, from the spreadsheet prior to the calculation. Such truncation does not impede the procedure significantly, since there are only five such contacts affecting the final E_R value, namely contacts between carbonyl carbon atoms of the $Cr(CO)_5$ moiety and nitrogen donor atom of a ligand. The vdW

interactions between carbonyl carbon atoms within the $Cr(CO)_5$ moiety do affect actual values of $E_{vdw}(rep)$ but are constant in all structures within the set, therefore not affecting the gradient in the eqn **4.5**.

4.2.3 Calculation of other parameters

The *logP* values and the vdW molecular volumes (*MV*) for the coordinated amine ligands were calculated using the QSAR module of ChemPlus [39], a set of extension modules to HyperChem. The calculation of *logP* values is described in more detail in the following chapter (section 5.2.1.4).

4.2.4 Auxiliary computing

Auxiliary computer programs are compiled in Appendix I.

QBASIC routines have been written for:

- calculation of dihedral angles (Appendix I(c))
- transformation of HyperChem output energy information required for the RE calculation in order to import it into external spreadsheet software; since there is no option which allows only the vdW information to be saved [37], the routine includes a means to extract this from the output file (Appendix I(d))
- transformation of HyperChem output energy information required for plotting rotational profiles in order to import it into external spreadsheet software (Appendix I(e))

HyperChem scripts have been written for:

- generating and saving output energy information necessary for the RE calculation (Appendix I(f))
- generating and saving output energy information necessary for plotting rotational profiles (Appendix I(g))

Origin [38] script has been written for the RE calculation (Appendix I(h)).

4.3 RESULTS AND DISCUSSION

4.3.1 Steric effects of carrier ligands

The values of *LRE* and *CRE* for the twenty-seven selected ligands and correspondent complexes are given in Table 4.3. In order to investigate the "purity" of steric information these parameters were pair-wise correlated with *MV* (molecular volume) and *LogP* (hydrophobicity) of the ligands (Fig. 4.10). These parameters are also given in Table 4.3.

Table 4.3. Steric and hydrophobicity parameters.

No.	LRE	CRE	logP	MV	No.	LRE	CRE	logP	MV
1	10.00	4.33	-1.96	22	15	32.66	12.96	2.57	125
2	23.28	8.67	-.67	67	16	24.13	14.16	2.33	130
3	26.57	9.83	.12	82	17	23.02	15.91	4.57	182
4	28.83	16.45	.91	97	18	30.55	19.55	-1.34	123
5	31.80	21.34	1.71	114	19	31.62	13.10	2.37	131
6	32.87	15.96	2.50	130	20	30.00	12.50	-1.46	40
7	37.02	13.59	3.94	147	21	31.00	15.69	-.78	57
8	32.85	11.21	1.37	121	22	31.00	14.15	.16	74
9	30.17	13.61	1.37	121	23	31.39	21.57	.95	91
10	42.93	21.73	1.98	159	24	32.26	23.83	1.74	108
11	35.70	17.25	2.62	159	25	30.94	19.07	2.54	125
12	30.60	17.41	.01	74	26	31.62	16.07	3.33	142
13	33.92	15.83	.96	91	27	33.99	27.63	4.12	159
14	32.17	16.14	1.61	108					

Units: *LRE*, *CRE* – kcal/mol, *MV* – Å^3.
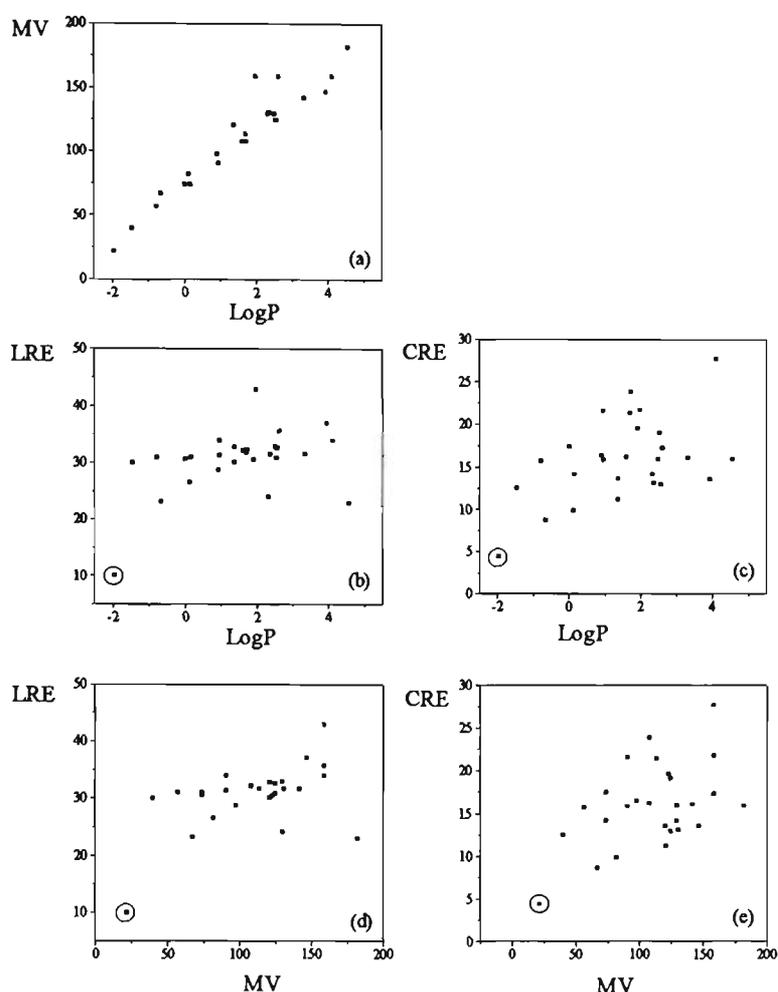
Figure 4.10. Relationships between steric and transport descriptors.

4.3.1.1 *LRE* as a steric parameter

Previously, to describe steric effects of carrier ligands in platinum-am(m)ine complexes, the *MV* parameter has been employed [40]. The use of *MV* as a steric descriptor oversimplifies steric effects since *MV* is a scalar quantity [41]. Besides, there is another consideration which must be taken into account.

Not surprisingly, for the ligands of this study *MV* correlates very well with hydrophobicity, represented by *logP*, as shown in Fig. 4.10a (R = 0.95). Hence *MV* carries with it an undesirable amount of transport information. The phenomenon and consequences of "overlapping information" in structural descriptors will be discussed in a greater detail in Chapter 5.

The steric descriptor *LRE* is much less correlated with hydrophobicity, as shown in Fig. 4.10b (R = 0.44). It should also be noted that when the sole ligand of the series which may be considered hydrophilic, i.e. NH_3 (circled data point), is omitted from the regression, the correlation coefficient drops to R = 0.22. Indeed, the horizontal array of Fig. 4.10b suggests a high degree of separation of transport and steric effects and bolsters confidence in *LRE* as a steric descriptor.

4.3.1.2 *CRE* as a steric parameter

Using *CRE* as an indicator of carrier ligand steric effects in platinum complexes is built upon the assumption that all of the platinum complexes referred to target the N7 position of guanine and that their individual biological profiles are a result of the structural variation of the amine ligand(s). Thus, the cisplatin analogues chosen for this study, feature standard modifications which are expected to influence biological outcomes [34].

Employing a rigid symmetrical steric probe such as $Cr(CO)_5$, the *LRE* method allows the relative repulsive energies of a series of ligands to be assessed and to be applied to other systems [17]. On the other hand, the *CRE* method developed in this work is tailored to a particular scenario. Namely, to the monodentate interaction of a platinum complex $Pt(NH_2R)_2Cl^+$ with the N7

position of a guanine moiety (Fig. 4.9). Thus guanine is a constant component in this interaction, while the carrier ligands are varied (Fig. 4.7, part (B)). Since 9-EtG is obviously not spherically symmetrical with respect to the direction of approach, it is first necessary for the approach to be optimised. This may be approximated by a global minimum conformation search for the postulated adduct. During this procedure the relative orientation of the nucleobase with respect to the coordination plane, the orientation of the 9-ethyl substituent relative to the nucleobase as well as the conformation of a carrier ligand itself are varied. Thus, the obtained *CRE* values represent relative measures of steric characteristics presented by each $Pt(NH_2R)_2Cl^+$ moiety towards a specific nucleoconstituent.

Like *LRE*, *CRE* correlates poorly with hydrophobicity (Fig. 4.10c, R = 0.5; with the NH_3 data point removed, R = 0.37). The slightly higher correlation for *CRE* compared to *LRE* is perhaps due to a shift of emphasis away from the symmetrical region around the donor nitrogen onto the bulk of the carrier ligand which may impinge on the opposing nucleobase.

Not unexpectedly, given the close relationship between hydrophobicity and *MV* for these ligands (Fig. 4.10a), the correlation of both *LRE* and *CRE* with *MV* is also poor (Fig. 4.10(d,e), R = 0.51 and 0.49 for *LRE* and *CRE* respectively). With the NH_3 data point removed, R = 0.29 and 0.49 for *LRE* and *CRE* respectively. For the ligands considered here, it is expected that a larger molecular volume will be associated with a more flexible molecule. In a platinum complex, ligand flexibility is expected to increase the likelihood of steric contacts with an opposing nucleobase.

The above considerations of *LRE* and *CRE* have the potential to provide insights

into the distribution of sterically significant bulk on the carrier ligand. This may be of particular importance in the design of systems whereby sterically restrictive carrier ligands are employed in an attempt to enforce or manipulate a particular orientation of nucleobase(s) such as HTT or HTH [42]. Thus, *LRE* and *CRE* parameters as applied to the approach of a metal complex to a target molecule allow for an improved description of steric effects in such systems. Where transport, electronic and steric effects are all to be considered, the *LRE* and *CRE* parameters have a higher degree of orthogonality than those employed previously.

4.3.2 Steric effects of nucleoconstituents

4.3.2.1 Use of the original RE strategy

For the common nucleobases studied here, the repulsive energy values have been calculated and are presented in Table 4.4 and Fig. 4.11. The numbers shown in Fig. 4.11, associated with each site assessed, represent a steric index, I_S:

$$I_S = \frac{E_R[N(nucleobase)]}{E_R[N7(guanine)]} \tag{4.6}$$

where N(nucleobase) is N1(guanine, hypoxanthine or adenine), N3(cytosine, thymine or uracil), or N7(guanine, hypoxanthine or adenine). This definition is predicated upon the N7 of guanine presenting the lowest relative steric hindrance to the probe, of all the sites compared in this study.

Table 4.4. Ligand Repulsive Energies Presented by Nucleobase Binding Site to the Metal Species $Cr(CO)_5$.

Base	Site	$E_R(Cr)^a$ (kcal/mol)
9-MeG	N1[b]	58
	N3	(76)[c]
	N7	36
9-MeH	N1[b]	42
	N3	(67)[c]
	N7	37
9-MeA	N1	57
	N3	(71)[c]
	N7	42
1-MeC	N3	56
1-MeT	N3[b]	52
1-MeU	N3[b]	50

[a]Estimated uncertainty ± 1 kcal/mol. [b]Deprotonated. [c]These values have higher uncertainty due to structural distortions in these adducts (see text).

An examination of the I_S values presented in Fig. 4.11 allows the following observations to be made. With respect to the N7 of purines (which are both intuitively the least sterically hindered as well as being the most accessible on a DNA duplex [22]) the I_S values are comparable for both guanine and hypoxanthine, but approximately 17% higher for adenine. This is in agreement with experimental evidence [9,10,29] which suggests a greater steric influence on metal coordination at N7 of an exocyclic amino compared to an exocyclic oxo substituent.
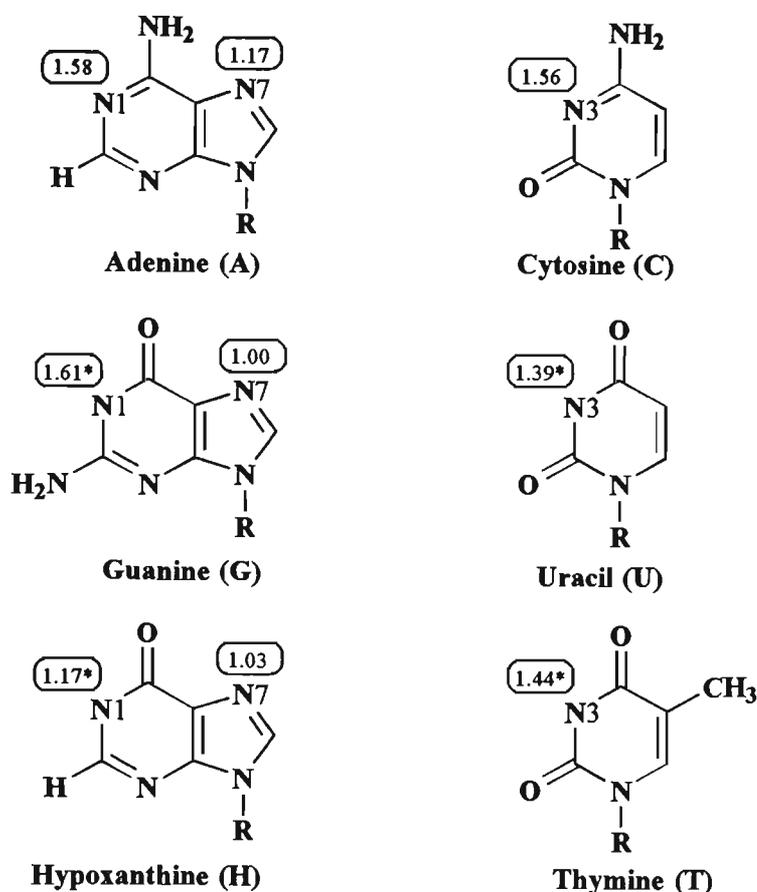
Figure 4.11. I_S values for potential metal binding sites on the common nucleobases; an asterisk denotes the value for a deprotonated site. For this study $R = CH_3$.

The sterically equivalent sites, N1 of guanine and N3 of cytosine, have comparable I_S values which are approximately 60% higher than the value for N7 of guanine. If attention is focused on the N1 position of adenine, the I_S value for this site is comparable to the values for the N1 site of guanine and the N3 site of cytosine, in spite of the absence of a concomitant exocyclic oxygen (replaced by hydrogen in the case of adenine). This is further evidence for the larger steric influence of the amino substituent and is consistent with the oxo substituent having only a relatively modest steric effect [10]. The same conclusion with respect to exocyclic oxo may be reached by comparing the I_S values of the N1 position of hypoxanthine with the N1 of adenine. For the $Cr(CO)_5$ moiety,

the steric influence of one oxo substituent ortho to a binding site (e.g. N1 of hypoxanthine) is equivalent to that of an amino substituent one atom removed from a binding site (e.g. N7 of adenine). The N3 of uridine and the N3 of thymine have intermediate I_S values, as might be expected.

In the purine systems, when N9 carries a substituent, the N3 position is rarely accessed by metal species due to the severe steric constraints [22]. Attempting to probe the N3 position by the method presented here results in an anomalous outcome reflected in inflated values of $E_R(Cr)$ and structural distortions; more specifically, the planarity of the nucleobase moiety is compromised and its orientation with respect to the $Cr(CO)_4$ (radial) plane is no longer close to 90°. Thus, it is possible that this method could be exploited to identify structural requirements for which coordination is precluded by steric factors. This could be useful in the design of metal complexes as site-specific reagents [43].

The steric parameters presented here for metal binding sites on nucleobases have sensible relative values and are in accord with deductions from reported experimental data where steric effects are considered to be operative. Thus they demonstrate the feasibility of quantifying relative steric effects in such systems and buttress the suggestion by Brown et al. [20] that the RE concept can be extended to ligands of nearly any kind. These workers also suggest that a variety of metal centers may also be considered. In this context, the metal species could equally well be varied, i.e. a series of modified platinum complexes, with a particular nucleobase binding site held constant. One would expect the steric parameters derived from such investigations to find particular application in quantitative structure activity/property relationship investigations (QSAR [44]/QSPR [45]) since, unlike the frequently employed molecular volume as a steric parameter [41], the E_R values represent steric effects at

the interface of the interaction and would be expected to carry little "transport" information relating to the hydrophobicity [46] or water solubility [7] of the metal complex.

4.3.2.2 Monofunctional platinum complexes as steric probes

4.3.2.2.1 Choice of platinum complexes as steric probes

The method described in the previous section, where a spherically-symmetrical $Cr(CO)_5$ probe was employed, may be cautiously extended to the interaction of more planar, non spherically-symmetrical moieties such as $Pt(A)_3$ (A ≡ am(m)ine) with nucleobase binding sites. For the present studies three carrier ligands were selected; namely, A ≡ NH_3 and $(A)_3$ ≡ dien and pmdien, where dien ≡ diethylenetriamine, pmdien ≡ 1,1,4,7,7-pentamethyldien. Thus, these platinum species now become steric probes.

H3N
|
H3N—Pt◄—N
|
H3N

Pt(NH3)3

NH2
|
HN—Pt◄—N
|
NH2

Pt(dien)

N(CH3)2
|
H3C—N—Pt◄—N
|
N(CH3)2

Pt(pmdien)

N represent
a target site
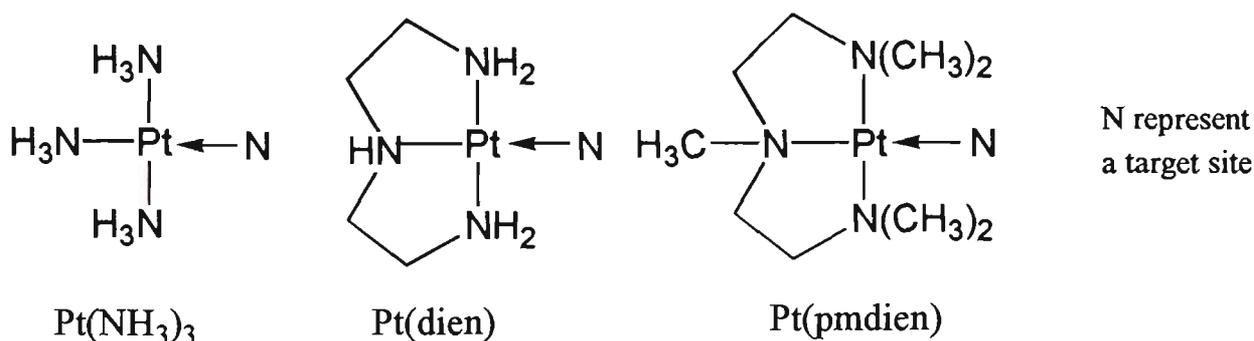
$Pt(NH_3)_3$ represents the least sterically restrictive probe in the above series. This metal carrier ligand system may be considered to model the monofunctional cisplatin binding to nucleobases [47].

Pt(dien) represents the intermediate steric probe in this series with respect to steric demands imposed on the nucleobase by the carrier ligand, primarily due to the ethylene (*en*) bridges. Pt(dien) – nucleoconstituent interactions have been widely studied and a broad range of thermodynamic and kinetic data is available, particularly associated with preferential binding to different endocyclic nitrogen sites [25]. Therefore, it may be possible to delineate the steric contribution to the binding for this system through QSAR/QSPR model building and a comparison of QSAR and QSPR results. Because of possible rotation about the Pt-N bond certain physical properties may not significantly reflect a steric aspect. However, this is not necessarily true of biological outcomes since steric aspects of the Pt-N rotation itself may be important, especially with respect to the consolidation of a particular structural motif on the DNA. Such consolidation may be achieved by non-covalent interactions (e.g. H-bonding, stacking, intercalation) in the case of monofunctional metal species and/or additional covalent linkages in the case of bifunctional species.

Pt(pmdien) is the most sterically restrictive probe in the above series. Its steric bulk around the metal center would be expected to force the nucleobases towards a perpendicular orientation with respect to the coordination plane. In regard to the binding to nucleoconstituents, this metal – carrier ligand system is expected to be similar to the Pt(TMED) system (TMED $\equiv$ N,N,N',N'-tetramethylethylenediamine). Crystal structure analysis of Pt(TMED) complexes with 9-MeG and 9-EtG reveals the values of nucleobase/PtN$_4$ dihedral angles to be in the range of 90° ± 4° [48].

4.3.2.2.2 Comparison of chromium and platinum steric probes

The repulsive energy values $E_R(Pt)$ for the three platinum steric probes and

twelve nucleobase binding sites (as defined in the Methods section) are presented in Table 4.5. As expected, examination of $E_R(Pt)$ values does not present as straightforward an account of the relative steric demands of nucleobase binding sites as does that of $E_R(Cr)$.

Table 4.5. Repulsive Energies $E_R(Pt)$ presented by nucleobase binding sites to metal species.

Base	Site	Pt(NH$_3$)$_3$		Pt(dien)		Pt(pmdien)[a]	
		$E_R(Pt)$	$E_R(Pt)_{90}$	$E_R(Pt)$	$E_R(Pt)_{90}$	*syn*	*anti*
9-MeG	N1[b]	6.85	5.22	6.76	5.03	39.7	39.1
	N3	6.35	6.09	8.55	8.55	47.5	47.6
	N7	6.76	2.94	6.68	2.89	34.6	32.8
9-MeH	N1[b]	6.89	4.44	6.66	4.04	35.7	35.2
	N3	5.46	5.34	7.63	7.74	43.1	46.9
	N7	6.54	2.89	7.07	2.89	34.8	32.6
9-MeA	N1	6.22	4.49	6.80	4.14	35.5	35.3
	N3	5.46	5.24	6.63	6.47	43.1	44.4
	N7	4.78	3.05	4.53	2.84	33.9	32.7
1-MeC	N3[b]	7.01	4.99	6.19	4.72	38.4	38.8
1-MeT	N3[b]	5.94	4.81	6.28	4.59	38.7	38.8
1-MeU	N3[b]	5.93	4.66	7.68	4.55	38.8	38.6

[a]Syn-anti classification is explained in section 4.3.2.2.5. [b]Binding sites requiring deprotonation prior to metal binding.

Correlations between $E_R(Cr)$ and $E_R(Pt)$ for three platinum moieties are shown in Fig. 4.12. These plots demonstrate that an increase of the steric bulk on the carrier ligand leads to more similarity between the chromium and platinum steric probes. With regard to the plots in Fig. 4.12 it should be kept in mind

that the $E_R(Cr)$ values associated with the N3 sites in purines are somewhat less reliable than the other values due to structural distortions of the corresponding adducts [49].
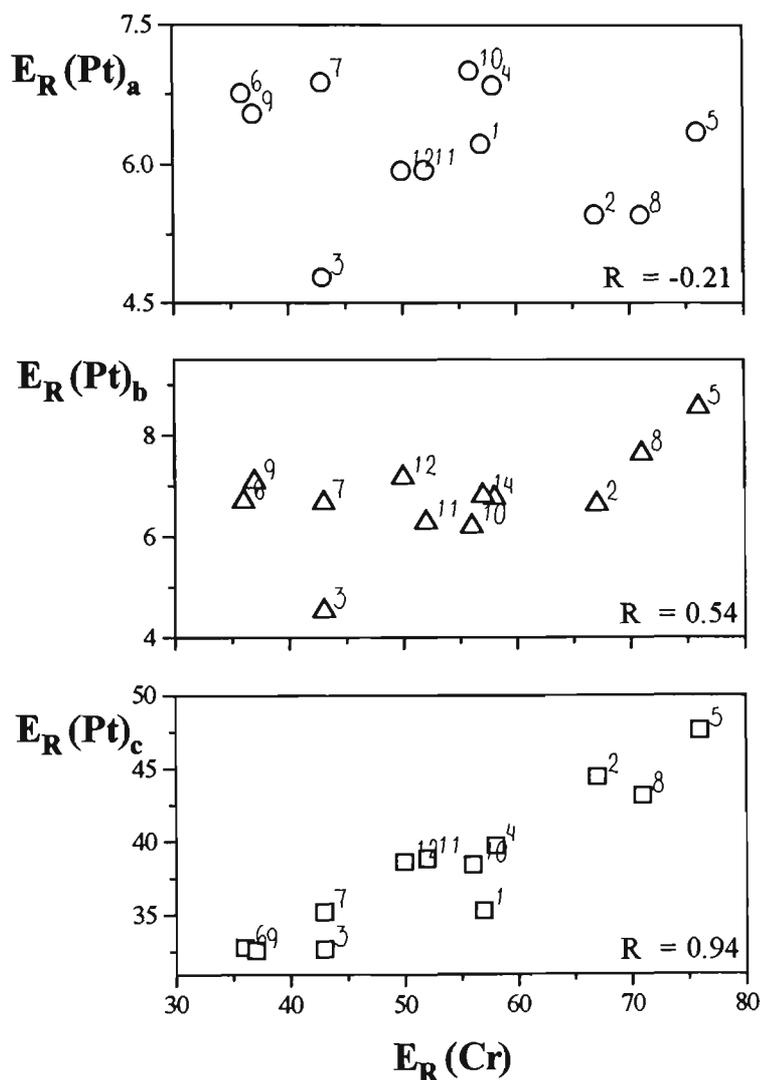
Figure. 4.12. Correlation between $E_R(Cr)$ and $E_R(Pt)$ for three platinum moieties: a - $Pt(NH_3)_3$, b - $Pt(dien)$, c - $Pt(pmdien)$. Binding site codes:
1 - N1(A), 2 - N3(A), 3 - N7(A), 4 - N1(G), 5 - N3(G), 6 - N7(G),
7 - N1(H), 8 - N3(H), 9 - N7(H), 10 - N3(C), 11 - N3(T), 12 - N3(U).

It must be appreciated that the more planar platinum species represent a different kind of probe than the chromium complex. In particular, in platinum probes there is a degree of conformational flexibility about the metal-nitrogen bond, related to the less encompassing frontier van der Waals surface of the probe and to the more variable orientation of the probe with respect to the target.

Fig. 4.13 shows the end-on views of the $Cr(CO)_5$ and $Pt(A)_3$ steric probes and reveals the "footprints" of the frontier van der Waals surfaces presented to nucleobase binding sites. In Fig. 4.14 the footprints of the frontier van der Waals surfaces of the $Cr(CO)_5$ and $Pt(NH_3)_3$ probes are overlayed onto typical nucleobase binding sites. Figs. 4.13 and 4.14 allow a better appreciation of how the planar nature of platinum probes, compared to the more encompassing nature of the chromium probe, may lead to less obvious outcomes for the relative steric demands of different binding sites.

$Pt(NH_3)_3$ $Pt(dien)$ $Pt(pmdien)$ $Cr(CO)_5$
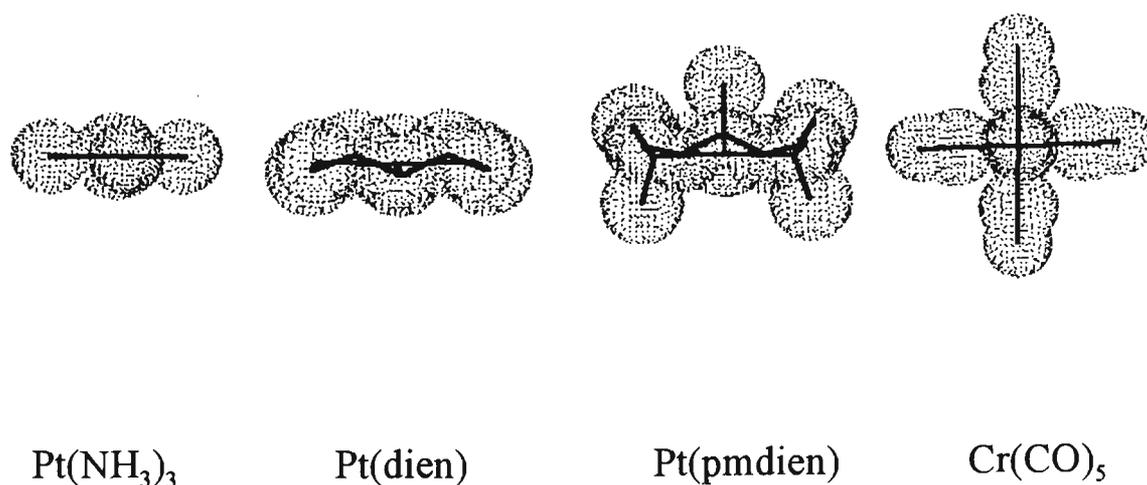
Figure 4.13. End-on views on $Pt(A)_3$ and $Cr(CO)_5$ steric probes showing vdW surfaces. Hydrogen atoms omitted for clarity.
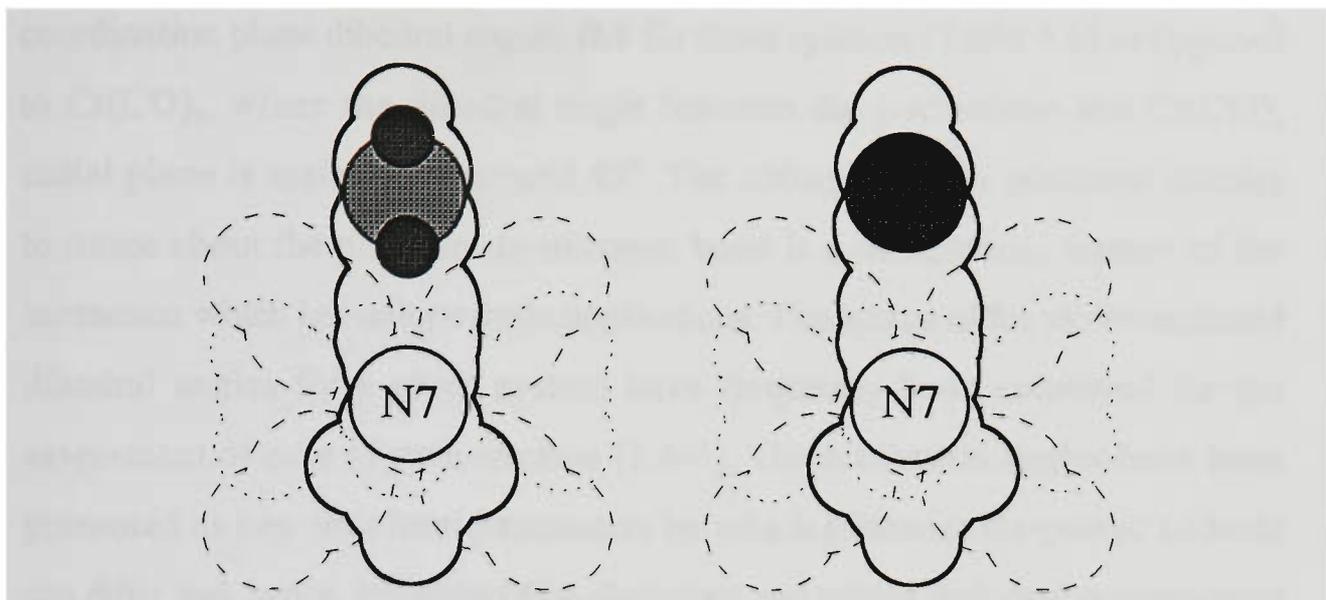
Figure 4.14. *An overlay of the frontier vdW surfaces of the steric probes Pt(NH₃)₃ (dotted line) and Cr(CO)₅ (dashed line) onto typical nucleobase binding sites (solid line), in this case – the N7 position of adenine (left) and guanine (right). Nitrogen, light gray; oxygen, dark gray (large); hydrogen, dark gray (small).*

Fig. 4.15 illustrates the approach of different steric probes to the N7 position of 9-methylguanine. It could be envisaged that in the $Pt(NH_3)_3$ and Pt(dien) systems the rotation of nucleobase with respect to the coordination plane is expected to be much less restricted than that in $Cr(CO)_5$ system.

$Pt(NH_3)_3$ Pt(dien)   Pt(pmdien) $Cr(CO)_5$

● carbon   ◉ nitrogen   ● oxygen   ● platinum   ○ chromium
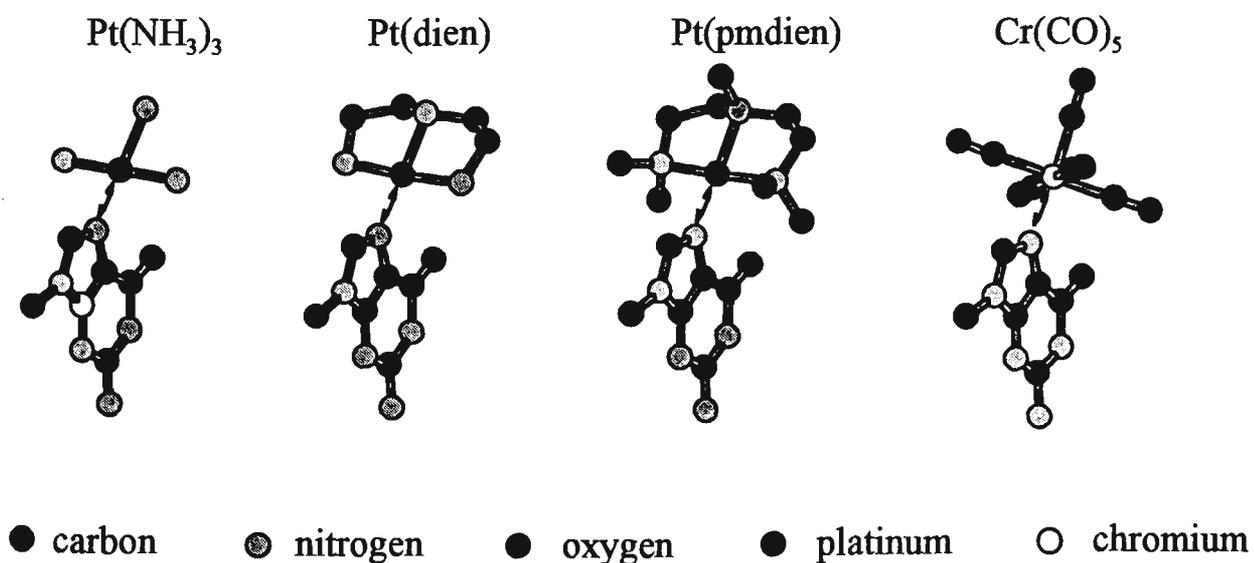
Figure 4.15. *The approach of steric probes to a typical binding site on a nucleobase, the N7 position on guanine.*

This relative conformational freedom allows a diversity of nucleobase/PtN$_4$ coordination plane dihedral angles *DA* for these systems (Table 4.6) as opposed to Cr(CO)$_5$, where the dihedral angle between the nucleobase and Cr(CO)$_4$ radial plane is maintained around 45°. The ability of many platinum species to rotate about the platinum-to-nitrogen bond is a mechanistic feature of the interaction which has unique steric implications. The spread of the aforementioned dihedral angles for a given system have frequently been employed for the assessment of ease of such rotation [1,4-6]. These dihedral angles have been presented as key structural parameters by which platinum bis-purine adducts can differ and, hence, influence DNA distortions and related biological consequences [50]. Dihedral angles have also been applied as a QSAR structural parameter for systems other than metal–nucleoconstituent, e.g. cyclic lactones [51].

Table 4.6. Nucleobase/PtN4 dihedral angles.

Base	Site	Pt(NH$_3$)$_3$	Pt(dien)
9-MeG	N1*	73	75
	N3	87	89
	N7	71	66
9-MeH	N1*	66	68
	N3	91	89
	N7	61	60
9-MeA	N1	70	68
	N3	92	94
	N7	65	65
1-MeC	N3*	71	72
1-MeT	N3*	102	74
1-MeU	N3*	104	68

Units - degrees. Asterisk denotes binding sites requiring deprotonation prior to metal binding.

4.3.2.2.3 Relationships between steric and conformational aspects in $Pt(NH_3)_3$- and Pt(dien)-nucleobase adducts

To explore the relationships between steric parameters and complex conformational flexibility in $Pt(NH_3)_3$ and Pt(dien) systems, values of repulsive energy $E_R(Pt)_{90}$ (Table 4.5) were calculated for a perpendicular approach of nucleobases with respect to the platinum coordination plane (Fig. 4.16). Relationships between repulsive energy values and the deviation of nucleobase orientation from 90° in optimised structures (ΔDA) are presented in Figs. 4.17 and 4.18. It may be seen that while the plots of $E_R(Pt)$ vs the deviation present scatters (Figs. 4.17 and 4.18, parts (a)), the data in the corresponding plots for the perpendicular approach ($E_R(Pt)_{90}$) display stronger relationships between the steric and conformational parameters (Figs. 4.17 and 4.18, parts (b)). It is worth noting that while latter more linear relationships could be fortuitous, the clustering of data points in these plots is not surprising. That is, the data points are clustered together on the basis of binding site type; namely: (i) N7 of purines, (ii) N1 of purines and N3 of pyrimidines, and (iii) N3 of purines.
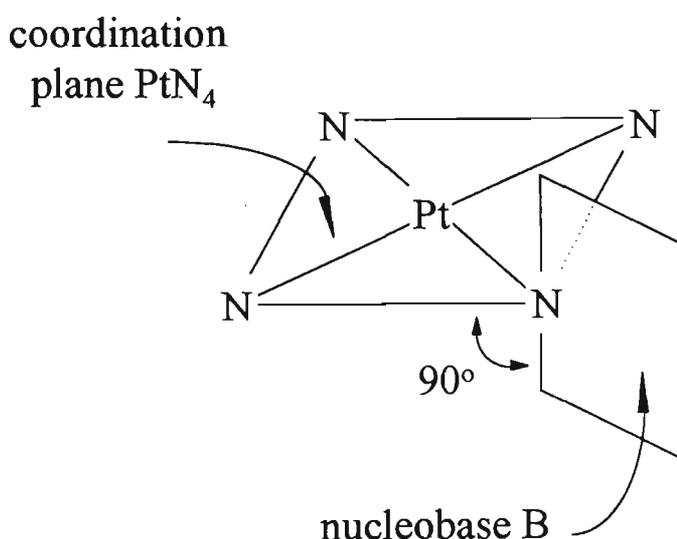
Figure 4.16. A schematic representation of the perpendicular orientation of a nucleobase with respect to the platinum coordination plane.
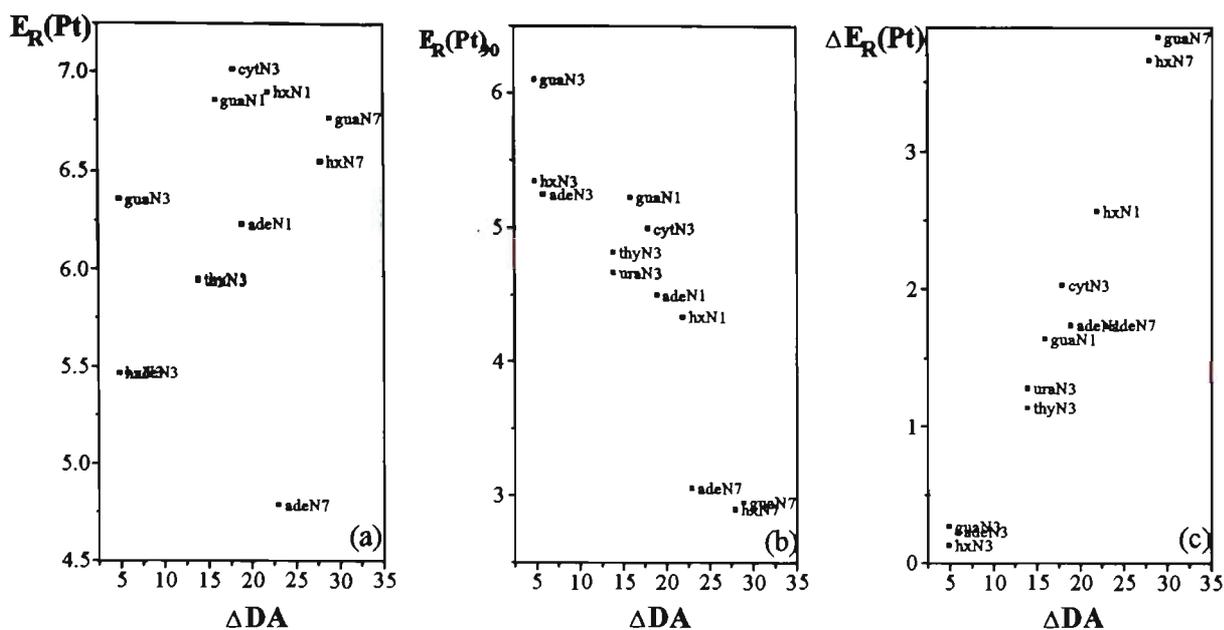
Figure 4.17. *Relationships between repulsive energies and dihedral angles for Pt(NH₃)₃(nucleobase) systems. (a) Optimised molecular systems. (b) Molecular systems opimised with the restriction of nucleobases to the perpendicular orientation with respect to platinum coordination plane.*
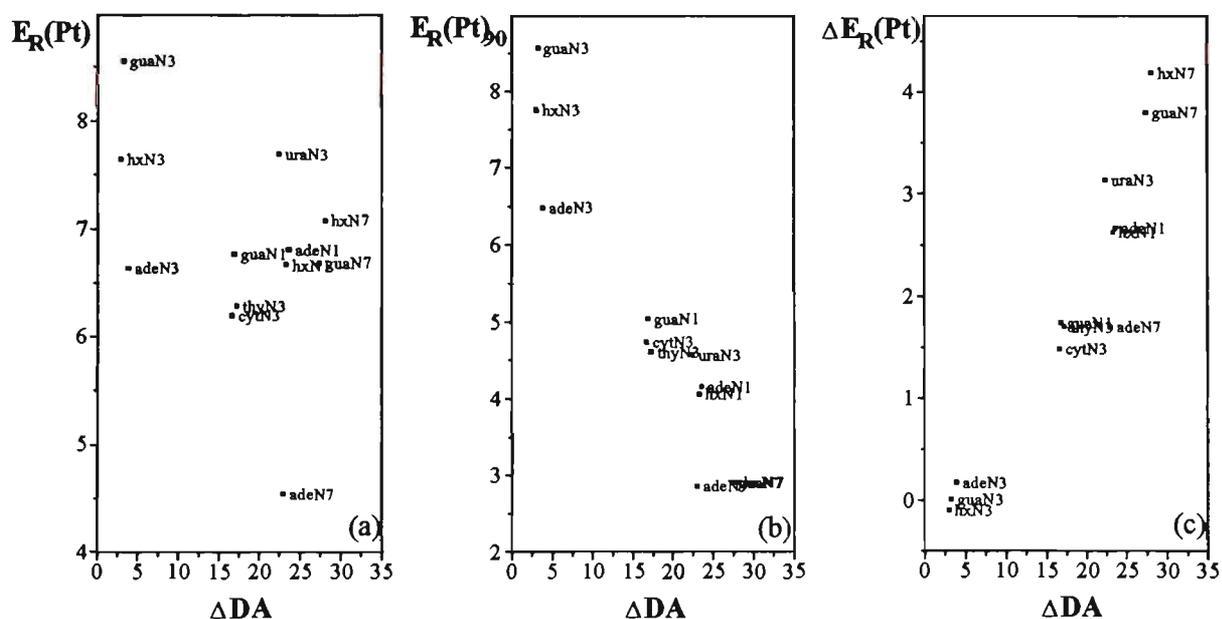
Figure 4.18. *Relationships between repulsive energies and dihedral angles for Pt(dien)(nucleobase) systems. (a) Optimised molecular systems. (b) Molecular systems opimised with the restriction of nucleobases to the perpendicular orientation with respect to platinum coordination plane..*

Although these data do not allow sharp differentiation between analogous binding sites with different neighbouring exocyclic groups, e.g. the N1 position in purines, the following examples warrant mentioning. The slightly higher value of $E_R(Pt)_{90}$ for the N7 position of adenine as compared to guanine and hypoxanthine may be due to the presence of the sterically more demanding NH_2 group [49]. The same argument applies to the comparison of $E_R(Pt)_{90}$ values for the N3 position of these nucleobases.

For $Pt(NH_3)_3$(nucleobase) adducts it can be seen that the steric bulk of the carrier ligand is insignificant and can not be used to distinguish between binding sites of the same type but with different adjacent exocyclic substituents, e.g. N7 position of purines. For Pt(dien)(nucleobase) adducts, incorporation of the ethylene-bridges could be expected to make $E_R(Pt)$ values more sensitive to such differences. However, the observation that analogous outcomes are found for the $Pt(NH_3)_3$ and Pt(dien) systems (Figs. 4.17 and 4.18) shows that the steric requirements of the ethylene-bridges incorporated in the **dien** carrier ligand, are not sufficient to impose a perpendicular orientation of a nucleobase with respect to the coordination plane, leaving the Pt(dien) moiety essentially flat. This theoretical observation is in line with the experimental finding [52] that the steric bulk of *en* in bis-platinum complexes is not sufficient enough to force DNA constituents into targeted head-to-head orientation.

One surprising result associated with repulsive energies for these systems is the reverse order, compared to $Cr(CO)_5$ steric probe, of the $E_R(Pt)$ values for the N7 position of purines; i.e., the value of $E_R(Pt)$ for adenine is lower than those in guanine and hypoxanthine (Figs. 4.17 and 4.18, parts (a)). Since in the Pt(dien) system the steric bulk of the carrier ligand is not evenly distributed on both sides of the coordination plane, in order to explain this observation

it was necessary, to investigate the dependence of repulsive energy on the orientation of nucleobase with respect to the coordination plane. The $E_R(Pt)$ values have been calculated for 9-MeA and 9-MeG nucleobases bound to Pt(dien) at N7 site and rotated around the N7-Pt bond in 10° steps. The resulting values, $E_R(Pt)_{rot}$, have been plotted against the nucleobase/coordination plane dihedral angle (DA) in the rotated structures (Fig. 4.19a).
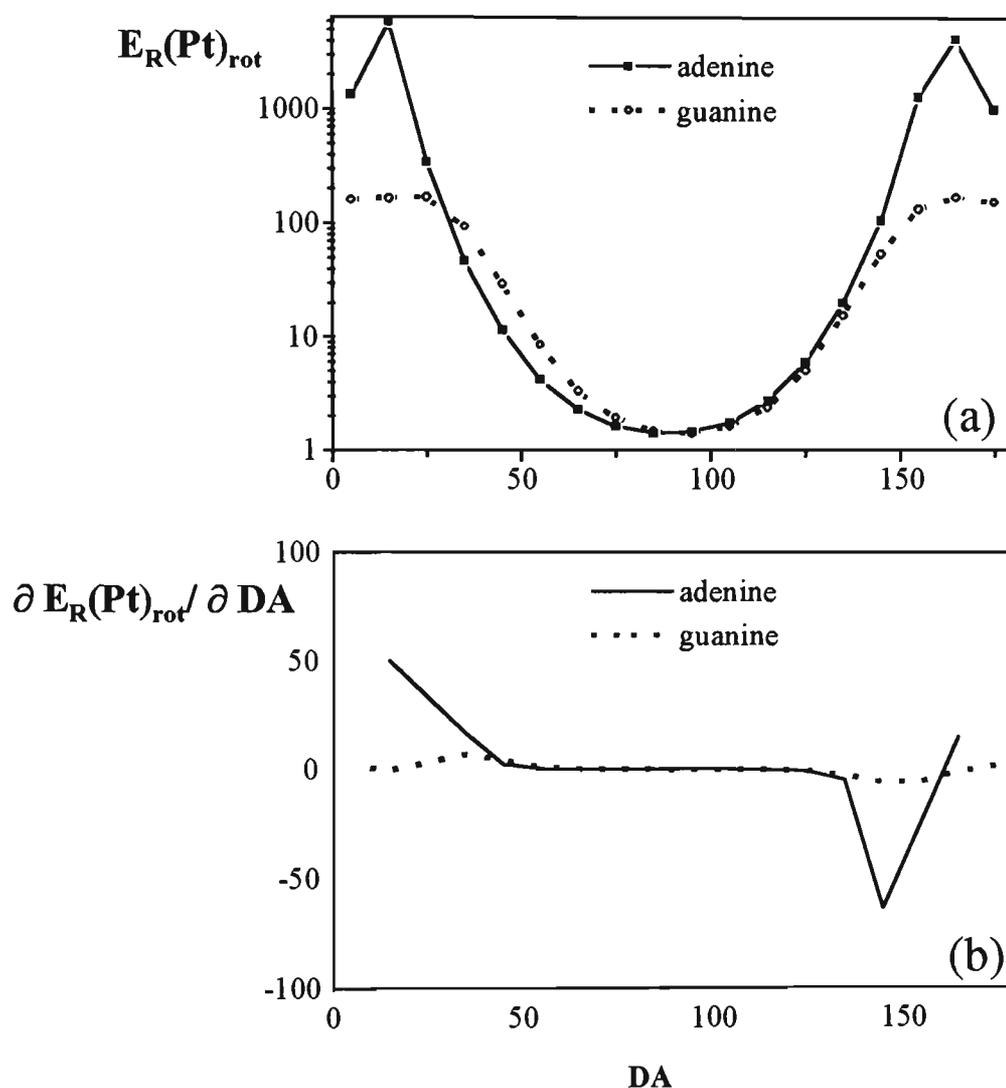
Figure 4.19. Rotational profiles of the Pt(dien)(nucleobase) adducts with N7-bound 9-methylguanine and 9-methyladenine.

The results demonstrate that the nature of exocyclic substituents, neighbouring a binding site (in this case – amino in adenine and oxo in guanine, both one

atom removed from the binding site), affects the steepness of the $E_R(Pt)_{rot}$ vs DA curve. More specifically, slightly higher values of $E_R(Pt)_{rot}$ are achieved for N7(G) as compared to N7(A) in the region around the minimum energy conformation for both adducts, $DA = \pm 40°$ to $90°$, where the steric repulsion between the carrier ligand and a nucleobase is also minimal. Much higher values of $E_R(Pt)_{rot}$ are achieved for N7(A) as compared to N7(G) in the region around $DA \sim 0°$ and $\pm180°$. The higher adenine values are caused by significantly stronger repulsion between the carrier ligand and the amino group in the adenine base, as compared to the oxo group in the guanine base. Estimation of the steepness of the curve in different DA regions may be inferred from the differential plot of $\partial E_R(Pt)_{rot}/\partial DA$ (Fig. 4.19b).

To further investigate the relationship between steric interactions and conformational flexibility in $Pt(NH_3)_3$(nucleobase) and Pt(dien)(nucleobase) systems the difference between repulsive energies for optimised and $90°$-restrained systems ($\Delta E_R(Pt)$) was also plotted against ΔDA (Figs. 4.17 and 4.18, parts (c)). The data in these plots presents even stronger (than parts (b)) linear relationships with a tighter clustering of data points according, not only to endocyclic binding sites, but also to their exocyclic environment. The data points requiring special attention in plots of $\Delta E_R(Pt)$ vs ΔDA (Figs. 4.17 and 4.18, parts (c)) are those which correspond to the N7 position of adenine. This binding site, with an amino substituent one atom removed from it, exhibits steric demands similar to that of the N1 positions of adenine and guanine and the N3 position of cytosine (binding sites with an amino substituent ortho to them). Notwithstanding the differences in the exocyclic environment, such correspondence of steric demands may be due to the fact that the orientation of nucleobases in these systems is similar, i.e. $\Delta DA = 16°$ to $19°$. On the contrary, the N7-bound guanine and hypoxanthine, which are less sterically restricted by the oxo substituent

ortho to a binding site, may rotate further away from the perpendicular, i.e. $ADA \sim 30°$, which leads to a larger difference between the $E_R(Pt)$ values for optimised and 90°-restrained structures. These observed differences in dihedral angles, rationalised in steric terms, are in accord with the conclusion made in Refs. [32,53] that steric effects of bulky exocyclic substituents can explain the enlargement of a dihedral angle in cases where electronic differences between binding sites are negligible.

Thus, it may be concluded from the above analysis of data that $\Delta E_R(Pt)$ is proportional to the deviation of nucleobase orientation from 90° approach in optimised systems. Therefore, it is reasonable to hypothesise that the rotation of nucleobases away from the perpendicular approach operates as a compensatory mechanism to relieve the steric strain. Such rotation may be of a particular relevance for the proposed consolidation of the interaction of cis-platinum complexes with nucleic acids. Thus, in the case of intrastrand crosslink formation, the rotation of nucleobases around the Pt-N bond of a monodentate adduct is widely recognised as being of particular importance.

4.3.2.2.4 Nucleobase orientation factors

The question arises as to which particular factors are determinative of the nucleobase orientation in these model systems and to what extent. In order to address this question the energy components[3] have been calculated for $Pt(NH_3)_3$(nucleobase) systems with respect to the rotation of nucleobases around the Pt-N axis. Plots of the total energy and calculated energy components against the nucleobase/coordination plane dihedral angle (DA) are presented in Figs. 4.20 – 4.23. These allow the following observations to be made.

---

[3] The AMBER energy terms comprising the overall molecular energy have been used, see section 3.2.2 for the definition of terms.
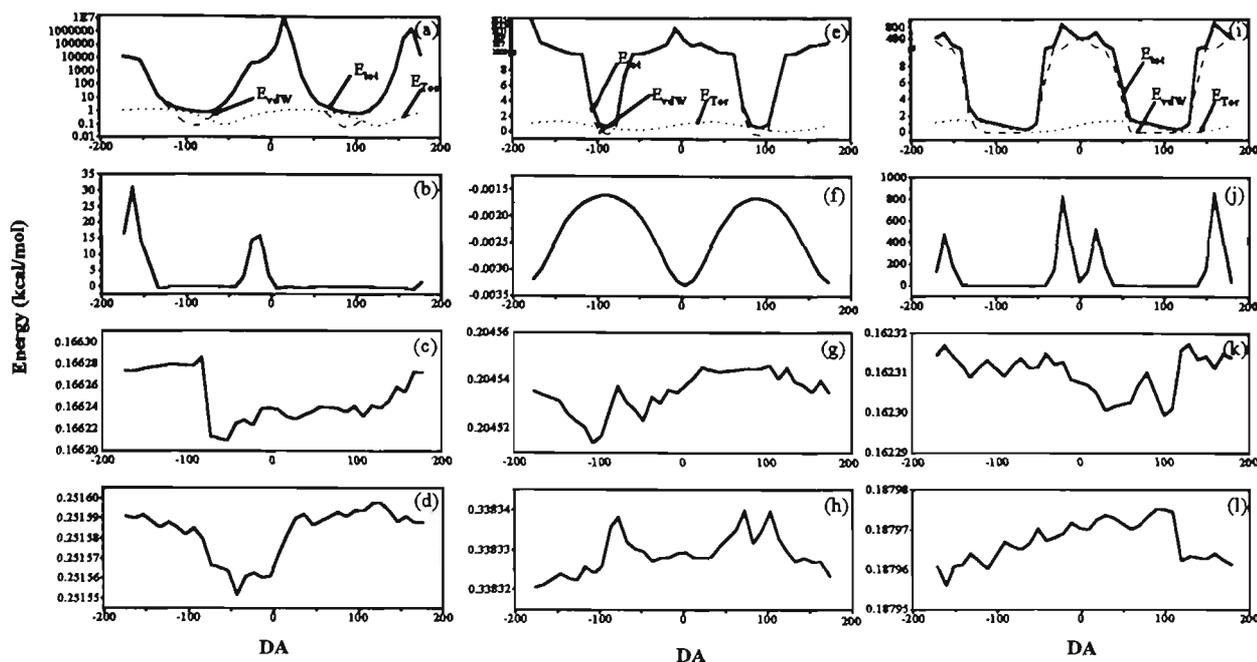
Figure 4.20. Rotational energy profiles for Pt(NH₃)₃(9-methylguanine) adducts.
a-d: N1-adduct; e-f: N3-adduct; i-l: N7-adduct
a,e,i: E_{total} (solid line); E_{vdW} (dashed line); E_{Tor} (dotted line);
b,f,j: E_{HB}; c,g,k: $E_{Stretch}$; d,h,l: E_{Bend}
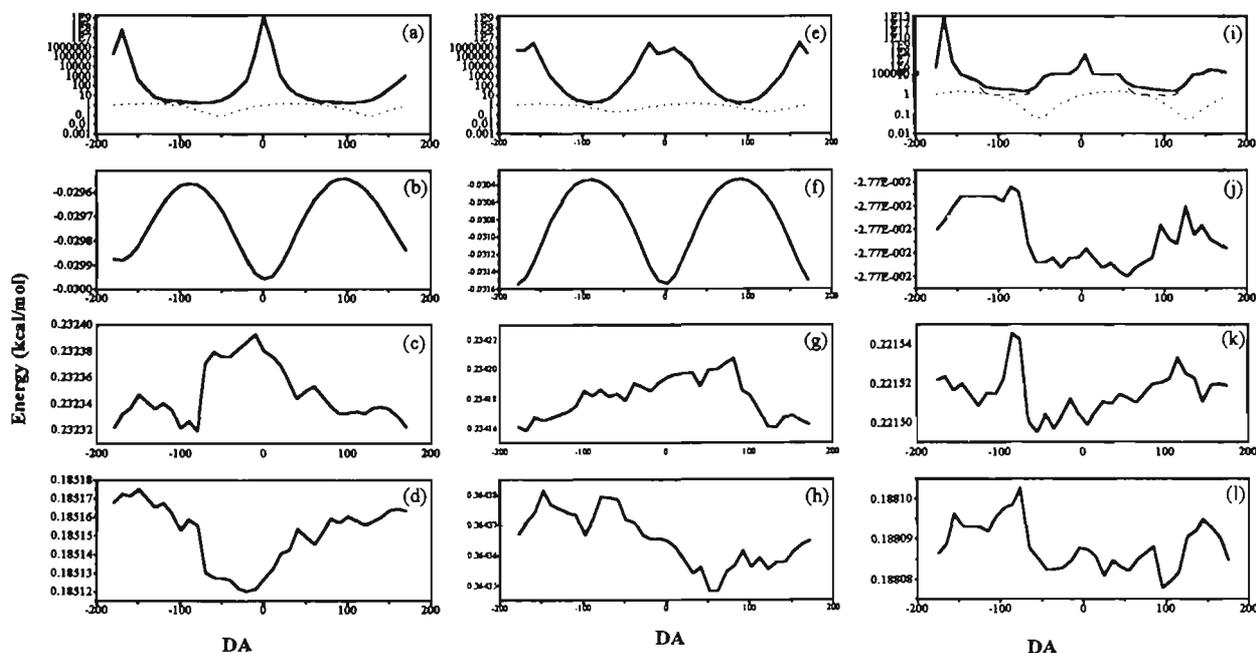
Figure 4.21. Rotational energy profiles for Pt(NH₃)₃(9-methyladenine) adducts.
a-d: N1-adduct; e-f: N3-adduct; i-l: N7-adduct
a,e,i: E_{total} (solid line); E_{vdW} (dashed line); E_{Tor} (dotted line);
b,f,j: E_{HB}; c,g,k: $E_{Stretch}$; d,h,l: E_{Bend}
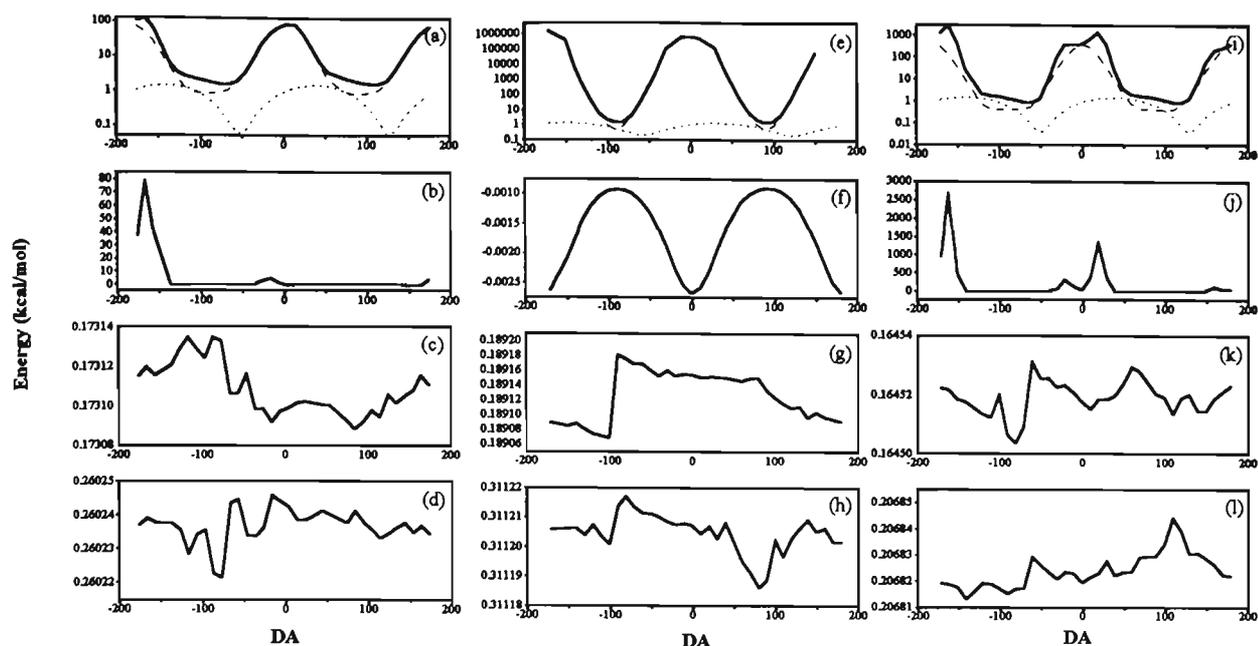
Figure 4.22. *Rotational energy profiles for Pt(NH₃)₃(9-methylhypoxanthine) adducts. a-d: N1-adduct; e-f: N3-adduct; i-l: N7-adduct*
a,e,i: E_{total} *(solid line);* E_{vdW} *(dashed line);* E_{Tor} *(dotted line);*
b,f,j: E_{HB}*; c,g,k:* $E_{Stretch}$*; d,h,l:* E_{Bend}
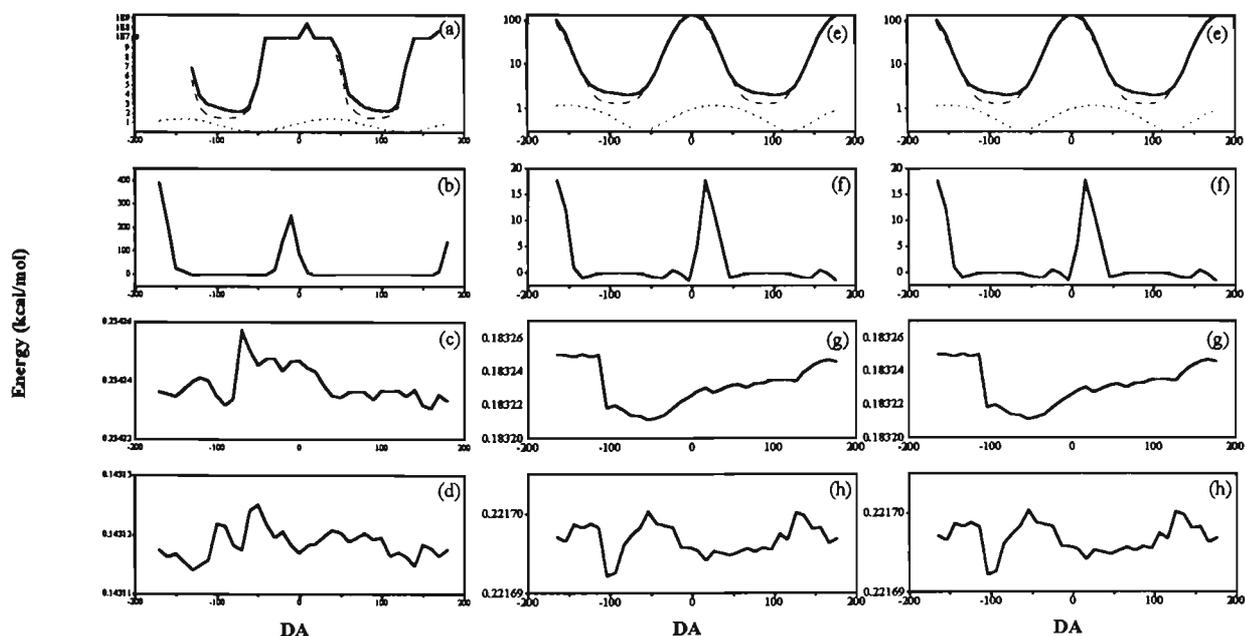
Figure 4.23. *Rotational energy profiles for Pt(NH₃)₃(1-methylpyrimidine-N3) adducts. a-d: 1-methylcytosine; e-f: 1-methyluridine; i-l: 1-methylthymine*
a,e,i: E_{total} *(solid line);* E_{vdW} *(dashed line);* E_{Tor} *(dotted line);*
b,f,j: E_{HB}*; c,g,k:* $E_{Stretch}$*; d,h,l:* E_{Bend}

It may be seen that in DA regions in the vicinity of $-180°$, $0°$ and $+180°$ the overall molecular energy E_{tot} is solely determined by strong vdW interactions E_{vdW} (Figs. 4.20 – 4.23, parts (a,e,i), dashed line). These arise, as expected, from non-bonded repulsion between the nucleobase exocyclic groups and the amine groups of the carrier ligand. It is obvious, therefore, why such an orientation of a nucleobase is prohibited.

The changes in the rotational energy E_{Tor} (Figs. 4.20 – 4.23, parts (a,e,i), dotted line) may be seen to be generally significantly smaller than those in E_{vdW}. However, in DA regions around $\pm$ ($60°$ to $90°$), which correspond to minimum energy conformations, the rotational profiles of E_{tot} are reflective of those of E_{Tor}. That is, the minima of E_{tot} correspond to or are shifted towards the minima of E_{Tor}. The N3-bound purines constitute the exception from the above observation (Figs. 4.20 – 4.23, parts (e – h)). Although the influence of E_{Tor} is still present, in these systems the non-bonded interaction energy also prevails in DA regions around $\pm$ ($60°$ to $90°$). Such a prevailing dependence of the total molecular energy on the non-bonded interactions may be explained by the severe coordination-obstructing substitution at the N9 position in purines. However, the difference between the inflated values of $E_R(Cr)$ (Table 4.4) and the "sensible" values of $E_R(Pt)$ associated with N3 position of purines (Table 4.5) suggests that the steric accessibility of this position for metal binding may depend on metal complex "flatness". The notion of metal complex flatness as an element of structural speciation is related to its symmetry and distribution of steric bulk in relation to a particular binding site. Such an attribute may be regarded as an important property of a metal complex since it is apparent that the binding of metal species to many sites on biomolecules may be controlled and even precluded by the degree of flatness (as defined here) of the metal complex.

With respect to the steric accessibility of the N3 position of purines the proposed attribute of metal complex flatness allows the following to be suggested. It may be that the potential for metal complex binding to this position, rather than being completely precluded by the N9-substitution [54], is actually controlled by the flatness of the attacking metal species. Indeed, in a number of studies such binding has been detected for some species [55,56]. Therefore it is feasible that binding selectivity based on this principle could be exploited in biological systems.

The energy landscape for hydrogen-bonding interaction with respect to nucleobase orientation (Figs. 4.20 – 4.23, parts (b,f,j)) shows that the change of E_{HB}, although consistent during the rotation, cannot be considered as a very significant contributing factor for the determination of minimum energy conformations. However, it is necessary to draw attention to certain aspects of hydrogen-bonding interactions in complexes with nucleobases containing oxo exocyclic substituents adjacent to the binding site. Namely, the change of E_{HB} in N3-bound pyrimidines and N1- and, particularly, N7-bound purines is appreciable in DA regions in the vicinity of $-180°$, $0°$ and $180°$, where the exocyclic oxo groups come into close vicinity to the amine groups of the carrier ligand. It is expected that the contribution from the hydrogen-bonding factor would increase for the carrier ligands containing hydrogen-donor atoms in the middle regions i.e. between the amine groups (e.g. on the carbon skeleton for the **dien** carrier ligand). However, for such carrier ligands the contribution from the non-bonded interaction could also increase in these conformationally relevant regions.

To conclude the analysis of factors which influence nucleobase orientation in $Pt(NH_3)_3(nucleobase)$ complexes, the bond-stretching and angle-bending energies, $E_{Stretch}$ and E_{Bend} respectively, have been plotted against DA (Figs.

4.20 – 4.23, parts (c,g,k) and (d,h,l) respectively). It may be seen from these rotational profiles that these factors are not significant contributors to the change of total energy during the rotation.

4.3.2.2.5 Steric and conformational aspects in Pt(pmdien)–nucleobase adducts

The steric bulk of the methyl substituents on the nitrogen atoms in the **pmdien** carrier ligand significantly lowers the conformational flexibility of Pt(pmdien)(nucleobase) adducts, compared to the $Pt(NH_3)_3$ and Pt(dien) analogs. Conformational sampling of the **pmdien** systems results in only two accessible conformations for the Pt(pmdien)(nucleobase) structures. These are called rotamers due to the restricted rotation of nucleobases around the Pt-N bond. Indeed, ^1H NMR studies show that, due to the steric bulk of the terminal methyl groups, the nucleobase rotation is significantly hindered and the rotamers peaks are well resolved [57]. Thus, perpendicular orientation of nucleobases with either of two arrangements of exocyclic functional groups with respect to the coordination plane results (Fig. 4.24). Since the steric bulk of the carrier ligand (*en*-bridges and methyl groups) is not symmetrically distributed on both sides of the coordination plane, these two conformations are not structurally equivalent. Therefore a stereochemical convention is necessary in order to systematically classify these structures. Three approaches to the development of such a convention are possible.

A convention has been proposed which is useful for the classification of rotamers on the basis of NMR signals [57]. In this *NMR-convention*, a *syn*-rotamer is defined as one for which the monitored nucleobase proton (adjacent to the binding site) is on the same side of the coordination plane as the 4-methyl group; an *anti*-rotamer is defined *vice versa* (Fig. 4.24).

NMR-convention:	syn	anti
steric convention:	A	B
absolute convention:	syn	anti
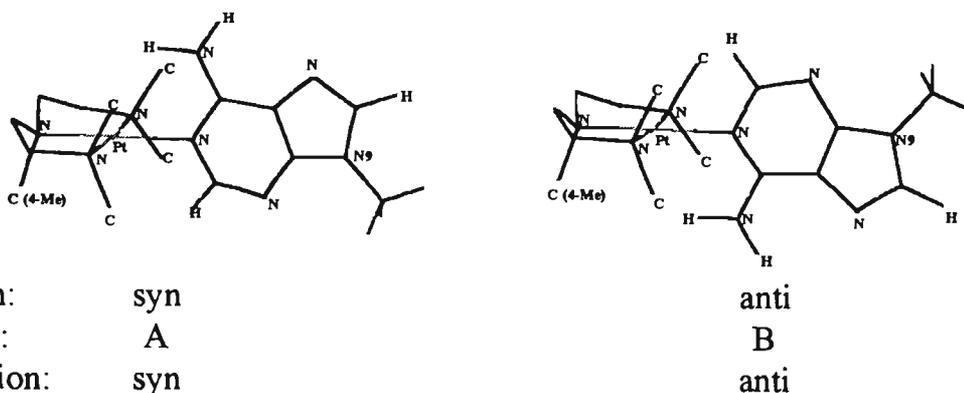
Figure 4.24. *Stereochemical conventions to distingtuish between rotamers for Pt(pmdien)(nucleoconstituent) adducts, e.g. Pt(pmdien)(9-methyladenine-N1).*

In the content of this project a *steric* convention may also be proposed based on a comparison of steric bulk of exocyclic functional groups (Fig. 4.24). In this convention a rotamer A may be defined as one for which a smaller exocyclic functional group is on the same side of the coordination plane as the 4-methyl group; a rotamer B may then be defined *vice versa*. Although the advantages of this convention are its simplicity and ease of visualisation, a classification of rotamers according to this convention depends on the assessment of steric effects of exocyclic functional groups. Therefore, this convention is not appropriate for studies where steric effects are being investigated, since a bias may be introduced.

An *absolute* convention may be advisable which is based on the orientation of a N9-substituent in purine nucleobases and N1-substituent in pyrimidine nucleobases with respect to the coordination plane (Fig. 4.24). In this convention a *syn*-rotamer may be defined as one for which a N9- or N1-substituent is on the same side of the coordination plane as the 4-methyl group; an *anti*-rotamer is defined *vice versa*. This convention is similar to the *NMR-convention*

[57], described above. However, the *absolute* convention allows to classify rotamers without regard for the direction of the *monitored* proton, which may vary for different nucleobases. The *absolute* convention may also be applied to modelling studies and extended to systems not yet investigated by NMR, including hypothetical structures. Finally, this convention allows *systematic* classification for adducts on different levels of nucleoconstituent complexity, from nucleobases to oligonucleotides.

With respect to the Pt(pmdien)(nucleobase) rotamers it has been found that, although not equivalent structurally, both conformations are very close energetically and have close values of $E_R(Pt)$. These are given in the Table 4.5, where they are classified on the basis of the proposed *absolute* stereochemical convention.

The analysis of $E_R(Pt)$ values, associated with the Pt(pmdien) probe, and their relationship with $E_R(Cr)$ values (Fig. 4.12c) allows the following observations to be made. With respect to the N7 position of purines, the $E_R(Pt)$ values are the lowest and comparable for all three nucleobases. Therefore, while the steric influence of this probe leads to a certain rigidity in the final adducts (making it similar to $Cr(CO)_5$ in this respect), its relative planarity still allows nucleobase binding sites with different neighbouring exocyclic substituents to produce steric effects of the same order of magnitude. The sterically equivalent sites, N1 of guanine and N3 of cytosine, yield close $E_R(Pt)$ values. These are also close to the values of N3 of uridine and thymine. The N1 position of adenine, displaying a $E_R(Cr)$ value comparable to those of the N1 position of guanine and the N3 position of cytosine, yields a lower $E_R(Pt)$ value, as compared to these sites, and similar to the $E_R(Pt)$ value of the N1 position of hypoxanthine. The N3 positions of purines, sterically crowded

by N9 substitution, yield the highest values of $E_R(Pt)$, mirroring the $Cr(CO)_5$ system. As expected, the N3 position of guanine gives rise to the largest $E_R(Pt)$ value due to the presence of the adjacent NH_2 group. However, there are no structural distortions which may indicate a preclusion of coordination.

A close examination of these similarities and differences suggests that the Pt(pmdien) steric probe, while not highly sensitive to the nature of exocyclic groups as, is the case for Cr, is more responsive to their number and arrangement. That is, on the basis of $E_R(Pt)$, the binding sites are clustered in the groups as shown in Table 4.7.

Table 4.7. Differentiation of binding sites by the Pt(pmdien) steric probe.

Number and arrangement of exocyclic groups	binding site(s)
one exocyclic group, two atoms removed from the site	N7 position of purines
one exocyclic group and a hydrogen atom, adjacent to the site	N1 position of hypoxanthine N1 position of adenine
two exocyclic groups, adjacent to the site	N1 position of guanine N3 position of pyrimidines
bulky methyl substituent, two atoms removed from the site and a hydrogen atom, adjacent to the site	N3 position of hypoxanthine N3 position of adenine
bulky methyl substituent, two atoms removed from the site and an exocyclic group, adjacent to the site	N3 position of guanine

4.3.2.2.6 Metal complex "flatness" as an element of structural speciation

Since it must be recognised that metal complex "flatness", as proposed above (section 4.3.2.2.4) is an element of structural speciation with important steric implications, it is desirable to devise a methodology for its quantification. This may be achieved by further exploiting the repulsive energy strategy. However, in this case, instead of using a metal complex to probe a series of different ligands as in Refs. [15-19] and in work [58][4], a series of metal complexes with varying degrees of flatness are individually used to probe a binding site on a substrate which is flanked by a variety of exocyclic substituents on both sides of a binding site. Although, nucleobases are adequate substrates for the assessment of flatness as shown previously in this chapter, one or more idealised substrates may be devised where the flanking substituents may be arbitrarily controlled. An additional requirement for a substrate is that it has a degree of planarity, typified by nucleobases. Such an idealised substrate is depicted in Fig. 4.25.

metal complex approach ⟶
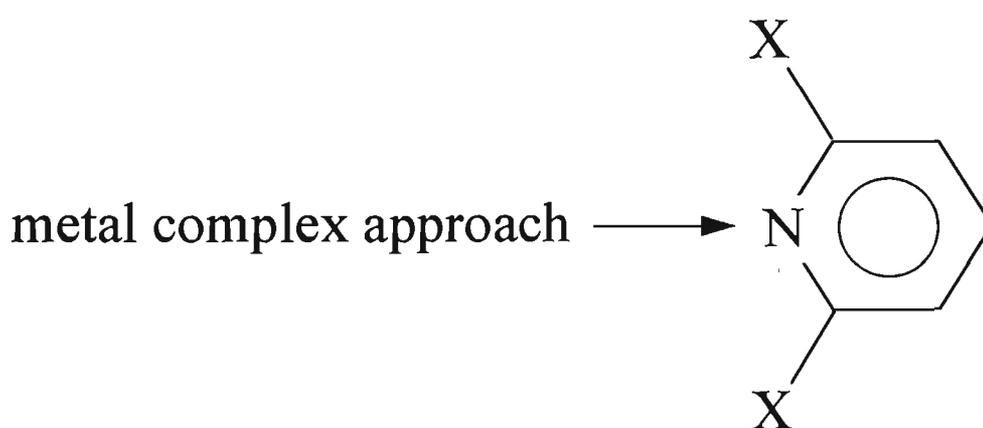
Figure 4.25. A schematic representation of flatness assessment. Substituent X may be O, NH_2, OH, alkyl, halogen, etc.

---

[4] A reprint of this work is bound in the thesis.

In order to effectively quantify the notion of flatness, the $Cr(CO)_5$ probe is used as a benchmark being defined as "non-flat". The reciprocal of the correlation coefficient R for the relationship between the E_R values for the complexes for which flatness is being assessed (with a series of substrates) and the $Cr(CO)_5$ complex may be taken as a flatness parameter. Furthermore, on the basis of such a parameter a scale may be introduced on which each complex i is assigned a relative flatness index:

$$F_i = \frac{R_{complex,\ least\ correlated\ with\ Cr(CO)_5}}{R_i}$$

This definition is predicated upon the following assumption. For a defined series of substrates, in a series of metal complexes, the complex with the E_R values least correlated with those of the $Cr(CO)_5$ probe is expected to demonstrate the highest degree of flatness for this series. For a specifically designed set of substrates these relative flatness indices may be expected to be transferable to other systems. As such, they should allow a comparison between structurally remote (unrelated) metal moieties with respect to their binding to a variety of binding sites.

4.4 REFERENCES

1. Orbell, J.D.; Marzilli, L.G.; Kistenmacher, T.J. *J. Am. Chem. Soc.* **1981**, *103*, 5126.

2. Saenger, W. *Principles of Nucleic Acids Structure;* Springer: New York, **1984**.

3. Sherman, S.E.; Lippard, S.J. *Chem. Rev.* **1987**, *87*, 1153.

4. Kozelka, J.; Fouchet, M.-H.; Chottard, J.-C. *Eur. J. Biochem.* **1992**, *205*, 895.

5. Schollhorn, H.; Randashl-Sieber, G.; Muller, G.; Thewalt, U.; Lippert, B. *J. Am. Chem. Soc.* **1985**, *107*, 5932.

6. Sherman, S.E.; Gibson, D.; Wang, A.H.-J.; Lippard, S.J. *J. Am. Chem. Soc.* **1988**, *110*, 7368.

7. Bloemink, M.J.; Engelking, H.; Karentzopoulos, S.; Krebs, B.; Reedijk, J. *Inorg. Chem.* **1996**, *35*, 619.

8. Kistenmacher, T.J.; Orbell, J.D.; Marzilli, L.G. In *Platinum, Gold and other Metal Chemotherapeutic Agents;* Lippard, S.J. (Ed.); ACS: Washington D.C. **1983**, 191-208.

9. Kinjo, Y.; Tribolet, R.; Corfu, N.A.; Sigel, H. *Inorg. Chem.* **1989**, *28*, 1480.

10. Kinjo, Y.; Ji, L.-N.; Corfu, N.A.; Sigel, H. *Inorg. Chem.* **1992**, *31*, 5588.

11. White, D.; Coville, N.J. *Adv. Organomet. Chem.* **1994**, *36*, 95.

12. Tolman, C.A. *J. Am. Chem. Soc.* **1970**, *92*, 2956.

13. Tolman, C.A. *Chem. Rev.* **1977**, *77*, 313.

14. White, D.; Taverner, B.C.; Leach, P.G.L.; Coville, N.J. *J. Comput. Chem.* **1993**, *14*, 1042.

15. Brown, T.L. *Inorg. Chem.* **1992**, *31*, 1286.

16. Choi, M.-G.; Brown, T.L. *Inorg. Chem.* **1993**, *32*, 1548.

17. Choi, M.-G.; Brown, T.L. *Inorg. Chem.* **1993**, *32*, 5603.

18. White, D.; Brown, T.L. *Inorg. Chem.* **1994**, *33*, 5591.

19. White, D.; Brown, T.L. *Inorg. Chem.* **1995**, *34*, 2718.

20. Caffery, M.L.; Brown, T.L. *Inorg. Chem.* **1991**, *30*, 3907.

21. Woo, T.K.; Ziegler, T. *Inorg. Chem.* **1994**, *33*, 1857.

22. Farrell, N. *Transition Metal Complexes as Drugs and Chemotherapeutic Agents;* Kluwer: Dordrecht, **1989**.

23. Tipson, R.S.; Townsend, L.B. *Nucleic Acid Chemistry;* Wiley: New York, **1991**.

24. Lippert, B. *Gazz. Chim. Ital.* **1988**, *118*, 153.

25. Martin, R.B. *Acc. Chem. Res.* **1985**, *18*, 32.

26. Pullman, A.; Pullman, B. In *Chemical Applications of Atomic and Molecular Electrostatic Potentials;* Politzer, P.; Truhlar, D.G. (Eds.); Plenum: New York, **1981**, 381-406.

27. Kistenmacher, T.J. In *Chemical Applications of Atomic and Molecular Electrostatic Potentials;* Politzer, P.; Truhlar, D.G. (Eds.); Plenum: New York, **1981**, 445-67.

28. Basch, H.; Krauss, M.; Stevens, W.J.; Cohen, D. *Inorg. Chem.* **1986**, *25*, 684.

29. Reily, M.D.; Marzilli, L.G. *J. Am. Chem. Soc.* **1986**, *108*, 6785.

30. Lavery, R.; Pullman, A.; Pullman, B. *Int. J. Quantum Chem.* **1981**, *20*, 49.

31. Braddock, P.D.; Connors, T.A.; Jones, M.; Khokhar, A.R.; Melzack, D.H.; Tobe, M.L. *Chem.-Biol. Interactions* **1975**, *11*, 145.

32. Orbell, J.D.; Solorzano, C.; Marzilli, L.G.; Kistenmacher, T.J. *Inorg. Chem.* **1982**, *21*, 3806.

33.   Lippert, B. *Progr. Inorg. Chem.* **1989**, *37*, 1.

34.   Silverman, R.B. *The Organic Chemistry of Drug Design and Drug Action;* Academic Press: San Diego, **1992**.

35.   HyperChem™ Version 4. Canada: Hypercube, Inc. **1994**.

36.   White, D. **1995**, (personal communication).

37.   Polowin, J. **1995**, (personal communication).

38.   ORIGIN, Technical Graphics and Data Analysis in Windows™ Version 2. MicroCal Software, Inc. **1992.**

39.   ChemPlus™ Version 1. Canada: Hypercube, Inc. **1993.**

40.   Abdul-Ahad, P.G.; Webb, G.A. *Int. J. Quantum Chem.* **1982**, *21*, 1105.

41.   Charton, M. In *Steric Effects in Drug Design;* Charton, M.; Motoc, I. (Eds.); Springer: Berlin, **1983**, 108-18.

42.   Baker, A.T.; Crass, J.; Kok, G.B.; Orbell, J.D.; Yuriev, E. *Inorg. Chim. Acta.* **1993**, *214*, 169.

43.   Dobi, A.L.; Matsumoto, K.; Agoston, D.V. *Nucl. Acids Res.* **1994**, *22*, 4846.

44.   Kubinyi, H. *QSAR: Hansch Analysis and Related Principles;* VCH: Weinheim, **1993**.

45.   Katritzky, A.R.; Lobanov, V.S.; Karelson, M. *Chem. Soc. Rev.* **1995**, *24*, 279.

46.   Ghose, A.K.; Pritchett, A.; Crippen, G.M. *J. Comput. Chem.* **1988**, *9*, 80.

47.   Hambley, T.W. *Inorg. Chem.* **1991**, *30*, 937.

48.   Orbell, J.D.; Taylor, M.R.; Birch, S.L.; Lawton, S.E.; Vilkins, L.M.; Keefe, L.J. *Inorg. Chim. Acta.* **1988**, *152*, 125.

49.   Cramer, R.E.; Dahlstrom, P.L. *J. Am. Chem. Soc.* **1979**, *101*, 3679.

50. Schroder, G.; Kozelka, J.; Sabat, M.; Fouchet, M.-H.; Beyerle-Pfnur, R.; Lippert, B. *Inorg. Chem.* **1996**, *35*, 1647.

51. Seidel, W.; Meyer, H.; Born, L.; Kazda, S.; Dompert, W. In *QSAR and Strategies in the Design of Bioactive Compounds;* Seydel, J.K. (Ed.); VCH: Weinheim, **1985**, 366-9.

52. Marcelis, A.T.M.; Korte, H.-J.; Krebs, B.; Reedijk, J. *Inorg. Chem.* **1982**, *21*, 4059.

53. Orbell, J.D.; Wilkowski, K.; Marzilli, L.G.; Kistenmacher, T.J. *Inorg. Chem.* **1982**, *21*, 3478.

54. pp. 201-19 in Ref. 2.

55. Massoud, S.S.; Sigel, H. *Eur. J. Biochem.* **1989**, *179*, 451.

56. Raudaschl-Sieber, G.; Schollhorn, H.; Thewalt, U.; Lippert, B. *J. Am. Chem. Soc.* **1985**, *107*, 3591.

57. Kim, S.-H.; Martin, R.B. *Inorg. Chim. Acta.* **1984**, *91*, 11.

58. Yuriev, E.; Orbell, J.D. *J. Comput.-Aided Mol. Design* **1996**, *10*, 589.

Chapter 5

Investigation of Quantitative Structure-Activity Relationships in platinum systems

5.1 INTRODUCTION

5.1.1 Preamble

The biological activity of a compound can be regarded as a function of its physicochemical and structural properties. A question that arises is which particular features of the molecule are responsible for its biological activity. The classic answer to this question lies in structure – activity analysis, the aim of which is to discover a structure-activity relationship (SAR), that is, to find if there is a pattern relating changes in activity to the changes in the molecular structure.

The development of a SAR ideally involves conceiving a hypothesis as to a mode of action, an important part of which is postulating a target molecule, e.g. protein or DNA, usually called *a receptor*. In addition, to explain the biological activity of a compound, 3D structural information is desirable with respect to the receptor site and the compound in question. Other desirable information includes the types of interactions between the receptor, the compound and the biological media. When such information is not available, which is often the case, a working model needs to be developed which allows correlation of biological activity with chemical structure with a view to predicting biological activity for novel compounds. **Quantitative structure-activity relationships** (QSARs) offer such models and give solutions in the fashion of "a black box" approach. These models are based on the assumption that biological activity is directly related to structural and physicochemical properties of compounds and the contributions of these properties are additive (to a certain degree). Thus, the overall paradigm of QSAR analysis can be expressed by presenting a biological activity as a linear combination of weighted

contributions:

$$A = \sum_i n_i a_i + b \qquad (5.1)$$

where A is a biological activity usually expressed as $\log(1/C)$, where C is a molar concentration required to produce some predetermined response. Sometimes logarithms of thermodynamic or kinetic constants are used as a monitored quantity in the above expression, producing quantitative structure-property relationships (QSPR). Symbols a_i and n_i are the contributing factor and its respective weight, b is a factor common for all members of a congeneric set.

The immediate objective of any QSAR study is to identify the structural and physicochemical contributions, which produce a statistically significant correlation with biological activity for a set of molecules, and to use the obtained relationship to calculate/predict the biological activity of new analogs. While the first part (correlation) appears to be straightforward, the second part (prediction) is more complex. Although sometimes the predictive power is taken as a criterion, prediction is not the primary goal of a QSAR study [1], rather a good description and understanding of the system. Therefore, QSAR should be considered not as an answering-all-questions magic wand, but as a cooperative tool in the rational design of biologically active compounds.

In QSAR studies models are derived using physicochemical properties (Hansch analysis), indicator variables (Free Wilson analysis), three-dimensional profiles (3D QSAR) or combinations of the above [1]. The descriptors utilised in QSAR analysis are usually divided into three principal groups. These correspond to the three principal kinds of interaction between the active molecule and the biological system which are considered to be important; namely: hydrophobic, electronic and steric interactions. The following sections overview

the structural descriptors traditionally used in QSAR studies. It has to be noted that only traditional steric parameters are described in this chapter; for the specific inorganic descriptors see Chapter 4.

5.1.2 Structural descriptors

5.1.2.1 Classification on the basis of descriptor generation method

Measured properties (e.g. molar refractivity, chemical shift, etc.) have the advantage of carrying clear physico-chemical meaning, but require a compound to already exist in order to make the necessary measurements. Apart from the necessity of carrying out laboratory experiments, which are usually time- and labour-intensive, the use of measured properties prevents using obtained relationships for the prediction of biological activity of hypothetical compounds.

Computed molecular properties originate from molecular mechanics (MM) and quantum mechanics (QM). MM methods permit estimation of routine descriptors such as molecular volume and molecular surface area, calculation of molecular energies and energy components, comparison between energies of different conformations and derivation of miscellaneous parameters based on the above. QM calculations are often employed to describe the electronic structure of a compound, e.g. point charges, polarizability.

Topological indices originate from mathematical theory, more specifically, from applied graph theory. Chemical structure is represented as a series of nodes (atoms) and edges (bonds). The individual arrays of atoms and bonds allow the calculation of connectivity patterns ("paths") and molecular connectivity indices, typically referred to as topological indices, including

three main types [2]: the chi indices ($^m\chi$, structural attributes), the kappa indices ($^m\kappa$, shape) and the topological equivalence states T (individual characteristics of atoms and groups of atoms in a molecule). The main advantages of the topological approach are its generality, ease of calculation and conformational independence. However, the topological indices are usually not informative enough, often lack any physical or chemical meaning and cannot be extrapolated to other compounds [3].

5.1.2.2 Electronic descriptors

Biological activity is affected by many electronic effects, occurring under a variety of conditions and through a variety of mechanisms and interactions. These include electrostatics, charge-transfer, dispersion, hydrogen bonding, polarisation, etc. Therefore it is not surprising that a very large number of electronic parameters has been developed for use in QSAR studies.

Linear Free Energy Relationships (LFER). The theoretical treatment of the *LFER* approach and its application to QSAR studies has been reviewed [4,5]. The *electronic substituent constant* σ (Hammet constant) is defined through its relationship with equilibrium (or rate) constants for the dissociation of benzoic acids:

$$\sigma_X = \log K_X/K_H \qquad (5.2)$$

where K_H is the dissociation constant of an unsubstituted parent compound and K_X is that of a X-substituted compound. The electronic substituent constant provides a measure of the electron-attracting power of a substituent relative to hydrogen. Hammet constants have been tabulated for a very large number of substituents [6].

Experimental quantities are successfully applied in QSAR as electronic parameters when they properly describe electronic molecular features, consequential for the biological activity of interest. They replace the electronic substituent constants with the advantage of characterisation of definite relationships, permitting clear physico-chemical interpretation. Furthermore, one experimental descriptor may replace a complex multiparameter relationship built on substituent constants. It should be noted that experimental quantities used as QSAR descriptors could be position dependent and involved in both linear and non-linear relationships. The experimental quantity most widely used as electronic descriptor is the ionisation constant pK_a (acidity/basicity) [7]. Other experimental quantities used as electronic descriptors include dipole moment μ and electronic polarizability α [8], chemical shift δ and IR frequency ν [9].

Quantum-mechanical parameters. Calculating and using these parameters in QSAR and drug design requires a description of the structural features of the molecules, their electronic distributions and the energies required to modify and perturb them during the interactions. Quantum-mechanical approaches to the electronic structure and activity of transition metal complexes have been reviewed [10].

The energies of frontier orbitals E_{LUMO} (lowest unoccupied molecular orbital) and E_{HOMO} (highest occupied molecular orbital) and their difference ΔE have been used as electronic descriptors in QSAR studies in the fields of medicinal chemistry and other areas [11-16]. The values of frontier orbital energies are commonly computed by semiempirical and less frequently – by *ab initio* calculations.

Molecular electrostatic potentials (MEP) are a powerful tool in the

rationalisation of chemical/biochemical reactivity in terms of locating the site of electrophilic attack. They also allow a quantification of electronic character of biologically active compounds in QSAR and rational drug design [10,17]. MEPs are also used for the derivation of atomic charges [18-20] as well as being one of the most important descriptors exploited in 3D-QSAR [21]. MEP is defined through the interaction between a molecule and a unit positive charge. Mathematically it is expressed as:

$$V(r) = \sum \frac{Z_A}{|R_A - r|} - \int \frac{\rho(r')}{|r' - r|} dr' \tag{5.3}$$

where the first term is a contribution from the nuclei (positive), the second term results from the continuous electronic charge density (negative). Thus, the sign of electrostatic potential value depends on which effects (nuclear or electronic) dominate at a particular position. In this expression the point r is the position of a positively charged probe, Z_A is the nuclear charge on atom A at position R_A and $\rho(r')$ is the electronic density.

Atomic point charges q are of importance in the application of QM to biological and pharmacological problems [22,23]. Furthermore, QM allows calculation of charge distribution not experimentally available.

5.1.2.3 Hydrophobicity descriptors

The hydrophobicity affects the biological activity of compounds through both hydrophobic interactions at the binding site and/or transport to the target site. While the hydrophobic interactions include drug–receptor or drug–enzyme interactions or purely chemical reactions stabilising transition states, transport involves transfer through aqueous and non-aqueous media in complex

biosystems and through physiological barriers such as cell membranes.

Hydrophobicity is modelled by a range of parameters which are related to each other [4]. The octanol-water partitioning system introduced by Hansch [24], is the most widely used model. This system simulates the blood/lipid partition. In this model the quantity π is defined as:

$$\pi_X = logP_X - logP_H \qquad (5.4)$$

where P_H is a partitioning coefficient of a non-substituted parent compound and P_X is a partitioning coefficient of a X-substituted compound in an octanol-water system. This parameter characterises the hydrophobicity of the substituent X relative to hydrogen and represents the hydrophobic analog of the Hammet constant σ. Not only substituents, but certain other structural subsets (e.g. chain branching, rings, multiple bonds) may be expressed in terms of π. Hansch et al. have determined and collected a vast amount of experimental π values [24]. Ideally, π possesses additivity so that for any compound $logP = logP_H + \Sigma\pi_i$. This provides the means for the calculation of $logP$ from π values of substituents and structural subsets and *vice versa*. A range of approaches has been developed for such calculations [1,25] and these have proven extremely useful in QSAR studies since they significantly reduce the need for the experimental determination of partition coefficients. While calculating $logP$ values using the additivity concept and substituent contributions, particular care should be taken since additivity may be lacking where special conditions occur; e.g. proximity of strongly polarising or strongly charged groups, intramolecular hydrogen bonding or intra- and intermolecular steric effects.

The relationship between biological activity and hydrophobicity can be

represented by a linear equation or a parabola, depending on the hydrophobicity range covered. If a relationship is parabolic then the *logP* value, optimal for the biological activity in question, may be predicted.

Other parameters used to describe hydrophobicity include: *log k'*, where *k'* is the HPLC retention index, *CSA* – the cavity size area, and *-log c_s*, where c_s is the solubility in water [1,4].

5.1.2.4 Steric descriptors

it may be argued that steric effects present the greatest challenge in the field of quantitative structure-activity relationships [1,4,6]. In order to correctly describe steric effects, 3D structures of both a compound and its binding site have to be known. However, even if they are known, which is less often than not, the steric effects also depend on conformation, binding mode, etc.

Application criteria for using steric parameters in QSAR studies have been developed [26,27]:

- A steric parameter should be a measure of substituent/ligand/molecule "size", not in terms of "volume" or "number of heavy atoms", but in terms of "radii". For a non- symmetric group the distance from the centre to the active site is of interest. It is obvious then that a steric parameter should be a <u>vector</u> rather than a <u>scalar</u> quantity in order to represent a steric *fit* as opposed to a steric *volume*.

- A steric parameter should be measurable or calculatable for all congeners in a series.

- A steric parameter should be as independent as possible from electronic

and transport parameters.

Classic steric parameters. The first parameter developed for intramolecular steric effects is the *steric substituent constant* E_s (Taft constant):

$$E_s = (\log k_X/k_0)_{XCOOR} \qquad (5.5)$$

where k_X and k_0 are the rates of the acid-catalysed hydrolysis of esters with substituents X and CH_3 respectively [28]. Values of E_s have been compiled [7,29]. Hancock suggested that E_s values represent, besides the steric features, a hyperconjugation effect and proposed the *modified steric substituent constant*, E_s^c, from which the hyperconjugation effect is eliminated [30].

Charton used *van der Waals radii*, R_v (by Pauling [31], Bondi [32], and Allinger [33]) as a measure of steric effects and proposed a new parameter – *the Charton constant*, $v_X = R_X - R_H = R_X - 1.20$, where R_X and R_H are the van der Waals radii of substituents X and hydrogen respectively [26].

The very widely used parameter for the characterisation of substituent size and intermolecular steric effects is the *molecular refractivity MR* [34]:

$$MR = \frac{n^2 - 1}{n^2 + 2} \times V \qquad (5.6)$$

where n is the refractive index and V is the molar volume. *MR* is proportional to polarizability, hence it carries some electronic information and does not have a purely steric nature. It is more characteristic of dispersive interactions.

The parameters described above are single values. But since substituents are three-dimensional they could be better described by a set of values

corresponding to different directions within a substituent. Verloop et al. introduced such set of STERIMOL parameters [35], denoting the length of the substituent in the direction it is attached to a parent molecule and the widths of substituent in four different directions.

"Bulk" parameters. Van der Waals molecular volume MV is the volume embodied by van der Waals surface area A of a molecule, which is determined as the surface of the intersection of spheres each of which is centred at the equilibrium position of the atomic nucleus, with vdW radius for each atom [36]. MV has been quite successfully employed in QSAR studies for a number of series of biologically active compounds [35], but there have been several controversies with respect to MV being a proper steric parameter [34,37,38]. Molecular weight, and MR are sometimes classified as bulk parameters [34].

Molecular shape parameters are based on the use of a reference structure to compare the molecular shapes [39]. Many modern approaches to the description of molecular shape and to the analysis of molecular similarity are reviewed in Ref. [40].

The Amoore's approach (the similarity index Δ) uses an average of absolute differences between correspondent radii in pairs of silhouettes (projections of CPK models on three orthogonal planes) [39]. This method has been applied to studies of ant alarm pheromone activity and of odour similarity and resulted in modest correlations.

In the Allinger's approach the A parameter is derived from the number of matches and mismatches of sub-cube occupation while comparing two cubes

within which compared molecule and reference structure are placed [39]. The A parameter has been used for studying an odour intensity of fatty acids and gave poor correlations.

In *the Hopfinger molecular shape analysis*, the shape reference indices are proposed for measuring the common volume of steric overlap V_O between the system being studied and an idealised molecule [41].

The Minimal Steric Difference (MSD) approach [42] permits the molecular shape dissimilarity to be expressed by a single number. From the set of molecules under study the most active one is selected as a reference structure (named "standard", S_0). Its shape is considered approximately complementary to the receptor cavity. Each molecule from the set under study is compared to this reference structure and the *MSD parameter*, defined as the number of heavy nonsuper(im)posable atoms of the two compared molecules, is calculated. Tested in QSAR, this parameter offered rather modest results [42].

In the Monte Carlo version of the MSD method, the *MCD parameter* [42] is obtained by superposition of not atoms, but vdW envelopes of two molecules in terms of their atomic cartesian coordinates and vdW radii. Application of MCD instead of MSD did not improve the above correlations.

The Minimal Topological Difference (MTD) approach is the derivation of the above [42]. The *MTD parameter* is calculated against a "hypermolecule" as a reference structure; hypermolecule H is obtained by atom per atom superposition of all molecules from the set starting with the most active one and seeking maximal superposition. Its vertices (atomic knots), j, correspond to the interior of the cavity ($\varepsilon_j = -1$), to its walls ($\varepsilon_j = 1$) or to sterically

irrelevant regions ($\varepsilon_j = 0$). The assignment of these ± 1 and 0 standard indicators is done subjectively [39]. Each molecule i is thus described by the vector $x_i = [x_{ij}]$. It is taken that $x_{ij} = 1$ if the vertex j of H is occupied by a non-hydrogen atom of the molecule i, otherwise $x_{ij} = 0$. Finally, MTD values are calculated over all the hypermolecule j vertices according to:

$$MTD_i = s + \sum \varepsilon_j x_{ij} \qquad (5.7)$$

where s is the number of cavity vertices. This parameter has found wide application in QSAR studies [43-45]. The defects of the MTD approach are discussed at length in Ref. [39].

5.1.2.5 Descriptor interrelationships

The reliable application of eqn 5.1 requires that different factors contributing to biological activity (parameters for steric, hydrophobic and electronic effects) are *separable*. While theoretically the concept of separability seems quite obvious, practically it is often difficult to ensure that the chosen parameters are indeed pure characteristics of the structural effects they are supposed to represent. It is often the case that these parameters are *interrelated*. The use of interrelated parameters often allows good correlation with the biological activity. However, the interpretative power of such relationships is often too vague, since it is unclear which particular effects contribute more strongly to the final activity. Considered below are important relations between molecular parameters:

1. Correlations between hydrophobic and electronic parameters are uncommon [4]. Nevertheless, it should be pointed out that the π value depends on electronic and steric interactions of a substituent with its

environment.

2. Correlations between π or *logP* and molar refractivity *MR* or molar volume *MV* are more frequent.

3. Correlations between steric and electronic parameters are relatively rare. However, one particular relationship to be borne in mind is the correlation between molar refractivity *MR* and polarizability α, manifesting the dualistic nature of the former.

In cases when different influences "overlap" in chosen parameters, a method to "separate" these influences is needed. This problem could be looked at from two different points of view, corresponding to two aspects of QSAR analysis: physico-chemical and mathematical. From the physico-chemical point of view, this problem can be minimised by a suitable choice of subsets, having for example, the same hydrophobic and/or electronic effects (e.g. alkyl groups with the same number of carbon atoms) and varying steric effects, depending on structure (e.g. branching). Another example of a suitable subset would be a homologous series having long alkyl chain groups, with a terminal active site; in such a series the steric effects are more or less constant, while hydrophobicity varies. From the mathematical, or statistical, point of view the problem of descriptor interrelationships is related to collinearity. This is discussed further in more detail (section 5.1.3.3).

5.1.2.6 3D QSAR analysis

The fact that 3D structures play an important role in the biological activity of chemical compounds seems quite obvious. An apparent example is biological activity of stereoisomers. Having identical physico-chemical characteristics, they are discriminated by a binding site, which is often a chiral environment

itself, as if they were completely different molecules. The discussed above shape analyses, MSD, MTD and related receptor mapping techniques present early approaches to **3D QSAR analysis**. Several methods and related programs were developed within the fast developing field of 3D QSAR analysis: e.g. GRID, LUDI, HINT, APEX-3D [46]. The most widely used method is *CoMFA* (Comparative Molecular Field Analysis) [21].

5.1.3 Methods of QSAR building

The purpose of any QSAR study is, primarily, a description of a system of interest and, desirably, a prediction of biological activity for new compounds. The molecular description involves the application of existing or elaboration of new parameters, most appropriate for a particular problem. Mathematical methods, although they play a secondary role in the QSAR field [22], provide a necessary tool for manipulating the biological data and descriptors.

5.1.3.1 Multiple Linear Regression (MLR)

MLR analysis has originally been used in early QSAR studies and still remains the most widely used technique of QSAR formulation. Regression analysis correlates independent variables **X** (structural descriptors) with dependent variables **Y** (biological activity). Mathematically, it represents an exact procedure, although it derives correlations from data which contains experimental error. The MLR apparatus is described in the Methods section (5.2.2) together with the associated statistical attributes.

In MLR analysis, out of all possible equations the "best" one must be selected. Five fundamental criteria for such selection are as follows [47]:

1. the broadest feasible range of independent variables must be considered: the selected variables should be orthogonal to each other

2. the choice of each variable must be validated by an adequate selection procedure (see below)

3. principle of parsimony should be followed

4. in order to avoid chance correlations there should be at least five data points for each variable

5. physico-chemical and biological considerations should be taken into account as much as statistical ones

5.1.3.2 Variable selection procedures

The best subset regression, i.e. the search over all possible subsets, permits to inspect all regression equations built out of all independent variables. This method has a voracious appetite for computer time, therefore most software packages permit the user to employ some criterion in order to exclude from the output those combinations of variables that can be ruled out *a priori* [48]. Some authors consider this procedure to be most useful, particularly if the number of independent variables is not too large [48].

The stepwise procedures are used if there is a large number of variables to select from. They involve the *backward elimination* or the *forward selection* or the combination of both. These methods are based on the deletion of a variable from the equation or inclusion of a variable into the equation on the basis of the *F*-ratio statistic and linear effects of other variables.

The stagewise procedure computes the plots of residuals against variables which are candidates for the inclusion. This procedure is useful for the identification of outliers and influential points.

The caveat. There are two problems related to variable selection in MLR analysis. Firstly, often there may be models clustered around the "chosen" one, which are nearly as good and often statistically indistinguishable. Secondly, the possibility to pick up a nonsense relationship as the best model could be fairly high as has been shown with an example of random numbers used as independent variables [49]. Notwithstanding these problems, the procedures for variable selection, if carried out intelligently, can be very useful. The precautions (the relative importance of which depends on the particular objective of the model) to take should be: not to drop an important variable, not to drop an easy-to-predict variable in favour of a more troublesome one. Finally, the results of variable selection based on statistical relations should be secondary to the theoretical considerations based on the underlying structure of the relationship.

5.1.3.3 Collinearity

As already mentioned (section 5.1.2.5), miscellaneous molecular parameters can exhibit various relationships between themselves. Such a phenomenon is called *collinearity*, and this presents certain problems in QSAR analysis:

- a secondary parameter may enter the equation, if it is correlated with the relevant one, leading to misinterpretations;

- faulty values of regression coefficients in the equation with correlated variables may lead to false conclusions;

- the correlation between parameters describing different types of structural effects makes it impossible to draw a conclusion about the presence of a certain effect and its relevant significance;

- if the relationship between parameters exists solely in the sample under study, the QSAR equation gives completely fruitless results with very limited predictive power;

- collinearity may also occur between more than two parameters, such a phenomenon is often (but arguably) called *multicollinearity*.

Thus, the problem of *collinearity* is related to the mechanistic interpretability of results and to the selection of the best equation. In regard to these aspects, two ways to avoid and/or break down collinearity could be followed. An optimal selection of training series and structural descriptors (the examples are presented in section 5.1.2.5) may be undertaken. Alternatively, or complementarily, the data preprocessing and reduction and the application of special mathematical techniques, such as principal components and partial least squares, may be employed.

5.1.3.4 Factor analysis approaches

The most important use of *principal components analysis* (PCA) is to reduce the *m*-dimensional data structure to a smaller number of dimensions [50]. This allows the detection of clustering objects and outliers and the results of PCA can be considered to be a model of the data. The mathematical description of PCA is provided in the Methods section (5.2.2).

The most encouraging new method in multivariate statistics is the *partial least squares* (PLS) regression [51]. In this method a very large number of

independent variables (the **X** block) can be correlated with one or more dependent variables (the **Y** block). Methodologically this method is similar to PCA, but differs from the latter in that the components are not independently deduced from the two blocks of variables. They are also constructed to maximise correlations between two blocks. The results of the PLS analysis can be represented as regression coefficients of the **X** variables. Although perfect correlations are often obtained in PLS analysis, mostly due to a large number of **X** variables, the goodness-of-fit is no criterion for the obtained model. Model selection is carried out on the basis of a cross-validation procedure, giving a criterion for the predictive power of the model. The cross-validation procedure is described in the Methods section (5.2.2) together with its associated statistical attributes. PLS, as well as PCA, has a range of considerable benefits in comparison with ordinary least squares regression; the ability to deal with collinearity and robustness being two of the most important. For application in 3D QSAR, where the number of independent variables reaches thousands, PLS is obviously a one and only solution. However, these methods also have certain drawbacks [52]. In particular, although the components could be converted into measures of **X** contributions, they have no physical meaning. Also, the result could be obscured by noise variables.

5.2 METHODS

5.2.1 Physico-chemical and structural descriptors

Twenty-seven platinum complexes of general formula cis-$[Pt(NH_2R)_2Cl_2]$ were chosen as a training set for model building, carrier ligands NH_2R are summarised in Fig. 4.8.

5.2.1.1 Biological activity parameters

The following biological activity parameters were used. pLD_{50} – acute toxicity, where LD_{50} is the minimum dose causing death in 50% of exposed animals: pID_{90} – antitumor potency, where ID_{90} is the minimum dose resulting in 90% tumor regression. $logTI$ – therapeutic index, the measure of the selectivity of the compound as an antitumor agent (TI = LD_{50}/ID_{90}). All of these were measured in the same manner [53] for all the compounds in the series. Units for LD_{50} and ID_{90} were converted from mg/kg to mol/kg. The values of biological activity parameters are compiled in Table 5.1.

5.2.2.2 Steric descriptors

To model carrier ligand steric effects in platinum complexes two descriptors were calculated, as described in the previous chapter (Table 4.3). These are as follows:

LRE – Ligand Repulsive Energy, expressed by the gradient of the vdW repulsive energy between the ligand and a $Cr(CO)_5$ fragment, a "steric probe", to which it binds (Fig. 4.7(A)).

Table 5.1. Biological activity of platinum complexes cis-[Pt(NH$_2$R)$_2$Cl$_2$].

No.	am(m)ine ligand NH$_2$R	pLD$_{50}$	pID$_{90}$	logTI
1	ammonia	1.36	2.27	.91
2	cyclopropylamine	.82	2.22	1.40
3	cyclobutylamine	.66	2.15	1.49
4	cyclopentylamine	-.04	2.26	2.30
5	cyclohexylamine	-.84[a]	1.59	2.43[b]
6	cycloheptylamine	-.31	1.81	2.11
7	cyclooctylamine	-.10	.35	.46
8	$endo$-2-aminonorbornane	-.14	1.61	1.74
9	exo-2-aminonorbornane	-.25[a]	1.19	1.44[b]
10	1-aminoadamantane	-.04[a]	-.04[a]	.00[b]
11	2-aminoadamantane	-.15[a]	-.15[a]	.00[b]
12	isopropylamine	1.06	2.63	1.57
13	isobutylamine	.69	1.82	1.13
14	isoamylamine	-.42	1.88	2.30
15	2-aminohexane	-.19	1.23	1.42
16	cyclohexylmethylamine	.74	1.16	.42
17	1-amino-4-cyclohexylbutane	.64	.64	.00[b]
18	4-aminocyclohexanol	1.35	1.71	.36
19	4-methyl-cyclohexylamine	-.30	-.38	-.08[b]
20	methylamine	1.25	1.43	.19[a]
21	ethylamine	1.13	1.47	.34
22	n-propylamine	1.16	1.50	.34
23	n-butylamine	.57	1.61	1.04
24	n-pentylamine	.68	1.07	.40
25	n-hexylamine	-.33	-.51	-.15[b]
26	n-heptylamine	-.26	-.26	.00[b]
27	n-octylamine	.42	.42	.00[b]

[a]These values, given in the literature as >, were arbitrarily assigned the minimum quoted value. [b]These values were calculated from the values for potency and toxicity.

CRE – Complex Repulsive Energy, expressed by the gradient of the vdW repulsive energy between the metal – carrier ligand moiety and 9-ethylguanine to which it binds (Fig. 4.7(B)).

5.2.1.3 Electronic descriptors

Six parameters were chosen to describe the electronic structure of amine ligands in platinum complexes (Table 5.2):

$q(N)_{PEOE}$ – the partial atomic charge on the amine nitrogen, calculated using the QSAR module of ChemPlus [54], a set of extension modules to HyperChem, which employs an empirical model built on the Partial Equalization of Orbital Electronegativity (PEOE) method [55]. According to this model, the electronegativity of an atom lessens as it accumulates negative charge flowing from less electronegative to more electronegative atoms. Thus, the electronegativity of each atom equalises stopping the charge flow at a point corresponding to the charge distribution. This model, if fully implemented, clearly results in a physically impossible outcome. That is, the partial charges of all atoms of a given element in the system become equal. Hence, in a PEOE scheme atoms are parameterised not only according to the element, but also according to hybridisation. The computation in this method is iterative and the electronegativity is precluded from completely equalising, so that atoms are more responsive to their immediate neighbours. This method depends only on atom type and connectivity, which gives it both advantages and disadvantages in comparison with semiempirical calculations. Calculations are much faster than by semiempirical methods. Since geometry is irrelevant for this calculation, atomic charges are not biased by it. This method, however, is insensitive to conformational variation.

$q(N)_{AMI}$ – the partial atomic charge on the amine nitrogen, calculated using the AM1 quantum-mechanical method [56]. This method is based on the NDDO (Neglect of Diatomic Differential Overlap) approximation [57] and is regarded to be the most accurate semiempirical method of HyperChem and the best method for collecting quantitative information.

Partial atomic charge/charge density at certain atoms or parts of a molecule can carry information on the reactivity for that part of a molecule [58]. In this study complexes were "split into amines" to obtain nitrogen partial charges for two reasons. First of all, it simplifies the calculation. Secondly, in this series of congeners structural changes occur only in carrier ligands and are not affected by the metal; that is not to say that the electronic characteristics of metal atom do not change with the structural variations. Finally, to the best of my knowledge to date there is no available semiempirical method properly treating transition metals beyond Cd. In order to check the suitability of using partial charges in "amines" for representation of charges in complexes, nitrogen charges in corresponding Pd complexes were calculated using the ZINDO/1 semiempirical method. A reasonable correlation with the above values for amines was found (Fig. 5.1).

Energies of the frontier orbitals, E_{HOMO} and E_{LUMO}, could be first approximations to the compound's nucleophilicity and electrophilicity, respectively [59]. ΔE is the difference in energy levels between E_{HOMO} and E_{LUMO}. There are two major objections to semiempirical calculations of these energies, mainly concerned with E_{LUMO}. The theoretical objection is that the validity of equating E_{LUMO} to the electron affinity is doubtful, because the energies of unoccupied orbitals are not optimised in the Hartree-Fock procedure used. However, the recent studies by B.W.Clare
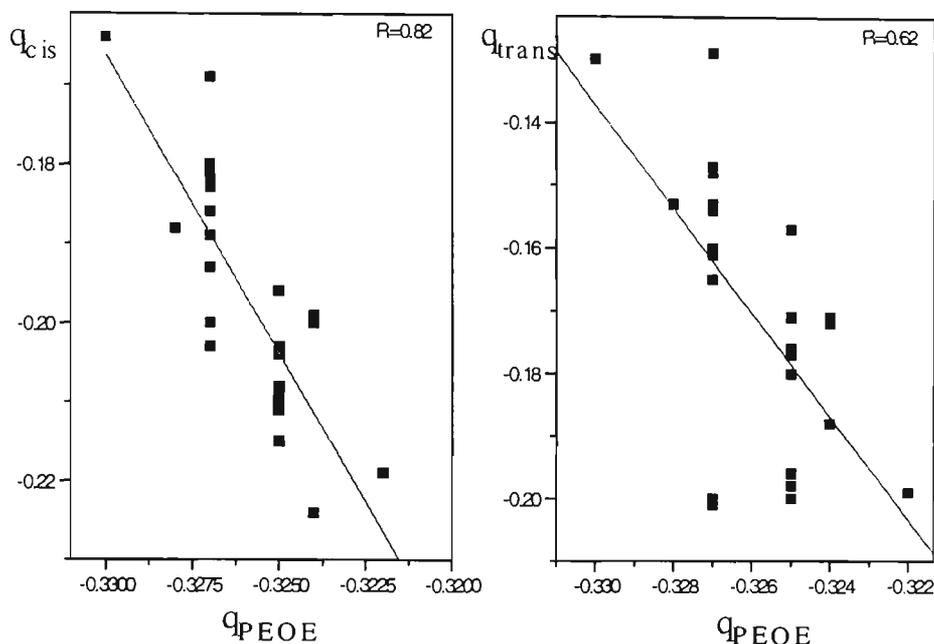
Figure 5.1. *Relationships between partial charges on the carrier ligand nitrogen atoms in [Pd(NH₂R)₂Cl(9-EtG)] complexes, calculated by ZINDO/1 semiempirical method, and partial charges on the nitrogen atoms in amines (NH₂R), calculated by the PEOE method. Cis and trans subscripts correspond to position of amine carrier ligands in Pd complexes with respect to the chlorine atom.*

[15,16] addressing the charge-transfer complexes have shown that E_{LUMO} is an extremely good measure of the ability of the donor-acceptor pair to form such complexes. The computational objection is based on a discrepancy, sometimes essential, between the frontier orbital energy values obtained by different methods (e.g. CNDO and PM3). However, it has been shown that the values obtained with more modern PM3, AM1 and MNDO methods correlate well with each other and with those obtained from *ab initio* calculations [60].

pK_a – the acidity constant of the amine [61-63]. pK_a values serve as a general measure of electronic influence and degree of ionisation [7]. pK_a values reflect electronic properties in a direct manner, as seen from the definition of the Hammett constant σ and, therefore, could be used instead of σ values.

Table 5.2. Electronic descriptors.

No.	q(N)$_{AM1}$	q(N)$_{PEOE}$	E$_{LUMO}$	E$_{HOMO}$	pK$_a$	ΔE
1	-.375	-.339	4.102	-10.530	4.76	-14.632
2	-.311	-.325	2.841	-9.568	9.10	-12.409
3	-.326	-.325	3.424	-9.720	10.64	-13.144
4	-.330	-.325	3.449	-9.760	10.65	-13.209
5	-.342	-.325	3.502	-9.594	10.64	-13.096
6	-.345	-.325	3.479	-9.583	10.67	-13.062
7	-.342	-.325	3.464	-9.499	10.65	-12.963
8	-.333	-.324	3.384	-9.672	10.67	-13.056
9	-.330	-.324	3.367	-9.670	10.67	-13.037
10	-.322	-.322	3.350	-9.616	10.14	-12.966
11	-.336	-.324	3.395	-9.603	10.67	-12.998
12	-.344	-.325	3.492	-9.689	10.67	-13.181
13	-.343	-.327	3.522	-9.695	10.48	-13.217
14	-.342	-.327	3.571	-9.617	10.64	-13.188
15	-.344	-.325	3.488	-9.648	10.63	-13.136
16	-.343	-.327	3.486	-9.671	10.67	-13.157
17	-.343	-.327	3.391	-9.730	10.57	-13.121
18	-.341	-.325	3.293	-9.707	10.58	-13.000
19	-.342	-.325	3.493	-9.600	10.67	-13.093
20	-.340	-.330	3.714	-9.819	10.64	-13.533
21	-.342	-.328	3.579	-9.740	10.64	-13.319
22	-.342	-.327	3.594	-9.703	10.60	-13.297
23	-.341	-.327	3.589	-9.692	10.78	-13.281
24	-.341	-.327	3.575	-9.695	10.71	-13.270
25	-.341	-.327	3.565	-9.697	10.63	-13.262
26	-.341	-.327	3.554	-9.696	10.67	-13.250
27	-.342	-.327	3.544	-9.699	10.65	-13.243

The amines of this investigation were geometry optimised utilising the MM+ force field of HyperChem; the conjugate gradient (Polak-Ribiere) algorithm was employed with the termination condition being RMS < 0.1 kcal/($\text{Å}\times$mol). The QM-derived descriptors used in this investigation were then obtained using AM1 Single Point calculations with the following set of options: total charge = 0; spin multiplicity = 1 (lowest state, Restricted Hartree-Fock calculation method for spin pairing); SCF controls: convergence limit = 0.01 kcal/mol and iteration limit = 50.

5.2.1.4 Hydrophobicity descriptors

Hydrophobicity, or transport, effects were modelled using the octanol/water partition coefficient *logP*. The *logP* values for coordinated amine ligands were calculated using the QSAR module of ChemPlus (Table 4.3). This calculation is based on an additive function of atomic contributions:

$$logP = \sum_i n_i a_i \qquad (5.8)$$

where n_i is the number of atoms of type i, and a_i is the contribution of the corresponding atom type. Atom type classification is based on valence geometry (hybridisation), formal charge density, approachability of the solvent molecule towards the atom and the linear independency of the columns of the data matrix arising from the above equation. Atomic parameters and the functionality for the *logP* calculation are taken from Ref. [64].

To verify that ligand *logP* values may be used to model complex hydrophobicity or transport effects, all available experimental chloroform/water partition coefficients for the complexes themselves were examined for correlation with the theoretical

octanol/water partition coefficients for the ligands themselves (Fig. 5.2). A reasonable correlation was found which is consistent with other workers [65] which showed a strong correlation between experimentally determined hydrophobicity of both complexes and single uncomplexed ligands. It should be noted that in the latter study the partition coefficient contribution of ligands to the hydrophobicity of their complexes do not equal the *logP* values of the free ligands.
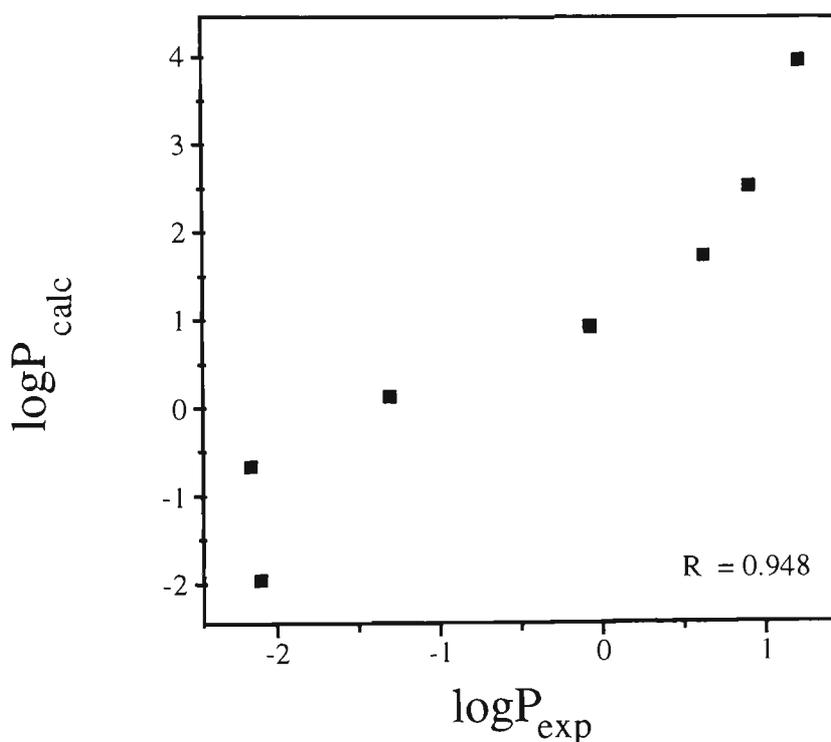
Figure. 5.2. Plot of calculated logP values (octanol/water) for the coordinated amine ligands in [Pt(NH₂R)₂Cl(9-EtG)] versus experimental logP values (chloroform/water) for complexes themselves, where R represents cycloalkyl substituents. Experimental values are taken from Ref. [53].

5.2.2 Multivariate analysis

Data analysis was performed using the SPSS [66] and SCAN [67] software packages.

5.2.2.1 Multiple Linear Regression analysis

A linear regression model for a data set containing multiple independent variables is derived on the basis of a least squares procedure [68]:

$$y = ax_1 + bx_2 + \dots + mx_m + k \tag{5.9}$$

The terms a, b, ... m are called *regression coefficients*. The power of the obtained model is characterised by a number of statistical parameters.

The *correlation coefficient* R is a measure of goodness-of-fit of the obtained model, the squared correlation coefficient R^2 is the measure of *explained variance* of the data:

$$R^2 = 1 - \frac{\sum \Delta^2}{S_{yy}} \tag{5.10}$$

In this formula S_{yy} is the *overall (total) variance*:

$$S_{yy} = \sum (y_{obs} - y_{mean})^2 = \sum y^2 - \frac{(\sum y)^2}{n} \tag{5.11}$$

$\sum \Delta^2$ is the sum of squared error, the *unexplained or residual variance*:

$$\sum \Delta^2 = \sum (y_{obs} - y_{calc})^2 \tag{5.12}$$

In the above equations n is the number of objects and y_{obs}, y_{mean} and y_{calc} represent observed, mean and calculated values of Y variable, respectively.

The *standard deviation* s is an absolute measure of quality of fit, since it takes into consideration the number of degrees of freedom:

$$s = \sqrt{\frac{\sum \Delta^2}{n - k - 1}} = \sqrt{\frac{(1 - R^2)S_{yy}}{n - k - 1}} \qquad (5.13)$$

where k is the number of variables. The *confidence intervals* for the required level of statistical significance can be calculated from the standard deviation.

The **F-ratio** is a measure of *statistical significance* of the model:

$$F = \frac{R^2(n - k - 1)}{k(1 - R^2)} \qquad (5.14)$$

Stepwise regression analysis. This procedure involves computerised selection of the best single variable and then considers the remaining variables one at a time until the best two (or more) - variable equation is arrived at. A variable enters the equation only if the significance probability, p, associated with the *F-test* (which tests the hypothesis that the correlation coefficient equals zero) is less than .05 (95% confidence level). The process continues until the addition of a variable is not justified by the "*F-statistic*".

Enter regression analysis. This alternative/complementary approach involves computation of multiple regression with a fixed set of variables.

Because situations are possible where significant variables show up only in

a certain combination of two or more [1,69], both of the above approaches were employed as follows. Firstly, *stepwise*-MLR was applied to pick up the "best" equations with one or more variables. Secondly, *enter*-MLR was carried out on intuitively chosen variables to see if an improvement with respect to correlation strength and statistical significance occurs. To avoid generating meaningless regression results, endpoint vs descriptor plots were checked for clusters, outliers, and parabolic behaviour, prior to the regression analysis.

5.2.2.2 Partial least squares analysis

This procedure constitutes of the iterative PLS runs in which certain number of objects are eliminated from the data set in some predefined manner and then predicted by the obtained model. The statistical characteristics of the PLS model are as follows. The standard deviation s_{PRESS} is a criterion of the optimal number of components:

$$s_{PRESS} = \sqrt{\frac{PRESS}{n - k - 1}} \tag{5.15}$$

The R^2 is defined similarly to regression analysis; the R^2_{cv}, or R^2_{PRESS}, is defined correspondingly for cross-validation runs. The latter is always smaller than the former and is a measure of goodness-of-prediction:

$$R^2_{cv} = 1 - \frac{PRESS}{\sum (y_{obs} - y_{mean})^2} \tag{5.16}$$

where *PRESS* is a predictive residual sum of squares and y_{pred} are the predicted values of Y variable.

$$PRESS = \sum (y_{obs} - y_{pred})^2 \tag{5.17}$$

5.2.2.3 Model interpretation

The following interpretation of correlation coefficients was implemented:
0 to 0.2 – very weak; 0.2 to 0.4 – weak; 0.4 to 0.7 – moderate; 0.7 to 0.9
– strong; 0.9 to 1 – very strong. The following criteria were implemented
for the choice of best model [70]: to be useful the model should be able to
explain the majority of biological variance, i.e., R^2 should be greater than
0.5; an increase of 0.05 to 0.1 in the value of the correlation coefficient R
and a similar decrease in the value of the standard error s are estimates of
a significant change. The cross-validated variance R^2_{cv} is a measure of the
predictive power of a model (the "leave-one-out" procedure was used).

5.3 RESULTS AND DISCUSSION

5.3.1 Analysis of reported QSAR studies in platinum systems

An extensive literature search revealed a limited number of attempts to correlate the biological activity of platinum-am(m)ine complexes with their structures; six groups being concerned with antitumor activity [43,71-77] and one with mutagenicity [78]. A search of the MedChem/Biobyte QSAR database [79,80] revealed an additional two QSARs involving platinum-am(m)ine complexes. These studies, for toxicity only, include cisplatin amongst a number of miscellaneous compounds [81], and a series of substituted dichloro(o-phenylenediamine)platinum(II) complexes [82].

Of the QSAR studies on platinum systems reported in the literature, Abdul-Ahad et al. [71] have investigated the correlation of standard indicators of biological activity, namely acute toxicity, pLD_{50}, antitumor potency, pID_{90}, and therapeutic index, $logTI$, with a variety of structural descriptors. These include a wide range of electronic parameters, and the molecular volume, MV, as a steric parameter. In some cases this has resulted in moderate to strong correlations with a high degree of statistical significance. However, in this study, parameters relating to transport phenomena (e.g. hydrophobicity) have not been taken into account; and the use of MV oversimplifies steric effects since it is a scalar quantity [34]. Steric effects are, in fact, often highly localised and directional.

Considering the use of MV (and other bulk parameters) in QSAR studies it must be reported that it often results in strong and significant correlations [34]. Which means that such use is quite justified for the prediction of biological activity, if this is considered as the sole goal for the building of QSAR in

a particular study. But if a question is asked about the physical interpretation of the obtained relationship and if an insight into the mechanism of biological activity is of interest, the use of bulk parameters poses a problem. Interpretations of such relationships are often vague and contradictory. Furthermore, *MV* is correlated with *MR* [37] and *logP* [38] and, hence, bears some electronic and transport information. Therefore, if it is used in combination with other parameters, a high degree of collinearity is to be expected, affecting the statistical significance of the final relationship.

Tang et al. [75] have used exclusively electronic descriptors to obtain regression equations (having more than three variables) for platinum complexes, which show moderate to strong correlation and high statistical significance. These authors report the synthesis of a complex with relatively low toxicity and high potency which possesses the electronic characteristics required by their regression results. But it is not clear how one goes about rationally designing compounds according to obtained requirements. Therefore, it is not clear whether this approach is generally applicable.

Another QSAR study, employing exclusively electronic parameters is reported in Ref. [76]. This study suffers from the same limitations specified for analysis carried out by Tang et al. [75]. Authors [76] deliberately omitted parameters, which describe transport effects, membrane permeability and metabolic processes, from the regression analysis based on the assumption that platinum complexes are stable and inert up until they get into the cell.

Regression analyses carried out in these studies [71,75,76] have been regarded not to be reliable for predicting compounds with lower toxicity [83]. Therefore, the results from these studies have only limited applicability.

Simon et al. [43,72] have sought correlations involving hydrophobicity, electronic and steric effects with respect to biological activity. These workers have suggested that the biological activity indicator *TI* depends predominately on steric features [72]. For the steric optimisation of *TI* a receptor mapping procedure was employed via minimal steric difference (MSD) technique resulting in quite strong and statistically significant correlation results. However, this approach has certain drawbacks and has received considerable criticism [39,84]. In particular, it should be noted that the parameter used can code both steric and non-steric factors, and the method does not account for a molecule's three-dimensional structure and conformational flexibility.

All of the researchers referred to so far used either measured physicochemical properties or computed descriptors representing classic QSAR studies (Hansch type analysis) [69].

A more recent attempt to quantify structure-activity relationships for platinum complexes involves the application of a graph-theoretical method and employs purely topological indices. Thus Romanowska et al. [73,74] have demonstrated the usefulness of molecular topology for structure-activity studies of platinum-am(m)ine complexes. In another study using topological indices [77] an arguable suggestion is brought forward that the use of topological descriptors enables the search of new antitumour platinum complexes without the exact knowledge of their mechanism of action and metabolic pathways. Unfortunately, this report [77] represents the only work by these authors in regard to Pt-QSAR found in the literature, therefore not allowing deeper analysis of their approach. Finally, with respect to their use for QSAR analysis, even though topological indices generally offer an almost universal means of representing a chemical structure, they usually lack physical or chemical meaning and cannot be extrapolated

to other compounds [3,85].

The major weakness of the MLR, performed in the above studies, and especially in Ref. [75], is that many of the descriptors are interrelated and the same information is carried by more than one parameter. Such collinearity diminishes the statistical significance, complicates the interpretation of the correlation and may lead to false conclusions and predictions [86]. Collinearity may be diagnosed from a correlation matrix. Thus, correlation matrices have been reported by some authors [43,73], although it was necessary to perform correlation matrix computations based on published data [71,75] where this has not been carried out by the authors themselves (Tables 5.3 and 5.4).

Table 5.3. Correlation matrix for the descriptors used in Ref. [75].

	q_{Pt}	$-q_{Cl}$	$-q_N$	q_{Am}	ΔE	Q_{Pt-Cl}	Q_{Pt-N}	ΔQ	N_H
q_{Pt}	1								
$-q_{Cl}$	0.470	1							
$-q_N$	-0.446	0.327	1						
q_{Am}	-0.161	0.791	0.682	1					
ΔE	0.201	-0.208	**-0.642**	-0.395	1				
Q_{Pt-Cl}	**-0.786**	**-0.81**	0.084	-0.357	-0.139	1			
Q_{Pt-N}	0.052	0.742	0.419	0.793	-0.002	-0.53	1		
ΔQ	-0.359	**-0.867**	-0.27	-0.719	-0.054	0.790	-0.938	1	
N_H	**-0.781**	-0.217	0.507	0.297	-0.058	0.619	0.166	0.132	1

Descriptors: q_{Pt}, $-q_{Cl}$, $-q_N$, q_{Am} - electronic charge on corresponding atom and moiety; ΔE - the gap between energies of LUMO and HOMO; Q_{Pt-Cl}, Q_{Pt-N} - overlap population on corresponding bonds; ΔQ - the difference between the above; N_H -the number of protons on N atom of an amine. Depicted in bold are the correlation coefficients between the descriptors, simultaneously used in equations obtained in Ref. [75] which show moderate to strong collinearity (correlation coefficient more than 0.6) .

Table 5.4. Correlation matrix for the descriptors used in Ref. [71].

	BE	E.F(P_7)	E.F^2(P_7)	E.F(P_9)	E$_x$(P_3)	E$_z$(P_9)	F^N(Pt)	MV	V$_2$(P_7)
BE	1								
E.F(P_7)	0.619	1							
E.F^2(P_7)	0.615	0.999	1						
E.F(P_9)	-0.35	-0.252	-0.25	1					
E$_x$(P_3)	0.566	0.343	0.336	-0.265	1				
E$_z$(P_9)	-0.376	-0.255	-0.252	**0.995**	-0.302	1			
F^N(Pt)	-0.075	-0.189	-0.191	-0.344	-0.073	-0.352	1		
MV	-0.982	-0.622	-0.618	0.393	-0.481	0.413	-0.004	1	
V$_2$(P_7)	**0.69**	0.414	0.407	-0.068	0.464	-0.114	0.128	**-0.663**	1

Descriptors: BE - molecule binding energy (the difference between the INDO total energy of the molecule and the sum of the energies of the isolated constituent atoms); E.F(P_7), E.F(P_9) - the modulus of the electric field at p (receptor point); E$_x$(P_3), E$_z$(P_9) - components of the electric field at p, calculated by the finite differentiation of the electrostatic potential; F^N(Pt) - frontier electron density for the nucleophilic attack at Pt; MV - molecular volume; V$_2$(P_7) - the energy of polarisation of the complex by a unit point charge placed at p (second order interaction energy, calculated by an uncoupled Hartree-Fock perturbation procedure). Depicted in bold are the correlation coefficients between the descriptors, simultaneously used in equations obtained in Ref. [71] which show moderate to strong collinearity (correlation coefficient more than 0.6).

In summary, research groups studying QSARs of platinum antitumour complexes have investigated the effects of a variety of structural descriptors on biological activity. The highest priority in these studies has been given to electronic descriptors. Hydrophobicity has been less considered, and a quantification of steric effects has lagged far behind. This could possibly be due to the lack of appropriate steric descriptors as applied to inorganic systems. Although special protocols have been developed for quantification of steric effects in inorganic systems (see section 4.1.2), there has been no attempt to extend these protocols to QSAR investigations relating to the biological profiles of

metal complexes, such as the *cis*-platinum derivatives.

Another limitation of the described studies is that all QSARs, discussed above, considered the structural features of unbound platinum complexes without a consideration of the structural features of the possible molecular target(s).

Following sections of this chapter describe the MLR analysis carried out in an attempt to properly account for all three kinds of effects (transport, electronic and steric). In particular, to describe steric effects in platinum complexes the repulsive energy strategy, described in the previous chapter is employed. The new parameter, *CRE*, is applied in an attempt to account for the steric features of the molecular targets of platinum complex binding.

5.3.2 Simple relationships

For all series combined (Fig. 4.8), an analysis of the relationship between the biological activity and the structural descriptors (Figs. 5.3 and 5.4) reveals the following trends.

1.   Not unexpectedly [87], the biological activity suggests a parabolic dependence on *logP*, especially the acute toxicity and, to a lesser degree, the therapeutic index (Fig. 5.3(a-c)). This observation has led to the inclusion of the quadratic term log^2P into the MLR analysis.

2.   With respect to *LRE* and *CRE*, the biological activities do not reveal any obvious correlations (Fig. 5.3(d-i)).

3.   A degree of independence of biological activity with respect to electronic
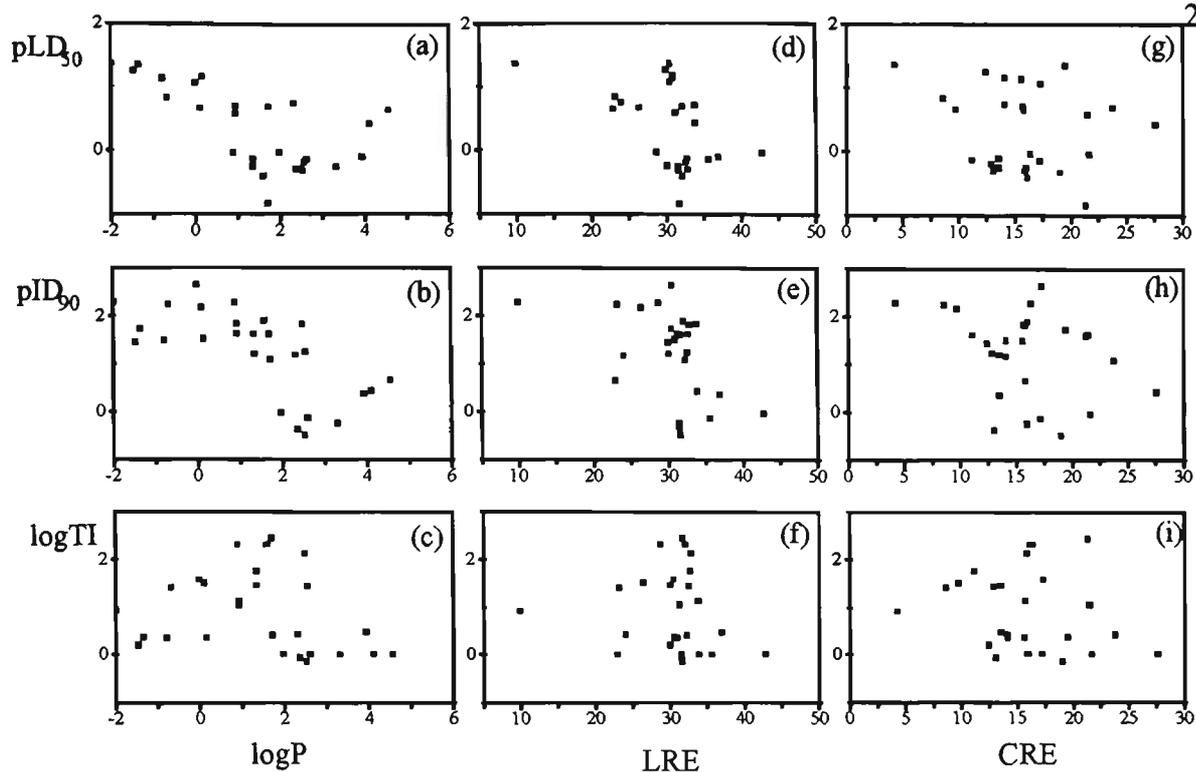
Figure 5.3. Biological activity parameters versus transport and steric descriptors.
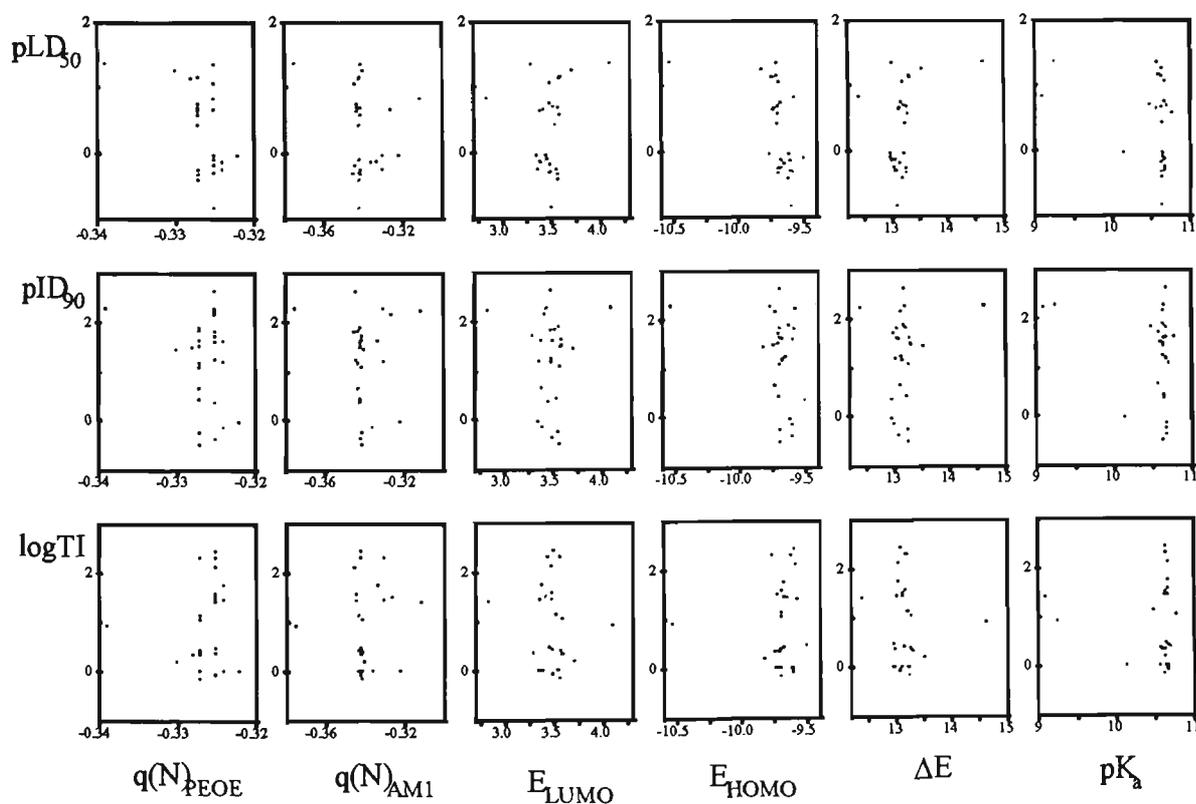
Figure 5.4. Biological activity parameters versus electronic descriptors.

descriptors is suggested, at least for the amines under study. This is manifested by the vertical arrays of Fig. 5.4. Notably, for all of the electronic descriptors, the ligand NH_3 does not conform to this trend. The ligand cyclopropylamine also lies off the trend for a number of descriptors, namely $q(N)_{AM1}$, E_{LUMO}, ΔE and pK_a. Subsequent MLR analysis was carried out both with and without these two complexes and results did not appear to be significantly different.

5.3.3 Collinearity

All structural descriptors used for the MLR analysis were checked for collinearity. The correlation matrix (Table 5.5) shows that the hydrophobicity does not correlate with any individual steric or electronic parameter used. Of the two steric parameters, *LRE* correlates with two out of the six electronic parameters, although not strongly. The steric parameters *LRE* and *CRE* are not strongly correlated with one another reflecting the different physical phenomena underlying their calculation. As expected, electronic descriptors are strongly interrelated, except for pK_a.

Table 5.5. Correlation matrix for the variables used for the MLR analysis.

	logP	log^{2}P	LRE	CRE	pK$_a$	q(N)$_{PEOE}$	q(N)$_{AM1}$	E$_{LUMO}$	E$_{HOMO}$	ΔE
logP	1									
log^{2}P	0.788	1								
LRE	0.416	0.055	1							
CRE	0.418	0.228	0.602	1						
pK$_a$	0.438	0.074	**0.723**	0.528	1					
q(N)$_{PEOE}$	0.401	-0.019	**0.741**	0.372	0.779	1				
q(N)$_{AM1}$	0.170	-0.143	0.547	0.233	0.683	0.845	1			
E$_{LUMO}$	-0.144	0.031	-0.292	-0.064	-0.423	-0.761	-0.848	1		
E$_{HOMO}$	0.495	0.103	**0.749**	0.413	0.862	0.897	0.798	-0.698	1	
ΔE	0.337	0.036	0.553	0.249	0.685	0.896	0.894	-0.929	0.913	1

5.3.4 Multiple relationships

5.3.4.1 Stepwise models

Table 5.6 (upper part) shows *stepwise* regression equations together with their goodness-of-fit (correlation strength), goodness-of-prediction and statistical significance parameters. The plots of the values of pLD_{50}, pID_{90} and *logTI* based on these equations are shown in Fig. 5.5(a-c). It appears from stepwise regression analysis that the biological activity of platinum complexes can be modelled quite well by carrier ligand hydrophobicity. That is, the most significant factor influencing biological activity in this case appears to be transport, i.e. accumulation and distribution. However, it must be realised that factors which are operative at the target site may well be masked by transport effects. That is, the biological activity of a complex often depends on its ability to bind to target sites in a specific way, not simply to get there. Developing QSARs for systems such as cisplatin, which manifest their biological activity through a specific mechanism of action (e.g. binding to DNA), requires that determinative steps of the mechanism such as accumulation and distribution, bioactivation (hydrolysis), DNA binding and finally deactivation should be accounted for in a model. This means that parameters describing all these processes should be generated and included into the model. This may be achieved by introducing a second (and if necessary – a third) terms into the QSAR, in particular those that account for steric and/or electronic effects.

5.3.4.2 Enter *models*

Table 5.6 (middle and bottom parts) shows "best" *enter* regression equations together with their goodness-of-fit (correlation strength), goodness-of-prediction and statistical significance parameters.

Table 5.6. Results of the MLR analysis.

equation	n	s	R	F(p)	R^2cv
stepwise MLR analysis					
$pLD_{50} = 0.08 \log^2 P - 0.45 \log P + 0.56$	27	.42	.78	19.10(.0000)	.51
$pID_{90} = -0.35 \log P + 1.69$	27	.68	.67	19.85(.0002)	.36
$\log TI = -0.07 \log^2 P + 1.20$	27	.77	.44	6.14(.0204)	.11
enter MLR analysis with *LRE* as steric parameter					
$pLD_{50} = 0.09 \log^2 P - 0.48 \log P - 0.01 \, LRE + 0.27 \, pK_a - 1.96$	27	.42	.80	10.49(.0001)	.53
$pID_{90} = -0.34 \log P - 0.067 \, LRE + 1.8 \, E_{HOMO} + 21^{*}$	27	.66	.72	8.48(.0006)	.42
$-0.31 \log P - 0.037 \, LRE - 0.63 \, E_{LUMO} + 5.0$	27	.68	.70	7.46(.0012)	.31
$\log TI = -0.13 \log^2 P + 0.27 \log P - 0.07 \, LRE + 72 \, q(N)_{PEOE} + 27^{*}$	27	.74	.59	2.92(.044)	.10
$-0.15 \log^2 P + 0.32 \log P - 0.05 \, LRE - 0.42 \, E_{LUMO} + 4.2$	27	.75	.57	2.72(.056)	.14
enter MLR analysis with *CRE* as steric parameter					
$pLD_{50} = 0.095 \log^2 P - 0.51 \log P - 0.02 \, CRE + 0.14 \, pK_a - 0.58$	27	.42	.80	19.10(.0000)	.52
$pID_{90} = -0.35 \log P - 0.02 \, CRE + 0.33 \, E_{HOMO} + 5.19$	27	.71	.67	19.85(.0002)	.33
$\log TI = -0.12 \log^2 P + 0.26 \log P - 0.04 \, CRE - 0.11 \, pK_a + 2.8$	27	.77	.54	6.14(.0204)	.10

Equations marked with an asterisk have a strong degree of collinearity (see Table 5.5), so the second equation is provided.
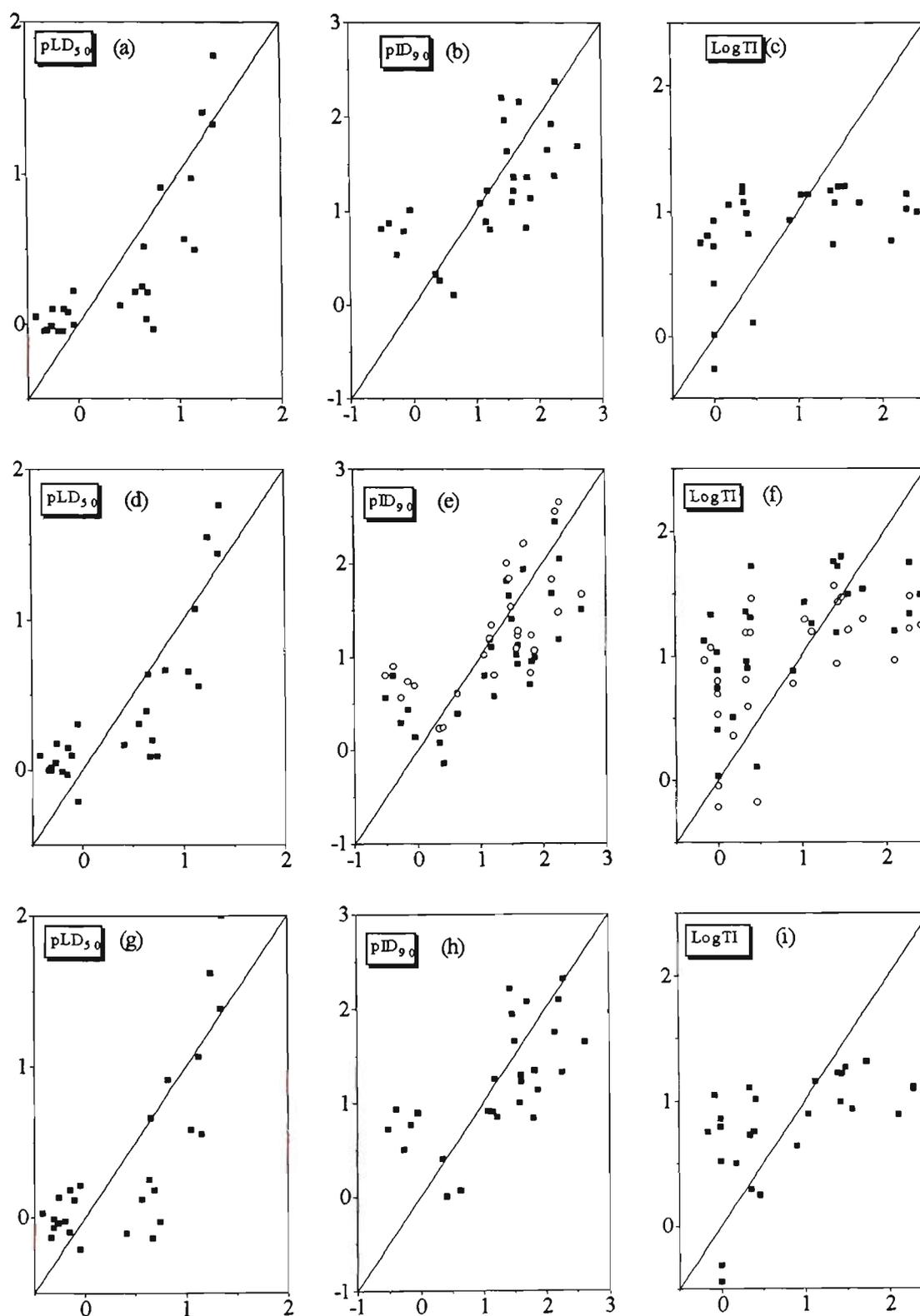
Figure 5.5. *Plots of calculated versus observed values for the biological activity indicators. Data in plots (a,b,c) calculated according to equations from Table 7a; (d,e,f) - equations from Table 7b; (g,h,i) - equations from Table 7c. In plots (e,f)* ■ *represent equations marked with an asterisk,* ○ *represent second equations provided for* pID_{90} *and* $logTI$.

The plots of the values of pLD_{50}, pID_{90} and $logTI$ based on "best" *enter* regression equations are shown in Fig. 5.5(d-i), and show that the equations obtained are fairly predictive of toxicity and potency. However, the QSAR models for the therapeutic index did not achieve the quality that was hoped for initially. Improved results may be achieved with a better choice of initial descriptors or a better formulation of the model (through different chemometrical methods).

Regression results presented here (using *LRE* and *CRE* as steric descriptors) show a marginal improvement (with respect to overall correlation strength and statistical significance) in comparison with the *stepwise* models. However, the partial correlation coefficients and statistical significance of variables describing steric and electronic effects are smaller than those of *logP* (Appendix IX). In some cases these regression equations show an improvement in goodness-of-fit compared to those obtained in previous studies [43,71-75]. However, it is impossible to compare their predictive power since no cross-validation has been carried out or reported in these works. In contrast to some previous studies [71,75], results presented here do not show a predominant dependency of biological activity on electronic structure. Rather, the importance of transport effects is emphasised. They also reveal a consistent contribution of steric effects to the biological outcomes.

5.3.5 PLS analysis

In order to circumvent problems related to collinearity in multivariate data analysis and to ascertain the predictive power of a proposed model the use of PLS is superior to that of MLR [52]. However, the use of MLR is more appropriate for the initial investigation. That is, although the data contains multiple responses (pLD_{50}, pID_{90} and $logTI$) and a certain degree of collinearity

exists in the data, factors that usually call for use of PLS, the small number of descriptors, unlike in 3D QSAR analysis (e.g. CoMFA [21]) limits the number of possible models. Therefore, MLR analysis with careful selection of variables in order to avoid collinearity appears to present a better choice. Apart from a predictive power, what is being looked for in this analysis is a mechanistic interpretation of data and PLS is believed to be poorer in this respect [52]. Furthermore, in order to get a more mechanistic insight, QSAR models are built separately for different biological responses in MLR; the power of PLS would not be exploited in this regard.

Having said all this, it has to be mentioned that PLS analysis has indeed been performed on the studied data set. The results show no significant improvement in goodness-of-fit and goodness-of-prediction in comparison with results from MLR for the models containing collinear descriptors. These results are compiled in Appendix X.

5.3.6 Relationship of biological profiles with repulsive energies

Whilst it is acknowledged that the biological activity of platinum-am(m)ine complexes is dominated by their transport properties, reflected by a strong dependency on hydrophobicity as revealed by QSAR results, *vide infra*, steric and electronic effects <u>after</u> the molecule has reached its target site may become determinative of their relative activities. Further information on steric effects may be extracted by looking for trends involving biological indicators and *LRE* and *CRE* within given series of ligands.

For example, the relationship between toxicity, pLD_{50}, and *LRE* is examined in Figs. 5.6 – 5.9 (parts (a)). It is apparent for cyclic, straight-chain and branched-chain
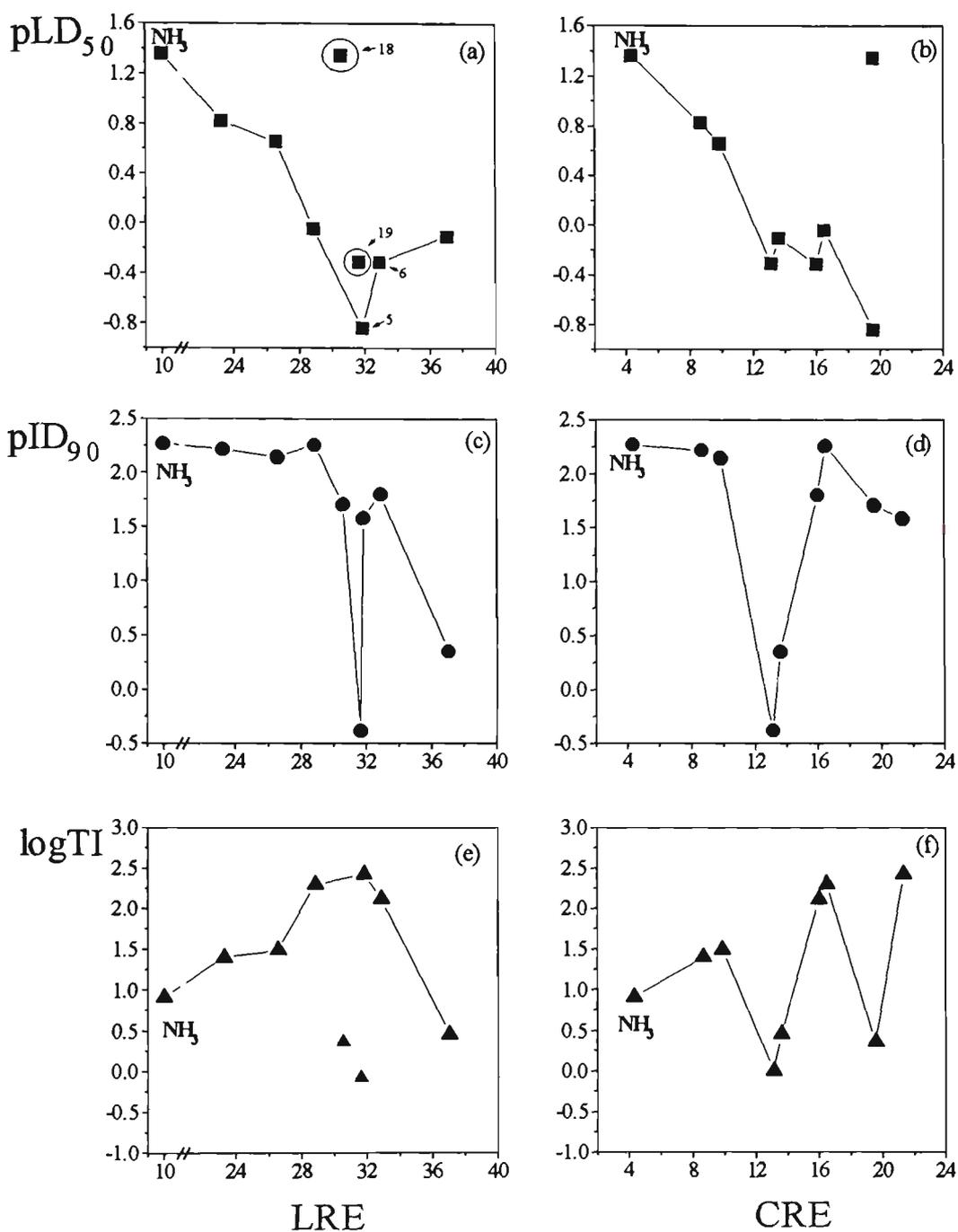
Figure 5.6. Biological activity indicators versus steric descriptors LRE and CRE for cyclic systems.
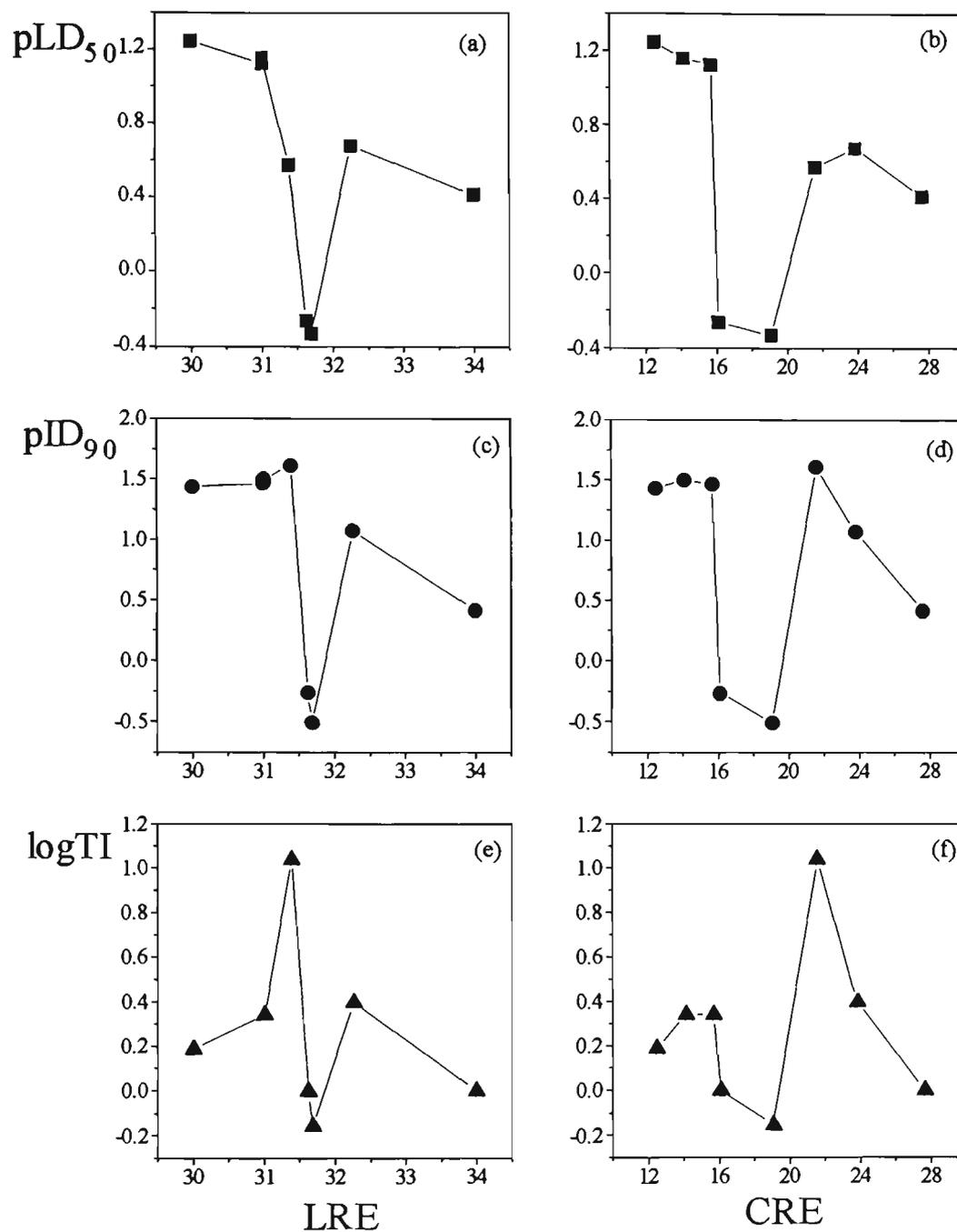
Figure 5.7. *Biological activity indicators versus steric descriptors LRE and CRE for straight-chain systems.*
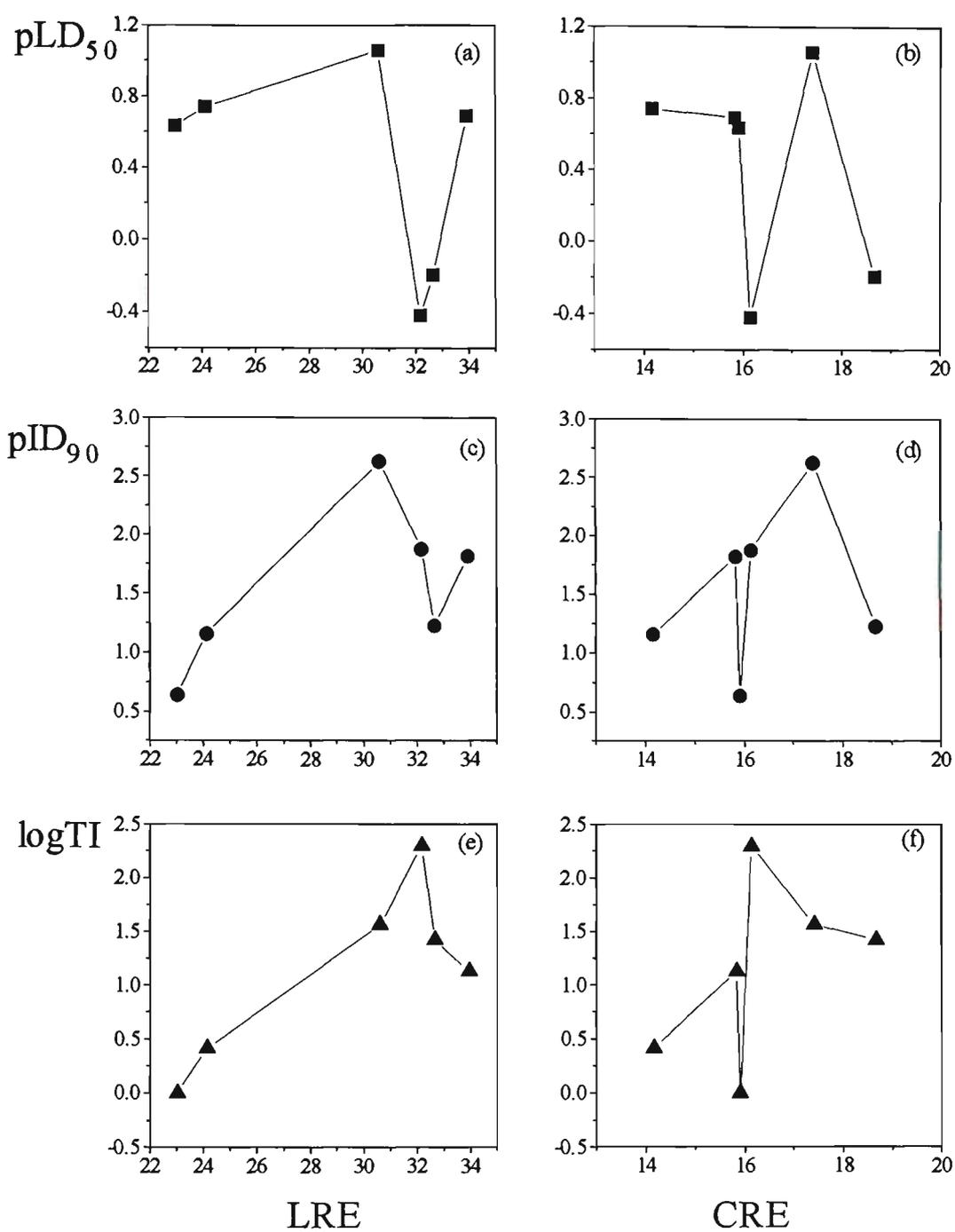
Figure 5.8. *Biological activity indicators versus steric descriptors LRE and CRE for branched-chain systems.*
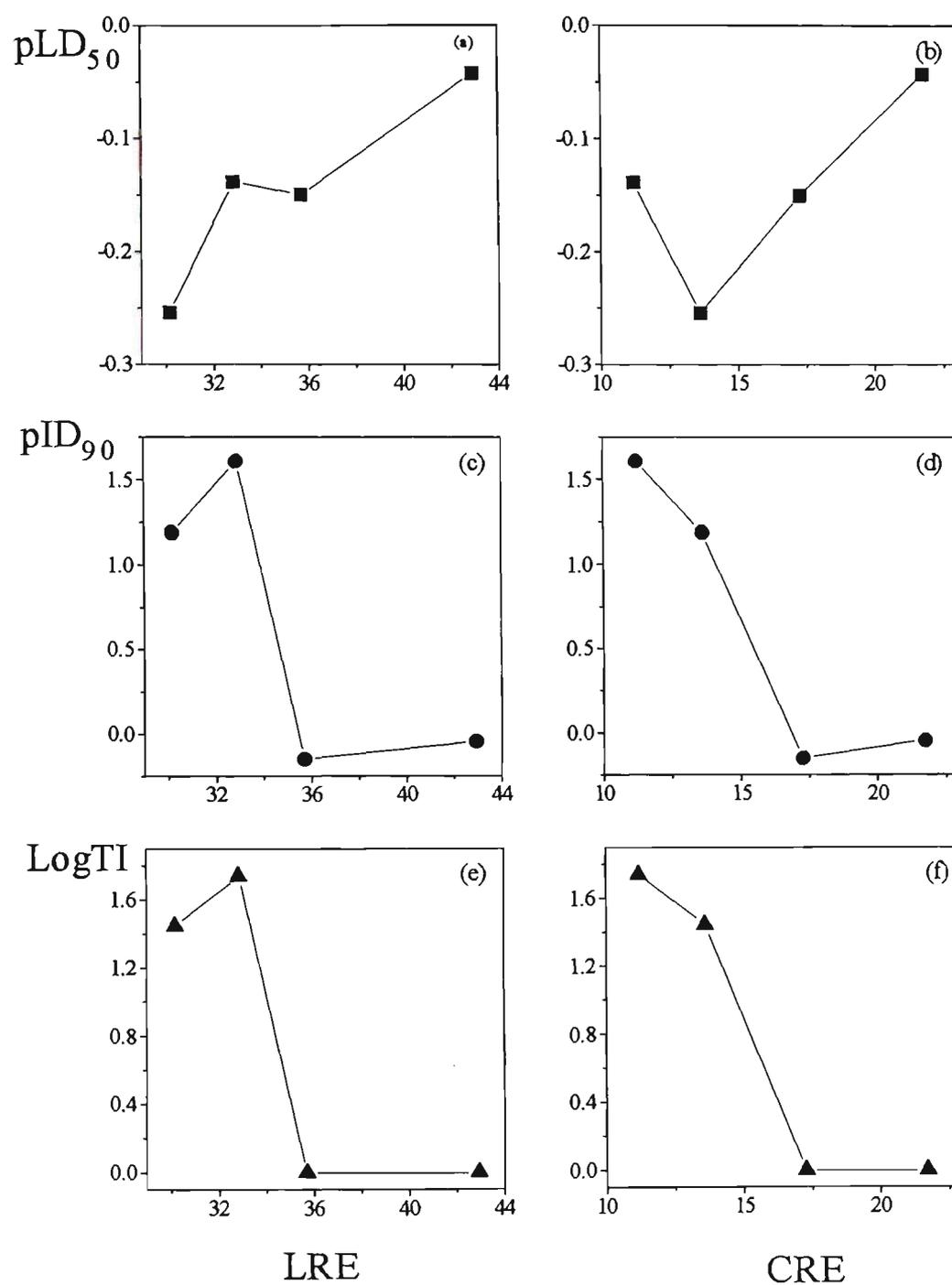
Figure 5.9. *Biological activity indicators versus steric descriptors LRE and CRE for polycyclic systems.*

systems, henceforth referred to as cycles, straights and branches respectively, that a pattern emerges. Interestingly, with increasing *LRE*, the toxicity tends to decrease to a minimum value around *LRE* = 32 kcal/mol for each series and then rises again. This might be suggestive of an optimal steric requirement for minimum toxicity within a given series. Note in Fig. 5.6a that the circled data points representing ligands **18** (and possibly **19**) do not conform to the described trend. These ligands are characterised by having OH and CH_3 substituents in the 4-position of the ring. The remote location of these substituents from the donor amino group does not significantly influence their *LRE* values, as expected. Thus, **5**, **18** and **19** are in an approximately vertical array. It is interesting to note, however, that the absolute toxicity relating to ligand **18** is enhanced, which may be due to the additional hydrogen-bonding capability of the hydroxyl group, stabilising a drug–DNA adduct [88]. The absolute toxicity relating to ligand **19** is of the same order of magnitude as that for **6**; this is anticipated, considering that hydrophobic character of both is very close. Complexes containing polycyclic ligands, henceforth referred to as polycycles, can be considered as presenting only the "right shoulder" of the above trend (Fig. 5.9a). This might be expected since there is a paucity of polycyclic hydrocarbons smaller than norbornane.

While toxicity vs *LRE* plots for individual series of compounds show that the toxicity passes through a minimum, all compounds combined do not display this trend (Fig 5.3d). Furthermore, the regression analysis with respect to toxicity for all compounds combined did not reveal a parabolic dependence on *LRE*. Thus, such trends appear to be series specific.

The observation of minimum values in pLD_{50} vs *LRE* plots, indicative of an optimal steric size for minimum toxicity, raises the following question. Through what mechanism does the steric bulk on carrier ligands, located on

the "ends" of a series of ligands, influence the biological outcome? Since the dependence of pLD_{50} on the steric bulk is not gradual, it is reasonable to suggest that steric interactions of platinum complexes with a biological target may be different along a series. Furthermore, the nearing of toxicity values for complexes on the opposite "ends" of the size scale (in scalar terms, e.g. MV, Fig. 5.10a) is suggestive of an increasing similarity of steric demands for these complexes. While smaller ligands are entirely expected to have low or no steric demands, for larger ligands two scenarios may be envisaged. Firstly, a ligand may present a "rigid bulk". Secondly, a ligand may present a "flexible bulk", which may tilt off or fold away into less sterically demanding conformations. The latter ligands are expected to be relatively more toxic. In order to be able to account for such conformational flexibility and to check the above hypothesis using the RE strategy, it is necessary to extend it to nucleoconstituents of higher organisation, up to oligonucleotides. In relation to a position of ligand **18** in Fig. 5.10a, it may be noted that again it represents an exception of the trend possibly due to the same reasons as proposed in the previous paragraph.

The relationship between the antitumor potency, pID_{90}, and *LRE* is shown in Figs. 5.6 – 5.9 (parts (c)) and reflects the trend observed for pLD_{50}. Closer investigation reveals that pID_{90} is approximately inversely related to the "size" of the ligands rather than depending on their *vector* steric requirements (Fig. 5.10b). That is, complexes with larger ligands which have low aqueous solubility and are highly hydrophobic show low antitumor potency. It can be proposed that the bioavailability of complexes on initial stages of their journey to the binding sites on DNA (that is, prior to crossing of cell membranes) influences the antitumor potency. This trend could be demonstrated by a dependency of pID_{90} on the experimentally determined solubility for cycles and branches and chloroform/water distribution for cycles only [53] (Fig. 5.11).
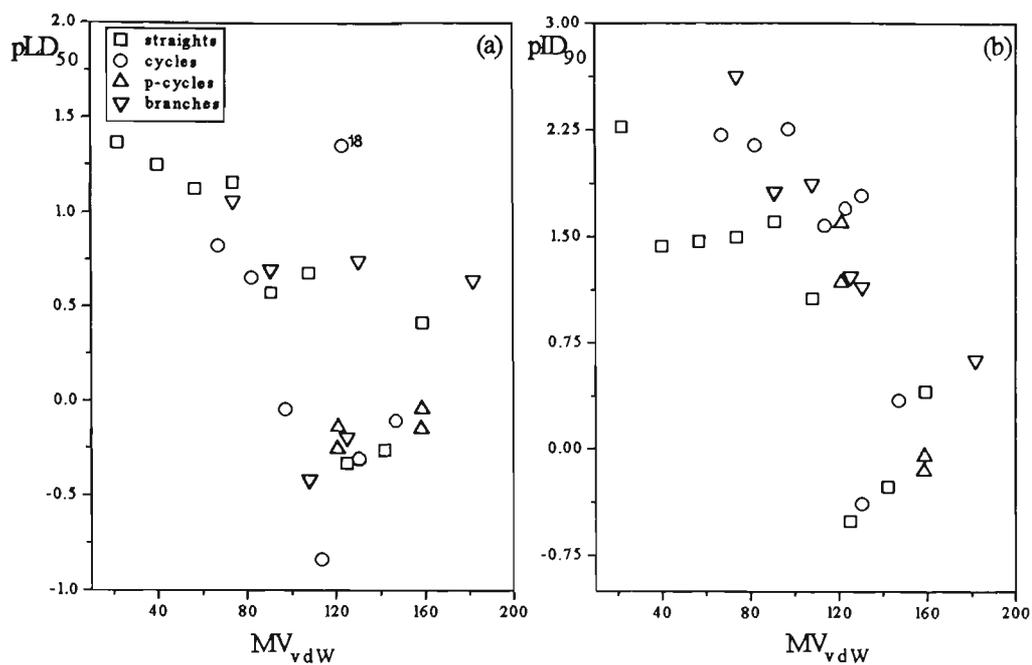
Figure 5.10. Relationships of biological activity parameters with scalar steric size of carrier ligands.
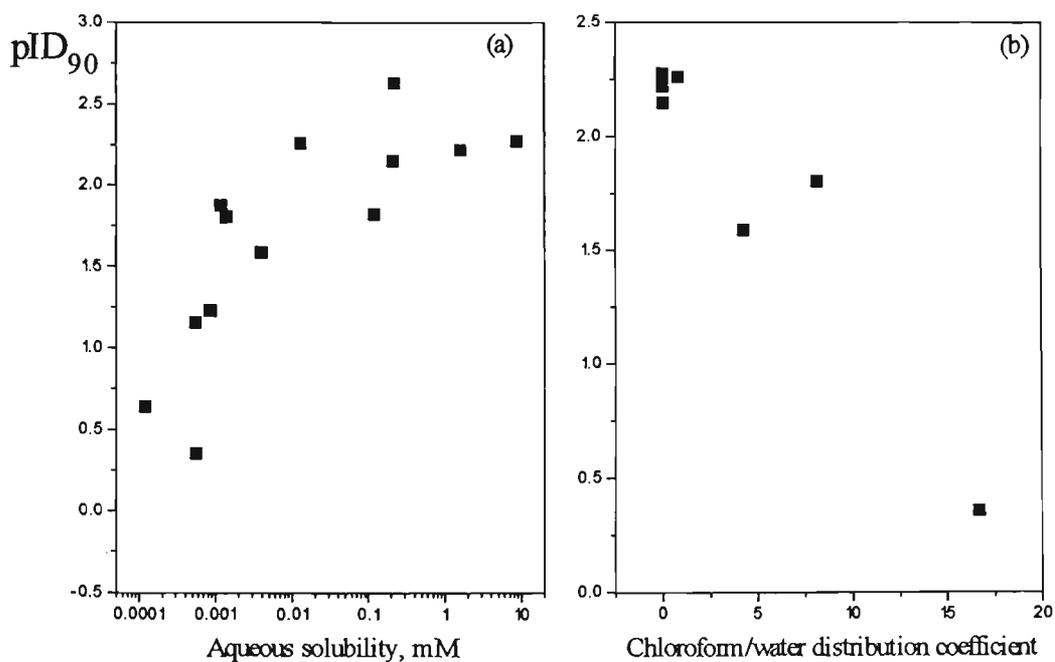
Figure 5.11. Relationships of antitumour potency with bioavailability indicators. (a) Cycles and branches. (b) Cycles only.

Unfortunately, the experimental data on aqueous solubility and/or aqueous/organic distribution for complexes with other ligands is not available, although the calculated ligand *logP* values can be used to show a similar trend (Fig. 5.3b). Once available for binding, complexes with larger ligands may bind quite strongly to DNA as a result of the "flexible bulk" suggested above.

These trends, for pLD_{50} and pID_{90} vs *LRE*, could perhaps be exploited for ligands of these types for rational drug design. A full explanation of such patterns of behaviour is more difficult to address.

The relationship between therapeutic index, *logTI*, and *LRE* is shown in Figs. 5.6 – 5.9 (parts (e)). Plots of *logTI* vs *LRE* do not reveal any obvious pattern of behaviour. However, it is of interest to note that for cycles and branches, the ligand responsible for the maximum *logTI* corresponds to the same ligand responsible for the minimum pLD_{50}. This is not the case for straights and polycycles. In the case of straights it can be rationalised by the observation that they are less sterically diverse as a series (in terms of *LRE*), *vide infra*. Consequently, from a steric point of view, they do not differ significantly from each other and the variation in the therapeutic index cannot be rationalised through steric effects. Generally speaking, changing the length of an alkyl chain which carries the terminal active group, in this case – amine, may not be an adequate choice of structural variation in structure-activity studies, aiming at the investigation of steric effects. In such a series scalar rather than vector steric requirements vary. An example of alternative structural variation, more useful for such studies is chain-branching (presented here) or active site-branching, e.g. series of ligands NH_3 – NH_2R – NHR_2 – NR_3. Biological activity data is available for such a series [89], and would present an interesting development from the present study, particularly in relation to the proposed optimal steric

size for minimal toxicity. Series of platinum complexes with polycyclic amine ligands, studied here, suffer from a lack of experimental data. Availability of biological data for a wider range of polycyclic ligand complexes could probably lead to some rationale of their *logTI* vs *LRE* profiles.

Biological profiles vs *CRE* (Figs. 5.6 – 5.9, parts (b,d,f)) show similar trends to those of *LRE* with some variations. In particular, for cycles the strongest repulsion between $Pt(NH_2R)_2Cl$ and 9-EtG corresponds to minimum toxicity, i.e. optimal steric size of ligand corresponds to minimum toxicity. But this trend is not general (e.g. straights). Presumably, other factors (both steric and non-steric) contribute to toxicity and may predominate in some cases. However, as in the case of *LRE*, optimal values of *CRE* (15 to 20 kcal/mol) for minimum toxicity can be cautiously suggested. The variations between *LRE* and *CRE* trends are to be expected since in $Pt(NH_2R)_2Cl(9\text{-EtG})$ systems the amine ligands are separated from the "steric probe" by platinum. Furthermore, in these systems the repulsive contributions of two amine moieties are embedded into the calculations. And these two moieties are non-equivalent, one being *cis* to the N7 position and the other *trans*, Fig. 4.9.

5.3.7 Overview

LRE and *CRE* parameters as applied to the approach of a metal complex to a target molecule allow for an improved description of steric effects in such systems. Where transport, electronic and steric effects are all to be considered, the *LRE* and *CRE* parameters have a higher degree of orthogonality than those employed previously. Future attempts to separate steric and transport effects and to expose purely a steric requirement may include a consideration of steric effects in complexes containing ligands of like hydrophobicity. Other

directions to investigate the scope of repulsive energy strategies for studying steric effects in biologically active metal complexes may well include quantitative structure property-relationship (QSPR) studies, allowing a focus on kinetic and/or thermodynamic properties. In addition, other chemometric techniques (e.g. neural nets [90]) may be more successful in finding a better QSAR/QSPR model in such systems. The repulsive energy strategy reported here may be extended to other systems for which biological activity information is available, namely mixed and bidentate carrier ligands, other possible targets on DNA, bisadducts and other biologically active metal complexes (e.g. [91]).

5.4 REFERENCES

1. Kubinyi, H. *QSAR: Hansch Analysis and Related Principles;* VCH: Weinheim, **1993**.

2. Hall, L.H.; Kier, L.B. In *Reviews in Computational Chemistry, Vol.2;* Lipkowitz, K.B.; Boyd, D.B. (Eds.); VCH: New York, **1991**, 367-422.

3. Bersuker, I.B.; Dimoglo, A.S. In *Reviews in Computational Chemistry, Vol.2;* Lipkowitz, K.B.; Boyd, D.B. (Eds.); VCH: New York, **1991**, 423-60.

4. Franke, R. *Theoretical Drug Design Methods;* Elsevier: Amsterdam, **1984**.

5. Wold, S.; Sjostrom, M. In *Correlation Analysis in Chemistry;* Chapman, N.B.; Shorter, J. (Eds.); Plenum: New York, **1978**, 1-54.

6. Hansch, C.; Leo, A. *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology;* **1995**.

7. Martin, Y.C. In *Physical Chemical Properties of Drugs;* Yalkowski, S.; Sinkula, A.A.; Valvani, S.C. (Eds.); Marcel Dekker: New York, **1980**, 49-110.

8. Hogberg, T.; Norinder, U. In *A Textbook of Drug Design and Developement;* Krogsgaard-Larsen, P.; Bundgaard, H. (Eds.); Harwood: Chur, **1992**, 55-91.

9. Ford, G.P.; Katritzky, A.R.; Topsom, R.D. In *Correlation Analysis in Chemistry;* Chapman, N.B.; Shorter, J. (Eds.); Plenum: New York, **1978**, 269-312.

10. Tsipis, C.A. *Coord. Chem. Rev.* **1991**, *108*, 163.

11. Hopfinger, A.J.; Cardozo, M.G.; Kawakami, Y. In *Nucleic Acid Targeted Drug Design;* Propst, C.L.; Perun, T.J. (Eds.); Marcel Dekker: New York, **1992**, 151-93.

12. Kim, K.H. In *3D QSAR in Drug Design;* Kubinyi, H. (Ed.); ESCOM: Leiden, **1993**, 619-42.

13. Verhaar, H.J.M.; Eriksson, L.; Sjostrom, M.; Schuurmann, G.; Seinen, W.; Hermens, J.L.M. *Quant. Struct.-Act. Relat.* **1994**, *13*, 133.

14. Venturelli, P.; Menziani, M.C.; Cocchi, M.; Fanelli, F.; De Benedetti, P.G. *J. Mol. Struct. (Theochem)* **1992**, *276*, 327.

15. Clare, B.W. *J. Mol. Struct. (Theochem)* **1995**, *331*, 63.

16. Clare, B.W. *J. Mol. Struct. (Theochem)* **1995**, *337*, 139.

17. Politzer, P; Truhlar, D.G. (Eds.) *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Plenum: New York, **1981**.

18. Ferenzy, G.G.; Reynolds, C.A.; Richards, W.G. *J. Comput. Chem.* **1990**, *11*, 159.

19. Reynolds, C.A.; Ferenzy, G.G.; Richards, W.G. *J. Mol. Struct. (Theochem)* **1992**, *256*, 249.

20. Tasi, G.; Palinko, I.; Nyerges, L.; Fejes, P.; Forster, H. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 296.

21. Kim, K.H. In *Molecular Similarity in Drug Design*; Dean, P.M. (Ed.); Chapman and Hall: London, **1995**, 291-331.

22. De Benedetti, P.G. *J. Mol. Struct. (Theochem)* **1992**, *256*, 231.

23. Kritzenberger, J.; Zimmermann, F.; Wokaun, A. *Inorg. Chim. Acta* **1993**, *210*, 47.

24. Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525.

25. van de Waterbeemd, H. *Programs and Methods for Calculation of logP Values: A Current Overview*; The QSAR and Modelling Society: http://www.pharma.ethz.ch/qsar/logp.html, **1996**.

26. Charton, M. In *Steric Effects in Drug Design*; Charton, M.; Motoc, I. (Eds.); Springer: Berlin, **1983**, 57-91.

27. White, D.; Taverner, B.C.; Leach, P.G.L.; Coville, N.J. *J. Comput. Chem.* **1993**, *14*, 1042.

28. Taft, R.W.J. In *Steric Effects in Organic Chemistry*; Newman, M.S. (Ed.); Wiley: New York, **1956**, 556-675.

29. Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley Interscience: New York, **1979**.

30. Hancock, C.K.; Meyers, E.A.; Yager, B.J. *J. Am. Chem. Soc.* **1961**, *83*, 4211.

31. Pauling, L. *The Nature of the Chemical Bond*; Cornell University: Ithaca, **1960**,

32. Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

33. Allinger, N.L. *Adv. Phys. Org. Chem.* **1976**, *13*, 1.

34. Charton, M. In *Steric Effects in Drug Design*; Charton, M.; Motoc, I. (Eds.); Springer: Berlin, **1983**, 108-18.

35. Fujita, T.; Iwamura, H. In *Steric Effects in Drug Design*; Charton, M.; Motoc, I. (Eds.); Springer: Berlin, **1983**, 119-57.

36. Pearlman, R.S. In *Physical Chemical Properties of Drugs*; Yalkowski, S.; Sinkula, A.A.; Valvani, S.C. (Eds.); Marcel Dekker: New York, **1980**, 321- 347.

37. Charton, M.; Charton, B.I. *J. Org. Chem.* **1979**, *44*, 2284.

38. Leo, A.; Hansch, C.; Jow, P.Y.C. *J. Med. Chem.* **1976**, *19*, 611.

39. Motoc, I. In *Steric Effects in Drug Design*; Charton, M.; Motoc, I. (Eds.); Springer: Berlin, **1983**, 93-105.

40. *Molecular Similarity in Drug Design*; Dean, P.M. (Ed.); Chapman and Hall: London, **1995**.

41. Hopfinger, A.J. *J. Am. Chem. Soc.* **1980**, *102*, 7196.

42. Balaban, A.T.; Chiriac, A.; Motoc, I.; Simon, Z. *Steric Fit in QSAR*; Springer: Berlin, **1980**.

43.   Simon, Z.; Holban, S.; Mracec, M.; Maurer, A.; Policec, S. *Rev. Roumaine Chim.* **1980**, *23*, 713.

44.   Muresan, S.; Bologa, C.; Mracec, M.; Chiriac, A.; Jastorff, B.; Simon, Z.; Naray-Szabo, G. *J. Mol. Struct. (Theochem)* **1995**, *342*, 161.

45.   Fabian, W.M.F.; Simon, Z. *Quant. Struct.-Act. Relat.* **1995**, *14*, 444.

46.   *3D-QSAR in Drug Design;* Kubinyi, H. (Ed.); ESCOM: Leiden, **1993**.

47.   Hansch, C.; Unger, S.H. *J. Med. Chem.* **1973**, *16*, 745.

48.   Srivastava, M.; Sen, A. *Regression Analysis: Theory, Methods, and Applications;* Springer: New York, **1990**.

49.   Topliss, J.G.; Costello, R.J. *J. Med. Chem.* **1972**, *15*, 1068.

50.   Meloun, M.; Militky, J.; Forina, M. *Chemometrics for Analytical Chemistry. Vol.1: PC Aided Statistical Data Analysis;* Ellis Horwood: New York, **1992**.

51.   Cramer, R.D.I. *Perspectives in Drug Discovery and Design* **1993**, *1*, 269.

52.   Franke, R.; Dove, S.; Gruska, A. In *Advanced Drug Design and Development. A Medicinal Chemistry Approach;* Kourounakis, P.N.; Rekka, E. (Eds.); Ellis Horwood: New York, **1994**, 43-97.

53.   Braddock, P.D.; Connors, T.A.; Jones, M.; Khokhar, A.R.; Melzack, D.H.; Tobe, M.L. *Chem.-Biol. Interactions* **1975**, *11*, 145.

54.   ChemPlus™: Extensions for HyperChem^R Version 1. Canada: Hypercube, Inc. **1993**.

55.   Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.

56.   Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

57.   Stewart, J.J.P. *J. Comput.-Aided Mol. Design* **1990**, *4*, 1.

58. Klein, C.L.; Stevens, E.D. In *Structure and Reactivity;* Liebman, J.F.; Greenberg, A. (Eds.); VCH: New York, **1988**, 25-64.

59. Fleming, I. *Frontier Orbitals and Organic Chemistry Reactions;* Wiley: Chichester, **1989**.

60. Rogers, D.W. In *Computational Chemistry Using the PC;* VCH: **1994**, 221-36.

61. Perrin, D.D. *Dissociation Constants of Organic Bases in Aqueous Solutions;* Butterworths: London, **1965**,

62. Martell, A.E.; Smith, R.M. *Critical Stability Constants, V.1-6;* Plenum: New York, **1989**.

63. Perrin, D.D.; Dempsey, B.; Serjeant, E.P. *pK$_a$ Prediction for Organic Acids and Bases;* Chapman and Hall: London, **1981**.

64. Ghose, A.K.; Pritchett, A.; Crippen, G.M. *J. Comput. Chem.* **1988**, *9,* 80.

65. Souchard, J.-P.; Ha, T.T.B.; Cros, S.; Johnson, N.P. *J. Med. Chem.* **1991**, *34,* 863.

66. Nie, N.H.; Bent, D.H.; Hull, C.H. *Statistical Package for the Social Sciences;* McGraw-Hill: New York, **1970**.

67. SCAN, Software for Chemometric Analysis Version 1. Minitab Inc. **1995.**

68. Edwards, A.L. *An Introduction to Linear Regression and Correlation;* W.H.Freeman & Co: San Francisco, **1976**.

69. Hansch, C. In *Correlation Analysis in Chemistry;* Chapman, N.B.; Shorter, J. (Eds.); Plenum: New York, **1978**, 397-438.

70. Plummer, E.L. In *Reviews in Computational Chemistry, Vol.1;* Lipkowitz, K.B.; Boyd, D.B. (Eds.); VCH: New York, **1990**, 119-68.

71. Abdul-Ahad, P.G.; Webb, G.A. *Int. J. Quantum Chem.* **1982**, *21,* 1105.

72. Simon, Z.; Maurer, A.; Mracec, M.; Policec, S.; Dragulescu, C. *Rev. Roumaine Biochim.* **1977**, *14*, 117.

73. Romanowska, K.; Kuduk-Jaworska, J. *Archiv. Immun. Therap. Exper.* **1991**, *39*, 75.

74. Romanowska, K. *Int. J. Quantum Chem.* **1992**, *43*, 175.

75. Tang, W.; Qu, Y.; Dai, A. *Pure Appl. Chem.* **1988**, *60*, 1271.

76. Dimoglo, A.S.; Choban, I.N.; Chumakov, Y.M.; Bersuker, I.B. *Khim.-Farm. Zh.* **1982**, *16*, 956.

77. Kiss, F.; Tvaruzek, P.; Hladka, E. In *Fifth International Symposium "Platinum and Other Metal Coordination Compounds in Cancer Chemotherapy"*; Nicolini, M.; Bandoli, G. (Eds.); Cleup Padua: Padua, **1988**, 369.

78. Hansch, C.; Venger, B.H.; Panthananickal, A. *J. Med. Chem.* **1980**, *23*, 459.

79. Hansch, C. *Acc. Chem. Res.* **1993**, *26*, 147.

80. Hansch, C.; Leo, A.; Hoekman, D.; Li, P. *A Comprehensive Approach to Structure-Activity Relationships*; Pomona College: Claremont, **1995.**

81. Quinn, F.R.; Milne, G.W.A. *Fund. and Appl. Toxicol.* **1986**, *6*, 270.

82. Panthananickal, A. *NCI Data, source a.panthananickal bio_1865*; **1986.**

83. Boudreaux, E.A. In *Metal-Based Anti-tumour Drugs*; Gielen, M.F. (Ed.); Freund: London, **1988**, 175-200.

84. Ghose, A.K.; Crippen, G.M. *J. Med. Chem.* **1985**, *28*, 333.

85. Maliski, E.G.; Bradshaw, J. In *Medicinal Chemistry: The Role of Organic Chemistry in Drug Research*; Ganellin, C.R.; Roberts, S.M. (Eds.); Academic Press: London, **1993**, 83-102.

86. Belsley, D.A.; Kuh, E.; Welsch, R.E. *Regression Diagnostics: Identifying Influencial Data and Sources of Collinearity*; Wiley: New York, **1988**.

87. Silverman, R.B. *The Organic Chemistry of Drug Design and Drug Action;* Academic Press: San Diego, **1992,**

88. Reedijk, J. *Inorg. Chim. Acta* **1992,** *198-200,* 873.

89. Beaumont, K.P.; McAuliffe, C.A.; Cleare, M.J. *Chem.-Biol. Interactions* **1976,** *14,* 179.

90. Burden, F.R. *Quant. Struct.-Act. Relat.* **1996,** *15,* 7.

91. Kuo, L.Y.; Kanatzidis, M.G.; Sabat, M.; Tipton, A.L.; Marks, T.J. *J. Am. Chem. Soc.* **1991,** *113,* 9027.

Chapter 6

Summary

Although this thesis focuses on the interaction of platinum complexes with nucleic acids and their constituents, a broader objective has also been to delineate steric effects in metal species in general with a view to relating these to their biological properties. In this regard an established repulsive energy methodology has been applied and extended to bioinorganic models and the notion of structural speciation has been introduced. Aspects of the structural speciation of metal complexes which have steric implications have been investigated both from experimental (case study and ratio-dependence studies) and theoretical (repulsive energy methodology) viewpoints. A number of unique molecular features which may be important for biological activity and for the understanding of the mechanism of action of biologically active compounds have been identified. The aim of this general discussion is to conceptually summarise these findings.

6.1 Carrier ligand steric requirements

The following properties of carrier ligands have been demonstrated to exert a specific steric influence.

Localised bulk. Carrier ligands may be judiciously designed to exploit this feature. In the presented case study, where this has been applied, evidence has been put forward for a unique phenomenon whereby metal coordination to nucleoconstituents may be precluded beyond a critical molar ratio. This has led to the suggestion of a "switching" event under stoichiometric control, which could have important biological implications (section 2.3.1).

Optimal bulk. Evidence have been presented which suggests that steric effects in the vicinity of the coordination metal, relating to the carrier

ligand, may be optimised with respect to certain biological outcomes. Thus an optimal steric requirement for minimal toxicity has been identified for several series of carrier ligands, coordinated to platinum (section 5.3.5).

Flexible bulk. For the series of ligands studied in this thesis an analysis of corresponding published biological data demonstrates that platinum complexes of smaller and larger (in scalar terms, e.g. MV) ligands within a given series show higher cytotoxicity and lower antitumour potency. In terms of steric effects it may be pointed out that both the smaller and the larger molecules might be expected to display minimal steric requirements with respect to adduct formation with DNA. Although this might be obvious for the smaller species, for the larger ones this may be achieved by rotation around the Pt-N bond or adopting a folded-back or any other accommodating conformation, minimising steric effects (section 5.3.5). This concept is related to the finding of optimal bulk, presented above.

Metal complex flatness. It has been shown that the flatness of a metal complex is related to its steric requirements and affects both site selectivity and the conformational flexibility of an adduct. The importance of this criterion has been suggested, in particular with respect to binding to sterically obstructed binding sites on biomolecular targets (section 4.3.2.2) [1].

It must be emphasised that the carrier ligand represents only a subset of the overall metal complex speciation and that other aspects, such as metal type and its oxidation state, the leaving group(s), coordination number and geometry, can also have steric connotations (section 1.3).

6.2 Binding site steric requirements

With respect to steric features of biomolecular binding sites it has been shown that:

- the small differences in steric demands of nucleobase binding sites resulting from the influence of exocyclic functional groups can be quantified and the generated parameters can be used to rationalise the metal binding preference order in steric terms only (section 4.3.2.1).

- the binding of metal complexes to the N3 site of purine nucleobases, which is severely sterically hindered by substitution at the N9 position, is less improbable than generally considered. The possibility of such binding exists for metal complexes with certain degree of flatness and binding to this position may have important biological consequences [2] (section 4.3.2.2).

6.3 Steric features of an adduct

To conclude this summary, the features of a consolidated adduct may be outlined as follows. It has been shown that the conformational aspects of a metal complex adduct with a nucleobase are related to steric demands exerted by both carrier ligand(s) and exocyclic functional groups in the vicinity of the target site. The NB/PtN$_4$ dihedral angles have been demonstrated to be a representative parameter of the adduct conformation with respect to steric effects. Although, while using these dihedral angles to interpret steric interactions it should be realised that such dihedral angles may also be influenced by other forces [3]. For example, the Pt...O axial

binding may favour large dihedral angles (~90°), maximising the overlap between platinum orbitals and lone pair density on oxygen. On the other hand, interligand nucleobase hydrogen bonding favours small dihedral angles (~50–60°) in order to minimise the distance between donor and acceptor atoms. Crystal packing forces (intercomplex, solvent or counter-ion derived) may also result in considerable variability in more flexible systems.

The NB/PtN_4 dihedral angles represent a conformational measure of adduct geometry in the immediate vicinity of the binding site. However, it is clear that steric interactions localised in the (carrier ligand)-Pt-(nucleobase) structural unit may affect the remote geometric features, although such a "steric vs conformational" relationship is less easy to elucidate than that within the unit. However, it should be understood that the dependency may exist between the two structural levels. Indeed, the extent to which a local distortion is transmitted to the wider molecular environment and *vice versa* is an important consideration in rational design.

6.4 References

1.   Yuriev, E.; Orbell, J.D. to be submitted to *J. Am. Chem. Soc.*

2.   Yuriev, E.; Orbell, J.D. to be submitted to *Inorg. Chem.*

3.   Orbell, J.D.; Wilkowski, K.; Marzilli, L.G.; Kistenmacher, T.J. *Inorg. Chem.* **1982**, *21*, 3478.

Sterically restrictive metal complexes. The synthesis and structural characterization of (1,2-bis(6-methylpyridin-2-yl)ethane-N,N')-(malonato)palladium(II) trihydrate. Proton NMR binding studies to cytosine and guanine derivatives

Anthony T. Baker* and Jeffrey K. Crass

Department of Chemistry, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007 (Australia)

Gaik B. Kok

Department of Pharmaceutical Chemistry, Victorian College of Pharmacy (Monash University), 381 Royal Parade, Parkville, Vic. 3053 (Australia)

John D. Orbell* and Elizabeth Yuriev

Department of Environmental Management, Victoria University of Technology, McKechnie Street, St. Albans, Vic. 3021 (Australia)

Abstract

Two different routes for the preparation of the bidentate ligand 1,2-bis(6-methylpyridin-2-yl)ethane (bmpe) have been explored and are described. 1:1 coordination of this ligand with palladium(II) and the isolation of crystals of the complex as the malonate trihydrate has allowed the crystal and molecular structure to be determined. The compound crystallizes in space group $P2_1/n$ with $a = 12.946(1)$, $b = 7.431(1)$, $c = 20.483(2)$ Å, $\beta = 91.941(3)°$, $Z = 4$. The structural analysis reveals that the 6-methyl groups of the bidentate bmpe ligand both project over one side of the square-plane, effectively blocking it off. The interaction of the Pd(bmpe)$^{2+}$ complex cation with cytosine and guanine derivatives has been investigated by proton NMR spectroscopy.

Introduction

The nature of the binding of a metal complex to a nucleic acid may be strongly influenced by steric factors. Structural studies on model systems for the interaction of the cisplatinum antitumour drugs with DNA demonstrate that, in some cases, intramolecular steric effects may even be determinative of the molecular conformation of the adduct [1]; although difficulties remain in deducing to what extent correlations arising from model systems can be carried over with fidelity to a metal–nucleic acid complex.

Recent studies on the mechanism of the cytotoxicity of the anticancer compound *cis*-diamminedichloroplatinum(II) (cisplatin) and its active analogs suggest that their mode of action may be related to the recognition by a key cellular protein of a structural motif on the DNA induced by metal coordination [2]. Uniquely distorted DNA and RNA conformations have also been implicated as being important in ribozymes [3] and in

genetic signalling [4]. Thus the manipulation of the steric parameters involved in the binding of metal species to nucleic acids may hold promise for the rational design of distortions in nucleic acids and may provide a means for the control of various biological functions related to nucleic acid topology.

For the cisplatinum systems, such steric demands are dictated by the bulk on the coordinated amine and by the DNA-base exocyclic groups which are contiguous to the metal binding site [5]. Thus there is scope for influencing the molecular conformation of an adduct by a judicious design of the metal complex via the coordinated ammine or other appropriate 'carrier ligands' [6]. In addition, since the aforementioned exocyclic functional groups vary with DNA base type, there is also the possibility of manipulating site specificity. This may be brought about, for example, by selectively precluding coordination at a particular base site. Such reagents would have obvious applications as biological probes [7].

An important mode of coordination for cisplatinum derivatives to DNA is an intrastrand crosslinkage be-

---

*Authors to whom correspondence should be addressed.

tween the N(7) positions of adjacent guanines. Up to 60% of cisplatin DNA adducts may be of this type [8]. Crystal structure analyses of systems modelling the guanine[N(7)]–Pt–guanine[N(7)] intrastrand crosslinkage reveal that in the majority of cases the bases are arranged in the head-to-tail (HT) conformation [5]. In only four cases has a head-to-head (HH) conformation be found [9], and these are of a fortuitous nature. An HH conformation has also been demonstrated in several complexes involving di- and trinucleotides [10, 11]. In the latter cases the HH conformation is dictated by the polymeric nature of the nucleotide ligand. The predominance of the HT conformation in the crystal structures of simple model systems has allowed the flexibility of this adduct to be systematically examined [12]; although the biological relevance of this geometry is not considered by some researchers to be as important as HH; this being the orientation of the bases in the nucleic acids themselves (albeit prior to any disruptive influence). A specific goal of our research is to develop metal complexes which sterically enforce head-to-head (HH) conformations of cis bis-coordinated 6-oxopurine bases, nucleosides and nucleotides. This would then allow the flexibility of the HH adduct to be explored further and to be manipulated by rational design of the carrier ligand.

It has been realized for some time that a coordinated ligand which selectively blocks off one side of a square-plane, has the potential for enforcing an HH geometry of cis bis-coordinated 6-oxo purines. Reedijk and co-workers [13] employed the ligand 1,2-bis(pyridin-2-yl)ethane, bpe, Fig. 1(a), in an attempt to achieve this goal. It was confirmed by these workers via an X-ray analysis of the complex Pt(bpe)(9-MeHX)$_2$(NO$_3$)$_2$, 9-MeHX = 9-methylhypoxanthine, that even though the bpe ligand provides different environments above and below the platinum coordination plane (by pitching the hydrogen atoms *ortho* to the pyridinyl nitrogens over one side of the coordination plane whilst maintaining the rigid ethylene bridge on the other side) it does
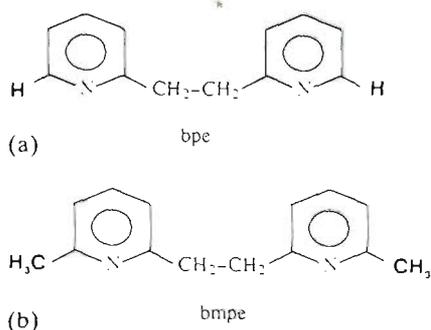
not influence the N(7)-coordinated 9-MeHX ligands enough to force them into an HH arrangement. In order to further investigate the possibility of achieving an enforced HH configuration and influencing its geometry sterically (by further ligand modification) we have synthesized, by two different routes, the ligand 1,2-bis(6-methylpyridin-2-yl)ethane, bmpe, Fig. 1(b). When coordinated to a square-planar metal centre, the presence in bmpe of the bulky methyl groups adjacent to the pyridinyl donor nitrogens compared to just hydrogen atoms in bpe, should result in an even greater obstruction of one side of the plane in bmpe compounds compared to bpe. This additional bulk might be sufficient to enforce the required HH coordination upon subsequent reaction of the M(bmpe)$^{2+}$ moiety with two equivalents of an N(7)-coordinated 6-oxopurine ligand. Thus we have synthesized and characterized by X-ray analysis the 1:1 complex of the bidentate bmpe ligand with palladium(II) as the malonate trihydrate and we have investigated, both by attempting to isolate complexes and by proton NMR spectroscopy, the coordinating ability of the Pd(bmpe)$^{2+}$ moiety towards 9-ethylguanine (9-EtG), guanosine (Guo), the disodium salt of guanosine 5'-monophosphate (GMPNa$_2$) and 1-methyl cytosine (1-MeC).

Experimental

K$_2$PdCl$_4$ was supplied by Aldrich; 9-EtG, guanosine, GMPNa$_2$ and 1-MeC were supplied by Sigma. The ligand bmpe was synthesized in our laboratories by procedures described in this paper. Common chemicals were obtained from other scientific supply houses.

Synthesis of 1,2-bis(6-methylpyridin-2-yl)ethane, bmpe

Method 1. This method followed a procedure similar to the preparation of 1,2-bis(pyridin-2-yl)ethane (bpe) reported by Campbell and Teague [14]. To a solution of 2,6-lutidine (29 ml; 0.25 mol) in dry THF (120 ml) at −30 °C was added n-butyl lithium (100 ml of a 0.25 M solution in hexane; 0.25 mol). The resulting bright orange–red solution was then left to stir for 1 h while the temperature rose to ambient. The solution was stirred rapidly while bromine (22 g; 0.138 mol) was added dropwise over 1 h at −40 °C. After stirring for an additional 1 h, the mixture was treated with water (50 ml) and 6 N HCl (50 ml). The two layers were separated, the aqueous layer made alkaline with NaOH and extracted with chloroform (3 × 20 ml). The combined chloroform extracts were then dried. The residue, after solvent removal, was distilled. The desired product (b.p. 124–128 °C/5 mm Hg) was obtained as a colourless liquid which solidified on standing to an off-white solid, 9.6 g (36% yield). The identity and purity of the product

Fig. 1. The ligands 1,2-bis(pyridine-2-yl)ethane, bpe (a) and 1,2-bis(6-methylpyridin-2-yl)ethane, bmpe (b). The protons and the methyl substituents *ortho* to the ring nitrogens in bpe and bmpe, respectively, are highlighted.

was confirmed by proton NMR and by gas chromatography/mass spectroscopy.

Method 2. This method is a slight modification of the procedure of Bass *et al.* [15]. A solution of t-butyl peroxide (6 g) in 2,6-lutidine (100 g), under nitrogen, was heated under reflux in an oil bath for 72 h. The low boiling point products and the excess 2,6-lutidine were removed by distillation through a fractionating column. The residue was distilled under reduced pressure to give 4.66 g (53% yield) of an off-white solid. This material was recrystallized from absolute light petroleum (40–60 °C) and its identity and purity confirmed by proton NMR and FT-IR. The crystals melted over the range 49–50 °C; literature value, 50–54 °C.

Preparation of Pd(bmpe)Cl₂

Method 1. Pd(bmpe)Cl₂ was prepared by reacting [Pd(benzonitrile)₂Cl₂] with bmpe in acetonitrile in a reaction analogous to that used to prepare several Pd(II) diimine complexes [16]. *Anal.* Found: C, 42.97; H, 4.35; N, 7.14; Pd, 27.1. Calc. for $C_{14}H_{16}Cl_2PdN_2$: C, 43.16; H, 4.14; N, 7.19; Pd, 27.3%.

Method 2. To a solution of 0.4 g of K₂PtCl₄ (1.2 mmol) in 5 ml of deionized water was added a solution of 0.26 g (1.2 mmol) of bmpe in 2 ml of acetone. A yellow precipitate formed immediately. The resulting mixture was stirred overnight, the precipitate filtered, and washed with water, ice-cold ethanol and ether, and dried in a dessicator over silica gel. The yield was 0.44 g (92%). *Anal.* Found: C, 43.0; H, 4.3; N, 7.3. Calc.: as given in Method 1.

Preparation of Pd(bmpe)(malonato)·3H₂O

To a stirred suspension of Pd(bmpe)Cl₂ in a dark bottle was added 1.95 equiv. of AgNO₃ as a concentrated solution in water. The mixture was stirred at a temperature of 60 °C for 2 h. The AgCl was filtered off over celite giving a pale yellow filtrate which is assumed to contain the diaquo species. To this solution was added a small excess of sodium malonate in a minimum amount of water. A slightly lighter yellow solution resulted which was reduced in volume at 70 °C and allowed to cool slowly. Yellow needles suitable for X-ray analysis were harvested and air dried. Drying over silica gel *in vacuo* caused the crystals to deteriorate; this is reflected in the analytical figures which correspond to the anhydrous compound. *Anal.* Found: C, 48.2; H, 4.7; N, 6.9. Calc. for $C_{17}H_{18}PdN_2O_4$: C, 48.5; H, 4.3; N, 6.7%.

Crystallography

Crystallographic data for [Pd(bmpe)(mal)]·3H₂O are given in Table 1. The diffraction data were collected on a Huber four-circle diffractometer in a $\theta/2\theta$ scan mode using graphite-crystal monochromated Mo Kα

TABLE 1. Crystallographic data for Pd(bmpe)(malonato)·3H₂O

Formula	$C_{17}H_{24}N_2O_7Pd$
Formula weight	474.79
a (Å)	12.946(1)
b (Å)	7.431(1)
c (Å)	20.483(2)
β (°)	91.941(3)
V (Å³)	1969.4
Z	4
Space group	$P2_1/n$ (No. 14)
T (°C)	22
λ (Å)	0.71069
ρ_{calc} (g/cm³)	1.60
μ(Mo Kα) (cm⁻¹)	9.66
Transmission coefficient	0.826–0.872
R^a	0.027
R_w^b	0.038

$^aR = \Sigma\|F_o| - |F_c\|/\Sigma|F_o|.$ $^bR_w = (\Sigma\|F_o| - |F_c\|^2/\Sigma w|F_o|^2)^{1/2}.$

radiation. The unit cell parameters were obtained from least-squares refinement on the setting angles of 37 reflections ($10 < 2\theta < 20°$). Of 5548 reflections ($h, k, \pm l$) collected to $2\theta_{max}$ of 57°, 4215 were considered observed with $I \geqslant 3\sigma(I)$. The space group $P2_1/n$ was assigned on the basis of systematic absences ($0k0$, k odd absent, and $h0l$, $h + l$ odd absent). The structure was solved by direct phasing and Fourier methods. Ligand hydrogen atoms were included in calculated positions after confirming their presence in difference Fourier maps. In particular, the geometry of the methyl groups was checked to confirm they were not disordered. The hydrogen atoms of the water molecules were located in a difference Fourier map. All hydrogen atoms were assigned fixed thermal parameters and ligand hydrogen atom positions were updated after each cycle of refinement to maintain geometry. Positional and anisotropic thermal parameters for the non-hydrogen atoms were refined using full-matrix least-squares. Atomic scattering factors and anomalous dispersion parameters were from International Tables for X-ray Crystallography. Crystallographic programs included in the UCLA crystallographic computing package include modified versions of the following: REDUCE (Coppens, Becker, Blessing and Broach), peak profile analysis, Lorenz and polarization corrections; MULTAN (Main), direct methods, Fourier analysis and map searching; ORFLS (Busing, Martin and Levy), structure factor calculations and least-squares refinement; ORFFE (Busing, Martin and Levy), distances, angles and error calculations; ABSORB (Coppens, Edwards and Hamilton), absorption corrections; ORTEP-II (Johnson), figure plotting; HYDROGEN (Trueblood) calculation of hydrogen atom positions. All calculations were performed using DEC VAX computers. Fractional coordinates of the non-hydrogen atoms are given in Table 2.

TABLE 2. Atomic coordinates and equivalent isotropic temperature factors for Pd(bmpe)(malonato)·3H₂O

Atom	x	y	z	$U^a \times 10^4$
Pd	0.23946(1)	0.17648(2)	0.56415(1)	284(1)
O(1)	0.0965(1)	0.0965(3)	0.5848(1)	386(9)
O(2)	0.3016(1)	-0.0244(3)	0.6174(1)	436(10)
O(3)	-0.0100(2)	-0.0425(3)	0.6482(1)	590(13)
O(4)	0.3142(2)	-0.1662(4)	0.7107(1)	805(18)
O(5)	-0.1326(2)	0.0136(3)	0.7533(1)	660(15)
O(6)	0.5318(2)	-0.1384(4)	0.7231(2)	1020(23)
O(7)	0.6735(2)	0.1427(4)	0.7149(2)	884(20)
N(1)	0.3827(1)	0.2367(3)	0.5355(1)	308(10)
N(2)	0.1841(1)	0.3859(3)	0.5094(1)	335(10)
C(1)	0.4041(2)	0.1890(3)	0.4734(1)	373(13)
C(2)	0.5032(2)	0.2090(4)	0.4505(1)	482(16)
C(3)	0.5806(2)	0.2777(4)	0.4920(2)	540(18)
C(4)	0.5571(2)	0.3251(4)	0.5548(2)	485(16)
C(5)	0.4574(2)	0.3042(3)	0.5759(1)	380(13)
C(6)	0.4287(2)	0.3577(5)	0.6440(1)	546(18)
C(7)	0.3158(2)	0.1065(4)	0.4335(1)	493(16)
C(8)	0.2367(3)	0.2364(5)	0.4059(1)	564(18)
C(9)	0.1947(2)	0.3916(4)	0.4438(1)	409(13)
C(10)	0.1635(2)	0.5438(5)	0.4091(2)	582(19)
C(11)	0.1194(3)	0.6865(4)	0.4395(2)	630(21)
C(12)	0.1059(2)	0.6776(4)	0.5057(2)	562(19)
C(13)	0.1402(2)	0.5276(4)	0.5406(1)	423(14)
C(14)	0.1306(2)	0.5206(4)	0.6135(2)	564(18)
C(15)	0.0767(2)	0.0176(3)	0.6385(1)	383(13)
C(16)	0.1604(2)	-0.0006(4)	0.6918(1)	472(15)
C(17)	0.2656(2)	-0.0701(4)	0.6726(1)	433(15)

$U_{eq} = [1/(6\pi^2)]\Sigma\Sigma\beta_{ij}a_ia_j$.

Description of the structure

The Pd(bmpe)(malonato) complex is shown, together with the labelling of the atoms, in Fig. 2. The coordination geometry about Pd(II) is square-planar within experimental error; the four equatorial positions are occupied by the nitrogen atoms of the bidentate bmpe ligand (labelled N(1) and N(2)) and by the oxygen donors of the bidentate malonate anion (labelled O(1) and O(2)). The bmpe ligand forms a boat-like seven-membered chelate ring and the two pyridine moieties of this ligand are twisted from the square-plane with angles of 66.2 and 73.0°. These angles compare well with 67.9° for both pyridine moieties in Pt(dmdap)(bpe)Cl₂·H₂O [17] and the 62.5 and 69.4° reported for the pyridine moieties in the complex Pt(bpe)(9-MeHX)₂(NO₃)₂ [13] (the asymmetry observed in the latter may be attributed to the bulk of the two coordinated 9-methylhypoxanthine ligands). The similarity in these angles between the coordinated bpe and bmpe systems reflects the rigidity of ligands of this type. Consistent with the observations of Reedijk and co-workers for the bpe system [13], we also observe distinct chemical shifts for the ethylene bridge protons in our NMR studies of the Pd(bmpe)²⁺ system in
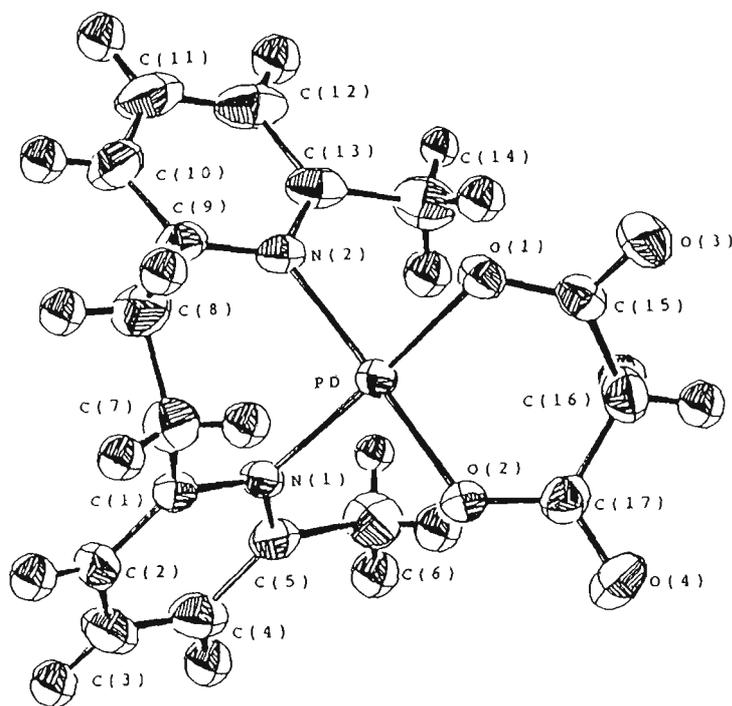
Fig. 2. View of the Pd(bmpe)(mal) molecule perpendicular to the coordination plane together with the atom numbering scheme. Anisotropic ellipsoids represent 50% probability boundaries. H atoms are represented as spheres of arbitrary radii.

solution. This is further evidence for the rigidity of the seven-membered chelate ring.

In the title compound, the bidentate coordination of the malonate to the palladium may be compared with its coordination to other metals [18]. Generally, the coordinated malonate ligand displays a high degree of conformational flexibility and can adopt a variety of conformations including chair, boat, envelope, or flattened with distortions towards 'skewboat' configurations. In the present case the palladium–malonate ring exhibits a boat conformation. Bond lengths and angles are consistent with those observed in other malonato complexes [18].

An interesting feature of the structure is the intramolecular approach of one of the ethylene bridge protons to the palladium atom, Fig. 3(a) and (b). The H(7a) to Pd distance of 2.436(5) Å (Pd–H(7a)–C(7) angle 109.0(3)°) is considerably shorter than the sum of the van der Waals radii $(r_H + r_{Pd} = 1.2 + 1.9 = 3.1$ Å) [19], and it is tempting to assume a weak attractive interaction. Agostic intramolecular interactions of the kind C–H–Pd have been reported in a number of structures [20–22]. Those representing a 'strong' interaction have H to Pd distances under 2 Å (e.g. 1.874 Å [23]), whereas the 'weaker' interactions are between 2.5 and 3.0 Å (e.g. 2.57 and 2.84 Å [20], 2.84 Å [21], 2.98 Å [22]). Only one structure [22] has a contact involving a methylene bridge, the other structures [20, 21, 23] involve α-hydrogens of phenyl rings. The presence of such an attractive interaction in the title complex
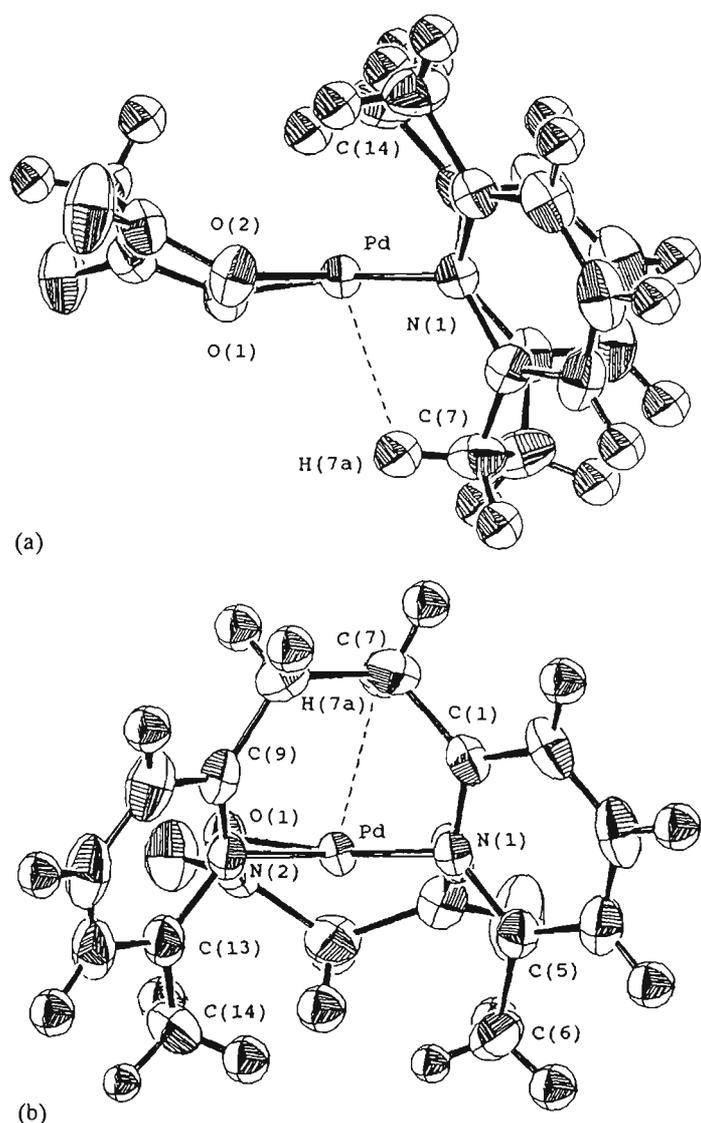
Fig. 3. (a) A side view and (b) an end view of the Pd(bmpe)(mal) molecule.

could contribute to the asymmetry in the dihedral angles between the pyridine moieties and the coordination plane (66.2 and 73.0°).

Selected bond distances and angles for the complex are given in Table 3. Bond lengths and angles associated with the coordination plane are typical of those reported in related systems. We do note however, a slight downward distortion of the Pd–O(1) bond away from the bulk of the C(14) methyl substituent, Fig. 3(a) and (b). This is reflected in the N(1)–Pd–O(1) angle of 173.7(7)°. The influence of the bulk of the C(14) methyl substituent is also reflected in the angles at N(2) which are consistent with a 'pushing-down' effect on the Pd–O(1) bond. A compendium of the angles around both N(1) and N(2), and a comparison with the corresponding angles in Pt(bpe)(9-MeHX)₂(NO₃)₂ are given Table 4. Other bond lengths and angles in the complex have expected values within experimental error.

Side and end views of the complex, which are presented in Fig. 3(a) and (b), respectively, clearly show

TABLE 3. Molecular geometry for Pd(bmpe)(malonato)·3H₂O: bond lengths (Å), bond angles (°)

Primary coordination sphere about the Pd atom			
Pd–N(1)	2.013(2)	Pd–O(1)	2.003(2)
Pd–N(2)	2.034(2)	Pd–O(2)	2.002(2)
N(1)–Pd–N(2)	88.90(8)	N(2)–Pd–O(2)	176.96(7)
N(1)–Pd–O(2)	88.14(8)	N(2)–Pd–O(1)	91.86(8)
N(1)–Pd–O(1)	173.76(7)	O(1)–Pd–O(2)	91.16(7)

1,2-Bis(6-methylpyridin-2-yl)ethane ligand			
N(1)–C(5)	1.348(3)	C(5)–C(6)	1.509(4)
N(1)–C(1)	1.358(3)	C(7)–C(8)	1.504(4)
N(2)–C(9)	1.356(3)	C(8)–C(9)	1.502(4)
N(2)–C(13)	1.366(3)	C(9)–C(10)	1.389(4)
C(1)–C(2)	1.389(4)	C(10)–C(11)	1.365(5)
C(1)–C(7)	1.512(4)	C(11)–C(12)	1.375(5)
C(2)–C(3)	1.389(5)	C(12)–C(13)	1.389(4)
C(3)–C(4)	1.378(5)	C(13)–C(14)	1.503(4)
C(4)–C(5)	1.384(4)		
C(5)–N(1)–C(1)	120.48(22)	N(1)–C(5)–C(6)	118.13(23)
C(5)–N(1)–Pd	123.54(17)	C(4)–C(5)–C(6)	121.38(26)
C(1)–N(1)–Pd	115.70(17)	C(8)–C(7)–C(1)	115.83(26)
C(9)–N(2)–C(13)	119.83(23)	C(9)–C(8)–C(7)	123.53(23)
C(9)–N(2)–Pd	121.61(17)	N(2)–C(9)–C(10)	119.69(27)
C(13)–N(2)–Pd	118.47(18)	N(2)–C(9)–C(8)	122.59(24)
N(1)–C(1)–C(2)	120.56(26)	C(10)–C(9)–C(8)	117.70(27)
N(1)–C(1)–C(7)	115.94(22)	C(11)–C(10)–C(9)	121.22(30)
C(2)–C(1)–C(7)	123.43(25)	C(10)–C(11)–C(12)	118.81(27)
C(1)–C(2)–C(3)	119.22(26)	C(11)–C(12)–C(13)	119.81(28)
C(4)–C(3)–C(2)	119.20(26)	N(2)–C(13)–C(12)	120.59(27)
C(3)–C(4)–C(5)	120.05(28)	N(2)–C(13)–C(14)	119.20(24)
N(1)–C(5)–C(4)	120.49(26)	C(12)–C(13)–C(14)	120.21(27)

The malonato ligand			
O(1)–C(15)	1.280(3)	O(4)–C(17)	1.217(3)
O(2)–C(17)	1.283(3)	C(15)–C(16)	1.518(4)
O(3)–C(15)	1.230(3)	C(16)–C(17)	1.521(4)
C(15)–O(1)–Pd	122.08(16)	C(15)–C(16)–C(17)	117.91(22)
C(17)–O(2)–Pd	121.79(17)	O(4)–C(17)–O(2)	121.66(28)
O(3)–C(15)–O(1)	121.02(25)	O(4)–C(17)–C(16)	118.78(26)
O(3)–C(15)–C(16)	118.97(24)	O(2)–C(17)–C(16)	119.55(23)
O(1)–C(15)–C(16)	120.00(22)		

TABLE 4. A compendium of the angles (°) around N(1) and N(2) in Pd(bmpe)(mal)·3H₂O and the corresponding angles in Pt(bpe)(9-MeHX)₂(NO₃)₂

Angle	Pd(bmpe)-(mal)·3H₂O	Pt(bpe)-(9-MeHX)₂(NO₃)₂
Pd–N(1)–C(5)	123.54(17)	120.2(8)
Pd–N(1)–C(1)	115.70(17)	119.7(8)
Pd–N(2)–C(13)	118.47(18)	115.1(7)
Pd–N(2)–C(9)	121.61(17)	124.8(8)

both of the methyl substituents of the bmpe ligand projecting to approximately the same extent over one side of the coordination plane, effectively blocking it off. Thus, as anticipated, the methyl substituents *ortho* to the pyridinyl nitrogens of the bmpe ligand project preferentially over one side of the coordination plane
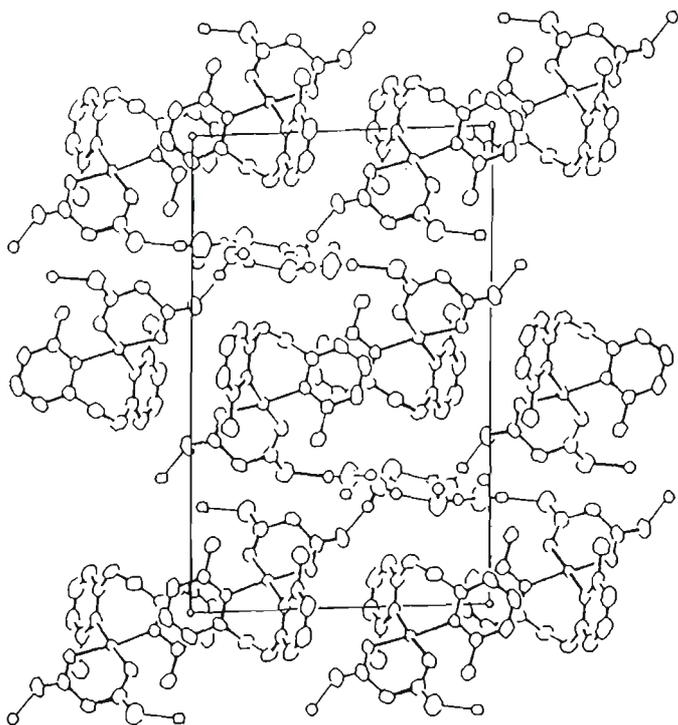
Fig. 4. The unit cell contents for Pd(bmpe)(mal)·3H₂O (a, to the right; b, to the viewer).

TABLE 5. hydrogen bonding distances (Å) within Pd(bmpe)-(malonato)·3H₂O

-O-H---	---O-	Distance
O(5)	O(3)	2.749(3)
O(6)	O(4)	2.828(4)
O(7)	O(6)	2.789(5)
O(5)ᵃ	O(7)	2.776(4)
O(7)ᵇ	O(5)	2.885(4)

Equivalent positions: ᵃ$x+1, y, z$; ᵇ$\frac{1}{2}-x, y-\frac{1}{2}, \frac{3}{2}-z$.

in the same way as the *ortho* hydrogen atoms of the bpe ligand in Pt(bpe)(9-MeHX)₂(NO₃)₂. As in the reported bpe system the ethylene bridge also projects to some extent over the other side of the plane, Fig. 3(a), and this is also expected to influence the coordinating properties of the Pd(bmpe)²⁺ cation.

The unit cell contents is depicted in Fig. 4. Intermolecular contacts less than 3.2 Å are listed in Table 5. The three water molecules of crystallization H₂O(5), H₂O(6) and H₂O(7) are all involved in a hydrogen bonding network with each other and with the uncoordinated oxygen atoms, O(3) and O(4), of the malonato ligand. The bmpe pyridine ring systems of adjacent molecules stack in a head-to-tail fashion at distances of approx. 3.4 Å apart.

Proton NMR results

All reagents used in the NMR experiments were of the highest purity. In D₂O medium, a stock solution

of Pd(bmpe)(D₂O)₂²⁺ was prepared as described previously for the diaquo species. An appropriate amount of this solution was mixed with two equivalents each of 1-MeC, Guo and 5'-GMPNa₂. The pHs of these solutions were 6.6, 5.5 and 6.5, respectively (pD values may be obtained by adding 0.4 to the pHs). Similarly, Pd(bmpe)Cl₂ was dissolved in Me₂SO-d₆ and combined with two equivalents of 9-EtG. All solutions had been standing for at least 24 h before their proton NMR spectra were recorded. The spectra were recorded on a Bruker WB AMX300 (300 MHz) spectrometer at 25 °C. All chemical shifts are reported downfield from DSS (2,2-dimethyl-2-silapentane-5-sulfonate) for the samples in D₂O, and downfield from TMS (tetramethylsilane) for the sample in DMSO-d₆. Spectra were also recorded under the same conditions for all of the free (potential) ligands. Evidence for coordination of the palladium complex to the N(7) position of the guanine base was monitored by observing the chemical shift of the H(8) resonance in the presence and absence of metal complex. Evidence for coordination to the N(3) position of the cytosine moiety was monitored by observing the chemical shifts for the H(5) and H(6) resonances in the presence and absence of metal complex. Results are summarized in Tables 6 and 7.

Coordination by Pt(II) or Pd(II) to the N(7) position of guanine is expected to result in a downfield shift for the H(8) resonance of at least 0.4 ppm compared to the free ligand [24]. Similarly, coordination to the N(3) position of cytosine is expected to result in a downfield shift for the H(5) and H(6) resonances of

TABLE 6. H(8) proton chemical shifts (δ) for guanine derivatives in the absence and presence of Pd(bmpe)²⁺

	9-EtGᵃ	Guo	5'-GMPNa₂
Free guanine derivative	7.70	7.98	8.20
Guanine derivative (2 equiv.) + Pd(bmpe)²⁺	7.70	8.03	8.23
Difference in shift	0.00	+0.05	+0.03

ᵃDownfield from TMS in Me₂SO-d₆. Other values are from DSS in D₂O.

TABLE 7. H(5) and H(6) chemical shifts (δ) for 1-methylcytosine in the absence and presence of Pd(bmpe)²⁺

	H(5)	H(6)
Free 1-methylcytosine	5.94	7.55
1-methylcytosine (2 equiv.) + Pd(bmpe)²⁺	5.87	7.51
Difference in shift	-0.07	-0.04

All values are downfield from DSS in D₂O. Values are averages of the doublet signals for H(5) and H(6).

at least 0.2 ppm [1]. The results summarized in Tables 6 and 7 show that the observed differences in the chemical shift values in the presence and absence of metal complex are considerably less in magnitude than what would be expected if binding had occurred and, in the case of the 1-methylcytosine signals, are in the opposite direction (upfield) than expected. Therefore, we conclude that binding of the Pd(bmpe)$^{2+}$ moiety to the N(7) and N(3) positions of guanine and cytosine, respectively, does not occur in any of these experiments under these conditions. This is consistent with our inability to isolate any complexes preparatively.

Discussion

To our knowledge, this is the first structural characterization of a coordination compound with the bmpe ligand. As expected, bmpe introduces an asymmetric environment with respect to the coordination plane in the same way as bpe [13]. Thus by replacing the hydrogen atoms *ortho* to the donor nitrogen atoms in the bpe ligand with methyl substituents, Fig. 1, we have succeeded in building up the bulk preferentially on one side of the coordination plane, Fig. 3(a). Attempts to introduce N(7)-bound 6-oxo purine ligands and N(3)-bound 1-methylcytosine into the coordination sphere by reacting the Pd(bmpe)$^{2+}$ moiety with two equivalents of the appropriate nucleoligand have not been successful, as evidenced by proton NMR studies and by our inability to isolate any 1:2 complexes preparatively. This is in contrast to the reported reaction of two equivalents of 9-methylhypoxanthine (9-MeHX) with Pt(bpe)(H$_2$O)$_2$$^{2+}$ which results in the ready isolation of [Pt(bpe)(9-MeHX)$_2$](NO$_3$)$_2$ [13]. Those workers have characterized that complex by X-ray crystallography and have demonstrated the bases to be coordinated in an HT arrangement via the N(7) positions. Thus it is surprising, when Pd(bmpe)(H$_2$O$_2$$^{2+}$ is exposed to two equivalents of a 6-oxopurine derivative, that at least one such ligand is not introduced into the coordination sphere with the 6-oxo moiety on the same side of the coordination plane as the ethylene bridge. This arrangement is readily accommodated for one of the two 9-MeHXs in the aforementioned [Pt(bmpe)(9-MeHX)$_2$](NO$_3$)$_2$ complex and there is no evidence to suggest that the ethylene bridge is significantly more obstructive in the Pd(bmpe)$^{2+}$ moiety than in the Pt(bpe)$^{2+}$ moiety, see Table 8. Indeed, it is remarkable how rigid ligands of this type appear to be. We have noted, however, the slight distortion in the molecule attributed to a 'pushing down' of one of the methyl substituents onto part of the coordination plane. This is reflected in the distortion of the square-plane itself and in the angles at the N(2) atom of the bmpe ligand.

TABLE 8. A comparison of the geometric parameters associated with the projection of the ethylene bridge across the coordination plane in Pd(bmpe)(mal)·3H$_2$O and Pt(bpe)(9-MeHX)$_2$(NO$_3$)$_2$. Values for the latter compound are given in parentheses

Contacts	Distance (Å)	Δ (Å)
C(8)–Pd (C(11)–Pt)	2.906 (3.028)	0.122
C(7)–Pd (C(31)–Pt)	3.270 (3.327)	0.057
C(8)–O(1) (C(11)–N(27))	3.893 (4.150)	0.257
C(7)–O(2) (C(31)–N(17))	4.204 (4.420)	0.216

This effect may be visualized in Fig. 3(a) and (b). Further details of these distortions are discussed under a previous heading. Although one must attribute the complete hindering of nucleobase binding in the bmpe system to the influence of the steric bulk of the exocyclic methyl substituents, we are not convinced that the subtle change in coordination geometry as described above, by itself, could have such a dramatic effect on the relative coordinating abilities of the bpe and bmpe systems. Investigations are continuing into other factors which could be operative.

Supplementary material

Further details of the crystal structure investigation are available on request from the Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK, on quoting the full journal citation.

Information available from the authors upon request includes GC-MS peaks (with $m/z > 3\%$) for bmpe, proton NMR data for bmpe, and FT-IR data for bmpe, Pd(bmpe)malonate and Pd(bmpe)Cl$_2$.

Acknowledgements

References

1 J.D. Orbell, L.G. Marzilli and T.J. Kistenmacher, *J. Am. Chem. Soc.*, *103* (1981) 5126.

176

2 P.M. Pil and S.J. Lippard, *Science*, *256* (1992) 234.
3 T.R. Cech, *Science*, *236* (1987) 1532.
4 J. Haseloff and W.L. Gerlach, *Nature (London)*, *334* (1988) 585.
5 T.J. Kistenmacher, J.D. Orbell and L.G. Marzilli, in S.J. Lippard (ed.), *Platinum, Gold and Other Metal Chemotherapeutic Agents*, Vol. 209, ACS Symposium Series, Washington, DC, 1983, p. 191.
6 (a) Y. Kidani, *Yakugaku Zasshi*, *105* (1985) 909; (b) G. Trovo and B. Longato, *Inorg. Chim. Acta*, *192* (1992) 13.
7 A.M. Pyle and J.K. Barton, in S.J. Lippard (ed.), *Progress in Inorganic Chemistry: Bioinorganic Chemistry*, Vol. 38, Wiley, New York, 1990, p. 413.
8 B. Lippert, in S.J. Lippard (ed.), *Progress in Inorganic Chemistry*, Vol. 37, Wiley, New York, 1989, p. 3.
9 H. Schöllhorn, G. Raudaschl-Sieber, G. Muller, U. Thewalt and B. Lippert, *J. Am. Chem. Soc.*, *107* (1985) 5932.
10 S.E. Sherman, D. Gibson, A.H.-J. Wang and S.J. Lippard, *J. Am. Chem. Soc.*, *110* (1988) 7368.
11 G. Admiraal, J.L. van der Veer, R.A.G. de Graaf, J.H.J. den Hartog and J. Reedijk, *J. Am. Chem. Soc.*, *109* (1987) 592.
12 J.D. Orbell, K. Wilkowski, B. de Castro, L.G. Marzilli and T.J. Kistenmacher, *Inorg. Chem.*, *21* (1982) 813.
13 A.T.M. Marcelis, H. Korte. B. Krebs and J. Reedijk, *Inorg. Chem.*, *21* (1982) 4059.
14 P.G. Campbell and P.C. Teague, *J. Am. Chem. Soc.*, *76* (1954) 1371.
15 K.C. Bass, B. McIntosh and A.G. Osborne, *Spectrosc. Lett.*, *18* (1985) 35.
16 G.R. Newkome, V.K. Gupta, H.C.R. Taylor and F.R. Fronczek, *Organometallics*, *3* (1984) 1549.
17 W.K. Kleibohmer, B. Krebs, A.J.M. Marcelis, J. Reedijk and J.L. van der Veer, *Inorg. Chim. Acta*, *75* (1983) 45.
18 C.G. van Kralingen, J. Reedijk and A.L. Spek, *Inorg. Chem.*, *19* (1980) 1481, and refs. therein.
19 L. Pauling, *The Nature of The Chemical Bond*, Cornell University Press, Ithaca, NY, 2nd edn., 1944, p. 80.
20 N.A. Bailey and R. Mason, *J. Chem. Soc. A*, (1968) 2594.
21 G.R. Clark and J.D. Orbell, *J. Organomet. Chem.*, *215* (1981) 121.
22 G. Minghetti, M.A. Cinella, A.L. Bandini, G. Banditelli, F. Demartin and M. Manaserro, *J. Organomet. Chem.*, *215* (1986) 387.
23 R.K. Brown, J.M. Williams, A.J. Schultz, G.D. Ittel and R.L. Harlow, *J. Am. Chem. Soc.*, *102* (1980) 981.
24 J.D. Orbell, M.R. Taylor, S.L. Birch, S.E. Lawton, L.M. Vilkins and L.J. Keefe, *Inorg. Chim. Acta*, *152* (1988) 125.

Determination of platinum and palladium in chelated systems by GFAAS

Peter Watkins and Elizabeth Yuriev
Department of Environmental Management
Victoria University of Technology
Melbourne Victoria Australia 3000

Introduction

With the discovery of the anti-tumour properties of cis-dichlorodiammineplatinum(II) (cisplatin)[1], intensive research has been completed which examines the role of platinum in biological systems. One major field of interest has been the interaction of cisplatin and related compounds with biological systems, especially with DNA and its constituents. The interactions of palladium complexes and DNA have also been examined. Principally, this is due to the labile nature of palladium. Platinum substitution reactions are known to be slow. Palladium proves to be an ideal model because it has similar coordination properties and substitution reactions occur generally in the order of 10^5 times faster when compared with platinum.

The technique of atomic absorption spectrometry (AAS) has been employed in this area of research to measure levels of platinum. It has been used to measure the distribution of Pt in biological fluids and systems[2] as well as used to study platinum coordination reactions[3,4].

It has been found that the species active toward the tumour is diaquodiammineplatinum(II). Generally, it is synthesized by the removal of the chloride ions from the complex using silver nitrate[5]. This allows the aquated form to be present in solution. Though this process is almost always quantitative, there are circumstances when accurate levels of platinum (and complex) need to be known. One such case is the determination of the stability constants for Pt-DNA and Pd-DNA complexes. In this case, it is necessary to use graphite furnace AAS[6]. This paper presents two methods which are suitable for the determination of platinum and palladium in systems such as cisplatin and related compounds.

Experimental

Method

A Varian SpectrAA 400 Atomic Absorption Spectrometer and GTA96 Graphite Tube Atomizer were employed for the measurements.

Platinum

A platinum SpectrAA hollow cathode lamp was used and the samples dispensed into pyrolytically coated graphite tubes. The Pt resonance line at 265.0 nm was used with a slit width of 0.2 nm. The lamp current setting was 5 mA.

A 1000 mg/L Pt solution (BDH Ltd, Spectrosol grade) was used as the stock standard solution. An intermediate solution of 10 mg/L Pt was prepared by serial dilution using 0.1% v/v HNO_3. Working standards were also prepared by serial dilution from the intermediate solution using 0.1% v/v HNO_3.

Palladium

A palladium SpectrAA hollow cathode lamp was used with pyrolytically coated tubes. The Pd resonance line at 244.8 nm was used with a slit width of 0.3 nm. The lamp current setting was 5 mA.

A 1000 mg/L Pd (Aldrich, USA) was used as the stock standard solution. An intermediate solution of 10 mg/L Pd as well as the working standards were prepared as for platinum.

Results and discussion

Ashing and atomization studies were performed in order to determine the optimal furnace parameters. A Pt solution was made in the presence of KNO_3 and ethylenediamine for this purpose. The results are shown in figure 1. It is clear from figure 1 that the best region for ashing lies between 800-1000 °C and a temperature of 2700 °C was chosen as the atomization temperature. The furnace parameters used for the analysis are shown in table 1.

Pt

Legend:
- ■ ashing curve
- ○ atomization curve
- ▲ background
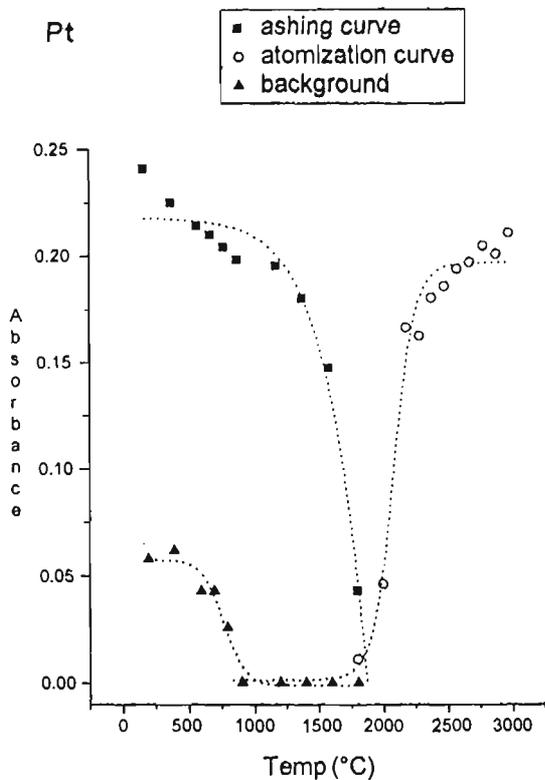
Figure 1: Optimal furnace parameter results for Pt

Table 1: Furnace parameters

Step No	Temperature (°C)	Time (sec)	Gas Flow (L/min)	Gas Type	Read Command
1	85	5.0	3.0	Normal	No
2	95	40.0	3.0	Normal	No
3	120	10.0	3.0	Normal	No
4	700	5.0	3.0	Normal	No
5	700	2.0	3.0	Normal	No
6	800	5.0	3.0	Normal	No
7	800	2.0	3.0	Normal	No
8	800	2.0	0.0	Normal	No
9	2700	1.3	0.0	Normal	Yes
10	2700	2.0	0.0	Normal	Yes
11	2700	2.0	3.0	Normal	No

A comparison was performed on the use of premixed standards and those made using the automix facility of the autosampler. The results are shown in table 2. It can be seen that the automix facility provides a calibration graph which is comparable with that from the premixed standards. This is a useful feature as it helps to minimize operator error as well as reduce consumption of the stock solution. The autosampler parameters are shown in table 3. A 100 μg/L Pt solution was used as the standard solution. Figure 2 shows a representative calibration graph.

The precision of the determination was established by analyzing a series of five solutions containing approximately 20 mg of K_2PtCl_4 in 100 mL. After serial dilution, the Pt concentration was calculated. Table 4 shows these results and it can be seen that the recovery is very good.

Table 2: Comparison of standard solutions

Conc (μg/L)	Abs Premixed	Abs Automixed
0.0	0.000	0.011
50.0	0.044 (7.7)	0.044 (3.3)
100.0	0.092 (5.2)	0.089 (1.0)
150.0	0.142 (1.0)	0.135 (1.1)

The values in parentheses are the relative standard deviations for 3 replicates

Table 3: Sampler parameters

	Volumes (fL) Solution	Blank	Modifier
Blank	--	20	
Standard 1	6	15	
Standard 2	10	10	
Standard 3	15	5	
Sample	10	10	

Recalibration Rage 0
Reslope Rate 0
Multiple Inject NO   Hot Inject NO   Pre Inject NO

Table 4: Platinum recovery

Solution	Expected Result (mg/L)	Found Result (mg/L)	%
1	94.5	89.8	95.0
2	93.5	91.5	97.9
3	98.7	97.4	98.7
4	96.8	93.7	96.8
5	99.6	99.9	99.4

2
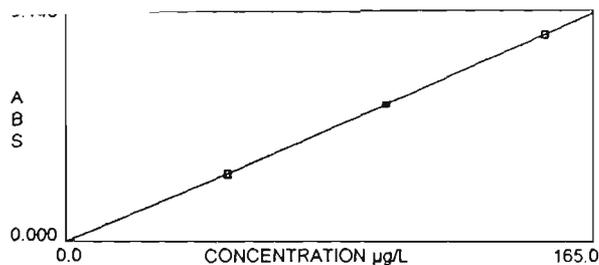
Figure 2: Representative calibration for Pt

Similar experiments were carried out for palladium. Figure 3 shows the ashing and atomization curves and table 5 shows the furnace parameters used for the analysis.
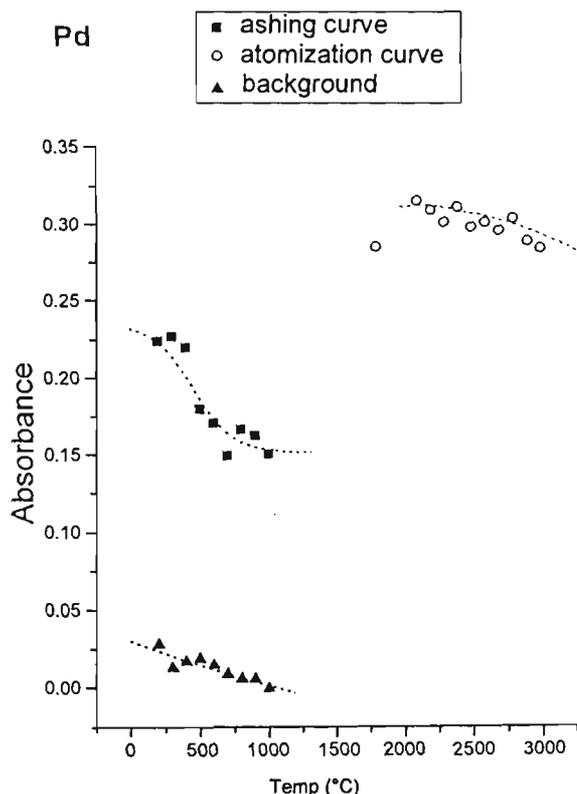
Figure 3: Ashing and Atomization curves for Pd

Table 5: Furnace parameters

Step No	Temperature (°C)	Time (sec)	Gas Flow (L/min)	Gas Type	Read Command
1	85	5.0	3.0	Normal	No
2	95	40.0	3.0	Normal	No
3	120	10.0	3.0	Normal	No
4	800	5.0	3.0	Normal	No
5	800	1.0	3.0	Normal	No
6	800	2.0	0.0	Normal	No
7	2600	1.0	0.0	Normal	Yes
8	2600	2.0	0.0	Normal	Yes
9	2600	2.0	3.0	Normal	No

A comparison was again made on the use of the automix facility for preparing the calibration curve, and, as for platinum, the graph compared very favourably with that obtained from premixed standards. Table 6 shows the autosampler parameters and a representative curve is shown in Figure 4. A 100 μg/L Pd solution was used as the standard solution. The precision of the analysis was also found to be comparable with that obtained for platinum.
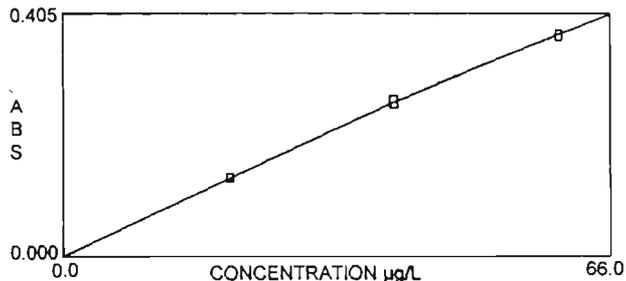
Figure 4:   Representative calibration for Pd

Table 6: Sampler Parameters

	Volumes (fL)		
	Solution	**Blank**	**Modifier**
Blank	—	20	
Standard 1	2	18	
Standard 2	4	16	
Standard 3	6	14	
Sample	10	10	
Recalibration Rage	0		
Reslope Rate	0		

Multiple Inject NO   Hot Inject NO   Pre Inject NO

3

Our work has been involved with the determination of formation constants for Pt/Pd complexes of bidentate ligands. For this task, it is therefore important to have a reliable means of accurately measuring the concentrations of each of the reaction components.

The systems which have been examined are complexes of Pt and Pd with bidentate ligands. Tables 7 and 8 show some results of the analyses. Some of this work has been presented elsewhere[7].

Table 7: Results of Analysis

Complex	Expected Result (g/L)	Experimental Result (g/L)	% Yield	Solvent
Pd(bmpe)$^{2+}$	17.5	9.9	57	water
Pd(en)$^{2+}$	17.8	17.5	98	water
Pd(bmpe)$^{2+}$	9.9	8.3	84	DMSO
Pd(en)$^{2+}$	10.0	8.1	81	DMSO

Table 8: Results of Analysis

Complex	Expected Result (g/L)	Experimental Result (g/L)	Percentage Yield
Pd (TMED)$^{2+}$	2.8	2.4	86
Pd (dmp)$^{2+}$	6.36	6.36	100
	2.2	1.7	77
Pd$_2$(bispep)$^{4+}$	4.4	1.2	27
Pd (bpe)$^{2+}$	3.9	2.7	70
Pd (Ie)$^{2+}$	7.0	5.62	80
Pt (NH$_3$)$_2$$^{2+}$	9.95	8.58	86
Pt (en)$^{2+}$	7.22	4.1	57

The constituents of DNA shows a wide variation in their solubilities in a number of solvents. It was therefore important to establish the efficiency of the removal of chloride from the complex and, consequently, the yield of the solvated species. Table 7 shows the results for two complexes which were dissolved in water and dimethylsulfoxide. Once established, proton NMR spectroscopy can be used to study the interactions between the metal complexes and DNA. A number of other ligands were also investigated and these results are shown in table 8.

References

1. B. Rosenberg, L. van Camp, J.E. Trosko and V.H. Mansour, Nature (London), 222, 385 (1969).

2. (a) A.F. LeRoy, M.L. Wehling, H.S. Sponseller, W.S. Friauf, R.E. Solomon, R.L. Dedrick, C.L. Litterst, T.E. Gram, A.M. Guarino and D.A. Becker, Biochemical Medicine, 18, 184 (1977).

   (b) M.C. McGahan and K. Tyczkowska, Spectrochimica Acta, 42B, 665 (1987).

3. J.P. Macquet and T. Theophanides, Atomic Absorption Newsletter, 14, 23 (1975).

4. J.P. Macquet and T. Theophanides, Biochimica et Biophysica Acta, 442, 142 (1976).

5. B. Lippert, Progress in Inorganic Chemistry, 37, 1 (1989).

6. (a) T. Ren, D.P. Bancroft, W.I. Sundquist, A. Masschelein, M.V. Keck and S.J. Lippard, J. Am. Chem. Soc., 115, 11341 (1993).

   (b) K. Inagaki and Y. Kidani, Inorg. Chim. Acta, 80, 171 (1983).

7. A.T. Baker, J.K. Crass, G.B. Kok, J.D. Orbell and E. Yuriev, Inorg. Chim. Acta, 214, 169 (1993).

Appendix 1:

Key to ligand abbreviations

en	=	ethylenediamine
TMED	=	tetramethylethylenediamine
bmpe	=	1,2-bis(6-methylpyridin-2-yl)ethane
bispep	=	1,2-di-(4-methyl-1-piperazinyl)ethane
dmp	=	1,4-dimethylpiperazine
Ie	=	1,2-bis(2-imidazolin-2-yl)ethane
bpe	=	1,2-bis(pyridin-2-yl)ethane

Steric Parameters for Metal Binding Sites on Nucleobases

Elizabeth Yuriev and John D. Orbell

Department of Environmental Management,
Victoria University of Technology, St. Albans Campus,
P.O. Box 14428, MCMC, Melbourne, Victoria 8001, Australia

Inorganic

Chemistry®

Notes

Steric Parameters for Metal Binding Sites on Nucleobases

Elizabeth Yuriev and John D. Orbell*

Department of Environmental Management,
Victoria University of Technology, St. Albans Campus,
P.O. Box 14428, MCMC, Melbourne,
Victoria 8001, Australia

Introduction

The binding of metal species to nucleic acids may influence various biological functions, including those related to structural motifs induced by coordination.[1] The endocyclic nitrogens of the common nucleobases (Figure 1) are of particular interest with respect to metallo-chemotherapeutics, such as the antitumor drug cisplatin (*cis*-diamminedichloroplatinum(II)), since these sites are considered to be key molecular targets for such compounds.[2]

For isolated nucleobases, factors which favor one such site over another include basicity[3] and steric considerations.[4] With respect to basicity, preferential binding (in particular to the N7 position of guanine) has been quantitatively rationalized through pK_a values,[3] and relative coordination strengths of different binding sites have also been rationalized on the basis of electrostatic potential energy distributions.[5] Steric influences prior to, during and after coordination[6] are dictated by the nature of the metal complex (usually by the carrier ligand) and by the features of the binding site itself, such as the neighbouring exocylcic substituent(s) in the case of the nucleobases.[5,7] The relative steric demands of such binding sites have been considered by some workers to be difficult to quantify.[4] In an attempt to achieve this goal, we have extended to nucleobase ligands the repulsive energy methodology developed by Brown et al.[8] Thus each nucleobase endocyclic nitrogen site is probed by a Cr(CO)₅ moiety to which it is hypothetically bound (Figure 2).

* Author for correspondence.

(1) (a) Takahara, P. M.; Rosenzweig, A. C.; Frederick, C. A.; Lippard. S. J. *Nature* **1995**, *377*, 649–652. (b) Huang, H.; Zhu, L.; Reid, B. R.; Drobny, G. P.; Hopkins, P. B. *Science* **1995**, *270*, 1842–1845.

(2) (a) Farrell, N. *Transition Metal Complexes as Drugs and Chemotherapeutic Agents*; Kluwer Academic: Dordrecht, The Netherlands, 1989. (b) Sundquist, W. I.; Lippard, S. J. *Coord. Chem. Rev.* **1990**, *100*, 293. (c) Tipson, R. S.; Townsend, L. B. *Nucleic Acid Chemistry*; Wiley: New York, 1991, pp 201–219.

(3) Martin, R. B. *Acc. Chem. Res.* **1985**, *18*, 32–38.

(4) Reily, M. D.; Marzilli, L. G. *J. Am. Chem Soc.* **1986**, *108*, 6785–6793.

(5) Kistenmacher, T. J.; Orbell, J. D.; Marzilli, L. G. In *Platinum, Gold and other Metal Chemotherapeutic Agents*; Lippard, S. J. Ed.; ACS Symposium Series 209; American Chemical Society: Washington DC, 1983, pp 191–208.

(6) Bloemink, M. J.; Engelking, H.; Karentzopoulos, S.; Krebs, B.; Reedijk, J. *Inorg. Chem.* **1996**, *35*, 619–627.

(7) (a) Kinjo, Y.; Tribolet, R.; Corfu, N. A.; Sigel H. *Inorg. Chem.* **1989**, *28*, 1480–1489. (b) Kinjo, K.; Liang-nian, J.; Corfu, N. A.; Sigel, H. *Inorg. Chem.* **1992**, *31*, 5588–5596.

(8) (a) Brown, T. L. *Inorg. Chem.* **1992**, *31*, 1286–1294. (b) Caffery, M. L.; Brown, T. L. *Inorg. Chem.* **1991**, *30*, 3907–3914. (c) Choi, M.-G.; Brown, T. L. *Inorg. Chem.* **1993**, *32*, 1548–1553. (d) Choi, M.-G.; Brown, T. L. *Inorg. Chem.* **1993**, *32*, 5603–5610. (e) Choi, M.-G.; White, D.; Brown, T. L. *Inorg. Chem.* **1994**, *33*, 5591–5594. (f) White, D.; Brown, T. L. *Inorg. Chem.* **1995**, *34*, 2718–2724.

Figure 1. I_S values for potential metal binding sites on the common nucleobases. An asterisk denotes the value for a deprotonated site. For this study $R = CH_3$.

Figure 2. Disk representation of the approach of the steric probe Cr(CO)₅ to the N7 position of 9-methylguanine. Pattern coding: light gray (small), hydrogen; light gray (large), oxygen; dark gray, nitrogen; white, carbon; black, chromium.

There are a number of criteria for choosing the Cr(CO)₅ moiety, not the least being that it is a good representative of a transition metal species in terms of the degree of crowding about the metal center.[8a] The resulting steric parameter, the ligand repulsive energy, E_R, represents the gradient of the van der Waals repulsive energy for the energy-minimized structure with respect to the Cr–N distance, scaled by the equilibrium Cr–N distance.

Results and Discussion

For the common nucleobases studied here, these E_R values are presented in Table 1.

The numbers shown in Figure 1, associated with each site assessed, represent a steric index, I_S, which may be defined as follows:

$$I_S = E_R[\text{N(nucleobase)}]/E_R[\text{N7(guanine)}]$$
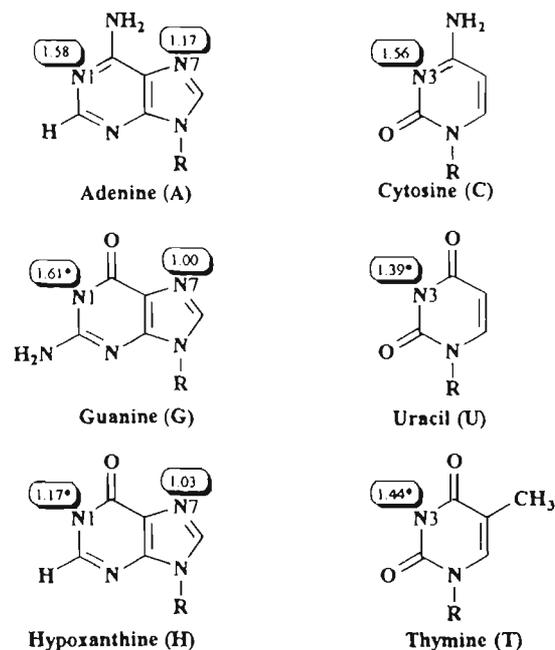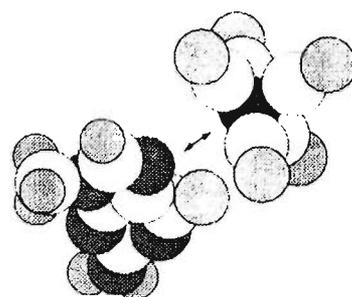
N(nucleobase) = N1(guanine, hypoxanthine, or adenine), N3-(cytosine, thymine, or uracil), or N7(guanine, hypoxanthine, or adenine).

The above definition is predicated upon the N7 of guanine presenting the lowest relative steric hindrance to the probe of all the sites compared in this study.

Table 1. Ligand Repulsive Energies Presented by Nucleobase Binding Site to the Metal Species Cr(CO)5 (Estimated uncertainty ± 1 kcal/mol)

base	site	E_R (kcal/mol)
9-MeG	N1[a]	58
	N7	36
9-MeH	N1[a]	42
	N7	37
9-MeA	N1	57
	N7	42
1-MeC	N3	56
1-MeT	N3[a]	52
1-MeU	N3[a]	50

[a] Deprotonated.

An examination of the I_S values presented in Figure 1 allows the following observations to be made. With respect to the N7 of the purines (which are both intuitively the least sterically hindered and being the most accessible on a DNA duplex[2a]) the I_S values are comparable for both guanine and hypoxanthine, but approximately 17% higher for adenine. This is in agreement with experimental evidence[4,7] which suggests a greater steric influence on metal coordination at N7 of an exocyclic amino compared to an exocyclic oxo substituent. The sterically equivalent sites, N1 of guanine and N3 of cytosine, have comparable I_S values which are approximately 60% higher than the value for N7 of guanine. If attention is focused on the N1 position of adenine, the I_S value for this site is comparable to the values for the N1 site of guanine and the N3 site of cytosine, in spite of the absence of a concomitant exocyclic oxygen (replaced by hydrogen in the case of adenine). This is further evidence for the larger steric influence of the amino substituent and is consistent with the oxo substituent having only a relatively modest steric effect.[7b] The same conclusion with respect to exocyclic oxo may be reached by comparing the I_S values of the N1 position of hypoxanthine with the N1 of adenine. For the Cr(CO)5 moiety, the steric influence of one oxo substituent ortho to a binding site (e.g. N1 of hypoxanthine) is equivalent to that of an amino substituent one atom removed from a binding site (e.g. N7 of adenine). The N3 of uridine and the N3 of thymine have intermediate I_S values, as might be expected.

In the purine systems, when N9 carries a substituent, the N3 position is rarely accessed by metal species due to the severe steric constraints.[2a] Attempting to probe the N3 position by the method presented here results in an anomalous outcome reflected in inflated values of E_R and structural distortions; more

specifically, the planarity of the nucleobase moiety is compromised, and its orientation with respect to the Cr(CO)4 (radial) plane is no longer close to 90°. Thus it is possible that this method could be exploited to identify structural requirements for which coordination is precluded by steric factors. This could be useful in the design of metal complexes as site-specific reagents.[9]

The steric parameters presented here for metal binding sites on nucleobases have sensible relative values and are in accord with deductions from reported experimental data where steric effects are considered to be operative. Thus they demonstrate the feasibility of quantifying relative steric effects in such systems and buttress the suggestion by Brown et al.[8b] that the E_R concept can be extended to ligands of nearly any kind. These workers also suggest that a variety of metal centers may also be considered. In this context, the metal species could equally well be varied (for example, a series of modified platinum complexes) with a particular nucleobase binding site held constant. One would expect the steric parameters derived from such investigations to find particular application in quantitative structure activity/property relationship investigations (QSAR[10]/ QSPR[11]) since, unlike the frequently employed molecular volume as a steric parameter,[12] the E_R values represent steric effects at the interface of the interaction and would be expected to carry little "transport" information relating to the hydrophobicity[13] or water solubility[6] of the metal complex.

Supporting Information Available: Text giving computational details including a table of MM+ added force field parameters for Cr-(CO)5/nucleobase complexes (3 pages). Ordering information is given on any current masthead page.

IC9610896

(9) Dobi, A. L.; Matsumoto, K.; von Agoston, D. *Nucleic Acids Res.* **1994**, 22, 4846−4847.
(10) Kubinyi, H. *QSAR: Hansch Analysis and Related Approaches*; VCH: Weinheim, Germany, 1993.
(11) Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *Chem. Soc. Rev.* **1995**, 24, 279−287.
(12) Charton, M. In *Steric Effects in Drug Design*; Charton, M., Motoc, I. Eds.; Springer-Verlag: Berlin, 1983; pp 107−118.
(13) Ghose, A. K.; Pritchett, A.; Gordon, M. C. *J. Comput. Chem.* **1988**, 9, 80−90.

589

Modelling steric effects in DNA-binding platinum(II)–am(m)ine complexes

Elizabeth Yuriev and John D. Orbell*

Department of Environmental Management, Victoria University of Technology, St. Albans Campus, P.O. Box 14428, MCMC, Melbourne, VIC 8001, Australia

Summary

A repulsive energy strategy has been employed in an attempt to delineate the steric contribution to the biological profile of a variety of platinum–am(m)ine complexes. Thus, relative steric descriptors have been calculated for the amine ligands themselves by the Ligand Repulsive Energy (LRE) methodology. This has been extended to a Complex Repulsive Energy (CRE) strategy whereby the steric requirements of the approach of a metal complex to a site on a target molecule may be evaluated. Specifically, the monodentate approach of a variety of platinum–am(m)ine complexes to the N7 site of a guanine moiety has been considered. The steric descriptors thus obtained have been used in QSAR analysis, resulting in improved regression equations. Attempts have also been made to relate the above descriptors to various biological indicators for given series of complexes. These investigations suggest an optimum steric requirement for minimum toxicity, which could aid in the rational design of such agents.

Introduction

The binding of metal species to nucleic acids is expected to affect a wide range of biological outcomes [1]. More specifically, by binding to DNA some metal species, such as the now familiar cisplatin (*cis*-dichlorodiammine-platinum(II)), show dramatic cytotoxic and antineoplastic activity [2]. Attempts to relate aspects of structure to activity for these compounds, such as the steric influence of the amine carrier ligand(s), have been primarily qualitative in nature [3]. As far as a quantitative treatment is concerned, an extensive literature search revealed a limited number of attempts to correlate the biological activity of platinum–am(m)ine complexes with their structures, four groups being concerned with antitumour activity [4–9] and one with mutagenicity [10]. A search of the MedChem/Biobyte QSAR database [11] revealed an additional two QSARs involving platinum–am(m)ine complexes. These studies, for toxicity only, include cisplatin amongst a number of miscellaneous compounds [12], and a series of substituted dichloro(o-phenylenediamine)platinum(II) complexes [13].

In one of the QSAR studies on platinum systems reported in the literature, Abdul-Ahad and Webb [4] have investigated the correlation of standard indicators of biological activity, namely, acute toxicity (pLD_{50}), antitumour potency (pID_{90}) and therapeutic index (log TI), with a variety of structural descriptors. These include a wide range of electronic parameters, and the molecular volume (MV) as a steric parameter. In some cases this has resulted in moderate to strong correlations with a high degree of statistical significance. However, in this study, parameters relating to transport phenomena (e.g. hydrophobicity) have not been taken into account, and the use of MV oversimplifies steric effects since it is a scalar quantity [14]. Steric effects are, in fact, often highly localised and directional.

Tang et al. [9] have used exclusively electronic descriptors to obtain regression equations (having more than three variables), which also show moderate to strong correlation and high statistical significance. Furthermore, these authors report the synthesis of a complex designed to possess the electronic characteristics required by their regression results. This complex is reported to show rela-

---

*To whom correspondence should be addressed.

TABLE 1
CORRELATION MATRIX FOR THE DESCRIPTORS USED IN REF. 9

	q_{Pt} (e)	$-q_{Cl}$ (e)	$-q_N$ (e)	q_{Am} (e)	ΔE (a.u.)	$Q_{Pt\text{-}Cl}$ (e)	$Q_{Pt\text{-}N}$ (e)	ΔQ (e)	N_H
q_{Pt}	1								
$-q_{Cl}$	0.4702	1							
$-q_N$	-0.446	0.3272	1						
q_{Am}	-0.1619	0.791	0.6817	1					
ΔE	0.2019	-0.2084	**-0.6416**	-0.3946	1				
$Q_{Pt\text{-}Cl}$	**-0.7859**	**-0.8099**	0.0843	-0.3569	-0.139	1			
$Q_{Pt\text{-}N}$	0.0517	0.742	0.4194	0.7933	-0.0016	-0.5301	1		
ΔQ	-0.3586	**-0.8667**	-0.2702	-0.7186	-0.0535	0.7904	-0.9384	1	
N_H	**-0.7809**	-0.2174	0.5071	0.2975	-0.0583	0.6186	0.1663	0.1317	1

Descriptors: q_{Pt}, $-q_{Cl}$, $-q_N$ and q_{Am} are electronic charges on corresponding atoms and moieties; ΔE is the gap between energies of LUMO and HOMO; $Q_{Pt\text{-}Cl}$ and $Q_{Pt\text{-}N}$ are overlap populations on corresponding bonds; ΔQ is the difference between the above; N_H is the number of protons on the N atom of the amine.

tively low toxicity and high potency. It is not clear, however, whether this approach is generally applicable.

Simon et al. [5,6] have sought correlations involving hydrophobicity, electronic and steric effects with respect to biological activity. These workers [6] have suggested that the biological activity indicator TI depends predominantly on steric features. For the steric optimisation of TI, a receptor mapping procedure was employed via the minimal steric difference technique, resulting in quite strong and statistically significant correlation results. This approach, however, has certain drawbacks and has received considerable criticism [15,16]. In particular, it should be noted that the parameter used can code both steric and nonsteric factors, and the method does not account for the molecule's three-dimensional structure and conformational flexibility.

All of the researchers referred to so far used either measured physicochemical properties or computed descriptors representing classic QSAR studies (Hansch-type analysis) [17].

A more recent attempt to quantify structure–activity relationships for platinum complexes involves the application of a graph-theoretical method and employs purely topological indices. Thus, Romanowska and co-workers [7,8] have demonstrated the usefulness of molecular topology for structure–activity studies of platinum–am(m)ine complexes. However, even though topological indices generally offer an almost universal means of representing a chemical structure, they usually lack physical or chemical meaning and cannot be extrapolated to other compounds [18,19].

The major weakness of the multiple linear regression (MLR) analysis performed in the above studies, and especially in Ref. 9, is that many of the descriptors are interrelated and the same information is carried by more than one parameter. Such collinearity diminishes the statistical significance, complicates the interpretation of the correlation and may lead to false conclusions and predictions [20]. Collinearity may be diagnosed from a correlation matrix. Thus, correlation matrices have been

TABLE 2
CORRELATION MATRIX FOR THE DESCRIPTORS USED IN REF. 4

	BE (eV)	$EF(P_7)$ (V/Å)	$EF^2(P_7)$ ((V/Å)²)	$EF(P_9)$ (V/Å)	$E_x(P_3)$ (V/Å)	$E_z(P_9)$ (V/Å)	$F^N(Pt)$ (-e)	MV (Å³)	$V_2(P_7)$ (V/e)
BE	1								
$EF(P_7)$	0.6187	1							
$EF^2(P_7)$	0.6146	0.9999	1						
$EF(P_9)$	-0.3498	-0.2518	-0.2496	1					
$E_x(P_3)$	0.5656	0.343	0.3364	-0.2655	1				
$E_z(P_9)$	-0.3757	-0.255	-0.2523	**0.9952**	-0.3024	1			
$F^N(Pt)$	-0.0747	-0.1888	-0.191	-0.3437	-0.073	-0.3518	1		
MV	-0.9821	-0.6219	-0.6182	0.3927	-0.4811	0.4128	-0.0037	1	
$V_2(P_7)$	**0.6895**	0.4143	0.4073	-0.0675	0.4642	-0.1137	0.1276	**-0.6625**	1

Descriptors: BE is the molecule binding energy (the difference between the INDO total energy of the molecule and the sum of the energies of the isolated constituent atoms); $EF(P_7)$ and $EF(P_9)$ are the moduli of the electric field at p (receptor point); $E_x(P_3)$ and $E_z(P_9)$ are the components of the electric field at p, calculated by the finite differentiation of the electrostatic potential; $F^N(Pt)$ is the frontier electron density for the nucleophilic attack at Pt; MV is the molecular volume; $V_2(P_7)$ is the energy of polarisation of the complex by a unit point charge placed at p (second-order interaction energy, calculated by an uncoupled Hartree–Fock perturbation procedure). The bold numbers in Tables 1 and 2 depict the correlation coefficients between the descriptors, simultaneously used in the equations obtained in Refs. 4 and 9, which show moderate to strong collinearity (correlation coefficient more than 0.6).
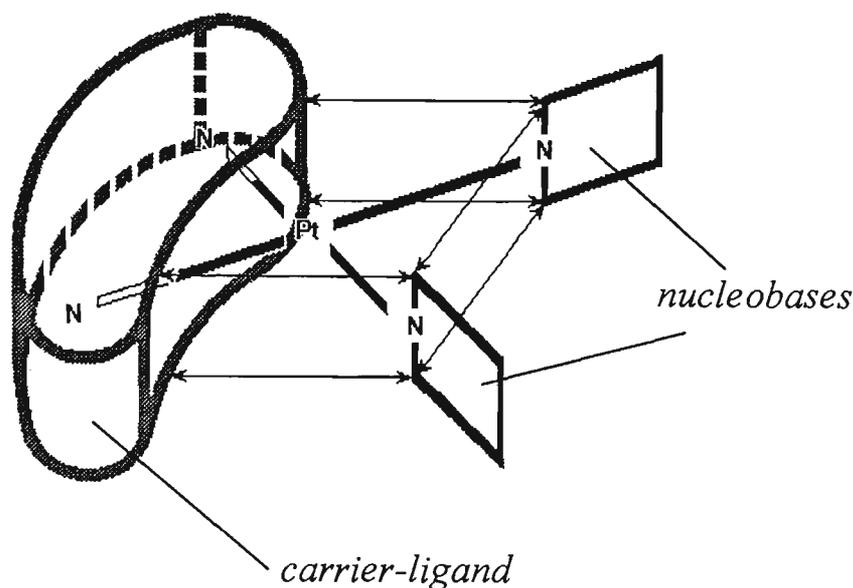
Fig. 1. Potential intramolecular steric interactions in the vicinity of the binding site for a bidentate, metal-mediated, intrastrand cross-linkage. Arrows depict potential steric contacts between the carrier ligand and nucleobases above and below the coordination plane.

reported by some authors [5,7], although we have found it necessary to perform our own computations based on published data [4,9] where this has not been carried out by the authors themselves, Tables 1 and 2.

In summary, research groups studying QSARs of platinum antitumour complexes have investigated the effects of a variety of structural descriptors on biological activity. The highest priority in these studies has been given to electronic descriptors. Hydrophobicity has been less considered, and a quantification of steric effects lags far behind. This could possibly be due to the lack of appropriate steric descriptors as applied to inorganic systems. A number of protocols have been developed in an attempt to quantify such steric effects. These include cone angles Θ [21], solid angles Ω [22], ligand repulsive energy E_r [23], modified vdW energy [24] and stereochemical conventions [25]. However, to our knowledge, there has been no attempt to extend these protocols to QSAR investigations relating to the biological profiles of metal complexes, such as the cis-platinum derivatives.

Our analysis of previous work on the QSARs of platinum complexes prompted us to carry out MLR analysis in an attempt to properly account for all three kinds of effects (transport, electronic and steric). In particular, to describe steric effects in platinum complexes, we have applied and extended a recently developed [23] repulsive energy strategy to calculate new steric descriptors. We have also attempted to ascertain the scope of these descriptors both in terms of quantitative structure–activity relationships and particularly with respect to the trends which may be observed in the biological profiles.

Steric effects in platinum systems

All the QSARs discussed above considered the structural features of unbound platinum complexes without a

consideration of the structural features of the possible molecular target(s). Today it is widely accepted that DNA is a major intracellular target for cisplatin binding and that its biological activity is manifested through such binding. The multiple binding sites on DNA present the possibility of a diversity of significant steric interactions [25]. Our initial goal is to develop a systematic method for the quantitative assessment of steric effects in platinum(II)–am(m)ine complex/nucleobase systems. For square-planar complexes which may coordinate to nucleobases via a bidentate intrastrand cross-linkage [25], potential intramolecular steric interactions in the immediate
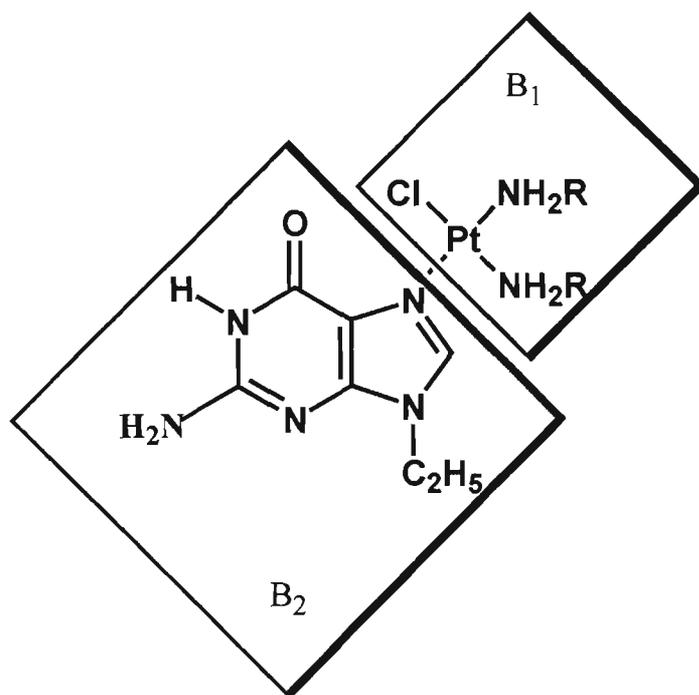
Fig. 2. A representative monoadduct between cis-platinum complexes and 9-ethyl guanine.

592

STRAIGHT CHAINS:

$H_3C{-}NH_2$ 20

$C_2H_5{-}NH_2$ 21

$n{-}C_3H_7{-}NH_2$ 22

$n{-}C_4H_9{-}NH_2$ 23

$n{-}C_5H_{11}{-}NH_2$ 24

$n{-}C_6H_{13}{-}NH_2$ 25

$n{-}C_7H_{15}{-}NH_2$ 26

$n{-}C_8H_{17}{-}NH_2$ 27

CYCLES:

$\triangleright{-}NH_2$ 2

$\diamondsuit{-}NH_2$ 3

$-NH_2$ 4

$-NH_2$ 5

$-NH_2$ 6

$-NH_2$ 7

$HO{-}{-}NH_2$ 18

$H_3C{-}{-}NH_2$ 19

BRANCHED CHAINS:

$-NH_2$ 12

$-NH_2$ 13

$-NH_2$ 14

$-NH_2$ 15

$-NH_2$ 16

$-NH_2$ 17

POLYCYCLES:

$-NH_2$ 8

$-NH_2$ 9

$-NH_2$ 10

$-NH_2$ 11
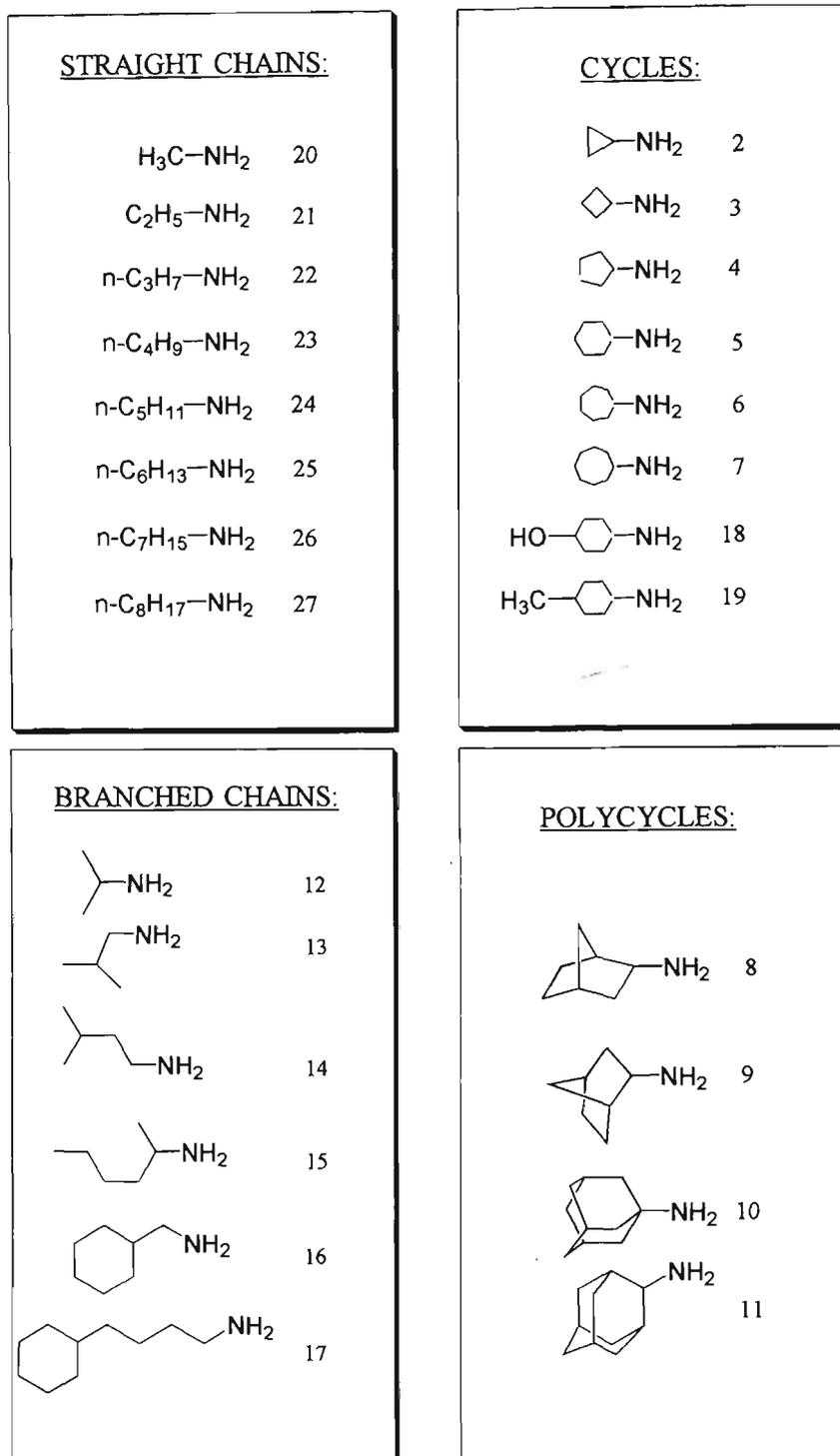
Fig. 3. Amine ligand series employed in this work.

vicinity of the binding site may be characterised schematically as in Fig. 1.

In this work we have directed our initial attention to steric aspects of a potential monodentate precursor [26] represented by the approach of B_1 to B_2, Fig. 2. In choosing this adduct, we assume an initial monodentate attack on the N7 position of the guanine moiety without a replacement of the remaining chloro ligand [27].

It is possible, of course, that the steric demands of both initial and consolidated interactions are relevant to the biological outcomes. In order to investigate the in-

fluence of steric properties on biological activity in the above systems, we have adopted two approaches. Firstly, relative quantitative steric descriptors for the ligands (NH_2R) themselves were calculated. Secondly, analogous steric descriptors were developed for the $PtCl(NH_2R)_2$ moiety, B_1 of Fig. 2, with respect to its approach to the nucleobase, B_2.

The first of these two approaches simplifies the calculation of parameters and, in its nature, parallels the use of traditional substituent constants (steric, electronic and hydrophobic) in classical QSAR studies of organic com-

pounds [28]. The second approach (with respect to steric effects only) is more reflective of ligand/receptor modelling in drug/protein studies and may also provide insight into the mechanism of action of the platinum drugs.

Methods

Compounds

Twenty-seven platinum complexes of general formula cis-[PtL$_2$Cl$_2$], where L is an amine ligand, were chosen as the training set for model building, see Fig. 3. These compounds were selected because their biological profiles have been well characterised [3], Table 3.

Endpoints (dependent variables)

The following biological activity parameters were used for building the model: (i) LD$_{50}$ – acute toxicity, the minimum dose to cause 50% animal death; (ii) ID$_{90}$ – antitumour potency, the minimum dose to cause 90% tumour regression; and (iii) TI – therapeutic index, the measure of the selectivity of the compound as an antitumour agent

TABLE 3
BIOLOGICAL ACTIVITY OF PLATINUM COMPLEXES OF THE TYPE cis-[PtL$_2$Cl$_2$]

No.	Amine ligand L	pLD$_{50}$	pID$_{90}$	log TI
1	Ammonia	1.36	2.27	0.91
2	Cyclopropylamine	0.82	2.22	1.40
3	Cyclobutylamine	0.66	2.15	1.49
4	Cyclopentylamine	−0.04	2.26	2.30
5	Cyclohexylamine	−0.84[a]	1.59	2.43[b]
6	Cycloheptylamine	−0.31	1.81	2.11
7	Cyclooctylamine	−0.10	0.35	0.46
8	endo-2-Aminonorbornane	−0.14	1.61	1.74
9	exo-2-Aminonorbornane	−0.25[a]	1.19	1.44[b]
10	1-Aminoadamantane	−0.04[a]	−0.04[a]	0.00[b]
11	2-Aminoadamantane	−0.15[a]	−0.15[a]	0.00[b]
12	Isopropylamine	1.06	2.63	1.57
13	Isobutylamine	0.69	1.82	1.13
14	Isoamylamine	−0.42	1.88	2.30
15	2-Aminohexane	−0.19	1.23	1.42
16	Cyclohexylmethylamine	0.74	1.16	0.42
17	1-Amino-4-cyclohexylbutane	0.64	0.64	0.00b
18	4-Aminocyclohexanol	1.35	1.71	0.36
19	4-Methyl-cyclohexylamine	−0.30	−0.38	−0.08[b]
20	Methylamine	1.25	1.43	0.19[a]
21	Ethylamine	1.13	1.47	0.34
22	n-Propylamine	1.16	1.50	0.34
23	n-Butylamine	0.57	1.61	1.04
24	n-Pentylamine	0.68	1.07	0.40
25	n-Hexylamine	−0.33	−0.51	−0.15[b]
26	n-Heptylamine	−0.26	−0.26	0.00[b]
27	n-Octylamine	0.42	0.42	0.00[b]

[a] These values, given in the literature as >, were arbitrarily assigned the minimum quoted value.
[b] These values were calculated from the values for potency and toxicity.

(TI = LD$_{50}$/ID$_{90}$). These parameters were measured in the same manner [3] for all the compounds in the series. The units for LD$_{50}$ and ID$_{90}$ were converted from mg/kg to mol/kg.

Descriptors (independent variables)

To model carrier-ligand steric effects in platinum complexes, two descriptors were calculated using the molecular mechanics facilities of HyperChem [29]. These are as follows: (i) LRE – ligand repulsive energy, expressed by the gradient of the van der Waals repulsive energy between the ligand and a Cr(CO)$_5$ fragment, a 'steric probe', to which it binds [23]; and (ii) CRE – complex repulsive energy, expressed by the gradient of the van der Waals repulsive energy between the metal–carrier-ligand moiety and 9-ethylguanine to which it binds, Fig. 2.

Six parameters were chosen to describe the electronic structure of amine ligands in platinum complexes: (i) q(N)$_{PEOE}$ – the partial atomic charge on the amine nitrogen, calculated using the QSAR module of ChemPlus (a set of extension modules to HyperChem), which employs an empirical model built on the partial equalisation of orbital electronegativity (PEOE) method [30]. (ii) q(N)$_{AM1}$ – the partial atomic charge on the amine nitrogen, calculated using the AM1 quantum mechanical method of HyperChem. This method is based on the neglect of diatomic differential overlap (NDDO) approximation [31] and is regarded to be the most accurate semiempirical method of HyperChem and the best method for collecting quantitative information. For this parameter and for q(N)$_{PEOE}$, partial atomic charge/charge density at certain atoms or parts of the molecule can carry information on the reactivity for that part of the molecule [32]. In order to check the suitability of using partial charges in 'amines' for representation in complexes, we calculated nitrogen charges in corresponding Pd complexes using the ZINDO/1 semiempirical method of HyperChem and found a reasonable correlation with the above values. (iii) E$_{HOMO}$ – the energy of the highest occupied molecular orbital of the amine. E$_{HOMO}$ could be a first approximation to the compound's nucleophilicity [33]. (iv) E$_{LUMO}$ – the energy of the lowest unoccupied molecular orbital of the amine. E$_{LUMO}$ could be a first approximation to the compound's electrophilicity [33]. (v) ΔE – the difference in energy levels between E$_{HOMO}$ and E$_{LUMO}$. (vi) pK$_a$ – the acidity constant of the amine [34–36].

Transport effects were modelled using the octanol/water partition coefficient log P, which is a measure of hydrophobicity. The log P values for the coordinated amine ligands were calculated using the QSAR module of ChemPlus. This calculation is based on an additive function of atomic contributions. Atomic parameters and the functionality for the log P calculation are taken from Ref. 37.
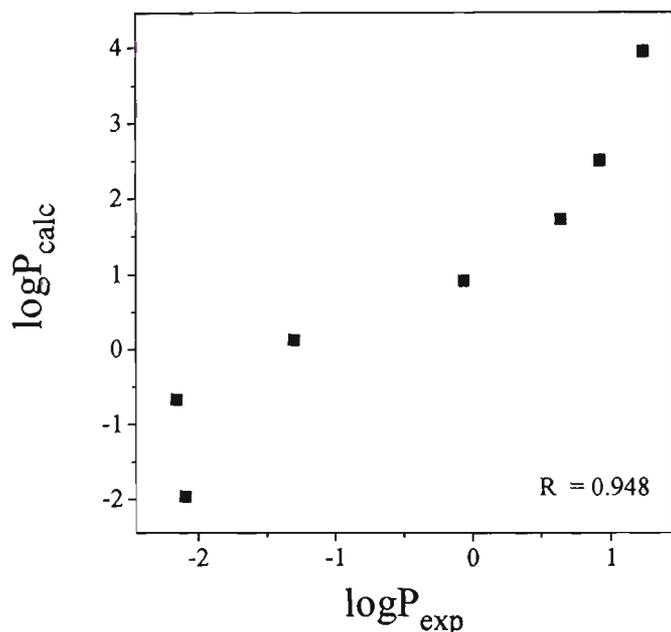
Fig. 4. Plot of calculated log P values (octanol/water) for the coordinated amine ligands in [Pt(NH₂R)₂Cl(9-EtG)] versus experimental log P values (chloroform/water) for the complexes themselves, where R represents cycloalkyl substituents. Experimental values are taken from Ref. 3.

To verify that ligand hydrophobicities may be used to model complex transport effects, all available experimental chloroform/water partition coefficients for the complexes themselves were examined for correlation with the theoretical octanol/water partition coefficients for the ligands themselves, Fig. 4. A good correlation was observed for the sample chosen. All of the above descriptors employed in the present work are listed in Table 4. The van der Waals molecular volumes for the amines were calculated using the QSAR module of ChemPlus, see Table 5.

Computational details

All molecular mechanics and quantum mechanics calculations were carried out using the HyperChem molecular modelling and simulation package, versions 3.0 and 4.0 [29], running on AST 486/33 or COMPAQ Pentium 5100 computers.

Semiempirical calculations

The amines of this investigation were geometry optimised utilising the MM+ force field [29]. The conjugate gradient (Polak–Ribiere) algorithm was employed with the termination condition being an rms of less than 0.1 kcal/(Å mol). The quantum-mechanics-derived descriptors used in this investigation were then obtained using AM1 single-point calculations.

TABLE 4
HYDROPHOBICITY, STERIC AND ELECTRONIC DESCRIPTORS

No.	log P	LRE (kcal/mol)	CRE (kcal/mol)	$q(N)_{AM1}$ (e)	$q(N)_{PEOE}$ (e)	E_{LUMO} (eV)	E_{HOMO} (eV)	pK_a	ΔE (eV)
1	−1.96	10.00	4.33	−0.375	−0.339	4.102	−10.530	9.25	14.632
2	−0.67	23.28	8.67	−0.311	−0.325	2.841	−9.568	9.10	12.409
3	0.12	26.57	9.83	−0.326	−0.325	3.424	−9.720	10.64	13.144
4	0.91	28.83	16.45	−0.330	−0.325	3.449	−9.760	10.65	13.209
5	1.71	31.80	21.34	−0.342	−0.325	3.502	−09.594	10.64	13.096
6	2.50	32.87	15.96	−0.345	−0.325	3.479	−09.583	10.67	13.062
7	3.94	37.02	13.59	−0.342	−0.325	3.464	−09.499	10.65	12.963
8	1.37	32.85	11.21	−0.333	−0.324	3.384	−9.672	10.67	13.056
9	1.37	30.17	13.61	−0.330	−0.324	3.367	−9.670	10.67	13.037
10	1.98	42.93	21.73	−0.322	−0.322	3.350	−9.616	10.14	12.966
11	2.62	35.70	17.25	−0.336	−0.324	3.395	−9.603	10.67	12.998
12	0.01	30.60	17.41	−0.344	−0.325	3.492	−9.689	10.67	13.181
13	0.96	33.92	15.83	−0.343	−0.327	3.522	−9.695	10.48	13.217
14	1.61	32.17	16.14	−0.342	−0.327	3.571	−9.617	10.64	13.188
15	2.57	32.66	12.96	−0.344	−0.325	3.488	−9.648	10.63	13.136
16	2.33	24.13	14.16	−0.343	−0.327	3.486	−9.671	10.67	13.157
17	4.57	23.02	15.91	−0.343	−0.327	3.391	−9.730	10.57	13.121
18	−1.34	30.55	19.55	−0.341	−0.325	3.293	−9.707	10.58	13.000
19	2.37	31.62	13.10	−0.342	−0.325	3.493	−9.600	10.67	13.093
20	−1.46	30.00	12.50	−0.340	−0.330	3.714	−9.819	10.64	13.533
21	−0.78	31.00	15.69	−0.342	−0.328	3.579	−9.740	10.64	13.319
22	0.16	31.00	14.15	−0.342	−0.327	3.594	−9.703	10.60	13.297
23	0.95	31.39	21.57	−0.341	−0.327	3.589	−9.692	10.78	13.281
24	1.74	32.26	23.83	−0.341	−0.327	3.575	−9.695	10.71	13.270
25	2.54	31.68	19.07	−0.341	−0.327	3.565	−9.697	10.63	13.262
26	3.33	31.62	16.07	−0.341	−0.327	3.554	−9.696	10.67	13.250
27	4.12	33.99	27.63	−0.342	−0.327	3.544	−9.699	10.65	13.243

TABLE 5
VAN DER WAALS MOLECULAR VOLUMES OF AMINES

No.	MV (Å³)	No.	MV (Å³)	No.	MV (Å³)
1	22	10	159	19	131
2	67	11	159	20	40
3	82	12	74	21	57
4	97	13	91	22	74
5	114	14	108	23	91
6	130	15	125	24	108
7	147	16	130	25	125
8	121	17	182	26	142
9	121	18	123	27	159

As the systems under study in this work present an appreciable conformational flexibility, two conformational search procedures for a global energy minimum were implemented, namely quenched dynamics followed by simulated annealing, and Monte Carlo conformational search. Both procedures resulted in almost identical final structures. Preference was given to the latter procedure because of its computational efficiency. A more detailed description of force field development, optimisation conditions, global energy minimum conformational searching and a comparison of molecular mechanics simulation results with X-ray analyses will be published separately.

Molecular mechanics

The structures of the chromium series, $Cr(CO)_5L$, used in the LRE calculations were optimised using the MM+ force field modified by adding the parameters presented in Table 6, in addition to those derived by Brown and co-workers [23]. For the CRE calculations the structures of the platinum series, $[PtL_2(9\text{-}EtG)Cl]$, were optimised using the AMBER force field modified by adding parameters developed for the modelling of monoadducts (Table 6), in addition to those derived by Yao et al. [38] for the modelling of bisadducts.

Repulsive energy strategy

To accommodate alternative software [29], the procedure as described in Ref. 23 was slightly modified. Information on these modifications is available on request. Using Ref. 23 as a control, identical outcomes were established for the LRE values calculated for selected amine ligands. The general strategy for the calculation of both LRE and CRE is similar and may be described as follows:

(1) Obtain the lowest energy structure for the complex ($Cr(CO)_5L$ or $[PtL_2(9\text{-}EtG)Cl]$ for LRE or CRE calculation, respectively, where L is an amine ligand).

TABLE 6
FORCE FIELD PARAMETERS

	r_0 (Å)	k_r (mdyn/Å)	k_r (kcal/(mol Å²))[a]	Θ_0 (°)	k_Θ (kcal/(mol rad²))[a]	ϕ_0 (°)	n	V/2 (kcal/mol)[a]	r* (Å)	ε (kcal/mol)
MM+ bond stretch parameters										
Cr-N(sp³)	2.14	2.625[b]								
AMBER parameters[c]										
Bond stretch parameters										
PT-CL	2.305[d]		366							
Angle bend parameters										
NB-PT-CL				90	42					
N3C-PT-CL				90	42					
N3T-PT-CL				180	42					
PT-N3-CT				109.47[d]	20					
PT-N3-H3				109.47[d]	20					
CB-NB-CK				106[e]	70					
Torsional parameters										
CK-NB-PT-CL						90	2	0.25		
CB-NB-PT-CL						90	2	0.25		
Nonbonded parameters[f]										
Atom type PT									1.75	0.1

[a] These force constants were set by analogy to those developed in Ref. 38.

[b] This force constant was set larger than that in Ref. 23 to adjust HyperChem for the inclusion of 1,3-metal centred contacts into the vdW interactions (J. Polowin, personal communication).

[c] Two new atom types were used: N3C (a Pt-bound ligand nitrogen, cis to Cl) and N3T (a Pt-bound ligand nitrogen, trans to Cl). All relevant parameters involving the N3 atom types available in the HyperChem AMBER force field were duplicated for these two atom types.

[d] These values were chosen upon analysis of available crystal structures.

[e] This angle value was found to give a better match to crystal structure data [27] than the angle value developed in Ref. 38.

[f] These parameters were set as suggested in Ref. 29.

(2) For a given complex, vary r_{Me-N} by ± 0.08 Å, with all the other internal coordinates frozen, to create a set of structures.

(3) Using the nonbonded parameters of a modified MM+ force field for both $Cr(CO)_5L$ and $[PtL_2(9\text{-}EtG)Cl]$ series, compute the repulsive portion of E_{vdw} for each structure in the above set according to

$$E_{vdw} \text{ (rep)} = \sum D_0 \exp\left[\gamma\frac{r_0 - r}{r_0}\right] \qquad (1)$$

(4) Calculate the repulsive energy according to

$$E_R = -r_e \left[\partial E_{vdw} \text{ (rep)}/\partial r_{Me-N}\right] \qquad (2)$$

where D_0 represents the potential well depth, γ is a scaling factor (typically 12.5), r is the interatomic distance in the energy-minimised structure and r_0 is the unstrained interatomic distance. r_{Me-N} represents a varied metal-to-nitrogen distance (the Cr-N distance in $Cr(CO)_5$ complexes and the Pt-N7 distance in $[PtL_2(9\text{-}EtG)Cl]$) and r_e is the metal-to-nitrogen distance in the energy-minimised structure.

Data analysis

Data analysis was performed using the SPSS [39] and SCAN [40] software packages. Multiple linear regression analysis was carried out in two ways, *stepwise* and *enter*.

Stepwise regression analysis

This procedure involves a computerised selection of the best single variable and then considers the remaining variables one at a time until the best two (or more)-variable equation is arrived at. A variable enters the equation only if the significance probability, p, associated with the F-test (which tests the hypothesis that the correlation coefficient equals zero) is less than 0.05 (95% confidence level). The process continues until the addition of a variable is not justified by the 'F-statistic'.

Enter regression analysis

This alternative/complementary approach involves the computation of multiple regression with a fixed set of variables. Because situations are possible where significant variables show up only in a certain combination of two or more [17], both of the above approaches were employed as follows. Firstly, stepwise-MLR was applied to pick up the 'best' equations with one or more variables. Secondly, enter-MLR was carried out on intuitively chosen variables to see if an improvement with respect to correlation strength and statistical significance occurs. To avoid generating meaningless regression results, endpoint versus descriptor plots were checked for clusters, outliers, parabolic behaviour, etc., prior to the regression analysis.

The following interpretation of correlation coefficients was implemented: 0–0.2, very weak; 0.2–0.4, weak; 0.4–0.7, moderate; 0.7–0.9, strong; 0.9–1, very strong. The following criteria were implemented for the choice of the best model [41]: to be useful the model should be able to explain the majority of biological variance, i.e. R^2 should be greater than 0.5; an increase of 0.05 to 0.1 in the value of the correlation coefficient R and a similar decrease in the value of the standard error s are estimates of a significant change. The cross-validated variance R_{cv}^2 is a measure of model predictive power (the leave-one-out procedure was used).

Results and Discussion

In this study we adopt the assumption that all of the platinum complexes referred to target the N7 position of guanine. Their individual biological profiles are attributed to the transport, steric and electronic consequences resulting from modification of the amine ligand(s). Thus, the cisplatin analogues chosen for this study feature standard modifications which are expected to influence biological outcomes [28].

LRE as a steric descriptor

It was mentioned previously (vide infra) that the use of MV as a steric descriptor oversimplifies steric effects since MV is a scalar quantity. In the context of this work, there is another consideration which must be taken into account. Not surprisingly, for the ligands of this study MV correlates very well with hydrophobicity as shown in Fig. 5a (R = 0.95). Hence, MV carries with it an undesirable amount of transport information. On the other hand, the steric descriptor LRE is much less correlated with hydrophobicity as shown in Fig. 5b (R = 0.44). It should also be noted that when the sole ligand of the series which may be considered hydrophilic, i.e. NH_3 (circled data point), is omitted from the regression, the correlation coefficient drops to R = 0.22. Indeed, the horizontal array of Fig. 5b suggests a high degree of separation of transport and steric effects, and bolsters our confidence in LRE as a steric descriptor.

CRE as a steric descriptor

Employing a rigid symmetrical steric probe such as $Cr(CO)_5$, the LRE method allows the relative repulsive energies of a series of ligands to be assessed and to be applied to other systems [42]. On the other hand, the CRE method developed in this work is tailored to a particular scenario. Thus, we consider the monodentate interaction of a platinum complex, namely $Pt(NH_2R)_2Cl^+$, with the N7 position of a guanine moiety, Fig. 2. If guanine is considered as the steric probe, since it is obviously not spherically symmetrical with respect to the direction of approach, it is first necessary for the approach to be
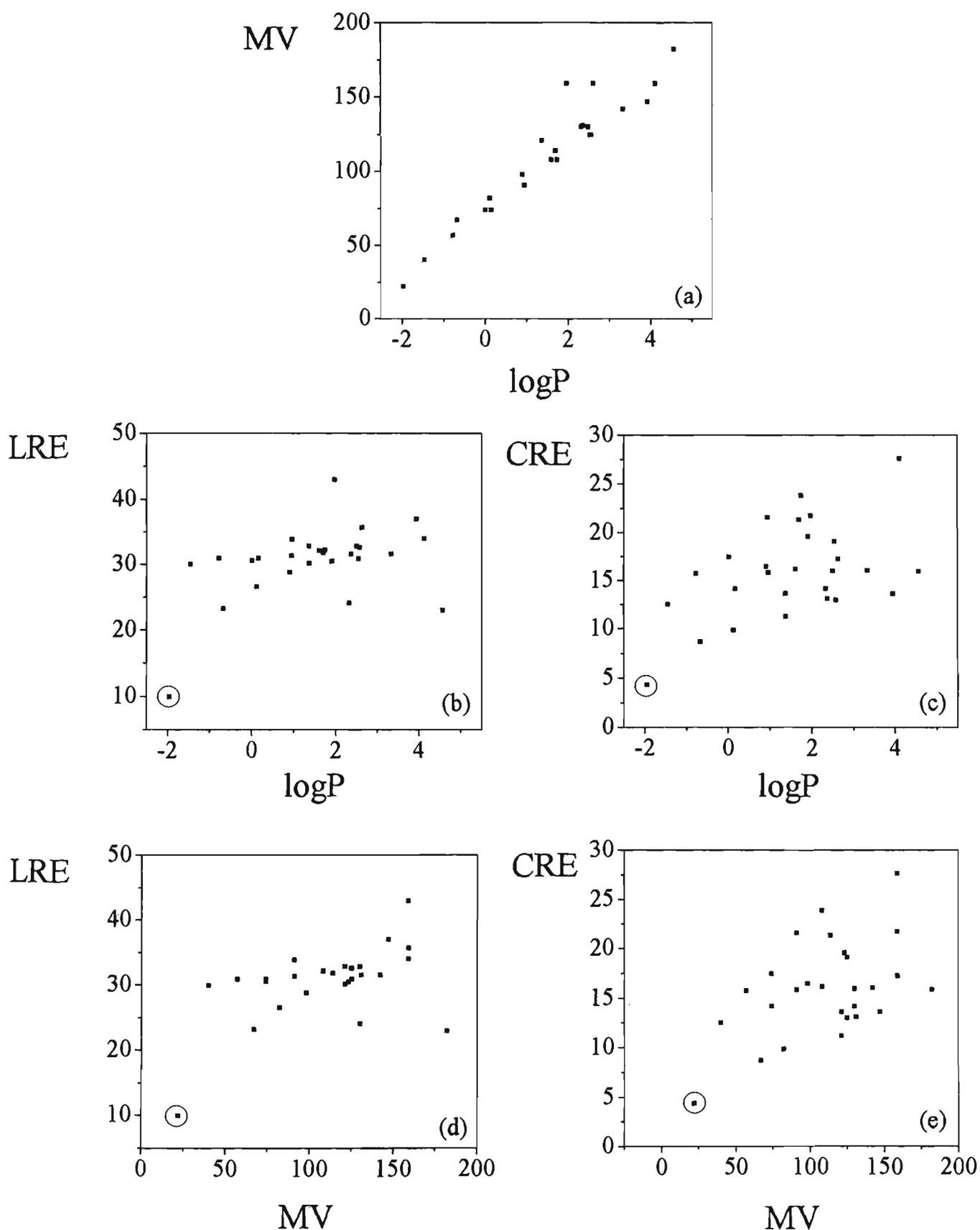
Fig. 5. Relationships between steric and transport descriptors.

optimised. This may be approximated by a global mini-
mum conformation search for the postulated adduct.
During this procedure, the orientation of the nucleobase
to the coordination plane, the orientation of the 9-ethyl
substituent relative to the nucleobase as well as the con-

formation of the carrier ligand are varied. Thus, the CRE
values obtained represent relative measures of steric char-
acteristics presented by each $Pt(NH_2R)_2Cl^+$ moiety to-
wards a specific DNA constituent.

Like LRE, CRE correlates poorly with hydrophobicity

598
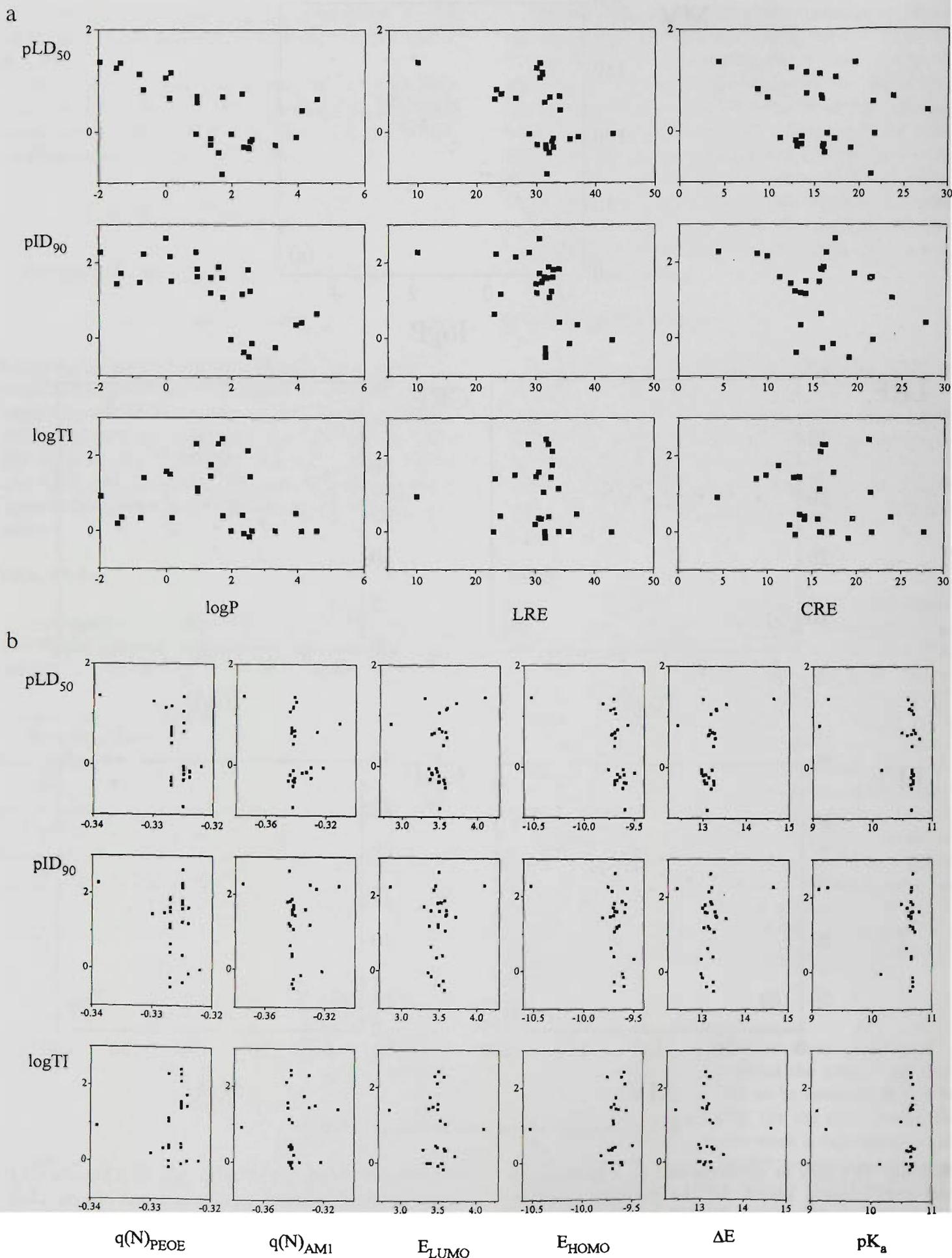
Fig. 6. Endpoints versus (a) transport and steric descriptors, and (b) electronic descriptors.

TABLE 7
CORRELATION MATRIX FOR THE VARIABLES USED FOR THE MLR ANALYSIS

	log P	LRE (kcal/mol)	CRE (kcal/mol)	pK_a	$q(N)_{PEOE}$ (e)	$q(N)_{AMI}$ (e)	E_{LUMO} (eV)	E_{HOMO} (eV)	ΔE (eV)
log P	1								
LRE	0.419	1							
CRE	0.418	0.605	1						
pK_a	0.428	0.560	0.499	1					
$q(N)_{PEOE}$	0.401	**0.739**	0.372	0.434	1				
$q(N)_{AMI}$	0.170	0.547	0.233	0.190	**0.845**	1			
E_{LUMO}	–0.144	–0.290	–0.064	0.114	**–0.761**	**–0.848**	1		
E_{HOMO}	0.495	**0.748**	0.413	0.493	**0.897**	**0.798**	**–0.698**	1	
ΔE	–0.337	–0.551	–0.249	–0.189	**–0.896**	**–0.894**	**0.929**	**–0.913**	1

(R = 0.5; with the NH_3 data point removed, R = 0.37), Fig. 5c. The slightly higher correlation for CRE compared to LRE is perhaps due to a shift of emphasis away from the symmetrical region around the donor nitrogen onto the bulk of the carrier ligand which may impinge on the opposing nucleobase.

Not unexpectedly, given the close relationship between hydrophobicity and MV for these ligands (Fig. 5a), the correlation of both LRE and CRE with MV is also poor (Figs. 5d and e) (R = 0.51 and 0.49 for LRE and CRE, respectively). With the NH_3 data point removed, R = 0.29 and 0.49 for LRE and CRE, respectively.

For the ligands considered here, it is expected that a larger molecular volume will be associated with a more flexible molecule. In a platinum complex, ligand flexibility is expected to increase the likelihood of steric contacts with an opposing nucleobase. The question of the rigidity of steric bulk will be addressed in future studies.

The above considerations of LRE and CRE have the

potential to provide insights into the distribution of sterically significant bulk on the carrier ligand. This may be of particular importance in the design of systems whereby sterically restrictive carrier ligands are employed in an attempt to enforce or manipulate a particular orientation of nucleobase(s) such as 'head-to-tail' or 'head-to-head' [43].

Analysis of simple endpoint–descriptor relationships

For all the series combined (Fig. 3), an analysis of the relationship between the biological activity and the structural descriptors (Figs. 6a and b) reveals the following trends:

(1) Not unexpectedly [28], the biological activity suggests a parabolic dependence on log P, especially the acute toxicity and, to a lesser degree, the therapeutic index, Fig. 6a. This observation prompted us to include the quadratic term $\log^2 P$ into the MLR analysis.

(2) With respect to LRE and CRE, the biological activities do not reveal any obvious correlations, Fig. 6a.

TABLE 8
RESULTS OF THE STEPWISE AND ENTER MLR ANALYSES

Equation	n	s	R	F(p)	R_{cv}^2
Stepwise					
$pLD_{50} = 0.08 \log^2 P - 0.45 \log P + 0.56$	27	0.42	0.78	19.10(0.0000)	0.51
$pID_{90} = -0.35 \log P + 1.69$	27	0.68	0.67	19.85(0.0002)	0.36
$\log TI = -0.07 \log^2 P + 1.20$	27	0.77	0.44	6.14(0.0204)	0.11
Enter with LRE as steric parameter					
$pLD_{50} = 0.09 \log^2 P - 0.48 \log P - 0.01 LRE + 0.27 pK_a - 1.96$	27	0.42	0.80	10.49(0.0001)	0.53
$pID_{90} =$	27	0.66	0.72	8.48(0.0006)	0.42
$-0.34 \log P - 0.067 LRE + 1.8 E_{HOMO} + 21*$	27	0.68	0.70	7.46(0.0012)	0.31
$-0.31 \log P - 0.037 LRE - 0.63 E_{LUMO} + 5.0$					
$\log TI =$	27	0.74	0.59	2.92(0.044)	0.10
$-0.13 \log^2 P + 0.27 \log P - 0.07 LRE + 72 q(N)_{PEOE} + 27*$	27	0.75	0.57	2.72(0.056)	0.14
$-0.15 \log^2 P + 0.32 \log P - 0.05 LRE - 0.42 E_{LUMO} + 4.2$					
Enter with CRE as steric parameter					
$pLD_{50} = 0.095 \log^2 P - 0.51 \log P - 0.02 CRE + 0.14 pK_a - 0.58$	27	0.42	0.80	19.10(0.0000)	0.52
$pID_{90} = -0.35 \log P - 0.02 CRE + 0.33 E_{HOMO} + 5.19$	27	0.71	0.67	19.85(0.0002)	0.33
$\log TI = -0.12 \log^2 P + 0.26 \log P - 0.04 CRE - 0.11 pK_a + 2.8$	27	0.77	0.54	6.14(0.0204)	0.10

Equations marked with an asterisk have a strong degree of collinearity (see Table 7), so the second equation is provided.
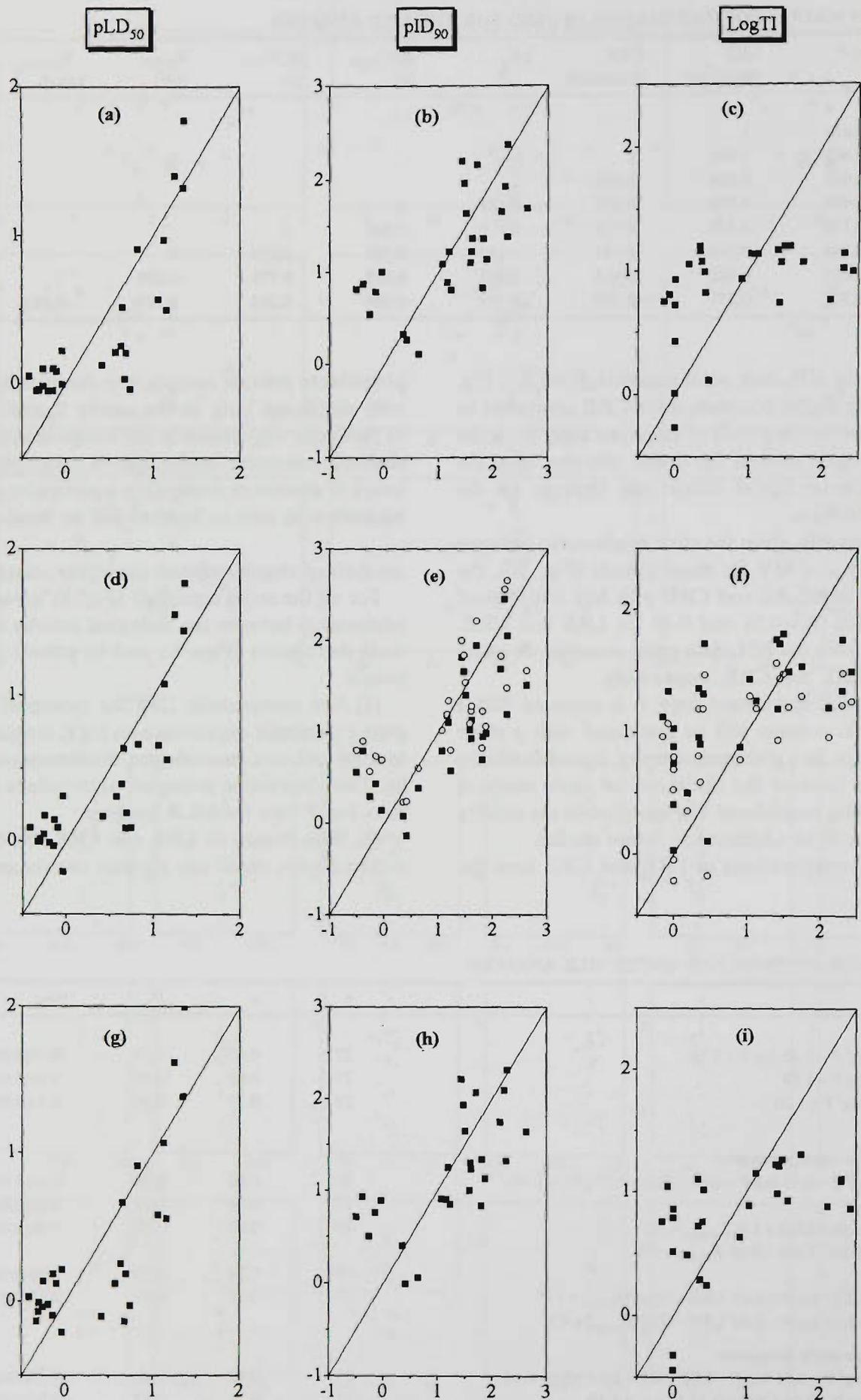
Fig. 7. Plots of calculated versus observed values for the biological activity indicators. Data in plots a, b and c: calculated according to equations from Table 8, upper part; d, e and f: equations from Table 8, middle part; g, h and i: equations from Table 8, lower part. In plots e and f, ■ represents equations marked with an asterisk and ○ represents second equations provided for pID_{90} and log TI.
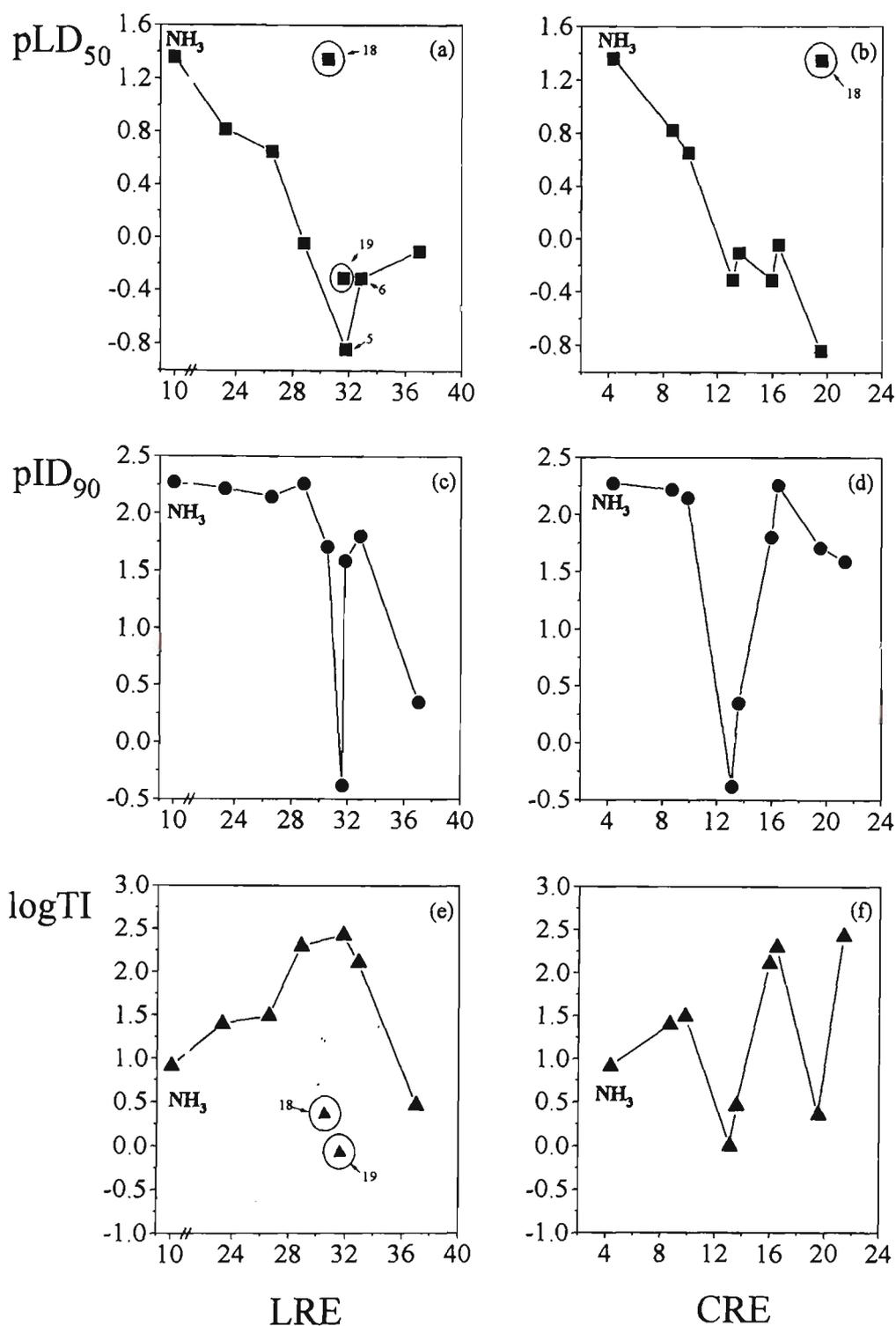
Fig. 8. Biological activity indicators versus steric descriptors LRE and CRE for cyclic systems.

(3) A degree of independence of biological activity with respect to electronic descriptors is suggested, at least for the amines under study. This is manifested by the vertical arrays of Fig. 6b. Notably, for all the electronic descriptors, the ligand NH_3 does not conform to this trend. The ligand cyclopropylamine also lies off the trend for a number of descriptors, namely $q(N)_{AMI}$, E_{LUMO}, ΔE and pK_a. Subsequent MLR analysis was carried out both with and without these two complexes and the results did not appear to be significantly different.

Collinearity

All structural descriptors used for the MLR analysis were checked for collinearity. The correlation matrix (Table 7) shows that the hydrophobicity does not correlate with any individual steric or electronic parameter used. For the two steric parameters, LRE correlates with two out of the six electronic parameters, although not strongly. The steric parameters LRE and CRE are not strongly correlated with one another, reflecting the different physical phenomena underlying their calculation. As
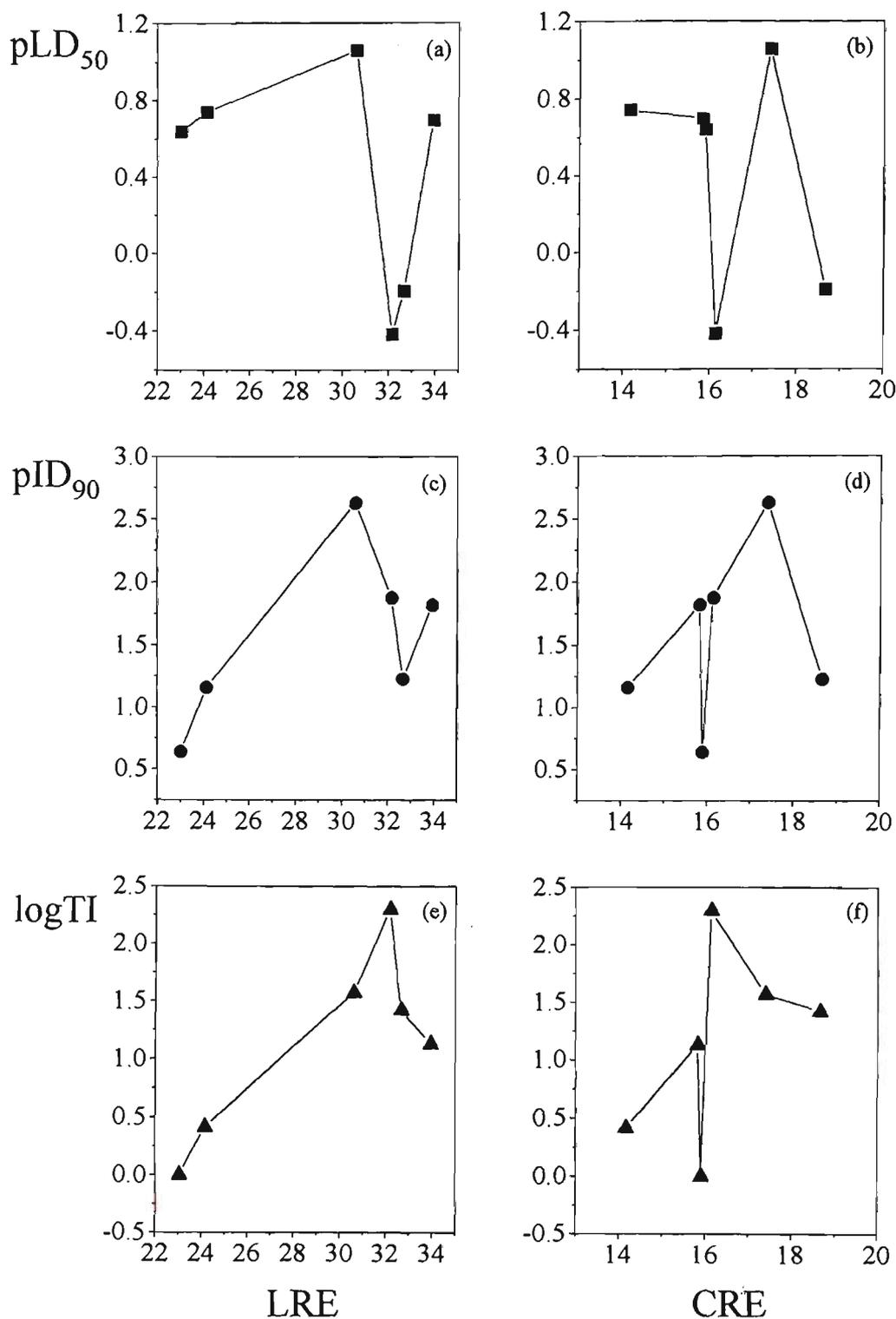
Fig. 9. Biological activity indicators versus steric descriptors LRE and CRE for branched-chain systems.

expected, electronic descriptors are strongly interrelated, except for pK_a.

Stepwise models

Table 8 (upper part) shows *stepwise* regression equations together with their goodness-of-fit (correlation strength), goodness-of-prediction and statistical significance parameters. The plots of the values of pLD_{50}, pID_{90} and $\log TI$ based on these equations are shown in Figs. 7a–c.

It appears from stepwise regression analysis that the biological activity of platinum complexes can be modelled quite well by carrier-ligand hydrophobicity. That is, the most significant factor influencing biological activity in this case appears to be transport, i.e. accumulation and distribution. However, it must be realised that factors which are operative at the target site may well be masked by transport effects. That is, the biological activity of a complex often depends on its ability to bind to target sites in a specific way, not simply to get there. This may
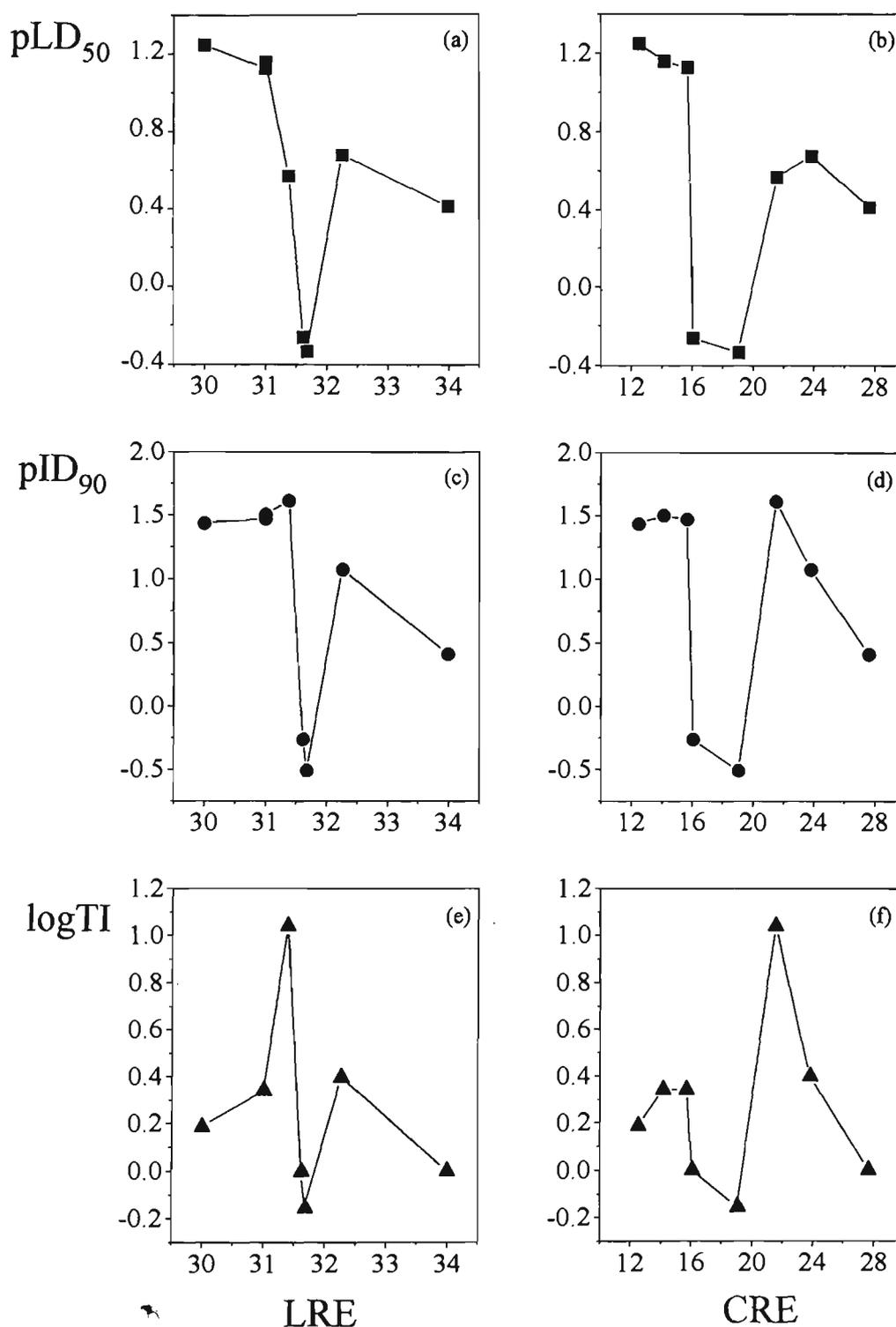
Fig. 10. Biological activity indicators versus steric descriptors LRE and CRE for straight-chain systems.

be accounted for by introducing second (and, if necessary, third) terms into the QSAR, in particular those that account for steric and/or electronic effects.

Enter models

Table 8 (middle and lower parts) show the 'best' *enter* regression equations together with their goodness-of-fit (correlation strength), goodness-of-prediction and statistical significance parameters. The plots of the values of pLD_{50}, pID_{90} and log TI based on these equations are

shown in Figs. 7d–i, and show that the equations obtained are fairly predictive of toxicity and potency. The QSAR models for the therapeutic index, however, did not achieve the quality that was hoped for initially. A better choice of initial descriptors or a better formulation of the model (through different chemometric methods) is suggested.

The regression results presented here (using LRE and CRE as steric descriptors) show a marginal improvement (with respect to overall correlation strength and statistical
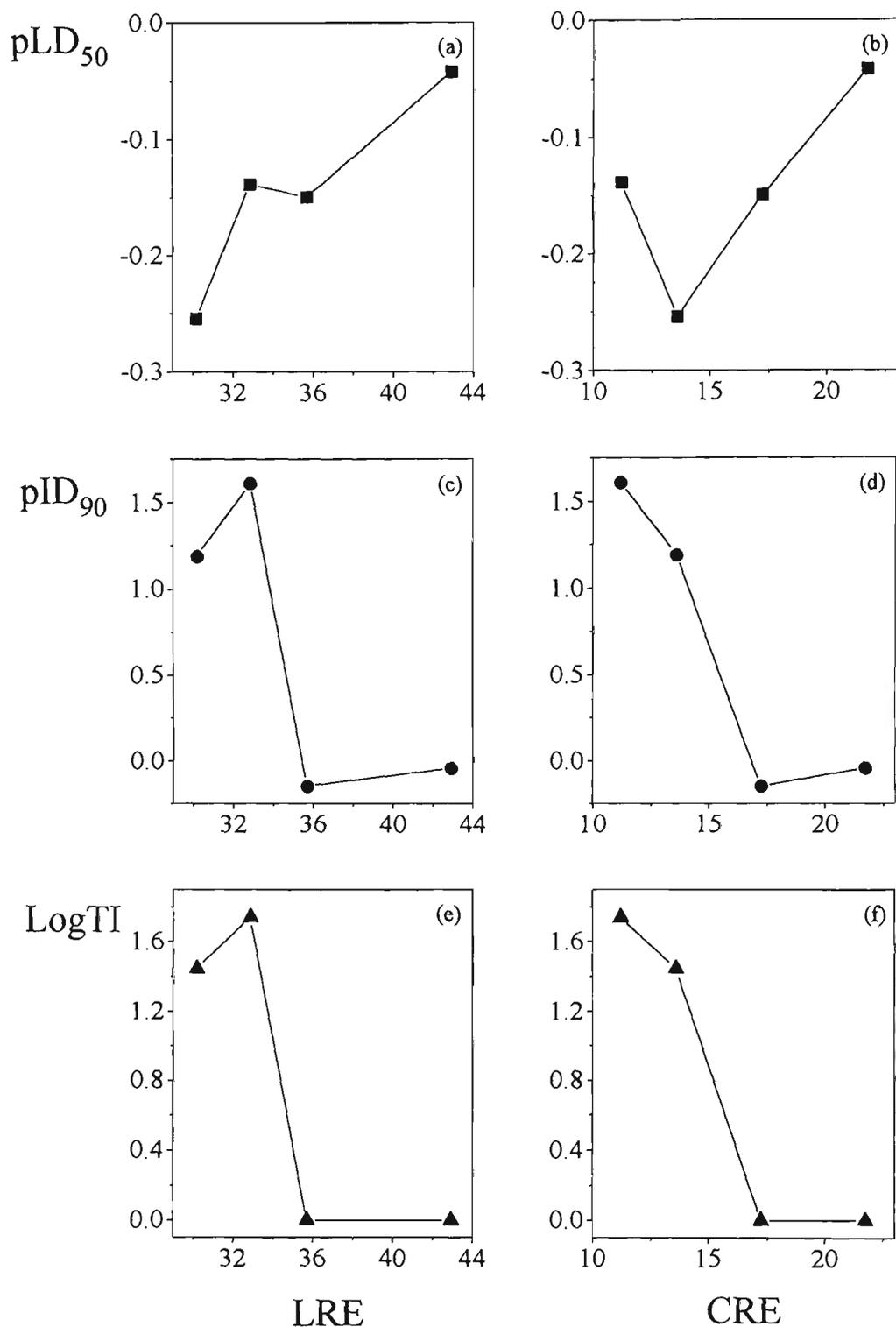
Fig. 11. Biological activity indicators versus steric descriptors LRE and CRE for polycyclic systems.

significance) in comparison with the *stepwise* models. However, the partial correlation coefficients and the statistical significance of variables describing steric and electronic effects are smaller than those of $\log P$ (data are not given here). In some cases these regression equations show an improvement in goodness-of-fit compared to those obtained in previous studies [4–9]; however, it is impossible to compare their predictive power since no cross-validation information is given in these works. In contrast to some previous studies [4,9], our results do not

show a predominant dependency of biological activity on electronic structure. Rather, the importance of transport effects is emphasised. They also reveal a consistent contribution of steric effects to the biological outcomes.

Relationship of biological profiles with repulsive energies

Whilst it is acknowledged that the biological activity of platinum–am(m)ine complexes is dominated by their transport properties, reflected by a strong dependency on hydrophobicity as revealed by QSAR results (vide

infra), steric and electronic effects *after* the molecule has reached its target site may become determinative of their relative activities. We have attempted to extract further information on steric effects by looking for trends involving biological indicators and LRE and CRE within given series of ligands.

For example, the relationship between toxicity (pLD_{50}) and LRE is examined in Figs. 8a, 9a, 10a and 11a. It is apparent for cyclic (cycles), straight-chain (straights) and branched-chain (branches) systems that a pattern emerges. Interestingly, with increasing LRE, the toxicity tends to decrease to a minimum value of around 32 kcal/mol for each series and then rises again. This might be suggestive of an optimal steric requirement for minimum toxicity within a given series. Note that in Fig. 8a the circled data points representing ligands **18** (and possibly **19**) do not conform to the described trend. These ligands are characterised by having OH and CH_3 substituents in the 4-position of the ring. The remote location of these substituents from the donor amino group does not significantly influence their LRE values, as expected. Thus, **5**, **18** and **19** are in an approximately vertical array. It is interesting to note, however, that the absolute toxicity relating to ligand **18** is enhanced, which may be due to the additional hydrogen-bonding capability of the hydroxyl group, stabilising a drug–DNA adduct [44]. The absolute toxicity relating to ligand **19** is of the same order of magnitude as that for **6**; this is anticipated, considering that the hydrophobic character of both is very close. Complexes containing polycyclic ligands (polycycles) (Fig. 11a) can be considered as presenting only the 'right shoulder' of the above trend. This might be expected since there is a paucity of polycyclic hydrocarbons smaller than norbornane.

While toxicity versus LRE plots for individual series of compounds show that the toxicity passes through a minimum, the regression analysis with respect to toxicity for all compounds combined did not reveal a parabolic dependence on LRE. Thus, such trends appear to be series specific.

The relationship between the antitumour potency (pID_{90}) and LRE is shown in Figs. 8c, 9c, 10c and 11c, and reflects the trend observed for pLD_{50}. These trends could, perhaps, be exploited for ligands of these types for rational drug design. A full explanation of such patterns of behaviour is more difficult to address.

The relationship between therapeutic index (log TI) and LRE is shown in Figs. 8e, 9e, 10e and 11e. Plots of log TI versus LRE do not reveal any obvious pattern of behaviour. However, it is of interest to note that for cycles and branches, the ligand responsible for maximum log TI corresponds to the same ligand responsible for minimum pLD_{50}. This is not the case for straights and polycycles. In the case of straights it can be rationalised by the obser-

vation that they are less sterically diverse as a series (in terms of LRE), vide infra. Consequently, from a steric point of view, they do not differ significantly from each other and a variation in therapeutic index cannot be rationalised through steric effects. Series of platinum complexes with polycyclic amine ligands suffer from the lack of data. The availability of biological data for a wider range of polycyclic ligand complexes could probably lead to some rationale.

Biological profiles versus CRE (Figs. 8–11; b, d and f) show similar trends to those of LRE with some variations. In particular, for cycles the strongest repulsion between $Pt(NH_2R)_2Cl$ and 9-EtG corresponds to minimum toxicity, i.e. the optimal steric size of the ligand corresponds to minimum toxicity. But this trend is not general (as in e.g. straight chains). Presumably, other factors (both steric and nonsteric) contribute to toxicity and may predominate in some cases. However, as in the case of LRE, optimal values of CRE (15–20 kcal/mol) for minimum toxicity can be cautiously suggested. The variations between LRE and CRE trends are to be expected, since in $Pt(NH_2R)_2Cl(9\text{-EtG})$ systems the amine ligands are separated from the 'steric probe' by platinum. Furthermore, in these systems the repulsive contributions of two amine moieties are embedded into the calculations, and these two moieties are nonequivalent, one being cis to the N7 position and the other trans, see Fig. 2.

Conclusions

LRE and CRE parameters as applied to the approach of a metal complex to a target molecule allow for an improved description of steric effects in such systems. Where transport, electronic and steric effects are all to be considered, the LRE and CRE parameters have a higher degree of orthogonality than those employed previously. Future attempts to separate steric and transport effects and to expose purely steric requirement may include a consideration of steric effects in complexes containing ligands of like hydrophobicity. Other directions to investigate the scope of repulsive energy strategies for studying steric effects in biologically active metal complexes may well include quantitative structure–property relationship (QSPR) studies, allowing a focus on kinetic and/or thermodynamic properties. In addition, other chemometric techniques (e.g. PLS, PCA, neural nets [45]) may be more successful in finding a better QSAR/QSPR model in such systems, although our preliminary results show no improvement in goodness-of-fit and goodness-of-prediction in comparison with the results from MLR for the models, containing log P and CRE, log P and LRE, log P and pK_a, where some degree of collinearity is present. The repulsive energy strategy reported here is to be extended to other systems for which biological activity information is available, namely mixed and bidentate carrier ligands,

other possible targets on DNA, bisadducts and other biologically active metal complexes (e.g. Ref. 46).

Acknowledgements

References

1 Barton, J.K., In Bertini, I., Gray, H.B., Lippard, S.J. and Valentine, J.S. (Eds.) Bioinorganic Chemistry, University Science Books, Mill Valley, CA, U.S.A., 1994, pp. 455–504.

2 Sherman, S.E. and Lippard, S.J., Chem. Rev., 87 (1987) 1153.

3 Braddock, P.D., Connors, T.A., Jones, M., Khokhar, A.R., Melzack, D.H. and Tobe, M.L., Chem.-Biol. Interact., 11 (1975) 145.

4 Abdul-Ahad, P.G. and Webb, G.A., Int. J. Quantum Chem., 21 (1982) 1105.

5 Simon, Z., Mracec, M., Maurer, A. and Policec, S., Rev. Roum. Chim., 25 (1980) 713.

6 Simon, Z., Holban, S., Mracec, M., Maurer, A., Policec, S. and Dragulescu, C., Rev. Roum. Biochim., 14 (1977) 117.

7 Romanowska, K. and Kudak-Jarowska, J., Arch. Immunol. Ther. Exp., 39 (1991) 75.

8 Romanowska, K., Int. J. Quantum Chem., 43 (1992) 175.

9 Tang, W., Qu, Y. and Dai, A., Pure Appl. Chem., 60 (1988) 1271.

10 Hansch, C., Venger, B.H. and Panthananickal, A., J. Med. Chem., 23 (1980) 459.

11 Hansch, C., Acc. Chem. Res., 26 (1993) 147. The public domain subset of the database can be reached at http://fox.pomona.claremont.edu/chem/qsar-db/index.html.

12 Quinn, F.R. and Milne, G.W.A., Fundam. Appl. Toxicol., 6 (1986) 270.

13 NCI Data, source a.panthananickal bio_1865. From Hansch, C., Acc. Chem. Res., 26 (1993) 147.

14 Charton, M., In Charton, M. and Motoc, I. (Eds.) Steric Effects in Drug Design, Springer, Berlin, Germany, 1983, pp. 107–118.

15 Motoc, I., In Charton, M. and Motoc, I. (Eds.) Steric Effects in Drug Design, Springer, Berlin, Germany, 1983, pp. 93-105.

16 Ghose, A.K. and Crippen, G.M., J. Med. Chem., 28 (1985) 333.

17 Hansch, C., In Chapman, N.B. and Shorter, J. (Eds.) Correlation Analysis in Chemistry, Plenum, New York, NY, U.S.A., 1978, pp. 397–438.

18 Bersuker, I.B. and Dimoglo, A.S., In Lipkowitz, K.B. and Boyd, D.B. (Eds.) Reviews in Computational Chemistry, Vol. 2, VCH, New York, NY, U.S.A., 1991, pp. 423–460.

19 Maliski, E.G. and Bradshaw, J., In Ganellin, C.R. and Roberts, S.M. (Eds.) Medicinal Chemistry: The Role of Organic Chemistry in Drug Research, Academic Press, London, U.K., 1993, pp. 83–102.

20 Belsley, D.A., Kuh, E. and Welsch, R.E., Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, Wiley, New York, NY, U.S.A., 1980.

21 Tolman, C.A., Chem. Rev., 77 (1977) 313.

22 White, D., Taverner, B.C., Leach, P.G.L. and Coville, N.J., J. Comput. Chem., 14 (1993) 1042.

23 Choi, M.-G. and Brown, T.L., Inorg. Chem., 32 (1993) 1548.

24 Woo, T.K. and Ziegler, T., Inorg. Chem., 33 (1994) 1857.

25 Orbell, J.D., Marzilli, L.G. and Kistenmacher, T.J., J. Am. Chem. Soc., 103 (1981) 5126.

26 Orbell, J.D., Solorzano, C., Marzilli, L.G. and Kistenmacher, T.J., Inorg. Chem., 21 (1982) 3806.

27 Lippert, B., In Lippard, S.J. (Ed.) Progress in Inorganic Chemistry, Vol. 37, Wiley, New York, NY, U.S.A., 1989, pp. 2–97.

28 Silverman, R.B., The Organic Chemistry of Drug Design and Drug Action, Academic Press, San Diego, CA, U.S.A., 1992.

29 HyperChem is a commercial molecular modelling package from Hypercube Inc., Waterloo, ON, Canada. The MM+ force field of HyperChem is an extension of the MM2 code developed by N. Allinger. It uses the latest MM2 (1991) parameters with the 1977 functional form. HyperChem supplements the standard MM2 by providing additional parameters.

30 Gasteiger, J. and Marsili, M., Tetrahedron, 36 (1980) 3219.

31 Stewart, J.J.P., J. Comput.-Aided Mol. Design, 4 (1990) 1.

32 Klein, C.L. and Stevens, E.D., In Liebman, J.F. and Greenberg, A. (Eds.) Structure and Reactivity, VCH, New York, NY, U.S.A., 1988, pp. 25–64.

33 Fleming, I., Frontier Orbitals and Organic Chemistry Reactions, Wiley, Chichester, U.K., 1989.

34 Perrin, D.D., Dissociation Constants of Organic Bases in Aqueous Solutions, Butterworths, London, U.K., 1965.

35 Smith, R.M. and Martell, A.E., Critical Stability Constants, Plenum, New York, NY, U.S.A., 1975.

36 Perrin, D.D., Dempsey, B. and Serjeant, E.P., pK_a Prediction for Organic Acids and Bases, Chapman and Hall, London, U.K., 1981.

37 Ghose, A.K., Pritchett, A. and Gordon, M.C., J. Comput. Chem., 9 (1988) 80.

38 Yao, S., Plastaras, J.P. and Marzilli, L.G., Inorg. Chem., 33 (1994) 6061.

39 Nie, N.H., Bent, D.H. and Hull, C.H., Statistical Package for the Social Sciences, McGraw-Hill, New York, NY, U.S.A., 1970. Version 6.1 for Windows was used in this work.

40 SCAN, Software for Chemometric Analysis, 1995 by Minitab Inc. Release 1 for Windows was used in this work.

41 Plummer, E.L., In Lipkowitz, K.B. and Boyd, D.B. (Eds.) Reviews in Computational Chemistry, Vol. 1, VCH, New York, NY, U.S.A., 1990, pp. 119–168.

42 Choi, M.-G. and Brown, T.L., Inorg. Chem., 32 (1993) 5603.

43 Baker, A.T., Crass, J.K., Kok, G.B., Orbell, J.D. and Yuriev, E., Inorg. Chim. Acta, 214 (1993) 169.

44 Reedijk, J., Inorg. Chim. Acta, 198–200 (1992) 873.

45 Haswell, S.J. (Ed.) Practical Guide to Chemometrics, Marcel Dekker, New York, NY, U.S.A., 1992.

46 Kuo, L.Y., Kanatzidis, M.G., Sabat, M., Tipton, A.L. and Marks, T.J., J. Am. Chem. Soc., 113 (1991) 9027.

APPENDICES

Auxilliary computing

Ia. Sample preparation sheet for ratio-dependent binding experiments

```
'Qbasic routines. Calculate required volumes of components and solvent and the
'resulting concentrations of components. The variations allow constant or gradually-
'changing increments in component ratio values, as well as other options.
'First program is used for continuous variation experiments.
DECLARE FUNCTION w! (x!, y!)
DateOfSamplePreparing$ = "25/6/96"
A$ = "Pd(en)CL2"
B$ = "GUA"
solvent$ = "DMSO"
method$ = "R-dep UV - cont. var."
SampleVolume = 5
MinR = .1
MaxR = 4
ConcA = .0001
ConcB = .0001
NSamples = 15
Excess = 1.05
TotalConc = .0001
PRINT "Ratio Dependent Calculations4 - Continuous variations"
PRINT "Date Of Samples Preparing - "; DateOfSamplePreparing$
PRINT "Metal Complex - "; A$; ","; " DNA constituent - "; B$
PRINT "Initial concentrations : "; A$; " -"; ConcA; "M, "; B$; " -"; ConcB; "M"
PRINT "Volume of Sample  "; " ="; SampleVolume; "ml"
PRINT "Total concentration of components in the sample -"; TotalConc; "M"
PRINT "Number of Samples ="; NSamples
PRINT
PRINT "N"; TAB(6); "Ratio", "Volume", "Volume", " Conc. in Sample"
PRINT TAB(15); A$, B$, A$, B$,
PRINT TAB(15); "ml", "ml", "mMol", "mMol"
StepR = (MaxR - MinR) / (NSamples - 1)
FOR R = MinR TO MaxR * 1.001 STEP StepR
    SampleVolumeA = TotalConc * SampleVolume / (R * ConcA + ConcA)
    SampleVolumeB = R * TotalConc * SampleVolume / (R * ConcB + ConcB)
    ConcAinSample = 1000 * ConcA * SampleVolumeA / SampleVolume
    ConcBinSample = 1000 * ConcB * SampleVolumeB / SampleVolume
    j = j + 1
    PRINT j; TAB(6); w(R, 2), w(SampleVolumeA, 3), w(SampleVolumeB, 3),
    w(ConcAinSample, 3), w(ConcBinSample, 3)
    VolumeA = VolumeA + SampleVolumeA * Excess
    VolumeB = VolumeB + SampleVolumeB * Excess
    NEXT R
PRINT
PRINT "N"; TAB(6); "RATIO", "Solvent Volumes For Dilution (ml)"
PRINT TAB(15); "to Sample", "to"; A$; "ref", "to"; B$; "ref"
```

```
FOR R = MinR TO MaxR * 1.001 STEP StepR
    h = h + 1
    SampleVolumeA = TotalConc * SampleVolume / (R * ConcA + ConcA)
    SampleVolumeB = R * TotalConc * SampleVolume / (R * ConcB + ConcB)
    VolumeAddTo = SampleVolume - SampleVolumeA - SampleVolumeB
    VolumeAddToA = SampleVolume - SampleVolumeA
    VolumeAddToB = SampleVolume - SampleVolumeB
    PRINT h; TAB(6); w(R, 2), w(VolumeAddTo, 3), w(VolumeAddToA, 3),
    "          ",
    w(VolumeAddToB, 3)
NEXT R
PRINT
PRINT "Volume of Solution of "; A$; " ="; w(VolumeA, 3); "ml"
PRINT "Volume of Solution of "; B$; " ="; w(VolumeB, 3); "ml"
END
FUNCTION w (x, y)
w = INT(x * 10 ^ y + .5) / 10 ^ y!
END FUNCTION


'Second program is used for molar ratio experiments.
DECLARE FUNCTION w! (x!, y!)
WantToPrint = 0
DateOfSamplePreparing$ = "  /7/96"
A$ = "Pd(en)CL2"
B$ = "GUA"
solvent$ = "H2O"
method$ = "R-dep NMR; constant conc. of metal complex"
SampleVolume = 10
MinR = .1
MaxR = 10
ConcA = .001
ConcB = .001
ConcAinSample = .0001
SampleVolumeA = ConcAinSample * SampleVolume / ConcA
NSamples = 25
Excess = 1.05
PRINT "Ratio Dependent Calculations6"
PRINT "Date Of Samples Preparing - "; DateOfSamplePreparing$
PRINT "Metal Complex - "; A$; ","; " DNA constituent - "; B$
PRINT "Method - "; method$
PRINT "Solvent - "; solvent$
PRINT "Initial concentrations: "; A$; " -"; ConcA; "M, "; B$; " -"; ConcB; "M"
PRINT "Concentration of "; A$; " in sample -"; ConcAinSample; "M"
PRINT "Volume of "; A$; " in sample -"; w(SampleVolumeA, 3); "ml"
PRINT "Volume of Sample  "; " ="; SampleVolume; "ml"
PRINT "Number of Samples ="; NSamples
PRINT
PRINT "N"; TAB(6); "Ratio", "Volume", "Volume", " Conc. in Sample"
PRINT TAB(15); B$, solvent$, B$
PRINT TAB(15); "ml", "ml", "mMol"
StepR = (MaxR - MinR) / (NSamples - 1)
```

```
FOR R = MinR TO MaxR * 1.001 STEP StepR
   SampleVolumeB = ConcAinSample * SampleVolume * R / ConcB
   SolventVolume = SampleVolume - SampleVolumeA - SampleVolumeB
   ConcBinSample = 1000 * ConcB * SampleVolumeB / SampleVolume
   j = j + 1
   PRINT j; TAB(6); w(R, 2), w(SampleVolumeB, 3), w(SolventVolume, 3),
   w(ConcBinSample, 2)
   VolumeB = VolumeB + SampleVolumeB * Excess
NEXT R
VolumeA = SampleVolumeA * NSamples * Excess
PRINT
PRINT "Volume of Solution of "; A$; " ="; w(VolumeA, 3); "ml"
PRINT "Volume of Solution of "; B$; " ="; w(VolumeB, 3); "ml"
END
FUNCTION w (x, y)
w = INT(x * 10 ^ y + .5) / 10 ^ y!
END FUNCTION
```

Ib. Transformation of a Varian format table into an ASCII file for importing into Origin

```
'Qbasic routine. Transforms Varian table format into ASCII format. Number of files
needs to be specified in the first line.
FOR NumRep = 1 TO 3
a$ = LTRIM$(STR$(NumRep))
OPEN "a:\varian\cary13e\report\report." + a$ FOR INPUT AS #1
OPEN "a:\repliza." + a$ FOR OUTPUT AS #2
CLS
FOR i = 1 TO 2
  INPUT #1, l$
  PRINT l$
NEXT i
DO UNTIL EOF(1)
  INPUT #1, x
  IF x = 0 THEN EXIT DO
  INPUT #1, y
  PRINT #2, x, y
  PRINT x, y
LOOP
CLOSE
NEXT NumRep
END
```

Ic. Dihedral angle calulation

```
'QBasic routine. Calculates dihedral angle between planes given parameters of two
'planes Ax_i + By_i + Cz_i +D = 0 (parameter D is not used in the calculation). Based 'on
"Modern 'Mathematical Analysis", M.H.Protter and C.B.Morrey, Jr, p.18.
'parameters of two planes (type in and press F5):
A1 = -1.12
```

```
B1 = -.71
C1 = -1
D1 = -24.86
A2 = -.53
B2 = 2.18
C2 = -1
D2 = -8.87
'Calculate the dihedral angle between two planes:
CONST PI = 3.141592654#
verch = A1 * A2 + B1 * B2 + C1 * C2
IF verch < 0 THEN
verch = -verch
END IF
cosinus = verch / (SQR(A1 ^ 2 + B1 ^ 2 + C1 ^ 2) * SQR(A2 ^ 2 + B2 ^ 2 + C2 ^ 2))
sinus = SQR(1 - cosinus ^ 2)
tangens = sinus / cosinus
angle = (ATN(tangens)) * 180 / PI
PRINT "Dihedral angle="; angle
```

Id. Extraction of vdW information from the HyperChem output file

*'QBasic routine. Cuts-out vdW data from a series of HyperChem output files (e.g.
'meoet#.log, where # is a file number) and creates a set of data-files for input into
'Origin for repulsive energy calculation*
```
FOR NumLog = 1 TO 9
a$ = LTRIM$(STR$(NumLog))
OPEN "C:\hyper\lre-cre\meoet" + a$ + ".log" FOR INPUT AS #1
OPEN "C:\hyper\lre-cre\meoet" + a$ + ".dat" FOR OUTPUT AS #2
CLS
   DO
   LINE INPUT #1, stroka$
   LOOP UNTIL LEFT$(stroka$, 5) = "Atom "
   DO
   LINE INPUT #1, stroka$
   PRINT #2, stroka$
   LOOP UNTIL LEFT$(stroka$, 5) = "Bond "
CLOSE #1
CLOSE #2
NEXT NumLog
END
```

Ie. Transformation of HyperChem output file into an ASCII file for importing into
Origin

*'Qbasic routine. Transforms a single column HyperChem output file into five
column 'table file for input into Origin for aset of rotated structures, e.g. nucleobase
in 'platinum complex is rotated along Pt-N bond: 1 - torsion angle, 2 - total-energy,
'3 - stretch-energy, 4 - bend-energy, 5 - torsion-energy, 6 - nonbond-energy,*

'7 - hbond-energy.
DIM A(7) AS SINGLE
OPEN "C:\hyper\rotate\ptnh3\dummy.log" FOR INPUT AS #1
OPEN "C:\hyper\rotate\ptnh3\dummy.dat" FOR OUTPUT AS #2
WHILE NOT EOF(1)
FOR i = 1 TO 7 STEP 1
INPUT #1, A(i)
NEXT i
PRINT #2, A(1), A(2), A(3), A(4), A(5), A(6), A(7)
WEND
CLOSE #1
CLOSE #2
END

If. Single point calculations for a set of structures with varied metal-ligand distance

*;HyperChem script. Sequentially opens # structures with varied metal-ligand
;distance, carries out single point calculation with the predetermined force field
;options and saves the created information into *.log output file.*
file-format hin
open-file MEOET#.hin
select-none
start-logging MEOET#.log,no
mechanics-print-level 3
do-single-point
stop-logging
next #

Ig. Rotation of a pre-selected part of the molecule around the metal-to-ligand bond

*;HyperChem script. Performs the rotation (molecule is pre-aligned), saves the value
;of the new torsion angle and information about total energy and energy-term values
;into a log-file; the output *.log without tags is created for input into the QBasic
;program (see Appendix Ie). When the script stops running because of the filling up
of ;the stack memory and HyperChem sends a warning message and a question about
;aborting the script, the answer should be "NO". It will then take through the
;sequence of dialog boxes and finally stop it. The answer "YES" will stop it
;immediately, but will add text lines at the end of log-file, which will affect its
;treatment in QBasic. The script should be initially started with 0° rotation to save
;information for the initial structure. This is not required when restarting the script
;after "stack memory stop".*

append-omsgs-to-file dummy.log
menu-edit-rotate
query-response-has-tag no
query-value named-selection-value 1
do-single-point
query-response-has-tag no

```
query-value total-energy
query-value stretch-energy
query-value bend-energy
query-value torsion-energy
query-value nonbond-energy
query-value hbond-energy
read-script rot-db1a.scr
```

Ih. Repulsive energy calculation

'Origin script.
'The first part is enabled only for Cr complexes where 1,3 vdW contacts are included
'in the calculation. Prior to the calculation the line numbers for relevant contacts are
'determined and assigned to variables a, b, c, d and e. The procedure sequentially
'handles # data-files.

```
a=218
b=176
c=130
d=80
e=37
open -w meoet$(#).dat
col(E)=col(Kv)*exp(12.5*(col(Rv)-col(R))/col(Rv))
import_E[a]=0
import_E[b]=0
import_E[c]=0
import_E[d]=0
import_E[e]=0
wo -s 6 0 6 0
Erep=sum(import_E)
type -a "$(sum.total)"
Evdwrepvs.R_EvdwREP[1]=(sum.total)
next #
```
'thus obtained values of Evdw (rep) are plotted against varied metal-ligand distance
'and the slope valeu is calculated by the following sub-script
```
Menu id = 3009
del %C.LINE;
LR %C;
type -a;
type Linear Regression for %C:;
type "Y = A + B * X";
ShowLinearFit;
```

Appendix II

Analytical treatment of the stepwise complex formation

Section 2.1.1.2 details miscellaneous approaches to the determination of formation constants, employing different means of simplification or approximation. However, a general analytical treatment of stepwise complex formation is a desirable goal, which may be achieved by employing the Newton formula for a binomial. Thus, the term $(b - nc)^n$ from eqn **2.1** may be expressed as:

$$(b - nc)^n = \sum (-1)C_{n,k} b^{n-k}(nc)^k$$

where $C_{n,k}$ are binomial coefficients from the Pascal triangle. The sum in this equation is opened as follows:

$$(b - nc)^n = C_{n,0} b^n(nc)^0 - C_{n,1} b^{n-1}(nc)^1 +$$

$$+ C_{n,2} b^{n-2}(nc)^2 - \ldots \pm C_{n,n} b^0(nc)^n$$

Two points should be considered at this stage. Firstly, since for weak complex(es) the concentration of a complex c has a very small value, the value of c^2 is negligible in comparison with b; hence, only the first two members of the above series are to be considered. Secondly, according to the Pascal triangle for binomial coefficients: $C_{n,0} = 1$ and $C_{n,1} = n$.
Thus,

$$(b - nc)^n = b^n - n^2 b^{n-1} c$$

Substituting this expression into eqn **2.1** gives the general equation for weak complexes:

$$\frac{1}{K} = \frac{ab^n \varepsilon^\star}{\Delta A} - b^n - an^2 b^{n-1} + \frac{n^2 \Delta A b^{n-1}}{\varepsilon^\star} \qquad (\text{II.1})$$

For 1:1 complexation (n = 1), this is identical to eqn **2.5**, but because $\Delta A/\varepsilon^\star$ (last member in eqn **2.7**) equals c, this term may be neglected, thus resulting in eqn **2.6**.

For 2:1 complexation (n = 2), the equation still contains the term c :

$$\frac{ab^2}{\Delta A} = \frac{1}{\beta_2 \varepsilon^\star} + \frac{4ab + b^2 - 4bc}{\varepsilon^\star} \qquad (\text{II.2})$$

Favourable experimental conditions and iterative techniques may also be applied to the general eqn **2.7** derived here. That is, for 2:1 complexation (n = 2) and $a \gg b$ or $a \gg b$,

the general equation becomes:

$$\frac{ab^2}{\Delta A} = \frac{1}{\beta_2 \varepsilon^\star} + \frac{4ab}{\varepsilon^\star} \qquad \text{(II.3)}$$

or

$$\frac{ab^2}{\Delta A} = \frac{1}{\beta_2 \varepsilon^\star} + \frac{b^2}{\varepsilon^\star} \qquad \text{(II.4)}$$

These are overall, not stepwise, formation constants. But under the specified conditions it is assumed that the reaction is forced to completion. Eqns II.3 and II.4 are analogous to those reported in Ref. [1].

References

1. Moriguchi, I.; Nayuki, M.; Kaneniwa, N. *Chem. Pharm. Bull.* 1969, *17*, 1339.

Cambridge Structural Database entries referred to throughout the thesis

REFCODE	COMPOUND (database nomenclature preserved)	REFERENCE (database referencing style preserved)
AFCYPT	Chloro-ethylenediamine-(1-(beta-D-arabino-furanosyl)cytosine) platinum (ii) chloride	Acta Cryst., B, 34, 1838,1978
AMPTMN	bis(bis(*mu*-1-Methylthyminato)-*cis*-diammine-platinum(ii))-manganese(ii) dichloride decahydrate	Inorg.Chim.Acta, 56, 15,1981
BAGKOJ10	Hydroxo-*cis*-diammine-(1-methylcytosine-N$3!)-platinum(ii) nitrate dihydrate at -65°C	Inorg.Chem., 21, 1936,1982
BAGKUP10	Aqua-*cis*-diammine-(1-methylcytosine-N$3!)-platinum(ii) dinitrate monohydrate	Inorg.Chem., 21, 1936,1982
BAHNUT	Diethylenetriamine-7,9-dimethylguanine-platinum(ii) bis(hexafluorophosphate)	Inorg.Chem., 20, 1835,1981
BAPKAE	*cis*-Diammine-(1-methylcytosine-N$3!)-(thyminato-N$1!)-platinum(ii) perchlorate	Inorg.Chem., 20, 2381,1981
BARZOJ	*cis*-Diammine-bis(1-methylcytosine)-platinum(ii) dinitrate 1-methylcytosine solvate	J.Am.Chem.Soc., 103, 5126,1981
BARZOJ01	*cis*-Diammine-bis(1-methylcytosine-N$3!)-platinum(ii) dinitrate 1-methylcytosine solvate	Inorg.Chem., 21, 3210,1982
BEKKOR	bis(bis(*mu*!3$-1-Methyluracil-N,O,O')-(tetra-ammine)-di-platinum)-silver pentanitrate tetrahydrate	Inorg.Chem., 21, 451,1982
BOHDAD	*cis*-Chlorodiammine-(N$2!,N$2!-dimethyl-9-methylguanine)-platinum(ii) hexafluorophosphate	Inorg.Chem., 21, 3806,1982
BOSSOR	bis(1-Methyluracilato-N$3!)-*cis*-diammine-platinum(ii) tetrahydrate	J.Am.Chem.Soc., 104, 6596,1982
CCENPT	*cis*-Dichloro-ethylenediamine-platinum(ii)	Acta Cryst., B, 31, 1672,1975

Refcode	Compound	Reference
CHLPTM	cis-Dichloro-bis(cyclohexylamine-N)-platinum(ii) hexamethylphosphoramide solvate	Acta Cryst., B, 36, 1789,1980
CIPMUJ	bis(mu!3$-1-Methyluracilato)-bis(cis-di-ammine-platinum(ii))-dinitrato-di-silver dinitrate dihydrate at -40°C	Inorg.Chem., 23, 1713,1984
COKRUP	bis((mu!2$-1-Methyluracilato-N$3!,O$4!)-(mu!2$-1-methylcytosine-N$3!,O$2!)-cis-diammine-platinum(ii))-copper(ii) tetranitrate hexahydrate	Inorg.Chem., 23, 2807,1984
CTSPTA	Chloro-cis-diammine-(1-methylcytosine-N!3$)-platinum(ii) nitrate (space group P21/c)	Inorg.Chem., 20, 335,1981
CTSPTA01	Chloro-cis-diammine-(1-methylcytosine-N!3$)-platinum(ii) nitrate (space group C2/c)	Inorg.Chem., 20, 335,1981
CTYEPT	Chloro-(thyminato)-(ethylenediamine)-platinum(ii)	Inorg.Chem., 19, 295,1980
CTYEPU	Chloro-(uracilato)-(ethylenediamine)-platinum(ii) hydrogen chloride dihydrate	Inorg.Chem., 19, 295,1980
CUKRAB	cis-Dichloro-diammine-platinum(ii) dimethylformamide solvate	Inorg.Chim.Acta, 106, 141,1985
CUKPUT	tris(1-Methylcytosine-N$3!)-ammine-platinum(ii) diperchlorate monohydrate	Inorg.Chim.Acta, 106, 75,1985
CUSRAJ	18-Crown-6 bis(dimethylacetamide) bis(cisplatin)	J.Chem.Soc.,Chem.Comm., 532,1985
CYTPTA10	cis-Diammine-(1-methylcytosine-N$3!)-(9-ethylguanine-N$7!)-platinum(ii) diperchlorate	Inorg.Chem., 21, 3216,1982
CYTPTD10	cis-Diammine-(1-methylcytosine-N$3!)-(9-ethylguaninato-N$7!)-platinum(ii) perchlorate tetrahydrate	Inorg.Chem., 21, 3216,1982
DASCUV	Nona-ammine-(mu!3$-9-ethylguanine-N$1!,N$3!,N$7!)-tri- platinum(ii) tetra-ammine-platinum(ii) hexanitrate monohydrate	J.Am.Chem.Soc., 107, 3591,1985
DEYXUA	Chloro-(1-methylthyminato-N$3!)-cis-diammine-platinum(ii) monohydrate	Inorg.Chim.Acta, 106, 177,1985
DODLAJ	cis-(9-Methyladenine-N$7!)-(1-methylcytosine-N$3!)-diammine-platinum(ii) diperchlorate monohydrate	Inorg.Chem., 24, 4001,1985

Code	Name	Reference
DODLEN	*trans*-(9-Methyladenine-N7!)-(1-methyl-cytosine-N3!)- diammine-platinum(ii) diperchlorate	Inorg.Chem., 24, 4001,1985
DOKJIW	bis(*mu*!2$-1-Methyluracil-N,O)-tetra-ammine-dichloro-di-platinum(iii) dichloride hydrate	Inorg.Chem., 25, 407,1986
DUJNIF	bis(*cis*-bis(Diammine-chloro-(1-methyl-cytosine-N))-platinum) tetracyano-platinum	Inorg.Chim.Acta, 124, 207,1986
DUJNOL	bis(*cis*-bis(Diammine-(1-methylcytosine-N))-platinum) bis(tetracyano-platinum) tetrahydrate	Inorg.Chim.Acta, 124, 207,1986
DUPGAW	*cis*-Ammine-dichloro-cyclobutylamine-platinum(ii)	Acta Cryst, C, 42, 1291,1986
DUXBED	Aqua-bis(*mu*!2$-(1-methyluracilato-O,N))-(nitrato-O)-tetra-ammine-di-platinum(iii) trinitrate trihydrate	Inorg.Chem., 25, 3384,1986
DUXBIH	Aqua-bis(mu!2$-(1-methyluracilato-O,N))-(nitrato-O)-tetra-ammine-di-platinum(iii) trinitrate dihydrate	Inorg.Chem., 25, 3384,1986
ENXPTA10	Diethylenetriamine-7,9-dimethylhypoxanthine-platinum(ii) bis(hexafluorophosphate) sesquihydrate	Inorg.Chem., 20, 1835,1981
ENXPTB10	(Ethylenediamine)-bis(7,9-dimethyl-hypoxanthine)-platinum(ii) bis(hexafluorophosphate)	J.Inorg.Biochem., 16, 33,1982
FATBIL	*cis*-bis((mu!2$-1-Methyluracilato-N$3!,O$4!)-(mu!2$-1-methylcytosine-N$3!,O$2!)-platinum(ii)-(mu!2$-aqua)-aqua-(mu!2$-nitrato-O,O')-(nitrato-O)-di-silver) dinitrate monohydrate	Inorg.Chim.Acta, 135, 155,1987
FAWYEH	Dichloro-(N,N-dimethylethylenediamine-N,N')-platinum(ii)	Acta Cryst., C, 43, 57,1987
FAZZAH	bis(Tetra-n-propylammonium) *cis*-dichloro- bis(methylamine)-platinum(ii) tetrachloro-platinum	J.Chem.Soc.,Chem.Comm., 443,1987
FEGBUO	*trans,trans,trans*-bis(Ammine-hydroxy-(1-methylcytosine-N$-3!)) - platinum(iv) dinitrate dihydrate	J.Am.Chem.Soc., 108, 3680,1986

Refcode	Compound	Reference
FITFUJ	(SP-4,2-(1alpha,2alpha,4alpha,5alpha))-Dichloro-(4,5-di-hydroxy-cyclohexane-1,2-diamine-N,N')-platinum(ii) monohydrate	J.Med.Chem., 30, 1327,1987
GALYAT	cis-Dichloro-trihydro-1,3-cyclopentadiamine-platinum(ii)	Jiegou Huaxue, 6, 240,1987
GEBVUE	trans-Diammine-bis(1-methylcytosine-N$3!)-platinum(ii) diperchlorate	Acta Cryst., C, 44, 611,1988
GECGOK	cis-Dichloro-(N,N'-di-t-butylethylenediamine-N,N')-platinum(ii)	Inorg.Chem., 27, 2422,1988
GECGUQ	cis-Dichloro-(N,N'-bis(R,S-alpha-methyl-benzyl)ethylenediamine-N,N')-platinum(ii)	Inorg.Chem., 27, 2422,1988
GEWGIY	(Dioxocyclam-N,N')-dichloro-platinum(ii) dihydrate	J.Chem.Soc.,Chem.Comm., 1166,1988
GIFVUM	cis-Dichloro-bis(dimethylamine-N)-platinum(ii)	Inorg.Chim.Acta, 153, 45,1988
GIFWAT	Cytosinium trichloro-(cytosine-N$3!)-platinum(ii) at -110°C	Inorg.Chim.Acta, 153, 31,1988
IPSMCP	trans-Dichloro-(bis(isopropyl)sulfoxide) (1-methylcytosine) platinum(ii)	Can.J.Chem., 54, 53,1976
JEHPOB	cis-Chloro-(1-methylcytosine-N)-(N,N,N',N'-tetramethyl- ethylenediamine)-platinum(ii) perchlorate hemihydrate	Acta Cryst., C, 46, 1326,1990
JIKJEP	cis-bis(1-Methylcytosine-N)-(N,N,N',N'-tetramethylethylenediamine-N,N')-platinum(ii) diperchlorate monohydrate	Acta Cryst.,C, 47, 852,1991
JILMIA	bis((mu!2$-1-Methylcytosine-N,N')-diammine-palladium(ii)) dinitrate dihydrate	Inorg.Chem., 30, 884,1991
JOMCOD	trans-bis(Methylamino)-(1-methylcytosine-N$3!)-(9-methyl-guanine-N$7!)-platinum cytosine chloride hydrate at -120°C	J.Am.Chem.Soc., 114, 357,1992
KABJUS	cis-(Chloro-diammine-(cytosine-N$3!)-platinum(ii)) chloride	J.Med.Chem., 32, 128,1989
KETDAO	Diaqua-diammine-bis(1-methyluracil-N)-platinum dinitrate	Z.Naturforsch., B, 45, 731,1990
KUKGUS	(Ethylenediamine)-bis(9-methylguanine)-platinum(ii) trihydrate	Inorg.Chem., 31, 2429,1992

Code	Name	Reference
MCPTEG10	cis-Diammine-(1-methylcytosine-N$3!)-(9-ethylguanine-N7!)-platinum(ii) cis-diammine-(1-methylcytosine-N$3!)-(9-ethylguaninato-N$7!)-platinum(ii) triperchlorate	Inorg.Chem., 21, 3216,1982
MCSPTA	trans-Dichloro-ammine-(1-methylcytosine-N$3!)-platinum(ii) hemihydrate	Inorg.Chem., 20, 808,1981
MCSPTB	trans-Diammine-bis(1-methylcytosine-N$3!)-platinum(ii) dinitrate	Inorg.Chem., 20, 808,1981
MZOEPT	(mu-3,3,6,6-Tetramethyl-3,6-diazaoctane-1,8-diyl)-bis (N,N,N',N'-tetramethylethylenediamine-chloro-platinum(ii)) diperchlorate	Inorg.Chim.Acta, 37, L545,1979
PASHIA	Chloro-(dimethylsulfoxide)-(N,N,N',N'-tetramethylethyl-enediamine)-platinum(ii) tetraphenylborate	Inorg.Chem., 31, 5061,1992
SEBVEA	Diammine-(1-methyl-2-oxo-4-hydroxy-pyrimidine-N$3!)-(1-methyl-2-oxopyrimidine-4-olate-N$3!)-platinum(ii) nitrate dihydrate at -100°C	J.Am.Chem.Soc., 111, 7213,1989
SEBVIE	Diammine-bis(1-methyl-2-oxo-4-hydroxy-pyrimidine-N$3!)-platinum(ii) hexachloro platinum dihydrate Diammine- bis(1-methyluracil-N$3!)-platinum(ii) hexachloro-platinum(ii) dihydrate	J.Am.Chem.Soc., 111, 7213,1989
SISCIG	(Diethylenetriamine-N,N',N'')-(9-methyladenine-N$1!)-platinum(ii) dinitrate monohydrate	Inorg.Chim.Acta, 176, 113,1990
VANPIJ	cis-Ammine-dichloro-cyclopentylamine-platinum(ii) hydrate	Acta Cryst, C, 45, 852,1989
VEPROX	trans-(bis(Aminomethyl)-(N-glycinate)-(1-methylcytosine-N)-platinum(ii) nitrate dihydrate	Inorg.Chim.Acta, 169, 195,1990
VERDUR	bis(Diethylenetriamine-N,N',N'')-(mu!2$-9-methylguanine-N,N')-di-platinum(ii) triperchlorate dihydrate	Inorg.Chem., 29, 1417,1990
VERFAZ	Diammine-(diethylenetriamine-N,N',N'')-(mu!2$-9-methyl-guanine-N,N')-(1-methyluracil-N)-di-platinum(ii) diperchlorate hydrate	Inorg.Chem., 29, 1417,1990
VOLCUU	cis(Diammine-(glycine-N)-(1-methylcytosine-N$3!)-platinum nitrate dihydrate	Inorg.Chim.Acta, 184, 209,1991

Appendix IV

Atomic coordinates of optimised structures (PDB format)

Table IV.1 [Pt(NH₃)₂(9-EtG)₂]

Unmodified MM+

```
HETATM    1 PT1  NUL     1       0.540    2.215    0.604
HETATM    2 N11  NUL     1      -1.456    2.165    0.450
HETATM    3 N10  NUL     1       0.313    1.654    2.514
HETATM    4 N11  NUL     1       2.584    1.102   -3.769
HETATM    5 C21  NUL     1       2.340    1.933   -4.692
HETATM    6 N21  NUL     1       2.834    1.698   -5.915
HETATM    7 N31  NUL     1       1.658    2.982   -4.506
HETATM    8 C41  NUL     1       1.187    3.225   -3.361
HETATM    9 C51  NUL     1       1.360    2.437   -2.293
HETATM   10 C61  NUL     1       2.108    1.319   -2.456
HETATM   11 O61  NUL     1       2.398    0.522   -1.595
HETATM   12 N71  NUL     1       0.717    2.938   -1.225
HETATM   13 C81  NUL     1       0.202    4.079   -1.711
HETATM   14 N91  NUL     1       0.470    4.307   -3.009
HETATM   15 C91  NUL     1       0.062    5.396   -3.860
HETATM   16 C91  NUL     1      -0.593    6.570   -3.121
HETATM   17 N12  NUL     1       4.521   -0.505    1.627
HETATM   18 C22  NUL     1       5.616    0.130    1.607
HETATM   19 N22  NUL     1       6.754   -0.550    1.797
HETATM   20 N32  NUL     1       5.683    1.378    1.412
HETATM   21 C42  NUL     1       4.626    2.036    1.211
HETATM   22 C52  NUL     1       3.400    1.497    1.198
HETATM   23 C62  NUL     1       3.294    0.168    1.438
HETATM   24 O62  NUL     1       2.263   -0.460    1.519
HETATM   25 N72  NUL     1       2.472    2.433    0.946
HETATM   26 C82  NUL     1       3.210    3.548    0.835
HETATM   27 N92  NUL     1       4.526    3.367    1.040
HETATM   28 C92  NUL     1       5.517    4.378    1.299
HETATM   29 C92  NUL     1       5.299    5.020    2.675
HETATM   30 H    NUL     1      -1.749    2.500   -0.490
HETATM   31 H    NUL     1      -1.798    1.191    0.587
HETATM   32 H    NUL     1      -1.881    2.778    1.176
HETATM   33 H    NUL     1       1.241    1.613    2.984
HETATM   34 H    NUL     1      -0.292    2.336    3.015
HETATM   35 H    NUL     1      -0.128    0.712    2.557
HETATM   36 H    NUL     1      -0.852    7.388   -3.833
HETATM   37 H    NUL     1      -1.538    6.261   -2.618
HETATM   38 H    NUL     1       0.093    6.999   -2.354
HETATM   39 H    NUL     1       0.962    5.771   -4.403
HETATM   40 H    NUL     1      -0.648    4.996   -4.621
HETATM   41 H    NUL     1      -0.381    4.756   -1.068
HETATM   42 H    NUL     1       3.122    0.289   -3.968
HETATM   43 H    NUL     1       2.613    2.410   -6.607
HETATM   44 H    NUL     1       3.365    0.835   -6.010
HETATM   45 H    NUL     1       6.538    3.934    1.244
HETATM   46 H    NUL     1       5.455    5.154    0.499
HETATM   47 H    NUL     1       6.072    5.796    2.881
HETATM   48 H    NUL     1       4.302    5.513    2.740
HETATM   49 H    NUL     1       5.357    4.257    3.486
HETATM   50 H    NUL     1       7.598    0.015    1.762
HETATM   51 H    NUL     1       6.640   -1.553    1.931
HETATM   52 H    NUL     1       4.524   -1.486    1.795
HETATM   53 H    NUL     1       2.758    4.534    0.639
HETATM   54 H    NUL     1       0.617    0.684    0.080
HETATM   55 H    NUL     1       0.463    3.748    1.130
CONECT    1    2    3   12   25   54   55
CONECT    2    1   30   31   32
CONECT    3    1   33   34   35
```

```
CONECT     4     5    10    42
CONECT     5     4     6     7
CONECT     6     5    43    44
CONECT     7     5     8
CONECT     8     7     9    14
CONECT     9     8    10    12
CONECT    10     4     9    11
CONECT    11    10
CONECT    12     1     9    13
CONECT    13    12    14    41
CONECT    14     8    13    15
CONECT    15    14    16    39    40
CONECT    16    15    36    37    38
CONECT    17    18    23    52
CONECT    18    17    19    20
CONECT    19    18    50    51
CONECT    20    18    21
CONECT    21    20    22    27
CONECT    22    21    23    25
CONECT    23    17    22    24
CONECT    24    23
CONECT    25     1    22    26
CONECT    26    25    27    53
CONECT    27    21    26    28
CONECT    28    27    29    45    46
CONECT    29    28    47    48    49
CONECT    30     2
CONECT    31     2
CONECT    32     2
CONECT    33     3
CONECT    34     3
CONECT    35     3
CONECT    36    16
CONECT    37    16
CONECT    38    16
CONECT    39    15
CONECT    40    15
CONECT    41    13
CONECT    42     4
CONECT    43     6
CONECT    44     6
CONECT    45    28
CONECT    46    28
CONECT    47    29
CONECT    48    29
CONECT    49    29
CONECT    50    19
CONECT    51    19
CONECT    52    17
CONECT    53    26
CONECT    54     1
CONECT    55     1
```

Modified MM+

```
HETATM     1  PT1 NUL      1      0.521   2.245   0.629
HETATM     2  N11 NUL      1     -1.501   2.177   0.470
HETATM     3  N10 NUL      1      0.318   1.648   2.559
HETATM     4  N11 NUL      1      2.576   1.107  -3.782
HETATM     5  C21 NUL      1      2.334   1.935  -4.709
HETATM     6  N21 NUL      1      2.833   1.697  -5.928
HETATM     7  N31 NUL      1      1.653   2.986  -4.529
HETATM     8  C41 NUL      1      1.188   3.232  -3.383
HETATM     9  C51 NUL      1      1.367   2.455  -2.309
HETATM    10  C61 NUL      1      2.107   1.332  -2.468
HETATM    11  O61 NUL      1      2.395   0.537  -1.603
HETATM    12  N71 NUL      1      0.708   2.960  -1.251
```

```
HETATM   13  C81  NUL      1       0.202    4.114   -1.709
HETATM   14  N91  NUL      1       0.487    4.311   -3.002
HETATM   15  C91  NUL      1       0.071    5.395   -3.860
HETATM   16  C91  NUL      1      -0.590    6.571   -3.130
HETATM   17  N12  NUL      1       4.520   -0.491    1.627
HETATM   18  C22  NUL      1       5.621    0.133    1.595
HETATM   19  N22  NUL      1       6.752   -0.560    1.784
HETATM   20  N32  NUL      1       5.702    1.380    1.397
HETATM   21  C42  NUL      1       4.647    2.044    1.207
HETATM   22  C52  NUL      1       3.416    1.521    1.216
HETATM   23  C62  NUL      1       3.296    0.193    1.453
HETATM   24  O62  NUL      1       2.260   -0.425    1.542
HETATM   25  N72  NUL      1       2.502    2.466    0.935
HETATM   26  C82  NUL      1       3.225    3.590    0.831
HETATM   27  N92  NUL      1       4.525    3.372    1.064
HETATM   28  C92  NUL      1       5.527    4.377    1.321
HETATM   29  C92  NUL      1       5.306    5.030    2.691
HETATM   30  H    NUL      1      -1.795    2.509   -0.472
HETATM   31  H    NUL      1      -1.835    1.201    0.606
HETATM   32  H    NUL      1      -1.933    2.787    1.195
HETATM   33  H    NUL      1       1.253    1.606    3.016
HETATM   34  H    NUL      1      -0.284    2.319    3.078
HETATM   35  H    NUL      1      -0.117    0.703    2.593
HETATM   36  H    NUL      1      -0.852    7.384   -3.846
HETATM   37  H    NUL      1      -1.535    6.261   -2.627
HETATM   38  H    NUL      1       0.093    7.008   -2.365
HETATM   39  H    NUL      1       0.967    5.771   -4.407
HETATM   40  H    NUL      1      -0.639    4.987   -4.618
HETATM   41  H    NUL      1      -0.371    4.798   -1.064
HETATM   42  H    NUL      1       3.109    0.289   -3.980
HETATM   43  H    NUL      1       2.619    2.410   -6.622
HETATM   44  H    NUL      1       3.364    0.834   -6.018
HETATM   45  H    NUL      1       6.545    3.923    1.280
HETATM   46  H    NUL      1       5.482    5.148    0.516
HETATM   47  H    NUL      1       6.081    5.806    2.895
HETATM   48  H    NUL      1       4.310    5.527    2.748
HETATM   49  H    NUL      1       5.359    4.274    3.509
HETATM   50  H    NUL      1       7.603   -0.003    1.755
HETATM   51  H    NUL      1       6.625   -1.560    1.920
HETATM   52  H    NUL      1       4.514   -1.472    1.795
HETATM   53  H    NUL      1       2.783    4.579    0.628
HETATM   54  H    NUL      1       0.418    3.772    1.166
HETATM   55  H    NUL      1       0.624    0.719    0.092
CONECT    1    2    3   12   25   54   55
CONECT    2    1   30   31   32
CONECT    3    1   33   34   35
CONECT    4    5   10   42
CONECT    5    4    6    7
CONECT    6    5   43   44
CONECT    7    5    8
CONECT    8    7    9   14
CONECT    9    8   10   12
CONECT   10    4    9   11
CONECT   11   10
CONECT   12    1    9   13
CONECT   13   12   14   41
CONECT   14    8   13   15
CONECT   15   14   16   39   40
CONECT   16   15   36   37   38
CONECT   17   18   23   52
CONECT   18   17   19   20
CONECT   19   18   50   51
CONECT   20   18   21
CONECT   21   20   22   27
CONECT   22   21   23   25
CONECT   23   17   22   24
CONECT   24   23
```

```
CONECT    25     1    22    26
CONECT    26    25    27    53
CONECT    27    21    26    28
CONECT    28    27    29    45    46
CONECT    29    28    47    48    49
CONECT    30     2
CONECT    31     2
CONECT    32     2
CONECT    33     3
CONECT    34     3
CONECT    35     3
CONECT    36    16
CONECT    37    16
CONECT    38    16
CONECT    39    15
CONECT    40    15
CONECT    41    13
CONECT    42     4
CONECT    43     6
CONECT    44     6
CONECT    45    28
CONECT    46    28
CONECT    47    29
CONECT    48    29
CONECT    49    29
CONECT    50    19
CONECT    51    19
CONECT    52    17
CONECT    53    26
CONECT    54     1
CONECT    55     1
```

Modified AMBER

```
HETATM     1  N11            1      2.561   1.048  -3.723
HETATM     2  C21            2      2.400   1.953  -4.760
HETATM     3  N21            3      2.956   1.632  -5.936
HETATM     4  N31            4      1.722   3.108  -4.629
HETATM     5  C41            5      1.210   3.293  -3.380
HETATM     6  C51            6      1.320   2.444  -2.303
HETATM     7  C61            7      2.044   1.219  -2.439
HETATM     8  O61            8      2.243   0.354  -1.590
HETATM     9  N71            9      0.656   2.991  -1.212
HETATM    10  C81           10      0.179   4.127  -1.637
HETATM    11  N91           11      0.475   4.377  -2.956
HETATM    12  C91           12      0.100   5.533  -3.806
HETATM    13  C91           13     -0.595   6.646  -3.017
HETATM    14  H             14     -0.827   7.471  -3.691
HETATM    15  H             15     -1.525   6.275  -2.584
HETATM    16  H             16      0.060   7.010  -2.226
HETATM    17  H             17      1.001   5.941  -4.266
HETATM    18  H             18     -0.570   5.188  -4.595
HETATM    19  H             19     -0.392   4.792  -1.007
HETATM    20  H             20      3.084   0.202  -3.901
HETATM    21  H             21      2.866   2.257  -6.724
HETATM    22  H             22      3.462   0.764  -6.035
HETATM    23  PT1  NUL       1      0.437   2.267   0.651
HETATM    24  N11  NUL       1     -1.560   2.056   0.358
HETATM    25  N10  NUL       1      0.223   1.568   2.544
HETATM    26  H    NUL       1     -1.870   2.382  -0.546
HETATM    27  H    NUL       1     -1.919   1.126   0.524
HETATM    28  H    NUL       1     -1.894   2.681   1.077
HETATM    29  H    NUL       1      1.153   1.510   2.934
HETATM    30  H    NUL       1     -0.358   2.138   3.141
HETATM    31  H    NUL       1     -0.171   0.640   2.475
HETATM    32  H    NUL       1      0.203   3.771   1.194
HETATM    33  H    NUL       1      0.659   0.765   0.109
```

```
HETATM   34  N12          34      4.476  -0.522   1.622
HETATM   35  C22          35      5.686   0.150   1.713
HETATM   36  N22          36      6.774  -0.594   1.954
HETATM   37  N32          37      5.799   1.483   1.570
HETATM   38  C42          38      4.611   2.101   1.320
HETATM   39  C52          39      3.374   1.509   1.214
HETATM   40  C62          40      3.248   0.094   1.381
HETATM   41  O62          41      2.227  -0.589   1.335
HETATM   42  N72          42      2.415   2.477   0.941
HETATM   43  C82          43      3.071   3.602   0.897
HETATM   44  N92          44      4.417   3.450   1.120
HETATM   45  C92          45      5.465   4.495   1.141
HETATM   46  C92          46      5.555   5.122   2.534
HETATM   47  H            47      6.431   4.065   0.875
HETATM   48  H            48      5.217   5.269   0.415
HETATM   49  H            49      6.324   5.895   2.534
HETATM   50  H            50      4.597   5.570   2.802
HETATM   51  H            51      5.815   4.359   3.269
HETATM   52  H            52      7.680  -0.154   2.032
HETATM   53  H            53      6.690  -1.595   2.055
HETATM   54  H            54      4.474  -1.525   1.741
HETATM   55  H            55      2.603   4.556   0.701
CONECT    1    2    7   20
CONECT    2    1    3    4
CONECT    3    2   21   22
CONECT    4    2    5
CONECT    5    4    6   11
CONECT    6    5    7    9
CONECT    7    1    6    8
CONECT    8    7
CONECT    9    6   10   23
CONECT   10    9   11   19
CONECT   11    5   10   12
CONECT   12   11   13   17   18
CONECT   13   12   14   15   16
CONECT   14   13
CONECT   15   13
CONECT   16   13
CONECT   17   12
CONECT   18   12
CONECT   19   10
CONECT   20    1
CONECT   21    3
CONECT   22    3
CONECT   23   24   25    9   42   32   33
CONECT   24   23   26   27   28
CONECT   25   23   29   30   31
CONECT   26   24
CONECT   27   24
CONECT   28   24
CONECT   29   25
CONECT   30   25
CONECT   31   25
CONECT   32   23
CONECT   33   23
CONECT   34   35   40   54
CONECT   35   34   36   37
CONECT   36   35   52   53
CONECT   37   35   38
CONECT   38   37   39   44
CONECT   39   38   40   42
CONECT   40   34   39   41
CONECT   41   40
CONECT   42   39   43   23
CONECT   43   42   44   55
CONECT   44   38   43   45
CONECT   45   44   46   47   48
```

```
CONECT    46    45    49    50    51
CONECT    47    45
CONECT    48    45
CONECT    49    46
CONECT    50    46
CONECT    51    46
CONECT    52    36
CONECT    53    36
CONECT    54    34
CONECT    55    43
```

Table IV.2 [Pt(NH₃)₂(N2-dimethyl-9-MeG)₂]

with dummy atoms

```
HETATM     1  O6           1    -1.072    2.051    5.547
HETATM     2  N1           2    -0.750    2.247    3.315
HETATM     3  N2           3    -0.548    2.537    0.996
HETATM     4  N3           4     0.971    1.129    2.080
HETATM     5  N7           5     1.397    0.162    5.504
HETATM     6  N9           6     2.367   -0.247    3.534
HETATM     7  C2           7    -0.091    1.956    2.126
HETATM     8  C4           8     1.328    0.622    3.291
HETATM     9  C5           9     0.731    0.867    4.506
HETATM    10  C6          10    -0.402    1.737    4.567
HETATM    11  C8          11     2.350   -0.479    4.887
HETATM    12  C20         12     0.080    2.304   -0.319
HETATM    13  C21         13    -1.654    3.514    0.992
HETATM    14  C9          14     3.281   -0.804    2.513
HETATM    15  H20         15    -0.581    2.583   -1.140
HETATM    16  H20         16     0.324    1.249   -0.441
HETATM    17  H20         17     0.995    2.892   -0.387
HETATM    18  H21         18    -1.852    3.909   -0.004
HETATM    19  H21         19    -1.401    4.352    1.642
HETATM    20  H21         20    -2.564    3.037    1.359
HETATM    21  H90         21     4.132   -1.283    2.996
HETATM    22  H91         22     3.641   -0.004    1.865
HETATM    23  H92         23     2.746   -1.540    1.913
HETATM    24  H1          24    -1.544    2.870    3.289
HETATM    25  H8          25     3.054   -1.125    5.390
HETATM    26  PT1  NUL     1     1.066    0.068    7.487
HETATM    27  CL1  NUL     1    -0.589   -1.496    7.126
HETATM    28  N10  NUL     1     2.522    1.451    7.778
HETATM    29  N11  NUL     1     0.732   -0.027    9.486
HETATM    30  H10  NUL     1     2.977    1.269    8.662
HETATM    31  H10  NUL     1     3.190    1.344    7.027
HETATM    32  H10  NUL     1     2.158    2.393    7.773
HETATM    33  H11  NUL     1    -0.225    0.277    9.599
HETATM    34  H11  NUL     1     1.344    0.597    9.992
HETATM    35  H11  NUL     1     0.831   -0.965    9.847
HETATM    36  H    NUL     1    -0.024    1.251    7.384
HETATM    37  H    NUL     1     2.159   -1.115    7.608
CONECT     1    10
CONECT     2     7    10    24
CONECT     3     7    12    13
CONECT     4     7     8
CONECT     5     9    11    26
CONECT     6     8    11    14
CONECT     7     2     3     4
CONECT     8     4     6     9
CONECT     9     5     8    10
CONECT    10     1     2     9
CONECT    11     5     6    25
CONECT    12     3    15    16    17
CONECT    13     3    18    19    20
CONECT    14     6    21    22    23
```

```
CONECT    15    12
CONECT    16    12
CONECT    17    12
CONECT    18    13
CONECT    19    13
CONECT    20    13
CONECT    21    14
CONECT    22    14
CONECT    23    14
CONECT    24     2
CONECT    25    11
CONECT    26    27    28    29    36    37     5
CONECT    27    26
CONECT    28    26    30    31    32
CONECT    29    26    33    34    35
CONECT    30    28
CONECT    31    28
CONECT    32    28
CONECT    33    29
CONECT    34    29
CONECT    35    29
CONECT    36    26
CONECT    37    26


without dummy atoms
HETATM     1 PT1  NUL     1      1.062    0.071    7.464
HETATM     2 CL1  NUL     1     -0.581   -1.506    7.110
HETATM     3 O6   NUL     1     -0.980    2.040    5.601
HETATM     4 N10  NUL     1      2.514    1.456    7.773
HETATM     5 N11  NUL     1      0.681    0.017    9.457
HETATM     6 N1   NUL     1     -0.716    2.238    3.352
HETATM     7 N2   NUL     1     -0.548    2.535    1.028
HETATM     8 N3   NUL     1      0.982    1.118    2.083
HETATM     9 N7   NUL     1      1.440    0.124    5.491
HETATM    10 N9   NUL     1      2.392   -0.276    3.512
HETATM    11 C2   NUL     1     -0.077    1.948    2.151
HETATM    12 C4   NUL     1      1.354    0.600    3.285
HETATM    13 C5   NUL     1      0.775    0.842    4.507
HETATM    14 C6   NUL     1     -0.347    1.722    4.597
HETATM    15 C8   NUL     1      2.383   -0.521    4.864
HETATM    16 C20  NUL     1      0.061    2.304   -0.296
HETATM    17 C21  NUL     1     -1.656    3.510    1.039
HETATM    18 C9   NUL     1      3.301   -0.821    2.483
HETATM    19 H10  NUL     1      2.957    1.285    8.664
HETATM    20 H10  NUL     1      3.202    1.389    7.036
HETATM    21 H10  NUL     1      2.106    2.380    7.769
HETATM    22 H11  NUL     1     -0.270    0.314    9.626
HETATM    23 H11  NUL     1      1.318    0.635    9.940
HETATM    24 H11  NUL     1      0.805   -0.926    9.796
HETATM    25 H20  NUL     1     -0.609    2.587   -1.108
HETATM    26 H20  NUL     1      0.303    1.249   -0.425
HETATM    27 H20  NUL     1      0.977    2.891   -0.375
HETATM    28 H21  NUL     1     -1.849    3.924    0.050
HETATM    29 H21  NUL     1     -1.409    4.337    1.706
HETATM    30 H21  NUL     1     -2.566    3.024    1.392
HETATM    31 H90  NUL     1      4.153   -1.307    2.959
HETATM    32 H91  NUL     1      3.663   -0.014    1.844
HETATM    33 H92  NUL     1      2.765   -1.550    1.875
HETATM    34 H1   NUL     1     -1.508    2.865    3.341
HETATM    35 H8   NUL     1      3.084   -1.177    5.358
CONECT     1     2     4     5     9
CONECT     2     1
CONECT     3    14
CONECT     4     1    19    20    21
CONECT     5     1    22    23    24
CONECT     6    11    14    34
```

```
CONECT     7    11    16    17
CONECT     8    11    12
CONECT     9     1    13    15
CONECT    10    12    15    18
CONECT    11     6     7     8
CONECT    12     8    10    13
CONECT    13     9    12    14
CONECT    14     3     6    13
CONECT    15     9    10    35
CONECT    16     7    25    26    27
CONECT    17     7    28    29    30
CONECT    18    10    31    32    33
CONECT    19     4
CONECT    20     4
CONECT    21     4
CONECT    22     5
CONECT    23     5
CONECT    24     5
CONECT    25    16
CONECT    26    16
CONECT    27    16
CONECT    28    17
CONECT    29    17
CONECT    30    17
CONECT    31    18
CONECT    32    18
CONECT    33    18
CONECT    34     6
CONECT    35    15
```

Table IV.3 [Pt(NH₃)₃(nucleobase)]

nucleobase - 1-MeT, N3-binding

HETATM	1	N1		1	31.070	1.684	-2.375
HETATM	2	C6		2	30.259	2.749	-2.690
HETATM	3	H6		3	29.375	2.927	-2.093
HETATM	4	C5		4	30.539	3.577	-3.728
HETATM	5	C5M		5	29.612	4.737	-4.040
HETATM	6	1H5M		6	29.000	4.987	-3.174
HETATM	7	2H5M		7	30.197	5.615	-4.317
HETATM	8	3H5M		8	28.961	4.464	-4.871
HETATM	9	C4		9	31.714	3.359	-4.555
HETATM	10	O4		10	32.041	4.045	-5.521
HETATM	11	N3		11	32.493	2.269	-4.175
HETATM	12	C2		12	32.223	1.414	-3.110
HETATM	13	O2		13	32.966	0.469	-2.850
HETATM	14	C		14	30.747	0.784	-1.241
HETATM	15	H		15	31.556	0.810	-0.510
HETATM	16	H		16	29.824	1.091	-0.748
HETATM	17	H		17	30.624	-0.236	-1.605
HETATM	18	Pt		18	34.165	1.889	-5.270
HETATM	19	N		19	35.833	1.511	-6.363
HETATM	20	N		20	35.309	2.665	-3.784
HETATM	21	H		21	35.644	1.691	-7.339
HETATM	22	H		22	35.259	2.061	-2.975
HETATM	23	H		23	34.964	3.582	-3.541
HETATM	24	H		24	33.606	0.736	-7.479
HETATM	25	H		25	36.100	0.544	-6.244
HETATM	26	H		26	36.585	2.108	-6.050
HETATM	27	H		27	36.267	2.736	-4.094
HETATM	28	N		28	33.015	1.115	-6.752
HETATM	29	H		29	32.431	1.844	-7.137
HETATM	30	H		30	32.433	0.381	-6.374
CONECT	1	2	12	14			
CONECT	2	1	3	4			
CONECT	3	2					
CONECT	4	2	5	9			
CONECT	5	4	6	7	8		
CONECT	6	5					
CONECT	7	5					
CONECT	8	5					
CONECT	9	4	10	11			
CONECT	10	9					
CONECT	11	9	12	18			
CONECT	12	1	11	13			
CONECT	13	12					
CONECT	14	1	15	16	17		
CONECT	15	14					
CONECT	16	14					
CONECT	17	14					
CONECT	18	19	20	28	11		
CONECT	19	18	21	25	26		
CONECT	20	18	22	23	27		
CONECT	21	19					
CONECT	22	20					
CONECT	23	20					
CONECT	24	28					
CONECT	25	19					
CONECT	26	19					
CONECT	27	20					
CONECT	28	18	24	29	30		
CONECT	29	28					
CONECT	30	28					

nucleobase - 1-MeC, N3-binding

HETATM	1	N1	1	16.064	2.695	-3.602
HETATM	2	C6	2	15.667	2.659	-2.286
HETATM	3	H6	3	16.406	2.746	-1.503
HETATM	4	C5	4	14.363	2.519	-1.946
HETATM	5	H5	5	14.083	2.495	-0.903
HETATM	6	C4	6	13.414	2.408	-3.023
HETATM	7	N4	7	12.108	2.272	-2.749
HETATM	8	1H4	8	11.793	2.245	-1.790
HETATM	9	2H4	9	11.439	2.195	-3.502
HETATM	10	N3	10	13.813	2.443	-4.307
HETATM	11	C2	11	15.131	2.584	-4.636
HETATM	12	O2	12	15.492	2.613	-5.811
HETATM	13	C	13	17.494	2.853	-3.967
HETATM	14	H	14	17.821	1.996	-4.556
HETATM	15	H	15	17.625	3.764	-4.552
HETATM	16	H	16	18.118	2.920	-3.075
HETATM	17	PT	17	12.500	2.289	-5.859
HETATM	18	N	18	11.192	2.135	-7.404
HETATM	19	N	19	12.773	4.255	-6.285
HETATM	20	H	20	10.248	2.223	-7.054
HETATM	21	H	21	13.725	4.513	-6.066
HETATM	22	H	22	12.137	4.811	-5.732
HETATM	23	H	23	12.034	-0.184	-6.284
HETATM	24	H	24	11.303	1.236	-7.851
HETATM	25	H	25	11.373	2.869	-8.073
HETATM	26	H	26	12.597	4.415	-7.267
HETATM	27	N	27	12.226	0.323	-5.433
HETATM	28	H	28	11.447	0.225	-4.798
HETATM	29	H	29	13.061	-0.045	-5.000
CONECT	1	2	11	13		
CONECT	2	1	3	4		
CONECT	3	2				
CONECT	4	2	5	6		
CONECT	5	4				
CONECT	6	4	7	10		
CONECT	7	6	8	9		
CONECT	8	7				
CONECT	9	7				
CONECT	10	6	11	17		
CONECT	11	1	10	12		
CONECT	12	11				
CONECT	13	1	14	15	16	
CONECT	14	13				
CONECT	15	13				
CONECT	16	13				
CONECT	17	10	18	19	27	
CONECT	18	17	20	24	25	
CONECT	19	17	21	22	26	
CONECT	20	18				
CONECT	21	19				
CONECT	22	19				
CONECT	23	27				
CONECT	24	18				
CONECT	25	18				
CONECT	26	19				
CONECT	27	17	23	28	29	
CONECT	28	27				
CONECT	29	27				

nucleobase - 1-MeU, N3-binding

HETATM	1	N1	1	11.895	3.256	-4.118
HETATM	2	C6	2	12.761	2.759	-5.063
HETATM	3	H6	3	12.432	2.660	-6.088
HETATM	4	C5	4	14.025	2.389	-4.742
HETATM	5	H5	5	14.677	2.003	-5.511
HETATM	6	C4	6	14.515	2.504	-3.381
HETATM	7	O4	7	15.646	2.200	-3.009
HETATM	8	N3	8	13.586	3.013	-2.476
HETATM	9	C2	9	12.284	3.397	-2.786
HETATM	10	O2	10	11.527	3.834	-1.921
HETATM	11	C	11	10.515	3.662	-4.483
HETATM	12	H	12	10.372	4.719	-4.257
HETATM	13	H	13	10.329	3.504	-5.545
HETATM	14	H	14	9.796	3.072	-3.912
HETATM	15	PT	15	14.180	3.217	-0.539
HETATM	16	N	16	14.772	3.422	1.391
HETATM	17	N	17	14.144	5.236	-0.743
HETATM	18	H	18	15.543	4.073	1.439
HETATM	19	H	19	13.249	5.515	-1.120
HETATM	20	H	20	14.879	5.524	-1.372
HETATM	21	H	21	14.705	0.949	0.510
HETATM	22	H	22	15.066	2.525	1.752
HETATM	23	H	23	14.002	3.772	1.943
HETATM	24	H	24	14.281	5.671	0.159
HETATM	25	N	25	14.216	1.197	-0.338
HETATM	26	H	26	14.691	0.794	-1.133
HETATM	27	H	27	13.269	0.847	-0.298
CONECT	1	2	9	11		
CONECT	2	1	3	4		
CONECT	3	2				
CONECT	4	2	5	6		
CONECT	5	4				
CONECT	6	4	7	8		
CONECT	7	6				
CONECT	8	6	9	15		
CONECT	9	1	8	10		
CONECT	10	9				
CONECT	11	1	12	13	14	
CONECT	12	11				
CONECT	13	11				
CONECT	14	11				
CONECT	15	8	16	17	25	
CONECT	16	15	18	22	23	
CONECT	17	15	19	20	24	
CONECT	18	16				
CONECT	19	17				
CONECT	20	17				
CONECT	21	25				
CONECT	22	16				
CONECT	23	16				
CONECT	24	17				
CONECT	25	15	21	26	27	
CONECT	26	25				
CONECT	27	25				

nucleobase - 9-MeG, N1-binding

HETATM	1	N9	1	21.882	2.314	-6.024
HETATM	2	C8	2	22.556	3.327	-6.660
HETATM	3	H8	3	22.419	3.569	-7.704
HETATM	4	N7	4	23.381	3.972	-5.882
HETATM	5	C5	5	23.250	3.348	-4.647
HETATM	6	C6	6	23.901	3.621	-3.405
HETATM	7	O6	7	24.734	4.493	-3.167
HETATM	8	N1	8	23.487	2.762	-2.383
HETATM	9	C2	9	22.553	1.748	-2.551
HETATM	10	N2	10	22.269	0.996	-1.478
HETATM	11	1H2	11	22.725	1.178	-0.596
HETATM	12	2H2	12	21.597	0.246	-1.552
HETATM	13	N3	13	21.942	1.494	-3.722
HETATM	14	C4	14	22.332	2.328	-4.725
HETATM	15	C	15	20.876	1.383	-6.580
HETATM	16	H	16	19.959	1.443	-5.994
HETATM	17	H	17	20.657	1.645	-7.615
HETATM	18	H	18	21.263	0.364	-6.542
HETATM	19	Pt	19	24.321	3.055	-0.550
HETATM	20	N	20	25.153	3.348	1.279
HETATM	21	N	21	22.585	3.875	0.108
HETATM	22	H	22	25.593	4.256	1.306
HETATM	23	H	23	21.895	3.147	0.225
HETATM	24	H	24	22.254	4.546	-0.570
HETATM	25	H	25	26.780	2.369	-0.519
HETATM	26	H	26	25.845	2.633	1.452
HETATM	27	H	27	24.434	3.300	1.986
HETATM	28	H	28	22.741	4.338	0.992
HETATM	29	N	29	26.057	2.239	-1.213
HETATM	30	H	30	26.330	2.692	-2.073
HETATM	31	H	31	25.922	1.252	-1.380
CONECT	1	2	14	15		
CONECT	2	1	3	4		
CONECT	3	2				
CONECT	4	2	5			
CONECT	5	4	6	14		
CONECT	6	5	7	8		
CONECT	7	6				
CONECT	8	6	9	19		
CONECT	9	8	10	13		
CONECT	10	9	11	12		
CONECT	11	10				
CONECT	12	10				
CONECT	13	9	14			
CONECT	14	1	5	13		
CONECT	15	1	16	17	18	
CONECT	16	15				
CONECT	17	15				
CONECT	18	15				
CONECT	19	20	21	29	8	
CONECT	20	19	22	26	27	
CONECT	21	19	23	24	28	
CONECT	22	20				
CONECT	23	21				
CONECT	24	21				
CONECT	25	29				
CONECT	26	20				
CONECT	27	20				
CONECT	28	21				
CONECT	29	19	25	30	31	
CONECT	30	29				
CONECT	31	29				

nucleobase - 9-MeA, N1-binding

```
HETATM     1  PT             1      33.283    2.250   -6.486
HETATM     2  N              2      34.668    2.047   -7.955
HETATM     3  N              3      32.362    3.710   -7.554
HETATM     4  H              4      34.215    2.116   -8.856
HETATM     5  H              5      32.804    4.597   -7.358
HETATM     6  H              6      31.388    3.754   -7.292
HETATM     7  H              7      34.139   -0.084   -5.920
HETATM     8  H              8      35.118    1.147   -7.872
HETATM     9  H              9      35.359    2.779   -7.869
HETATM    10  H             10      32.438    3.506   -8.540
HETATM    11  N             11      34.203    0.790   -5.417
HETATM    12  H             12      33.750    0.695   -4.520
HETATM    13  H             13      35.174    1.032   -5.282
HETATM    14  N9            14      29.226    3.018   -2.053
HETATM    15  C8            15      28.522    1.846   -2.196
HETATM    16  H8            16      27.681    1.579   -1.573
HETATM    17  N7            17      28.967    1.084   -3.158
HETATM    18  C5            18      30.037    1.793   -3.690
HETATM    19  C6            19      30.942    1.531   -4.744
HETATM    20  N6            20      30.896    0.415   -5.485
HETATM    21 1H6            21      30.192   -0.285   -5.300
HETATM    22 2H6            22      31.565    0.273   -6.229
HETATM    23  N1            23      31.892    2.453   -5.010
HETATM    24  C2            24      31.941    3.566   -4.278
HETATM    25  H2            25      32.721    4.267   -4.535
HETATM    26  N3            26      31.151    3.926   -3.268
HETATM    27  C4            27      30.204    2.980   -3.019
HETATM    28  C             28      29.014    4.110   -1.081
HETATM    29  H             29      28.840    5.045   -1.614
HETATM    30  H             30      28.150    3.889   -0.454
HETATM    31  H             31      29.897    4.213   -0.450
CONECT     1     2     3    11    23
CONECT     2     1     4     8     9
CONECT     3     1     5     6    10
CONECT     4     2
CONECT     5     3
CONECT     6     3
CONECT     7    11
CONECT     8     2
CONECT     9     2
CONECT    10     3
CONECT    11     1     7    12    13
CONECT    12    11
CONECT    13    11
CONECT    14    15    27    28
CONECT    15    14    16    17
CONECT    16    15
CONECT    17    15    18
CONECT    18    17    19    27
CONECT    19    18    20    23
CONECT    20    19    21    22
CONECT    21    20
CONECT    22    20
CONECT    23     1    19    24
CONECT    24    23    25    26
CONECT    25    24
CONECT    26    24    27
CONECT    27    14    18    26
CONECT    28    14    29    30    31
CONECT    29    28
CONECT    30    28
CONECT    31    28
```

nucleobase - 9-MeH, N1-binding

```
HETATM    1  N9           1     22.096   2.011  -6.032
HETATM    2  C8           2     23.271   2.149  -6.729
HETATM    3  H8           3     23.345   2.028  -7.800
HETATM    4  N7           4     24.293   2.441  -5.973
HETATM    5  C5           5     23.767   2.502  -4.689
HETATM    6  C6           6     24.409   2.784  -3.444
HETATM    7  O6           7     25.595   3.044  -3.250
HETATM    8  N1           8     23.522   2.748  -2.365
HETATM    9  C2           9     22.167   2.474  -2.481
HETATM   10  HN2         10     21.596   2.478  -1.564
HETATM   11  N3          11     21.564   2.210  -3.654
HETATM   12  C4          12     22.417   2.239  -4.715
HETATM   13  C           13     20.748   1.689  -6.549
HETATM   14  H           14     20.076   2.526  -6.358
HETATM   15  H           15     20.795   1.505  -7.622
HETATM   16  H           16     20.368   0.798  -6.049
HETATM   17  PT          17     24.262   3.129  -0.508
HETATM   18  N           18     24.998   3.510   1.345
HETATM   19  N           19     23.217   4.868  -0.451
HETATM   20  H           20     25.840   4.063   1.264
HETATM   21  H           21     22.269   4.680  -0.158
HETATM   22  H           22     23.209   5.284  -1.371
HETATM   23  H           23     25.900   1.322   0.249
HETATM   24  H           24     25.211   2.639   1.809
HETATM   25  H           25     24.311   4.019   1.882
HETATM   26  H           26     23.652   5.503   0.203
HETATM   27  N           27     25.306   1.389  -0.565
HETATM   28  H           28     25.874   1.372  -1.401
HETATM   29  H           29     24.662   0.611  -0.579
CONECT    1    2   12   13
CONECT    2    1    3    4
CONECT    3    2
CONECT    4    2    5
CONECT    5    4    6   12
CONECT    6    5    7    8
CONECT    7    6
CONECT    8    6    9   17
CONECT    9    8   10   11
CONECT   10    9
CONECT   11    9   12
CONECT   12    1    5   11
CONECT   13    1   14   15   16
CONECT   14   13
CONECT   15   13
CONECT   16   13
CONECT   17    8   18   19   27
CONECT   18   17   20   24   25
CONECT   19   17   21   22   26
CONECT   20   18
CONECT   21   19
CONECT   22   19
CONECT   23   27
CONECT   24   18
CONECT   25   18
CONECT   26   19
CONECT   27   17   23   28   29
CONECT   28   27
CONECT   29   27
```

nucleobase - 9-MeG, N3-binding

```
HETATM     1  N9              1      21.318    2.204   -1.944
HETATM     2  C8              2      20.269    1.853   -2.758
HETATM     3  H8              3      19.283    1.614   -2.387
HETATM     4  N7              4      20.576    1.830   -4.025
HETATM     5  C5              5      21.917    2.191   -4.061
HETATM     6  C6              6      22.794    2.331   -5.180
HETATM     7  O6              7      22.537    2.157   -6.369
HETATM     8  N1              8      24.082    2.708   -4.799
HETATM     9  H1              9      24.764    2.826   -5.534
HETATM    10  C2             10      24.476    2.933   -3.490
HETATM    11  N2             11      25.752    3.297   -3.301
HETATM    12 1H2             12      26.372    3.392   -4.093
HETATM    13 2H2             13      26.095    3.474   -2.368
HETATM    14  N3             14      23.648    2.806   -2.436
HETATM    15  C4             15      22.384    2.430   -2.788
HETATM    16  C              16      21.328    2.305   -0.469
HETATM    17  H              17      22.101    1.655   -0.060
HETATM    18  H              18      21.516    3.336   -0.171
HETATM    19  H              19      20.361    1.994   -0.072
HETATM    20  PT             20      24.357    3.178   -0.562
HETATM    21  N              21      25.063    3.548    1.305
HETATM    22  N              22      23.671    5.084   -0.682
HETATM    23  H              23      25.945    4.035    1.239
HETATM    24  H              24      23.042    5.264    0.088
HETATM    25  H              25      23.179    5.208   -1.555
HETATM    26  H              26      26.011    1.286   -0.124
HETATM    27  H              27      25.196    2.674    1.794
HETATM    28  H              28      24.399    4.119    1.808
HETATM    29  H              29      24.450    5.726   -0.639
HETATM    30  N              30      25.056    1.275   -0.454
HETATM    31  H              31      25.020    0.854   -1.371
HETATM    32  H              32      24.483    0.743    0.185
CONECT     1     2    15    16
CONECT     2     1     3     4
CONECT     3     2
CONECT     4     2     5
CONECT     5     4     6    15
CONECT     6     5     7     8
CONECT     7     6
CONECT     8     6     9    10
CONECT     9     8
CONECT    10     8    11    14
CONECT    11    10    12    13
CONECT    12    11
CONECT    13    11
CONECT    14    10    15    20
CONECT    15     1     5    14
CONECT    16     1    17    18    19
CONECT    17    16
CONECT    18    16
CONECT    19    16
CONECT    20    14    21    22    30
CONECT    21    20    23    27    28
CONECT    22    20    24    25    29
CONECT    23    21
CONECT    24    22
CONECT    25    22
CONECT    26    30
CONECT    27    21
CONECT    28    21
CONECT    29    22
CONECT    30    20    26    31    32
CONECT    31    30
CONECT    32    30
```

Table V.1. Experimental values of geometric parameters in platinum complexes containing chlorine ligand.

REFCODE	R_{Pt-Cl}	N3-PT-CL N3 oriented to CL		NB-PT-CL	N3-Pt-NB N3 oriented to CL		PT-N3-H3[a]		PT-N3-CT		$\omega_{N-N-N-Cl}$[b]
		cis	trans		trans	cis					
AFCYPT	2.299	91.5	172.7	91.7	93.5	171.2			107.3	104.0	9.7
BOHDAD	2.300	87.6	179.3	92.0	88.7	179.4	108.3 109.3 111.4	108.1 111.4 107.5			0.3
CCENPT	2.288 2.288	95.3 95.3	168.2 168.2				101.0 103.7	101.0 103.7	111.4	111.4	
CHLPTM	2.304 2.326	95.3 95.3	168.2 168.2						115.4	120.8	
CTSPTA	2.299	90.0	178.1	90.6	89.0	176.5					5.1
CTSPTA01	2.299	81.1	176.6	90.5	91.1	175.3					7.3
CTYEPT	2.298	91.4	175.4	91.5	93.0	172.5			109.8	110.1	7.2
CTYEPU	2.298	92.0	175.2	91.7	93.0	175.1			108.6	110.4	3.7
CUKRAB	2.308 2.315	86.8 89.9	176.5 178.8								
CUSRAJ	2.285 2.293	87.1 88.3	178.3 177.9								
DEYXUA	2.326	88.2	179.0	92.9	88.1	178.3					2.1
DUJNIF	2.291	87.7	178.9	91.4	87.5	178.2					1.0
DUPGAW	2.307 2.313	89.0 87.7	177.9 178.6				108.5	108.5	115.0		
FAWYEH	2.318 2.303	91.7 93.8	177.3 175.6				107.2	109.9	106.1 110.7	108.1 113.4	

Complex	Pt–N	∠	∠	∠	∠	∠	ω
FAZZAH	2.316	86.2 178.7			108.1 106.3	115.3 116.0	
	2.306	87.4 178.9			112.3 112.9		
FITFUJ	2.305	91.6 174.6			108.8 109.7	108.5 112.1	
	2.311	92.7 175.0			108.8 109.7		
GALYAT	2.320	86.3 178.7				118.5 119.6	
	2.317	85.7 179.1					
GECGOK	2.281	91.4 172.9			100.2 109.7	101.8 106.7	
	2.299	96.6 174.3				122.4 125.6	
GECGUQ	2.295	90.2 174.5			102.7 113.3	115.7 119.4	
	2.300	93.5 176.7					
GEWGIY	2.311	86.3 177.4				113.4 112.5	
	2.316	87.0 177.1				119.1 118.2	
GIFVUM	2.295	89.6 178.6				110.9 113.2	
	2.306	91.3 178.6				113.4 114.2	
	2.301	90.1 179.2				113.5 112.7	
	2.313	90.7 177.7				114.6 112.1	
JEHPOB	2.335	92.2 177.0	87.7	94.4 178.0		111.3 114.2	2.9
						111.5 110.7	
KABJUS	2.309	91.5 178.3	88.2	90.1 178.3	113.5 112.6		1.5
					112.8 112.6		
					113.5 112.5		
MZOEPT	2.285	92.5 117.2				113.3 111.8	
						111.4 110.4	
PASHIA	2.296	90.6 175.0				110.0 114.6	
						118.0 115.0	
						106.6 105.7	
VANPIJ	2.328	87.9 177.5			110.2 106.8	117.3	
	2.311	90.3 179.5			109.1 107.3		

Units: bond lengths - angstroms, angles - degrees. For AMBER atom types see Fig. 3.3. ªMeasured only for complexes containing NH_3 ligand. ᵇImproper torsion angle $\omega_{N\text{-}N\text{-}N\text{-}Cl}$ was measured only for monodentate complexes with DNA constituents.

Table V.2. Experimental values of geometric parameters in platinum complexes relevant to modelling of [Pt(A)₃(Nucleobase)].

REFCODE	Nucleobase	Binding site	Nitrogen atom type	R_{Pt-N}	PT-NC-CA	PT-NC-C	PT-NA-C
AFCYPT	C	N3	NC	1.983	123.3	118.6	
AMPTMN	1-MeT	N3	NA	2.016,2.023			117.5,124.8
BAGKOJ10	1-MeC	N3	NC	2.034	120.7	118.9	
BAGKUP10	1-MeC	N3	NC	2.036	122.5	116.5	
BAPKAE	1-MeC T	N3 N1	NC N*	2.026 2.040	123.8	116.3	
BARZOJ	1-MeC	N3	NC	2.031,2.039	120.8,122.7	117.5,116.0	
BARZOJ01	1-MeC	N3	NC	2.033,2.044	123.0,121.1	116.3,117.3	
BEKKOR	1-MeU	N3	NA	2.060,2.089			118.2,118.9 118.4,122.5
BOSSOR	1-MeU	N3	NA	2.033,2.050			117.4,119.3 119.2,118.4
CIPMUJ	1-MeU	N3	NA	2.025,2.026			118.3,118.2 120.9,121.0
COKRUP	1-MeU 1-MeC	N3	NA NC	2.045 2.037	119.7 119.7	120.7 120.7	114.3,114.3 121.6,121.6
CTSPTA	1-MeC	N3	NC	2.024	120.3	118.9	
CTSPTA01	1-MeC	N3	NC	2.058	118.5	117.9	
CUKPUT	1-MeC	N3	NC	2.059 2.052 2.038	122.7 122.8 124.7	117.0 118.2 115.2	
CYTPTA10	9-EtG 1-MeC	N7 N3	NB NC	2.01,2.02 2.03,204	123.2,124.2	114.3,114.5	

CYTPTD10	9-EtG	N7	NB	2.017			117.7
	1-MeC	N3	NC	2.00	123.1		
DASCUV	9-EtG	N1	NA	2.034			
		N7	NB	2.032		117.9	
		N3	NC	2.045			
DEYXUA	1-MeT	N3	NA	1.974	117.44		117.2,121.9
DODLAJ	9-MeA	N7	NB	2.012			
	1-MeC	N3	NC	2.049	121.1	117.6	
DODLEN	9-MeG	N7	NB	2.008			
	1-MeC	N3	NC	2.031	124.0	114.0	
DOKJIW	1-MeU	N3	NA	1.86,2.052			120.1,118.1
DUJNIF	1-MeC	N3	NC	2.018	121.3	117.1	
DUJNOL	1-MeC	N3	NC	2.010,2.042	121.2,122.1	115.5,118.7	
DUXBED	1-MeU	N3	NA	2.041,2.045			122.2,117.5
DUXBIH	1-MeU	N3	NA	2.022,2.048			119.9,117.4
ENXPTA10	7,9-diMeHX	N1	NA	2.052			118.9
ENXPTB10	7,9-diMeHX	N1	NA	2.020			122.7,122.8
FATBIL	1-MeU	N3	NA	2.035			
	1-MeC	N3	NC	2.046	122.5,122.5	116.0,116.0	119.7,119.9
FEGBUO	1-MeC	N3	NC	2.061,2.061	123.8,123.9	117.4,117.4	
GEBVUE	1-MeC	N3	NC	2.028,2.028	123.3,123.2	116.5,116.5	
GIFWAT	C	N3	NC	2.048	123.4	114.9	
IPSMCP	1-MeC	N3	NC	2.059	125.2	115.1	
JEHPOB	1-MeC	N3	NC	2.039	122.2	116.3	
JIKGEP	1-MeC	N3	NC	2.049,2.049	122.8,122.8	116.3,116.3	
JILMIA	1-MeC	N3	NC	2.032,2.032	122.5,121.5	116.6,117.2	

Code	Base	Atom	AMBER	Bond length	Angle	Angle	Angle
JOMCOD	9-MeG	N7	NB	2.010			
	1-MeC	N3	NC	2.051	119.0	118.1	
KABJUS	C	N3	NC	2.033	121.5	117.1	116.8,120.0
KETDAO	1-MeU	N3	NA	2.080,2.080			
KUKGUS	9-MeG	N1	NA	2.046,2.061			115.9,118.5
MCPTEGI0	9-EtG	N7	NB	2.050			
	1-MeC	N3	NC	2.014	122.9	118.7	
MCSPTA	1-MeC	N3	NC	2.030	122.0	117.8	
MCSPTB	1-MeC	N3	NC	2.026,2.026	124.0	115.3	
SEBVEA	1-MeU	N3	NA	2.035,2.039			117.4,116.4 119.7,123.0
SEBVIE	1-MeU	N3	NA	2.008,2.048			118.0,120.2 118.3,121.0
SISCIG	9-MeA	N1	NC	2.043	123.6		
VEPROX	1-MeC	N3	NC	2.044	121.4	117.0	
VERDUR	9-MeG	N1	NA	2.056			117.1
		N7	NB	2.015			
VERFAZ	9-MeG	N1	NA	2.04			
		N7	NB	2.02			117.7
	1-MeU	N3	NA	2.03			
VOLCUU	1-MeC	N3	NC	2.015	125.8	115.2	

Only binding modes other than N7(G) are considered. Angles not included in the table: PT-NC-CQ: 118.0 (JONLAZ), 114.4 (SISCIG); PT-NC-CB: 126.15 (DASCUV); PT-NA-CA: 118.9 (DASCUV), 119.2 (DASCUV); PT-NA-CB: 118.1 (ENXPTB10), 120.1, 122.9 (KUKGUS), 120.5 (VERDUR), 121.5 (VERFAZ). 21 complexes containing 1-MeT and 1-MeU have also been analysed for the values of PT-NA-C angles. However, they are not included in this table, since the nucleobases in these complexes demonstrate other binding modes than those relevant to modelling of the title complexes. Units: bond lengths - angstroms, angles - degrees. Multiple values of geometric parameters correspond to bis-adducts, dimeric and multinuclear structures. For definition of AMBER atom types and nucleobase numbering scheme see Fig. 3.5, PT - platinum atom type.

MM+ atom types assigned to nucleobase atoms for modelling $Cr(CO)_5(nucleobase)$ complexes
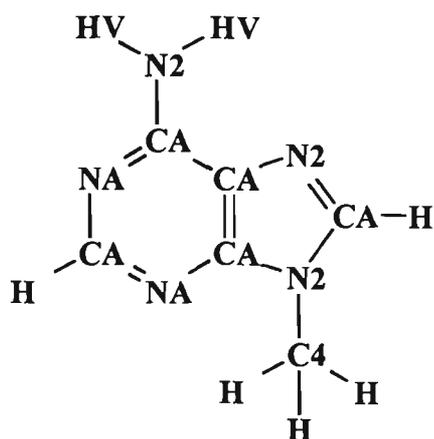
9-Methyladenine
9-MeA
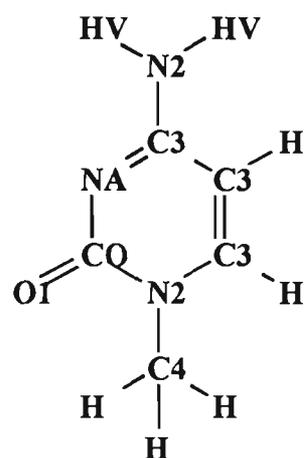
1-Methylcytosine
1-MeC
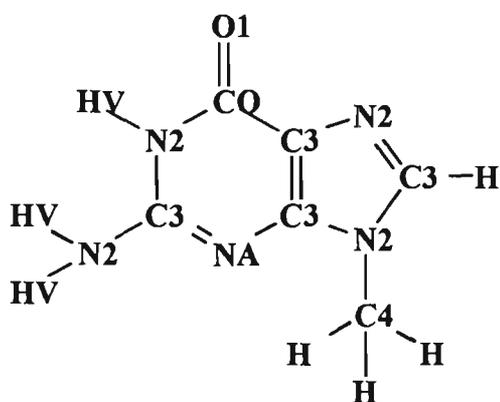
9-Methylguanine
9-MeG
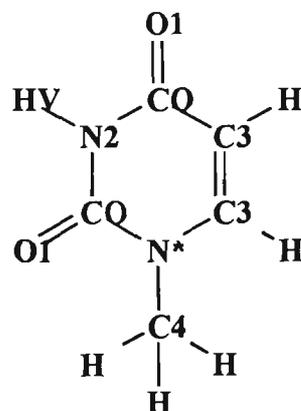
1-Methyluracil
1-MeU
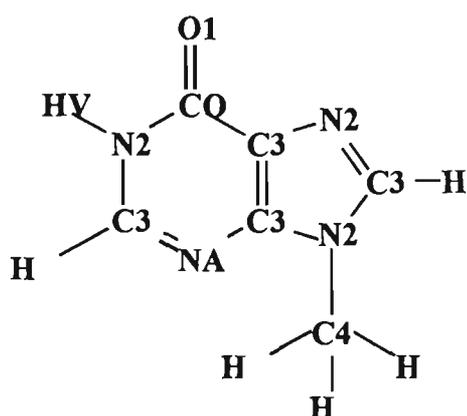
9-Methylhypoxanthine
9-MeH
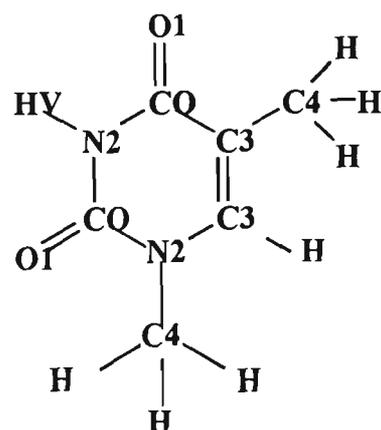
1-Methylthymine
1-MeT

Least-square-plane equations for the platinum coordination plane and the nucleobase plane in [Pt(NH$_3$)$_3$(nucleobase)] and [Pt(dien)(nucleobase)] optimised structures and in the experimental structures referred to in Table 3.12

The general equation[a]: z = Ax + By + D

		A	B	D	sd$_X$[b]	sd$_Y$[b]	sd$_Z$[b]	sd$_{intercept}$[b]	R^2
nucleobase					[Pt(NH$_3$)$_3$(nucleobase)]				
A(N1)	PtN$_4$	-1.29	-1.54	39.81	0.000	0.000	0.000	0.004	1
	base	-0.96	0.70	23.95	0.000	0.000	0.002	0.009	1
A(N3)	PtN$_4$	1.43	-0.47	-32.48	0.002	0.002	0.004	0.044	0.999
	base	0.40	5.46	-25.88	0.005	0.032	0.035	0.175	0.999
A(N7)	PtN$_4$	1.84	-1.53	-38.38	0.004	0.004	0.006	0.076	0.999
	base	0.41	3.60	-21.82	0.005	0.016	0.029	0.143	0.999
G(N1)	PtN$_4$	1.19	3.33	-39.73	0.056	0.126	0.104	1.668	0.997
	base	2.98	-2.59	-65.17	0.009	0.007	0.021	0.184	0.999
G(N3)	PtN$_4$	2.25	0.76	-57.70	0.014	0.007	0.017	0.371	0.999
	base	-3.21	11.80	40.50	0.061	0.231	0.110	0.838	0.995
G(N7)	PtN$_4$	1.61	-0.52	-19.39	0.002	0.001	0.003	0.017	1
	base	10.06	-37.8	-89.43	0.169	0.628	0.093	1.486	0.997
H(N1)	PtN$_4$	1.91	1.18	-50.43	0.001	0.001	0.001	0.020	1
	base	-1.53	7.91	11.82	0.006	0.022	0.019	0.105	0.999
H(N3)	PtN$_4$	2.33	0.86	-60.12	0.011	0.006	0.012	0.277	0.999
	base	-4.08	13.62	55.56	0.090	0.321	0.122	1.287	0.995
H(N7)	PtN$_4$	1.46	-0.57	-25.41	0.001	0.000	0.001	0.010	1
	base	8.27	-40.1	-36.31	0.208	0.971	0.125	0.969	0.994
C(N3)	PtN$_4$	1.13	-0.37	-19.18	0.048	0.035	0.093	0.592	0.996
	base	-14.3	131.5	-128.1	1.25	11.6	0.404	11.29	0.931
T(N3)	PtN$_4$	-0.17	2.16	-3.71	0.000	0.001	0.001	0.011	1
	base	-0.87	-0.97	26.39	0.004	0.003	0.012	0.113	0.999
U(N3)	PtN$_4$	3.11	-0.04	-44.55	0.080	0.025	0.071	1.135	0.998
	base	1.66	4.81	-39.49	0.009	0.023	0.021	0.181	0.999

nucleobase		A	B	D	sd_X^b	sd_Y^b	sd_Z^b	$sd_{intercept}^b$	R^2
		...							

nucleobase		[Pt(dien)(nucleobase)]							
A(N1)	PtN_4	-1.07	-0.47	30.17	0.040	0.030	0.079	1.357	0.997
	base	-0.73	1.37	15.08	0.001	0.002	0.006	0.040	0.999
A(N3)	PtN_4	1.29	0.14	-31.09	0.043	0.027	0.076	0.843	0.998
	base	-2.27	7.43	14.68	0.041	0.131	0.069	0.385	0.998
A(N7)	PtN_4	2.52	-4.13	-45.17	0.260	0.407	0.271	4.202	0.981
	base	0.90	1.88	-28.03	0.004	0.005	0.011	0.090	0.999
G(N1)	PtN_4	0.95	2.89	-33.03	0.083	0.183	0.171	2.238	0.992
	base	3.06	-2.73	-66.63	0.021	0.020	0.039	0.426	0.999
G(N3)	PtN_4	1.46	2.11	-43.12	0.070	0.092	0.114	1.946	0.996
	base	3.52	-3.29	-75.75	0.038	0.033	0.051	0.781	0.999
G(N7)	PtN_4	2.33	-2.71	-24.81	0.170	0.193	0.190	1.521	0.990
	base	0.45	2.46	-9.49	0.001	0.003	0.007	0.019	0.999
H(N1)	PtN_4	1.25	2.21	-38.00	0.076	0.114	0.136	2.121	0.995
	base	8.76	-13.2	-172.2	0.059	0.093	0.033	1.124	0.999
H(N3)	PtN_4	2.12	0.74	-54.82	0.086	0.045	0.104	2.178	0.997
	base	-4.82	15.48	67.83	0.144	0.485	0.156	2.134	0.993
H(N7)	PtN_4	3.38	-6.06	-40.87	0.458	0.798	0.365	4.927	0.967
	base	0.80	1.84	-21.45	0.003	0.003	0.008	0.051	0.999
C(N3)	PtN_4	1.36	0.79	-25.16	0.056	0.042	0.095	0.771	0.997
	base	1.64	-2.13	-22.31	0.003	0.003	0.005	0.033	0.999
T(N3)	PtN_4	-0.38	0.91	6.26	0.001	0.002	0.006	0.045	0.999
	base	-0.72	-0.74	21.23	0.032	0.032	0.074	1.120	0.997
U(N3)	PtN_4	1.37	4.33	-33.75	0.001	0.002	0.001	0.015	1
	base	5.48	-3.73	-66.55	0.028	0.019	0.022	0.325	0.999

refcode nucleobase		A	B	D	sd_X^b	sd_Y^b	sd_Z^b	$sd_{intercept}^b$	R^2
		experimental structures						
DEYXUA T(N3)	PtN$_4$	2.48	17.3	-31.10	0.259	1.734	0.281	3.177	0.980
	base	-1.66	-0.26	2.97	0.058	0.022	0.111	0.064	0.992
VEPROX C(N3)	PtN$_4$	-0.63	1.58	-1.10	0.016	0.026	0.040	0.066	0.999
	base	1.46	0.31	-2.95	0.051	0.037	0.114	0.321	0.993
SEBVIE U(N3)	PtN$_4$	0.00	0.26	1.88	0.036	0.037	0.104	0.070	0.958
	base	-0.34	-0.99	1.60	0.013	0.016	0.048	0.070	0.998
		-2.66	1.21	-1.95	0.085	0.061	0.125	0.214	0.994
BAHNUT G(N1)	PtN$_4$	-0.59	0.21	4.79	0.026	0.023	0.065	0.072	0.996
	base	0.58	1.06	1.67	0.018	0.021	0.090	0.075	0.997
SISCIG A(N1)	PtN$_4$	-0.17	-0.26	3.18	0.019	0.019	0.052	0.046	0.993
	base	1.79	-1.65	6.87	0.066	0.062	0.133	0.233	0.990
ENXPTA10 H(N1)	PtN$_4$	-0.44	-7.55	23.54	0.128	0.892	0.332	1.968	0.973
	base	0.29	-0.84	9.46	0.044	0.079	0.306	0.165	0.943
DASKUV G(N3)	PtN$_4$	-1.41	0.35	14.04	0.045	0.028	0.076	0.603	0.998
	base	0.28	-0.73	4.01	0.016	0.020	0.073	0.271	0.996

[a]Equations are generated on the basis of the atomic coordinates of heavy atoms of a nucleobase plane and Pt and nitrogen atoms of the coordination plane using the multiple regression option in Origin. [b]sd - standard deviation associated with the parameters A, B, C and D, respectively, in the general equation.

Appendix VIII

Charge distribution diagrams

VIIIa. Nucleobases and Pt(A)₃ moieties. Codes correspond to nucleobases and their binding sites in particular adducts.

A(N1)   A(N3)   A(N7)

G(N1)   G(N3)   G(N7)

H(N1)   H(N3)   H(N7)

C(N3)   T(N3)   U(N3)

Pt(NH₃)₃   Pt(dien)
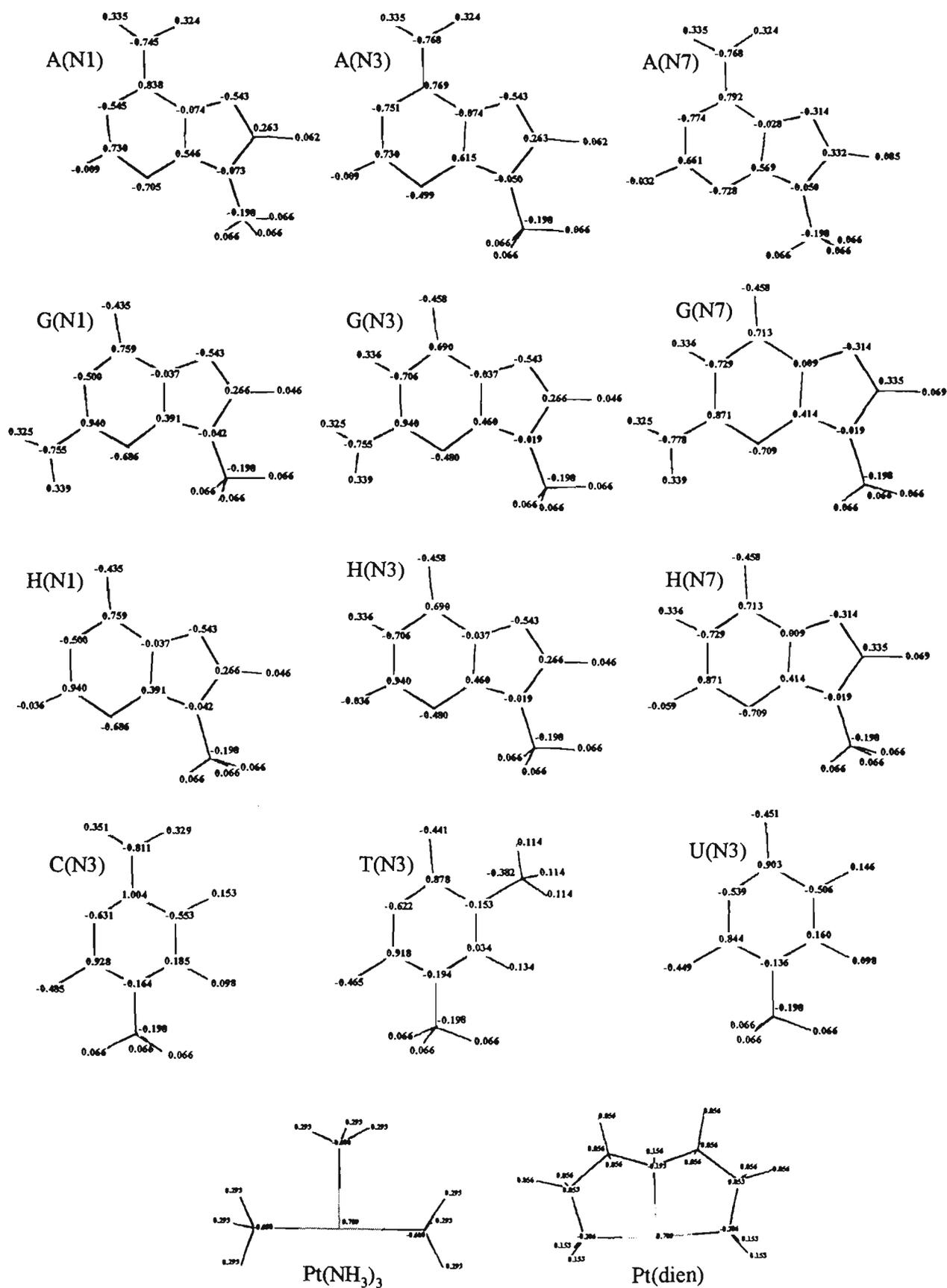
VIIIb. Carrier ligands. The numbers in brackets correspond to numbering of complexes in Fig. 4.8 and Table 5.1. The donor atoms are indicated by (N).

Cycles

Polycycles

Branches

(12)

0.025 0.025 0.142
-0.042 -0.240(N)
0.025 0.142
0.024
0.025 0.053
-0.042
0.025 0.025

(13)

0.023 0.023 0.051 -0.242(N) 0.142
-0.061 0.020 0.142
0.023 0.051
-0.027
0.023 0.031
-0.061
0.023 0.023

(14)

0.023 0.023
0.023 0.023
0.023
-0.063 -0.063 0.050 -0.242(N) 0.142
0.023 0.023 0.018 0.142
0.046
0.030 0.050
-0.030
0.028 0.028

(15)

0.025 0.025 0.142
0.023 0.023 0.027 0.027 -0.042 -0.240(N)
-0.065 -0.052 0.025 0.142
0.023 0.027
-0.056 -0.030 0.054
0.026 0.026 0.028 0.028

(16)

0.027 0.027
0.027 -0.053 0.027
-0.053 -0.049 0.051 -0.242(N) 0.142
0.027 0.027 0.021 0.142
0.027 -0.053 -0.021 0.051
0.027 0.031
-0.049
0.027 0.027

(17)

0.050 0.142
0.027 0.027 -0.018 -0.242(N) 0.142
0.027 -0.053 0.027 0.050 0.142
-0.053 0.027 0.052 -0.033 0.028
0.027 -0.050 0.027
-0.053 0.027 0.028
0.027 0.041 -0.050 0.027
-0.050 0.030 0.027
0.027 0.027

Straightes

(20)

0.047 0.141
0.006 0.141
0.047 -0.245(N)
0.047

(21)

0.024 0.050 0.142 0.142
0.024 -0.045 0.015 -0.243(N)
0.050
0.024

(22)

0.023 0.050
0.028 0.142 0.142
0.023 -0.064 -0.036 0.018 -0.242(N)
0.023 0.050
0.028

(23)

0.026 0.142 0.142
0.023 0.023 -0.242(N)
-0.065 0.055 0.028 0.018
0.026 -0.033 0.050 0.050
0.023 0.028

(24)

0.023 0.142 0.142
0.023 -0.065 0.023 0.027 0.142 -0.242(N)
0.026 -0.056 -0.052 0.028 0.018
0.026 -0.033 0.050
0.027 0.028 0.050

(25)

0.023 0.026
0.023 -0.065 -0.056 0.026 0.027 0.142 0.142
0.023 0.028 -0.242(N)
0.027 -0.053 -0.052 0.033 -0.018 0.050
0.027 0.027 0.028 0.050

(26)

0.023 0.026 0.027
0.023 -0.065 -0.056 -0.053 0.027 0.027 0.142 0.142
0.023 0.026 0.028 -0.242(N)
0.027 -0.053 -0.052 -0.033 0.018
0.027 0.027 0.028 0.050

(27)

0.023 0.026 0.027 0.027 0.142
0.023 -0.065 -0.056 -0.053 -0.053 0.027 0.027 0.028 0.142 -0.242(N)
0.023 0.026 0.027 -0.053 -0.052 -0.033 0.018
0.027 0.027 0.027 0.028 0.050

Appendix IX

Partial correlation coefficients

The result are presented in matrix form as follows: coefficient / (degrees of freedom) / 2-tailed significance; " . " is printed if a coefficient cannot be computed.

```
1.      Regression equations for toxicity

1a.     Controlling for..     PKA          LOGP       LOGP2

              PLD50          LRE
PLD50         1.0000        -.2007
              (    0)       (   22)
              P= .          P= .347

1b.     Controlling for..     LOGP         LOGP2      LRE

              PLD50          PKA
PLD50         1.0000         .3300
              (    0)       (   22)
              P= .          P= .115

1c.     Controlling for..     LOGP2        LRE        PKA

              PLD50          LOGP
PLD50         1.0000        -.7285
              (    0)       (   22)
              P= .          P= .000

1d.     Controlling for..     LRE          PKA        LOGP

              PLD50          LOGP2
PLD50         1.0000         .5833
              (    0)       (   22)
              P= .          P= .003

1e.     Controlling for..     PKA          LOGP       LOGP2

              PLD50          CRE
PLD50         1.0000         .1577
              (    0)       (   22)
              P= .          P= .462

1f.     Controlling for..     LOGP         LOGP2      CRE

              PLD50          PKA
PLD50         1.0000         .1936
              (    0)       (   22)
              P= .          P= .365

1g.     Controlling for..     LOGP2        CRE        PKA

              PLD50          LOGP
PLD50         1.0000        -.7619
              (    0)       (   22)
              P= .          P= .000

1h.     Controlling for..     CRE          PKA        LOGP

              PLD50          LOGP2
PLD50         1.0000         .6162
              (    0)       (   22)
              P= .          P= .001
```

```
1i.    Controlling for..    LOGP        LOGP2

              PLD50          LRE          PKA
PLD50         1.0000        -.0121        .2677
              (    0)       (   23)      (   23)
              P= .          P= .954      P= .196
LRE           -.0121         1.0000       .5553
              (   23)       (    0)      (   23)
              P= .954       P= .         P= .004


1j.    Controlling for..    LRE          PKA

              PLD50          LOGP         LOGP2
PLD50         1.0000        -.5747       -.2418
              (    0)       (   23)      (   23)
              P= .          P= .003      P= .244
LOGP          -.5747         1.0000       .8535
              (   23)       (    0)      (   23)
              P= .003       P= .         P= .000


1k.    Controlling for..    LOGP        LOGP2

              PLD50          CRE          PKA
PLD50         1.0000         .2440        .2677
              (    0)       (   23)      (   23)
              P= .          P= .240      P= .196
CRE            .2440         1.0000       .3882
              (   23)       (    0)      (   23)
              P= .240       P= .         P= .055


1l.    Controlling for..    CRE          PKA

              PLD50          LOGP         LOGP2
PLD50         1.0000        -.5988       -.2272
              (    0)       (   23)      (   23)
              P= .          P= .002      P= .275
LOGP          -.5988         1.0000       .8340
              (   23)       (    0)      (   23)
              P= .002       P= .         P= .000


2.     Regression equations for potency

2a.    Controlling for..    LOGP        LOGP2       EHOMO

              PID90          LRE
PID90         1.0000        -.4047
              (    0)       (   22)
              P= .          P= .050


2b.    Controlling for..    LOGP        LOGP2       LRE

              PID90          EHOMO
PID90         1.0000         .2258
              (    0)       (   22)
              P= .          P= .289


2c.    Controlling for..    LOGP2        LRE        EHOMO

              PID90          LOGP
PID90         1.0000        -.2344
              (    0)       (   22)
              P= .          P= .270


2d.    Controlling for..    LRE          EHOMO       LOGP

              PID90          LOGP2
PID90         1.0000        -.2204
              (    0)       (   22)
              P= .          P= .301
```

```
2e.    Controlling for..     LOGP        LOGP2

                PID90        EHOMO         LRE
PID90          1.0000       -.0269       -.3457
              (     0)      (    23)     (    23)
               P= .         P= .898      P= .091
EHOMO          -.0269        1.0000       .5780
              (    23)      (     0)     (    23)
               P= .898      P= .         P= .002


2f.    Controlling for..     EHOMO       LRE

                PID90        LOGP         LOGP2
PID90          1.0000       -.6371       -.6340
              (     0)      (    23)     (    23)
               P= .         P= .001      P= .001
LOGP           -.6371        1.0000       .8581
              (    23)      (     0)     (    23)
               P= .001      P= .         P= .000


2g.    Controlling for..     LOGP        LOGP2        ELUMO

                PID90        CRE
PID90          1.0000       -.1101
              (     0)      (    22)
               P= .         P= .609


2h.    Controlling for..     LOGP        LOGP2        CRE

                PID90        ELUMO
PID90          1.0000       -.0752
              (     0)      (    22)
               P= .         P= .727


2i.    Controlling for..     LOGP2       CRE          ELUMO

                PID90        LOGP
PID90          1.0000       -.3506
              (     0)      (    22)
               P= .         P= .093


2j.    Controlling for..     CRE         ELUMO        LOGP

                PID90        LOGP2
PID90          1.0000       -.1327
              (     0)      (    22)
               P= .         P= .536


2k.    Controlling for..     LOGP        LOGP2

                PID90        CRE          ELUMO
PID90          1.0000       -.1129       -.0793
              (     0)      (    23)     (    23)
               P= .         P= .591      P= .706
CRE            -.1129        1.0000       .0409
              (    23)      (     0)     (    23)
               P= .591      P= .         P= .846


2l.    Controlling for..     CRE         ELUMO

                PID90        LOGP         LOGP2
PID90          1.0000       -.6193       -.5564
              (     0)      (    23)     (    23)
               P= .         P= .001      P= .004
LOGP           -.6193        1.0000       .7967
              (    23)      (     0)     (    23)
               P= .001      P= .         P= .000
```

3.   Regression equations for therapeutic index

3a.   Controlling for..   LOGP   LOGP2   PC1

              LOGTI   LRE
LOGTI   1.0000   -.3546
             (0)   (22)
             P= .   P= .089

3b.   Controlling for..   LOGP   LOGP2   LRE

              LOGTI   PC1
LOGTI   1.0000   .1858
             (0)   (22)
             P= .   P= .385

3c.   Controlling for..   LOGP2   LRE   PC1

              LOGTI   LOGP
LOGTI   1.0000   .2852
             (0)   (22)
             P= .   P= .177

3d.   Controlling for..   LRE   PC1   LOGP

              LOGTI   LOGP2
LOGTI   1.0000   -.4541
             (0)   (22)
             P= .   P= .026

3e.   Controlling for..   LOGP   LOGP2

              LOGTI   LRE   PC1
LOGTI   1.0000   -.3088   -.0303
             (0)   (23)   (23)
             P= .   P= .133   P= .886
LRE   -.3088   1.0000   .5688
             (23)   (0)   (23)
             P= .133   P= .   P= .003

3f.   Controlling for..   LRE   PC1

              LOGTI   LOGP   LOGP2
LOGTI   1.0000   -.2538   -.4377
             (0)   (23)   (23)
             P= .   P= .221   P= .029
LOGP   -.2538   1.0000   .8694
             (23)   (0)   (23)
             P= .221   P= .   P= .000

3g.   Controlling for..   LOGP   LOGP2   PKA

              LOGTI   CRE
LOGTI   1.0000   -.1830
             (0)   (22)
             P= .   P= .392

3h.   Controlling for..   LOGP   LOGP2   CRE

              LOGTI   PKA
LOGTI   1.0000   -.1000
             (0)   (22)
             P= .   P= .642

3i.   Controlling for..   LOGP2   CRE   PKA

              LOGTI   LOGP
LOGTI   1.0000   .3220
             (0)   (22)
             P= .   P= .125

3j.   Controlling for..   CRE   PKA   LOGP

```
               LOGTI        LOGP2
LOGTI          1.0000      -.4815
               (     0)    (    22)
               P= .        P= .017
```

3k.   Controlling for..   LOGP   LOGP2

```
               LOGTI         CRE          PKA
LOGTI          1.0000      -.2362       -.1812
               (     0)    (    23)     (    23)
               P= .        P= .256      P= .386
CRE           -.2362        1.0000       .3882
               (    23)    (     0)     (    23)
               P= .256      P= .         P= .055
```

3l.   Controlling for..   CRE   PKA

```
               LOGTI        LOGP         LOGP2
LOGTI          1.0000      -.1843       -.4148
               (     0)    (    23)     (    23)
               P= .        P= .378      P= .039
LOGP          -.1843        1.0000       .8340
               (    23)    (     0)     (    23)
               P= .378      P= .         P= .000
```

Appendix X

Partial Least Squares Regression (PLS)

1. Multiple response (pLD50, pID90 and logTI included together), no steric or electronic descriptors

Number of Cross-validation Groups: 27
Optimal Number of Components: 2

Response Variable: pLD50
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	8.152	0.2436	10.20	0.0539
2	4.162	0.6138	5.288	0.5093

Regression Coefficients, Predictor Importances

	Intercept	logP	log2P
	0.5637	-0.4530	0.08468
Importance		-1.212	0.7168

Response Variable: pID90
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	11.70	0.4430	13.91	0.3378
2	11.52	0.4516	14.37	0.3160

Regression Coefficients, Predictor Importances

	Intercept	logP	log2P
	1.729	-0.2787	-0.02728
Importance		-0.5340	-0.1654

Response Variable: logTI
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	16.26	0.1113	18.20	0.0057
2	13.80	0.2458	16.54	0.0960

Regression Coefficients, Predictor Importances

	Intercept	logP	log2P
	1.166	0.1740	-0.1117
Importance		0.3573	-0.7257

2. Multiple response (pLD50, pID90 and logTI included together), LRE as steric parameter

Number of Cross-validation Groups: 27
Optimal Number of Components: 2

Response Variable: pLD50
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	6.981	0.3522	11.34	-0.0525
2	6.758	0.3729	22.54	-1.0913
3	4.573	0.5757	40.03	-2.7147
4	3.642	0.6621	44.23	-3.1046
5	3.437	0.6811	46.60	-3.3246
6	3.328	0.6912	38.97	-2.6161
7	3.219	0.7013	34.14	-2.1679
8	3.209	0.7022	44.75	-3.1523
9	3.209	0.7022	44.75	-3.1523

Regression Coefficients, Predictor Importances

Intercept	logP	LRE	q1	q2
-7.719	-0.08522	-0.01549	-0.3540	-15.56
Importance	-0.2280	-0.1386	-0.005866	-0.07250
LUMO	HOMO	pK	adeltaE	log2P
-0.03659	-0.3215	-0.03955	0.07367	-0.02125
-0.01122	-0.08928	-0.07129	0.03966	-0.1799

Response Variable: pID90
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	14.89	0.2915	27.15	-0.2918
2	10.59	0.4960	17.13	0.1848
3	10.58	0.4965	13.99	0.3344
4	10.55	0.4981	16.19	0.2297
5	9.187	0.5628	14.39	0.3150
6	9.095	0.5672	16.63	0.2087
7	9.086	0.5676	18.67	0.1116
8	9.029	0.5703	25.61	-0.2189
9	9.029	0.5703	25.61	-0.2189

Regression Coefficients, Predictor Importances

Intercept	logP	LRE	q1	q2
1.169	-0.1702	-0.02630	4.367	-7.443
Importance	-0.3260	-0.1685	0.05181	-0.02484
LUMO	HOMO	pKa	deltaE	log2P
-0.3866	-0.2828	-0.03483	-0.05005	-0.05033
-0.08485	-0.05624	-0.04496	-0.01930	-0.3051

Response Variable: logTI
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	18.02	0.0151	21.75	-0.1882
2	15.49	0.1539	20.46	-0.1178
3	12.98	0.2907	58.39	-2.1906
4	11.68	0.3615	73.20	-2.9998
5	11.18	0.3892	74.21	-3.0549
6	11.18	0.3893	68.02	-2.7164
7	11.00	0.3992	38.49	-1.1032
8	10.88	0.4053	35.21	-0.9239
9	10.88	0.4053	35.21	-0.9239

Regression Coefficients, Predictor Importances

Intercept	logP	LRE	q1	q2
8.849	-0.08465	-0.01078	4.701	8.069
Importance	-0.1738	-0.07400	0.05976	0.02885

LUMO	HOMO	pKa	deltaE	log2P
-0.3485	0.03826	0.004667	-0.1231	-0.02898
-0.08197	0.008154	0.006455	-0.05087	-0.1882

3. Multiple response (pLD50, pID90 and logTI included together), CRE as steric parameter

Number of Cross-validation Groups: 27
Optimal Number of Components: 1

Response Variable: pLD50
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	7.275	0.3250	10.24	0.0502
2	7.208	0.3311	22.94	-1.1283
3	4.152	0.6147	46.63	-3.3273
4	3.588	0.6670	34.92	-2.2406
5	3.256	0.6979	30.55	-1.8351
6	3.244	0.6989	38.84	-2.6037
7	3.204	0.7027	40.19	-2.7292
8	3.183	0.7046	40.95	-2.8001
9	3.183	0.7046	40.95	-2.8001

Regression Coefficients, Predictor Importances

Intercept	logP	CRE	q1	q2
-11.55	-0.06863	-0.01052	-1.922	-20.05
Importance	-0.1836	-0.07943	-0.03185	-0.09342

LUMO	HOMO	pKa	deltaE	log2P
0.07223	-0.3745	-0.04637	0.1230	-0.01474
0.02214	-0.1040	-0.08358	0.06625	-0.1248

Response Variable: pID90
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	14.85	0.2932	23.86	-0.1354
2	11.50	0.4527	19.13	0.0895
3	11.41	0.4571	15.61	0.2571
4	11.41	0.4572	17.64	0.1604
5	11.00	0.4764	18.15	0.1361
6	10.29	0.5104	22.09	-0.0512
7	9.953	0.5263	25.68	-0.2223
8	9.943	0.5268	26.65	-0.2684
9	9.943	0.5268	26.65	-0.2684

Regression Coefficients, Predictor Importances

Intercept	logP	CRE	q1	q2
-14.56	-0.09103	-0.01396	-2.549	-26.59
Importance	-0.1744	-0.07544	-0.03025	-0.08873

LUMO	HOMO	pKa	deltaE	log2P
0.09581	-0.4967	-0.06150	0.1632	-0.01956
0.02103	-0.09879	-0.07939	0.06293	-0.1186

Response Variable: logTI
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	17.93	0.0203	21.57	-0.1787
2	15.48	0.1543	19.91	-0.0881
3	13.41	0.2671	46.71	-1.5521
4	12.92	0.2939	49.46	-1.7025
5	12.92	0.2941	47.02	-1.5691
6	12.39	0.3230	75.90	-3.1472
7	12.25	0.3306	79.72	-3.3556
8	12.19	0.3341	82.75	-3.5216
9	12.19	0.3341	82.75	-3.5216

Regression Coefficients, Predictor Importances

Intercept	logP	CRE	q1	q2
-2.999	-0.02235	-0.003427	-0.6259	-6.529
Importance	-0.04589	-0.01985	-0.007958	-0.02335
LUMO	HOMO	pKa	deltaE	log2P
0.02352	-0.1220	-0.01510	0.04007	-0.004802
0.005533	-0.02599	-0.02089	0.01656	-0.03119

4. *Single response (pLD50, pID90 and logTI included separately), LRE as steric parameter*

4a. all descriptors

Response Variable: pLD50
Number of Cross-validation Groups: 27
Optimal Number of Components: 1
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	7.309	0.3218	22.85	-1.1206
2	6.085	0.4353	29.01	-1.6917
3	3.874	0.6406	53.44	-3.9585
4	3.490	0.6761	52.80	-3.8994
5	3.335	0.6905	42.01	-2.8980
6	3.263	0.6972	34.46	-2.1977
7	3.224	0.7009	37.23	-2.4544
8	3.209	0.7022	44.75	-3.1523
9	3.209	0.7022	44.75	-3.1523

Regression Coefficients, Predictor Importances

Intercept	logP	LRE	q1	q2
-13.11	-0.05194	-0.01127	-2.718	-22.85
Importance	-0.1390	-0.1008	-0.04504	-0.1065
LUMO	HOMO	pKa	deltaE	log2P
0.1233	-0.3770	-0.04128	0.1403	-0.006036
0.03780	-0.1047	-0.07441	0.07553	-0.05110

Response Variable: pID90
Number of Cross-validation Groups: 27
Optimal Number of Components: 5
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	15.56	0.2595	23.68	-0.1269
2	11.44	0.4557	18.73	0.1086
3	10.02	0.5231	17.91	0.1479
4	9.578	0.5442	18.02	0.1422
5	9.333	0.5558	16.43	0.2182
6	9.258	0.5594	26.53	-0.2627
7	9.225	0.5610	31.33	-0.4908

```
8            9.225      0.5610              31.33      -0.4908
Regression Coefficients, Predictor Importances
   Intercept        logP        LRE           q1          q2
     51.51       -0.3859     -0.08000      -28.04       149.4
Importance       -0.7395     -0.5127       -0.3328      0.4986
        LUMO        HOMO         pKa        deltaE
     0.05878      0.6741      0.03583      -0.1603
     0.01290      0.1341      0.04625      -0.06181
```

Response Variable: logTI
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	15.08	0.1759	20.88	-0.1409
2	14.59	0.2029	36.52	-0.9954
3	11.56	0.3683	78.18	-3.2718
4	11.27	0.3840	68.49	-2.7423
5	11.13	0.3917	58.29	-2.1847
6	10.98	0.4002	51.86	-1.8335
7	10.89	0.4052	38.24	-1.0895
8	10.88	0.4053	35.21	-0.9239
9	10.88	0.4053	35.21	-0.9239

```
Regression Coefficients, Predictor Importances
   Intercept        logP        LRE           q1          q2
     14.34      -0.05901     -0.008617      4.742       21.54
Importance       -0.1212     -0.05917       0.06029     0.07702
        LUMO        HOMO         pKa        deltaE       log2P
    -0.2834       0.1521     -0.008401     -0.1323     -0.03863
    -0.06666      0.03242    -0.01162      -0.05467    -0.2509
```

4b. only descriptors from "best" MLR equations

Response Variable: pLD50
Number of Cross-validation Groups: 27
Optimal Number of Components: 4
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	6.716	0.3768	9.143	0.1516
2	3.980	0.6307	7.064	0.3445
3	3.782	0.6490	5.702	0.4709
4	3.703	0.6564	5.104	0.5264

```
Regression Coefficients, Predictor Importances
   Intercept        logP        LRE          pKa        log2P
    -0.6210      -0.5052     -0.02020      0.1757      0.09610
Importance       -1.352      -0.1808       0.3166      0.8135
```

Response Variable: pID90
Number of Cross-validation Groups: 27
Optimal Number of Components: 3
Goodness-of-Fit and Goodness-of-Prediction.

Comp	Residual SS	R2	Resid PRESS	R2cv
1	12.97	0.3829	20.51	0.0240
2	10.61	0.4952	12.51	0.4045
3	10.12	0.5183	12.26	0.4165

```
Regression Coefficients, Predictor Importances
   Intercept        logP        LRE          HOMO
     20.39       -0.3448     -0.06496      1.724
Importance       -0.6606     -0.4162       0.3428
```

Response Variable: pID90
Number of Cross-validation Groups: 27
Optimal Number of Components: 10
Response Variable: pID90
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	11.50	0.4529	13.59	0.3535
2	10.76	0.4881	13.82	0.3421
3	10.74	0.4887	14.57	0.3065

Regression Coefficients, Predictor Importances

Intercept	logP	LRE	LUMO
2.876	-0.2492	-0.04872	0.04772
Importance	-0.4774	-0.3122	0.01047

Response Variable: logTI
Number of Cross-validation Groups: 27
Optimal Number of Components: 4
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	14.93	0.1842	17.61	0.0376
2	13.84	0.2440	19.59	-0.0702
3	12.39	0.3228	16.50	0.0983
4	12.05	0.3415	16.50	0.0984

Regression Coefficients, Predictor Importances

Intercept	logP	LRE	q2	log2P
26.17	0.2658	-0.06872	70.33	-0.1299
Importance	0.5457	-0.4718	0.2515	-0.8437

Response Variable: logTI
Number of Cross-validation Groups: 27
Optimal Number of Components: 4
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	15.24	0.1670	17.47	0.0457
2	14.15	0.2268	18.54	-0.0132
3	12.41	0.3220	17.37	0.0510
4	12.32	0.3268	15.74	0.1400

Regression Coefficients, Predictor Importances

Intercept	logP	LRE	LUMO	log2P
4.164	0.3247	-0.05226	-0.4161	-0.1458
Importance	0.6666	-0.3588	-0.09787	-0.9468

5. *Single response (pLD50, pID90 and logTI included separately), CRE as steric parameter*

5a. all descriptors

Response Variable: pLD50
Number of Cross-validation Groups: 27
Optimal Number of Components: 1
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	7.488	0.3051	23.01	-1.1351
2	6.178	0.4267	33.79	-2.1352
3	3.868	0.6411	48.42	-3.4931
4	3.384	0.6860	30.59	-1.8387
5	3.232	0.7001	30.99	-1.8759
6	3.188	0.7041	33.25	-2.0856
7	3.185	0.7045	37.40	-2.4701

Comp	Residual SS	R2	Resid PRESS	R2cv
8	3.183	0.7046	40.77	-2.7829
9	3.183	0.7046	40.77	-2.7829

Regression Coefficients, Predictor Importances

	Intercept	logP	CRE	q1	q2
	-14.39	-0.05586	-0.006356	-2.923	-24.58
Importance	-0.1494	-0.04797	-0.04844	-0.1145	

	LUMO	HOMO	pKa	deltaE	log2P
	0.1326	-0.4055	-0.04439	0.1509	-0.006492
	0.04065	-0.1126	-0.08002	0.08123	-0.05495

Response Variable: pID90
Number of Cross-validation Groups: 27
Optimal Number of Components: 3
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	15.60	0.2578	21.48	-0.0224
2	12.41	0.4094	20.89	0.0060
3	11.33	0.4610	17.81	0.1522
4	10.93	0.4796	20.66	0.0166
5	10.86	0.4834	24.66	-0.1737
6	10.80	0.4862	31.38	-0.4933
7	10.80	0.4862	40.75	-0.9393
8	10.80	0.4862	40.75	-0.9393

Regression Coefficients, Predictor Importances

	Intercept	logP	CRE	q1	q2
	-0.7354	-0.3552	-0.009002	-5.276	-0.01193
Importance	-0.6806	-0.04865	-0.06260	-0.0000398	

	LUMO	HOMO	pKa	deltaE
	-0.4181	-0.2593	0.05892	-0.06653
	-0.09178	-0.05157	0.07607	-0.02565

Response Variable: logTI
Number of Cross-validation Groups: 27
Optimal Number of Components: 1
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	15.02	0.1792	20.40	-0.1148
2	14.59	0.2029	32.29	-0.7644
3	13.01	0.2891	54.16	-1.9594
4	12.55	0.3144	60.83	-2.3236
5	12.35	0.3252	77.02	-3.2086
6	12.19	0.3338	86.02	-3.6999
7	12.19	0.3341	95.24	-4.2038
8	12.19	0.3341	102.7	-4.6100
9	12.19	0.3341	102.7	-4.6100

Regression Coefficients, Predictor Importances

	Intercept	logP	CRE	q1	q2
	13.49	-0.05510	-0.01809	4.428	20.11
Importance	-0.1131	-0.1048	0.05630	0.07191	

	LUMO	HOMO	pKa	deltaE	log2P
	-0.2646	0.1420	-0.007844	-0.1236	-0.03607
	-0.06224	0.03027	-0.01085	-0.05105	-0.2343

5b. only descriptors from "best" MLR equations

Response Variable: pLD50
umber of Cross-validation Groups: 27
Optimal Number of Components: 4
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	7.397	0.3136	9.520	0.1166
2	4.912	0.5442	7.241	0.3281
3	3.911	0.6371	6.671	0.3810
4	3.760	0.6511	5.175	0.5198

Regression Coefficients, Predictor Importances

Intercept	logP	CRE	pKa	log2P
-0.5814	-0.5415	0.01521	0.09102	0.1022
Importance	-1.449	0.1148	0.1641	0.8652

Response Variable: pID90
Number of Cross-validation Groups: 27
Optimal Number of Components: 2
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	13.19	0.3725	18.45	0.1222
2	11.59	0.4483	14.15	0.3266
3	11.59	0.4485	14.15	0.3264

Regression Coefficients, Predictor Importances

Intercept	logP	CRE	HOMO
4.592	-0.3460	-0.01330	0.2772
Importance	-0.6630	-0.07187	0.05512

Response Variable: logTI
Number of Cross-validation Groups: 27
Optimal Number of Components: 4
Goodness-of-Fit and Goodness-of-Prediction

Comp	Residual SS	R2	Resid PRESS	R2cv
1	15.47	0.1547	18.08	0.0123
2	14.36	0.2153	24.28	-0.3269
3	13.34	0.2709	17.37	0.0507
4	12.90	0.2950	16.60	0.0929

Regression Coefficients, Predictor Importances

Intercept	log2P	logP	CRE	pKa
2.508	-0.1326	0.2897	-0.03266	-0.08532
Importance	-0.8611	0.5947	-0.1892	-0.1180