

# Supplementary material

## Long range electromagnetic effects drive protein-protein approaches: an egg of Coulomb

Neri Niccolai<sup>1,2\*</sup>, Edoardo Morandi<sup>1</sup>, Davide Alloci<sup>3</sup>, Alberto Toccafondi<sup>4</sup> and Andrea Bernini<sup>1</sup>

<sup>1</sup> Department of Biotechnology, Chemistry and Pharmacy, University of Siena, 53100 Siena, Italy.

<sup>2</sup> SienaBioGrafix Srl, 53100 Siena, Italy.

<sup>3</sup> Swiss Institute of Bioinformatics, 1211 Geneva 4, Switzerland

<sup>4</sup> Department of Information Engineering, University of Siena, 53100 Siena, Italy

### Atom depth calculations

Atom depth is considered among the structural descriptors that correlate protein structure with folding and functional properties. The distance between an atom and the nearest water molecule or the closest surface dot has been proposed as a measure of the atom depth, but, in both cases, the three-dimensional character of depth is lost. To account for the overall molecular shape in atom depth calculations, measurements of intersections between the molecular volume and a sphere of a suitable radius, centered on the atom whose depth has to be quantified, have been proposed by the authors in a previous paper [1]. It is apparent that the smaller is the exposed volume, the deeper is the three-dimensional insertion of the investigated atom inside the molecular structure.

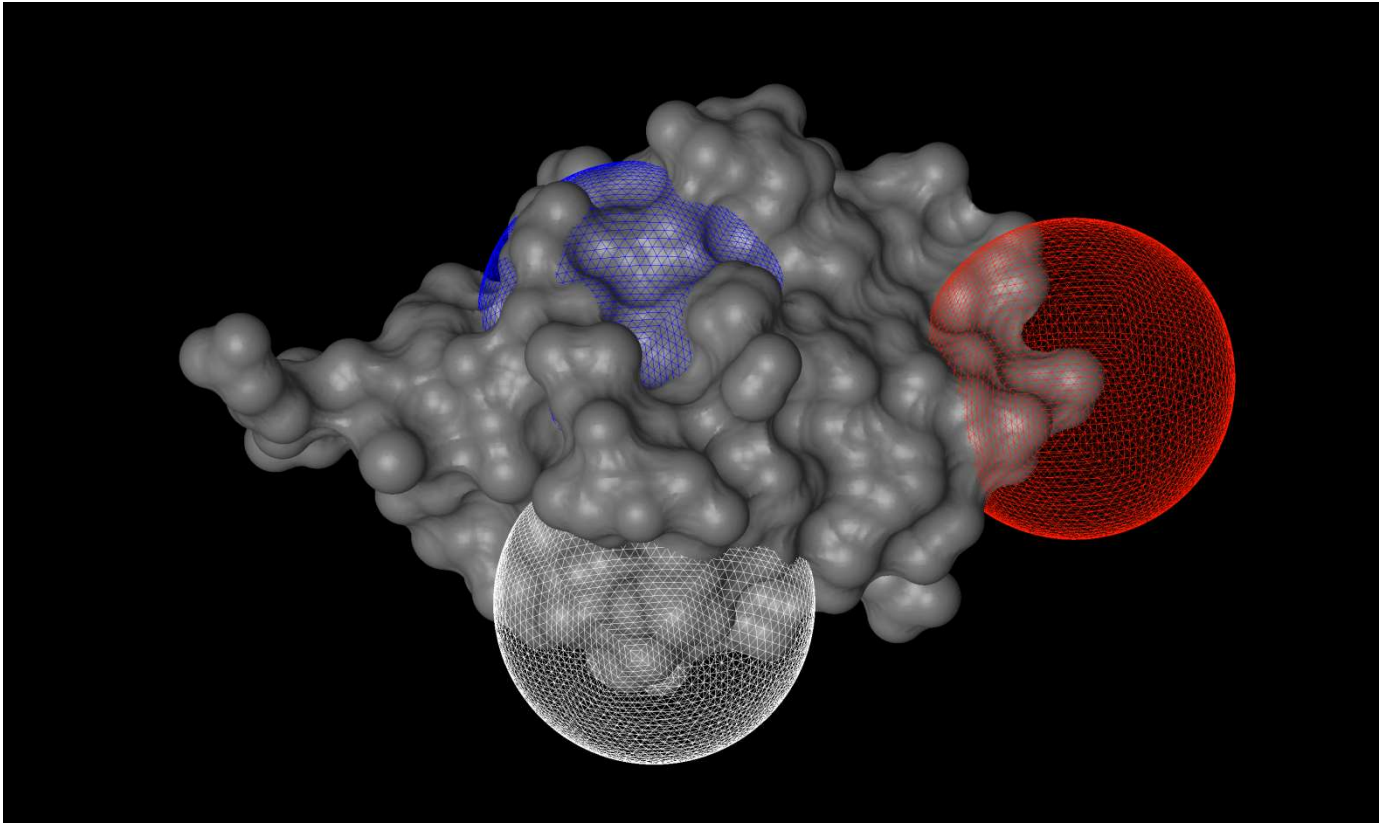
As, in general, depth is a very relative term which depends on the overall size and shape of the object under discussion, atom depth indexes,  $D_i$ 's are suggested as useful parameters to compare atom insertions within the considered molecular systems. Thus, for an atom  $i$  of a given molecule and a sampling radius  $r$ , a depth index  $D_{i,r}$  may be defined as

$$D_{i,r} = 2V_{i,r}/V_{0,r}$$

where  $V_{i,r}$  is the exposed volume of a sphere of radius  $r$  centered on atom  $i$  and  $V_{0,r}$  is the exposed volume of the same sphere when centered on an isolated atom. Thus, the limiting  $D_i$  values 0 and 2 are obtained for a totally buried or totally isolated atom, respectively.

---

[1] D. Varrazzo, A. Bernini, O. Spiga, A. Ciutti, S. Chiellini, V. Venditti, L. Bracci, N. Niccolai, Three-dimensional computation of atom depth in complex molecular structures, *Bioinformatics* 21 (2005) 2856-2860.



The exposed volumes of L43 CA (buried, color blue,  $D_i = 0.05$ ), E34 CD (intermediate, color white,  $D_i = 0.76$ ), K63 NZ (exposed, color red,  $D_i = 1.44$ ) are shown as mesh on Ubiquitin surface (PDB: 1UBQ) by drawing a sphere of 9 Å radius on each atom.

The algorithm for the calculation of  $D_i$ , extensively described in [1], has been implemented in the software Simple Atom Depth Index Calculator (SADIC), available at <http://www.sbl.unisi.it/prococoa>. The software outputs the  $D_i$  values calculated for the protein structure given in PDB format in the B-factor field of the same PDB file, allowing for coloring by depth index with the most common graphical viewers, like PyMol (<http://www.pymol.org>).

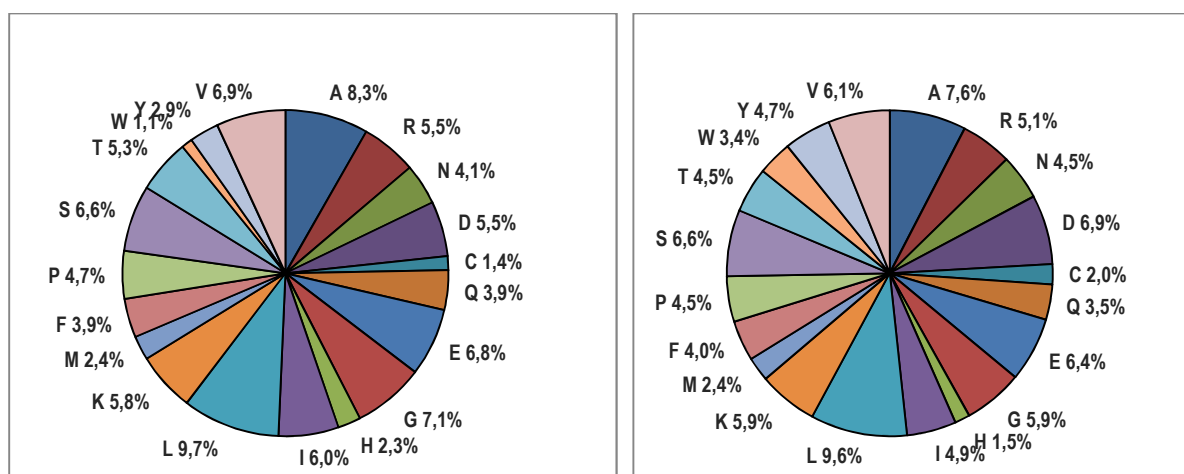
Table SI

Atom	# interactions	$D_i$	DSSP
NZ K6	1	1.29	E
NZ K11	2	1.06	
CD E16	2	1.22	E
CD E18	2	1.16	
CG D21	2	0.60	S
CD E24	3	1.17	H
NZ K27	6	0.35	H
NZ K29	4	0.77	H
CG D32	2	1.23	H
NZ K33	4	0.93	H
CD E34	3	0.76	H
CG D39	4	0.90	G
CZ R42	5	1.00	E
NZ K48	2	1.18	E
CD E51	5	1.19	
CG D52	4	0.66	T
CZ R54	3	1.21	S
CG D58	1	0.92	G
NZ K63	1	1.44	T
CD E64	1	1.10	T
CE1 H68	1	0.91	E
CZ R72	4	0.75	
CZ R74	2	1.42	S

Table SII

Atom	# interactions	$D_i$	DSSP
CD E23	6	1.55	
CG D24	3	1.33	
CZ R28	2	0.96	E
CG D32	1	1.30	E
NZ K35	4	1.17	E
NZ K38	0	1.00	E
NZ K40	1	1.42	E
CG D42	3	1.30	
CZ R44	2	1.77	S
CD E47	5	1.17	S
CD E53	3	0.19	H
NZ K56	3	0.21	H
CD E61	2	0.93	T
CZ R68	4	0.66	
CD E72	1	1.24	
CG D74	2	1.03	E
NZ K81	2	1.10	E
CD E88	5	1.10	E
NZ K90	3	1.16	E
NZ K93	1	1.25	S
NZ K94	2	1.43	
CG D98	3	1.10	H
CG D99	4	1.00	H
CD E102	4	1.10	H
NZ K106	5	1.28	H
CZ R107	6	0.43	T
NZ K109	3	1.00	H
CD E110	5	0.90	H
CZ R118	1	1.10	T
NZ K123	3	1.17	
NZ K127	2	0.99	E
CG D130	3	0.87	E
CD E136	1	1.11	S
CD E140	1	1.01	
NZ K141	2	0.98	
CD E146	1	1.03	T
CG D148	2	0.83	
CD E149	2	0.22	
CZ R154	0	1.06	E
CG D161	2	1.38	T
CZ R162	4	1.11	E
CZ R176	3	0.05	H

CE1 H177	4	0.05	H	CG D453	0	0.91	
CD E180	3	0.07	H	NZ K457	3	1.29	H
CG D189	1	0.54	E	CD E461	3	1.16	H
CG D190	2	0.51	T	NZ K462	4	1.16	T
CZ R192	5	0.78	T	CZ R465	4	0.64	E
CG D193	3	0.96	T	CG D472	1	0.74	H
CE1 H194	4	0.95	E	CG D477	2	0.26	T
CZ R199	2	1.13	S	NZ K499	2	0.91	H
CG D200	2	0.96	S	CG D502	2	0.80	H
CG D205	1	0.37	E	CD E511	3	1.31	G
NZ K206	1	1.19		CE1 H512	3	1.10	G
CZ R212	2	0.13	E	CZ R513	2	1.22	G
CG D217	3	0.14	E	CG D514	1	1.02	G
CZ R220	6	0.23	S	CD E520	1	0.01	E
CD E226	3	1.15	H	CD E526	3	0.36	T
NZ K229	2	0.94	H	CG D529	0	0.88	
CZ R230	2	0.74	H	CG D534	1	0.57	H
CD E233	3	0.51	H	CZ R536	1	0.64	H
NZ K239	3	0.03	T	CZ R540	1	0.34	H
CE1 H244	4	0.02	E	CD E546	2	0.18	H
CE1 H246	7	0.06	E	CZ R547	2	0.52	H
CG D249	6	0.14		CD E551	2	0.54	
NZ K259	5	1.05	S	CG D558	3	0.95	G
CZ R260	3	0.89	S	CD E560	1	0.58	H
NZ K265	0	1.20	H	CZ R562	2	0.25	H
NZ K273	0	1.15	S	CE1 H565	4	0.10	H
NZ K277	4	1.18	H	CZ R567	2	0.75	H
CZ R281	5	1.10	H	CD E568	3	0.53	H
CD E282	4	1.08	H	CZ R569	2	0.13	H
CZ R285	4	0.87	H	CZ R572	1	0.77	H
CD E289	4	0.68	H	CD E578	1	0.82	
CZ R290	4	0.49	T	CD E583	2	0.84	H
CZ R293	1	0.22	E	CD E589	2	0.33	T
CD E297	5	0.04	E				
CG D299	6	0.06	E				
CE1 H303	8	0.22	S				
CD E306	1	0.77	T				
CG D310	1	1.30	T				
CG D312	1	1.11					
NZ K318	3	1.24	T				
CD E320	1	1.30	S				
NZ K323	0	1.34	G				
CD E328	4	0.43	S				
NZ K338	2	1.01					
CG D339	2	1.09	T				
CD E340	3	1.05	H				
CD E346	3	0.95	H				
CG D347	3	0.80	H				
CG D351	3	0.65	H				
CD E354	4	1.04	H				
CG D357	3	0.77					
CG D360	2	0.45	S				
CE1 H363	3	0.04	E				
CG D367	3	0.18					
CD E368	4	0.56					
CD E371	2	0.83	H				
CG D378	0	1.25	H				
CZ R388	1	1.57	T				
CG D390	2	1.29					
CG D392	3	1.21	S				
NZ K393	4	1.19	S				
CD E394	5	1.18	S				
NZ K398	3	1.05	H				
NZ K406	3	0.93	H				
CG D409	3	0.97	H				
NZ K410	4	0.98	H				
NZ K413	3	1.36	H				
NZ K417	2	0.94	S				
NZ K418	2	1.12					
NZ K432	3	1.27	T				
CE1 H433	4	0.81	T				
CG D435	3	0.97	H				
CG D436	3	1.08	H				
NZ K440	4	1.21	T				
CG D441	4	1.15	T				
CG D442	2	0.88	T				



**Figure S1:** left panel: average amino acid composition of proteins contained in release 2014\_06 of UniProtKB/Swiss-Prot (<http://web.expasy.org>); right panel: amino acid composition of *Ostrinia furnacalis* chitinolytic enzyme (UniProtKB/Swiss-Prot ID: Q06GJ0)

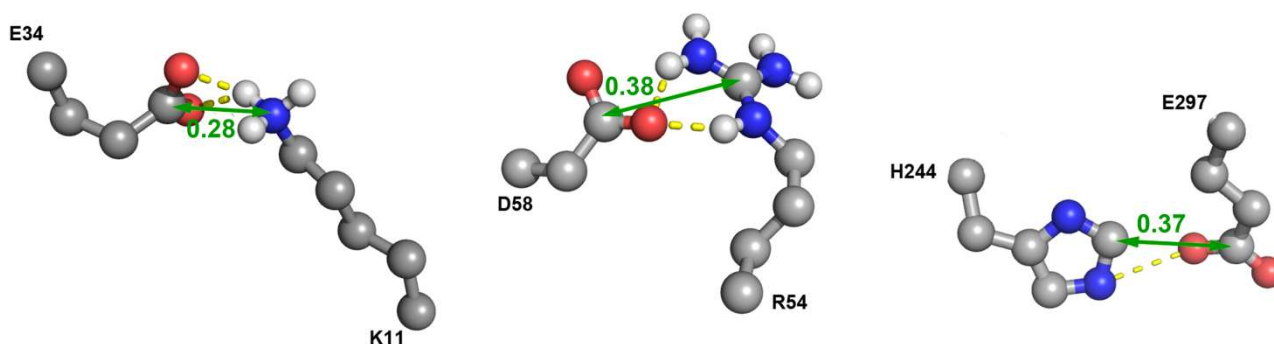
### Human ubiquitin

MQIFVKTLTGKITLEVEPSDTIENVKAKIQD<sup>33</sup>K<sub>H</sub>E<sub>H</sub>GIPPDQQRLLIFAGKQLEDGRTLSDYNIQ<sup>63</sup>K<sub>T</sub>E<sub>T</sub>STLHLVLRRLGG

### *Ostrinia furnacalis* chitinolytic enzyme

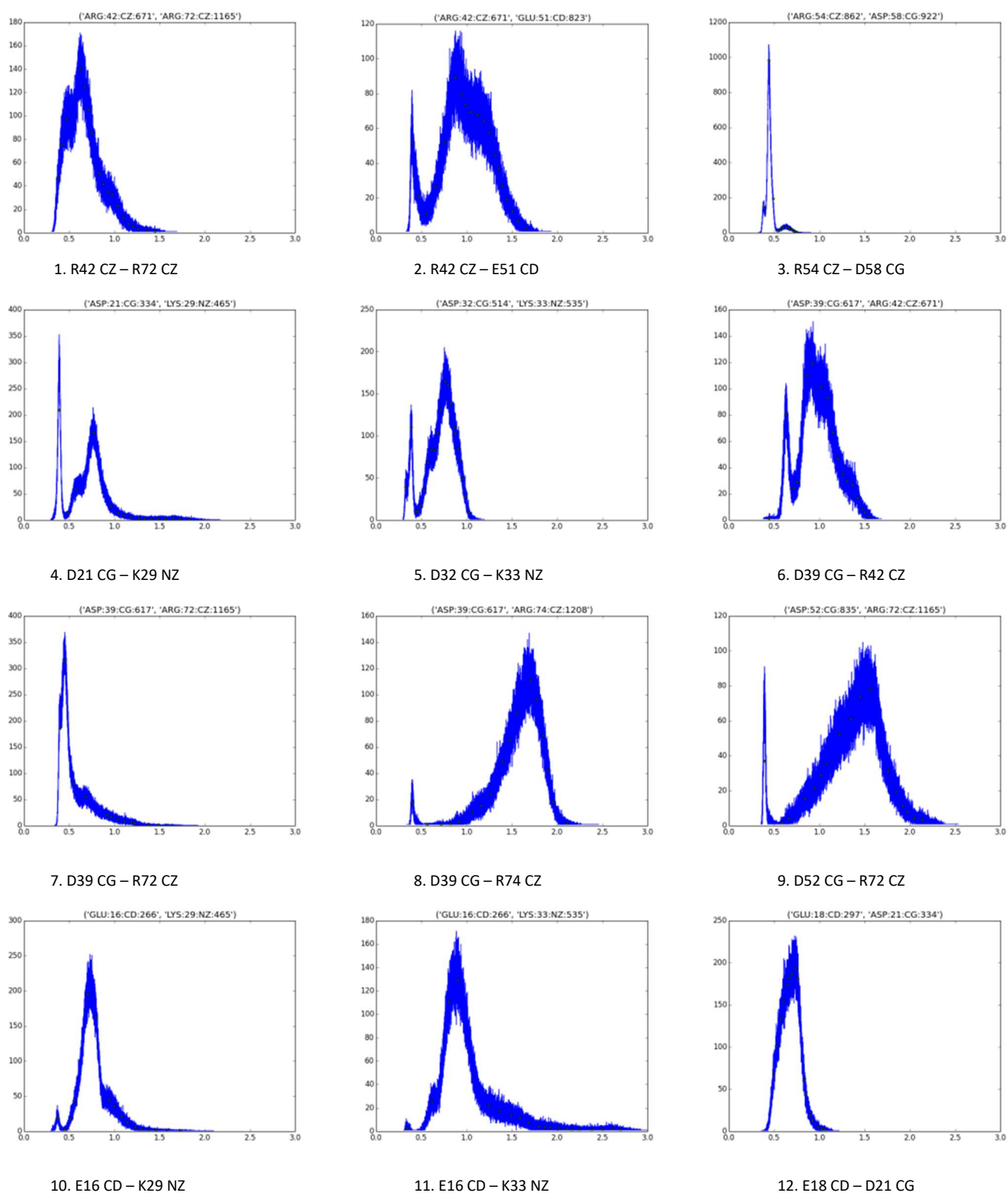
EDVVWRWSCDNGKCVKLNDRSSEPALSLEACKMFCNEYGLLWPRPTGEADLGNFLSKINLNSIEVKILKKGATDDLMEA  
 AAKRF<sup>109</sup>K<sub>H</sub>E<sub>H</sub>QVSLAIPRGSTPKLTGKAVDVVLVNNENPN<sup>140</sup>E<sub>K</sub>AFSLEMDESGLRVSPSGA<sup>161</sup>D<sub>T</sub>R<sub>E</sub>VNATITANSFFGMRH  
 GLETLSQLFVDDI<sup>192</sup>R<sub>T</sub>D<sub>T</sub>HLLMV<sup>199</sup>R<sub>S</sub>D<sub>S</sub>VNIS<sup>205</sup>D<sub>H</sub>K<sub>H</sub>PVYPYRGILLDTARNYYSIESIKRTIEAMAAVKLNTFHWHTDSQS  
 FPFVTTKRPNLYKFGALSPQKVYTKAAI<sup>281</sup>R<sub>H</sub>E<sub>H</sub>VVRFGI<sup>289</sup>E<sub>H</sub>R<sub>H</sub>GVRVLPEFDAPAHVGGEGWQDQDLDLTVCFKAEPWKSQCV  
 EPPCGQLNPT<sup>338</sup>K<sub>D</sub>T<sub>E</sub>ELYQYLEDIYSDMAEVFDTTDFIFHMGGDEVSEACWNSSDSIQNFMMQNRWDL<sup>393</sup>K<sub>H</sub>E<sub>H</sub>SFLKLW  
 NYFQKKAQ<sup>409</sup>D<sub>H</sub>K<sub>H</sub>AYKAFGKKLPLILWTSTLTNYKHIDDYLN<sup>440</sup>K<sub>T</sub>D<sub>T</sub>DIHQVWTTGVDPQIKGLL<sup>461</sup>E<sub>H</sub>K<sub>T</sub>GYRLIMSNYDA  
 LYFDCGYGAWVGAGNNWCSPYIGWQKVYDNSPAVIALEH<sup>513</sup>R<sub>G</sub>D<sub>G</sub>QVLGGEALWSEQSDTSTLDGRLWPRAAALA<sup>546</sup>E<sub>H</sub>  
 R<sub>H</sub>LWAEPATSWQDAEYRMLHIR<sup>568</sup>E<sub>H</sub>R<sub>H</sub>LVRMGIQAESLQPEWCYQNEGYCYS

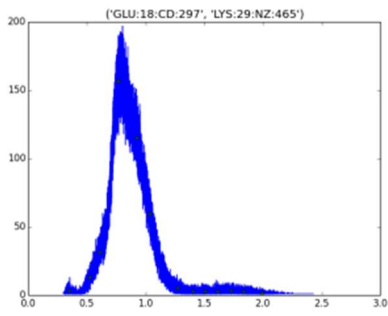
**Figure S2:** Dipeptide fragments of charged amino acids are highlighted in red; numbers refer to residue position and subscripts define DSSP assignments.



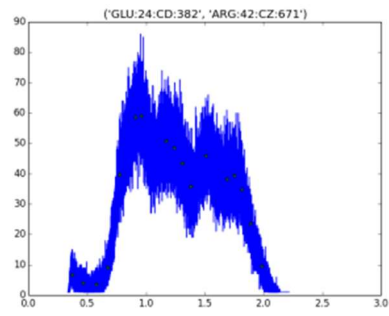
**Figure S3:** Closest approaches of charged side chains stabilized by the formation of hydrogen bonding. Green arrows highlight the distances  $r_C$  of E34 CG, D58 CD and E297 CG atoms respectively with K11 NZ, R54 CZ and H244 CE1 in *Ostrinia furnacalis* chitinolytic enzyme.

**Figure S4:** Binned distributions of close distance interactions between D, E, H, K and R side chains in hUBQ as calculated along the 500 ns MD trajectory with `g_distmap` tool. Only side chain-side chain approaches exhibiting more than 5 % occurrences at  $r_{IC} < 0.6$  nm have been considered.

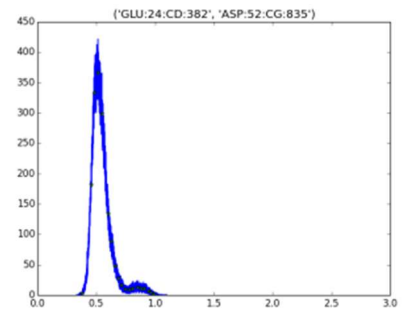




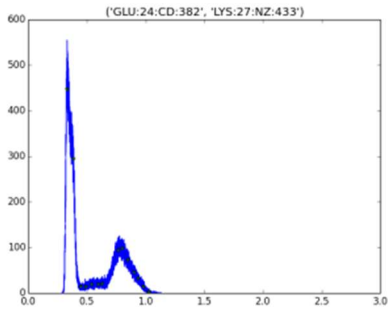
13. E18 CD - K29 NZ



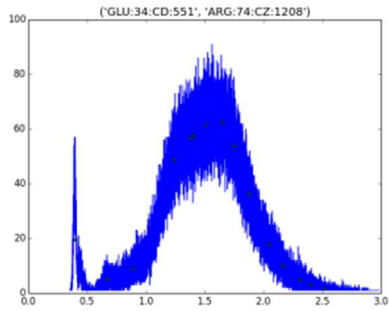
14. E24 CD - R42 CZ



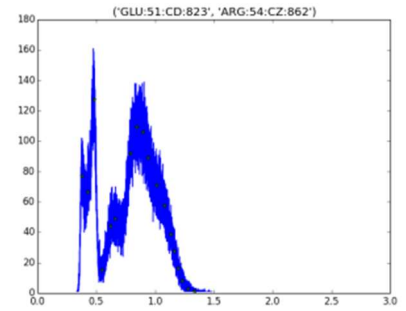
15. E24 CD - D52 CG



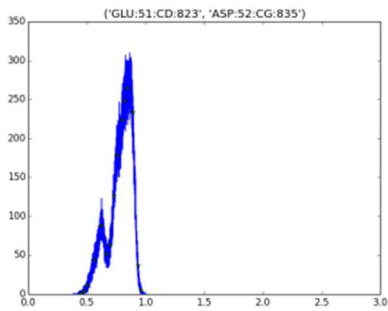
16. E24 CD - K27 NZ



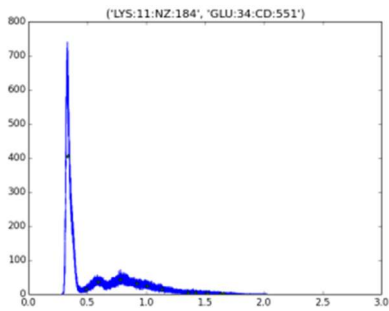
17. E34 CD - R74 CZ



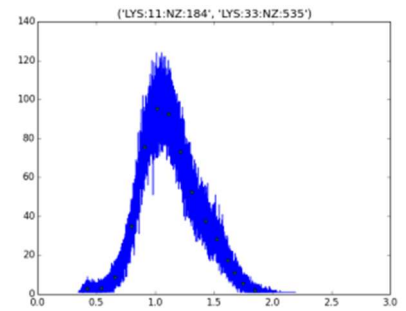
18. E51 CD - R54 CZ



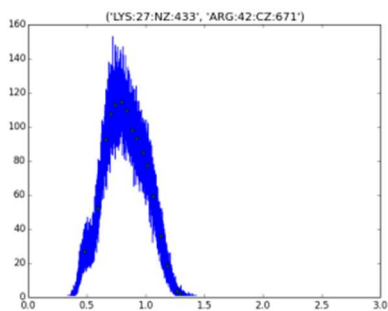
19. E51 CD - D52 CG



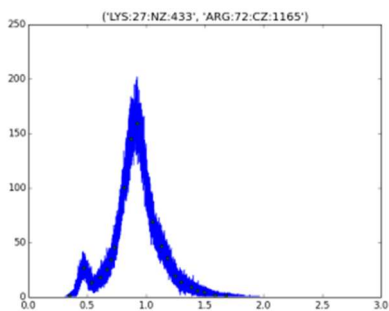
20. K11 NZ - E34 CD



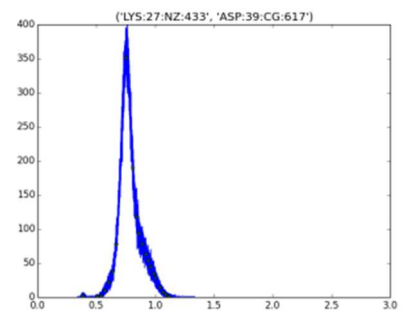
21. K11 NZ - K33 NZ



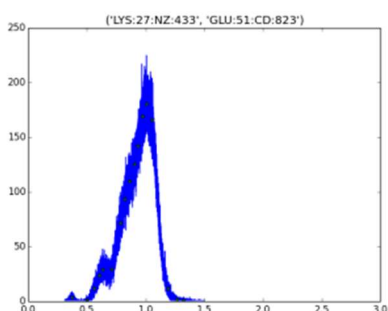
22. K27 NZ - R42 CZ



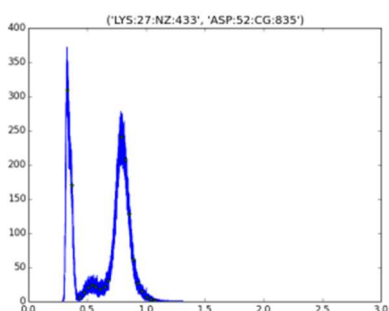
23. K27 NZ - R72 CZ



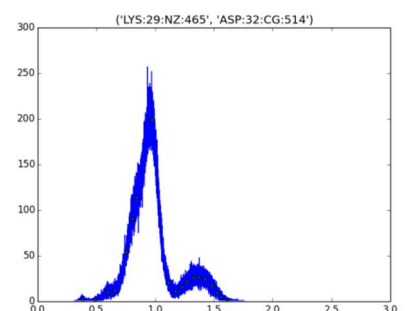
24. K27 NZ - D39 CG



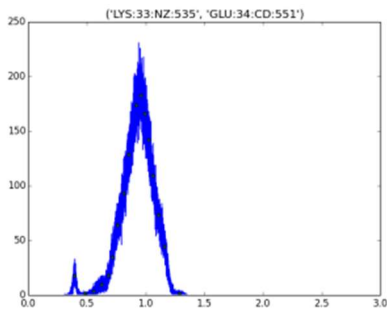
25. K27 NZ - E51 CD



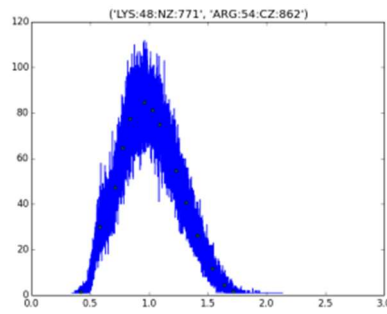
26. K27 NZ - D52 CG



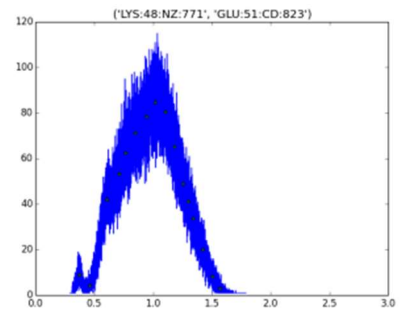
27. K29 NZ - D32 CG



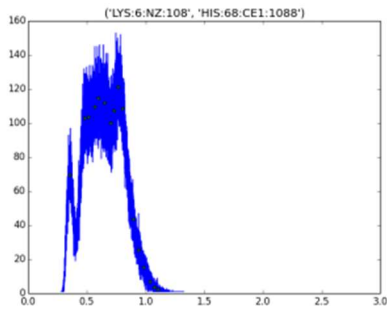
28. K33 NZ – E34 CD



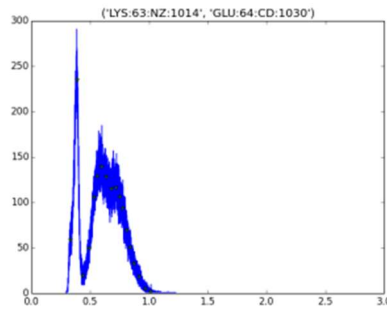
29. K48 NZ - R54 CZ



30. K48 NZ – E51 CD

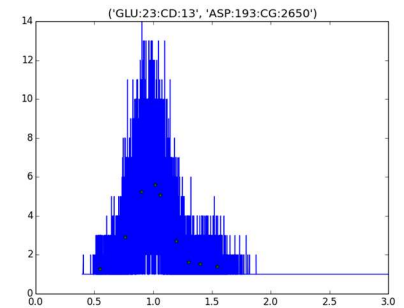
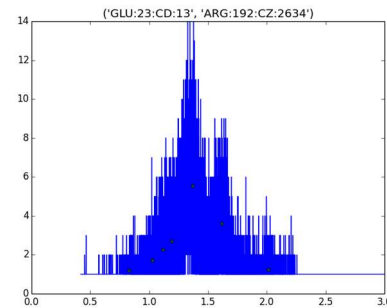
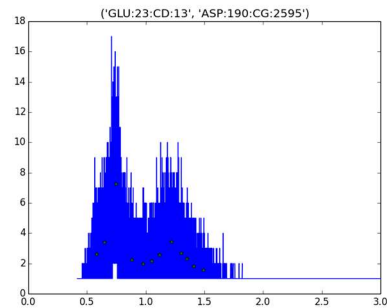
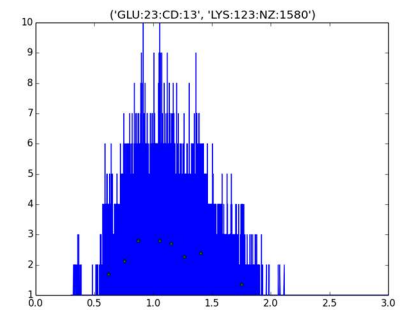
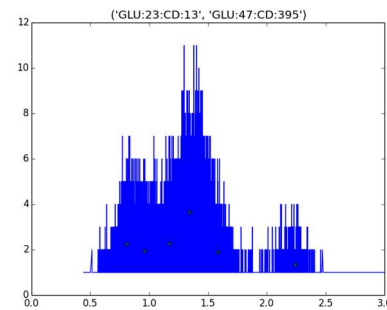
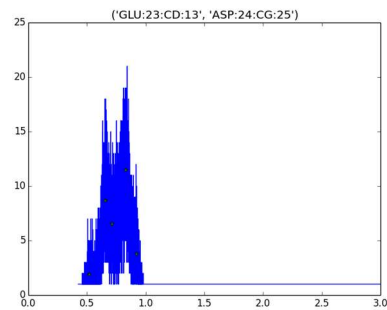


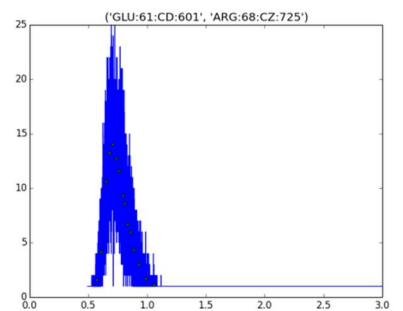
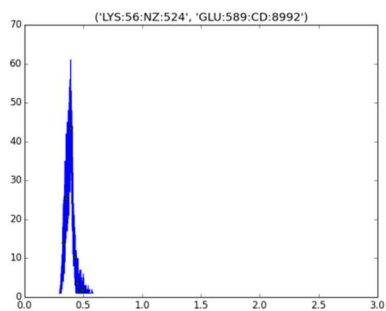
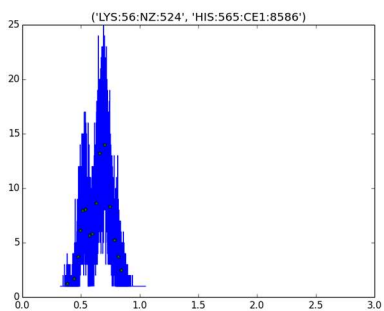
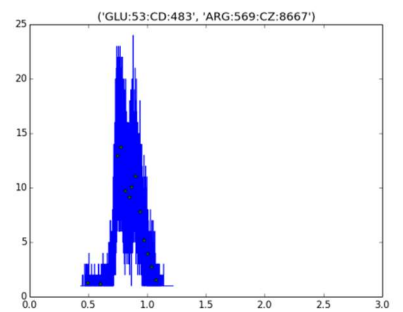
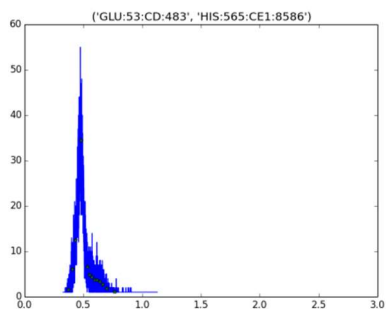
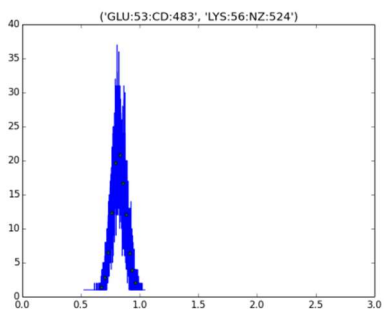
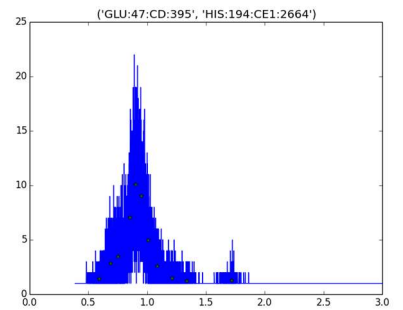
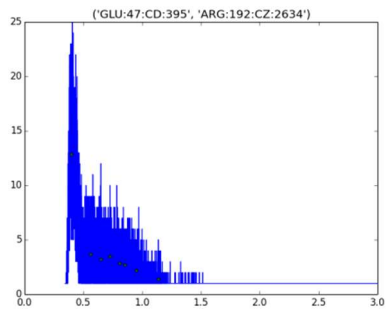
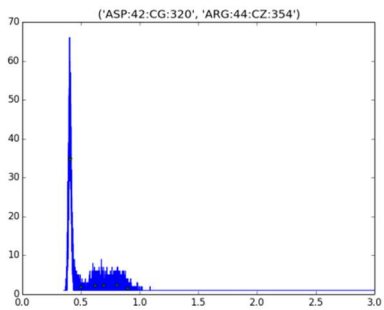
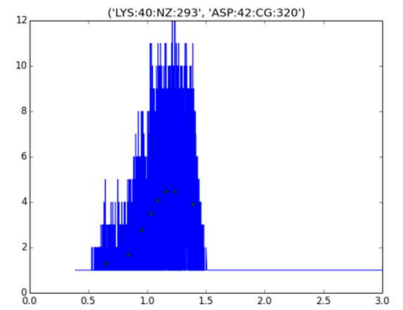
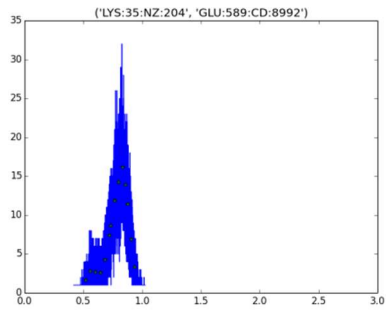
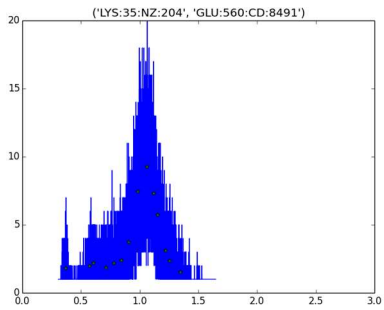
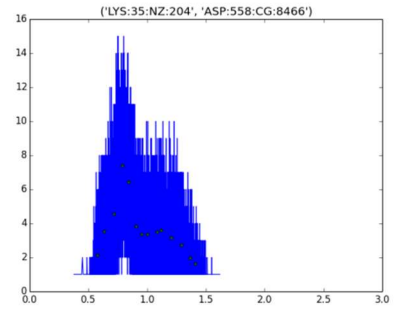
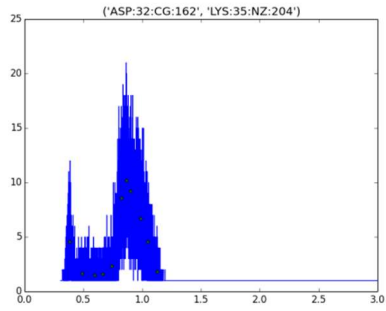
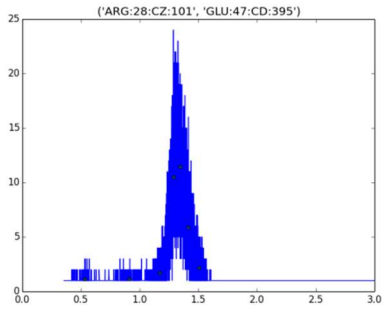
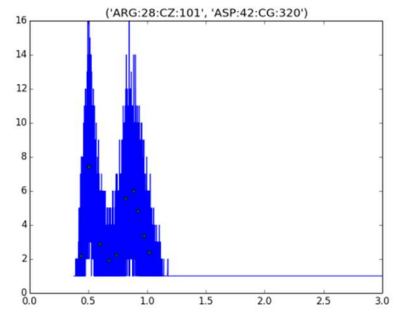
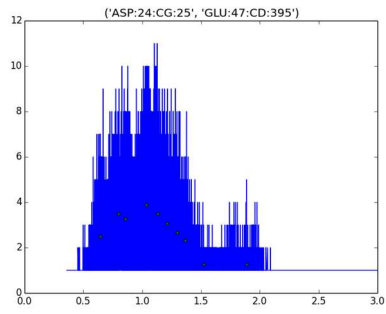
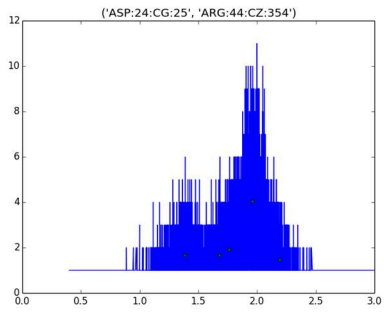
31. K6 NZ – H8 CE1



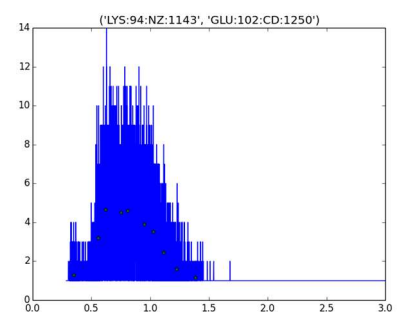
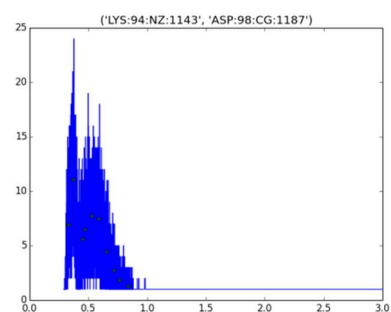
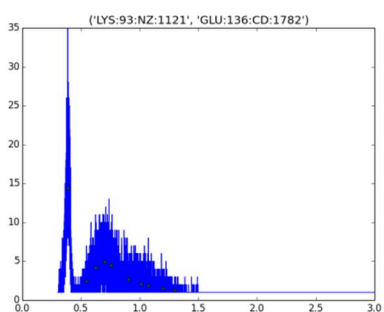
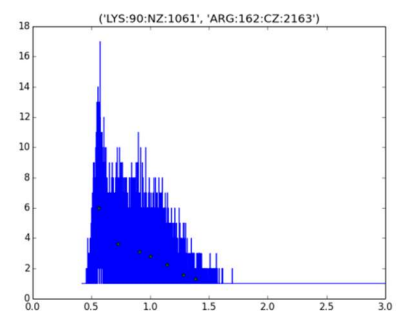
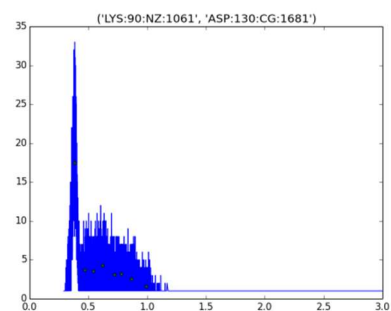
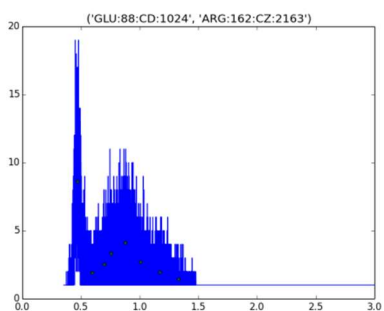
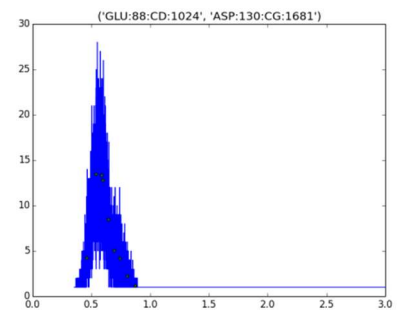
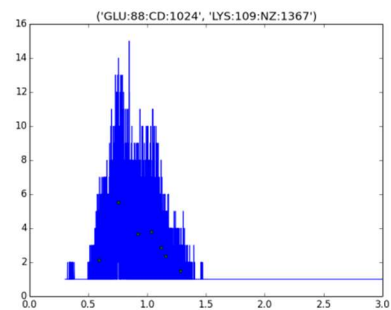
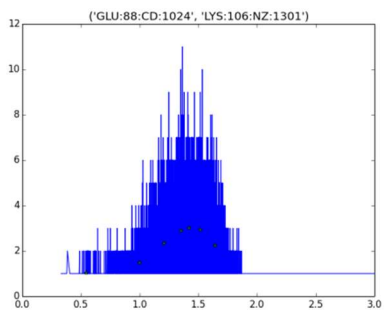
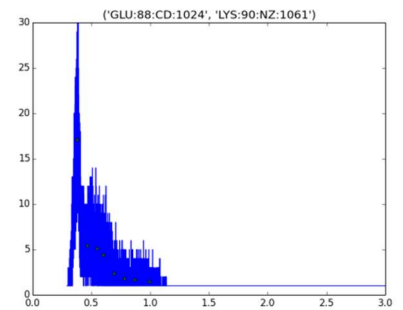
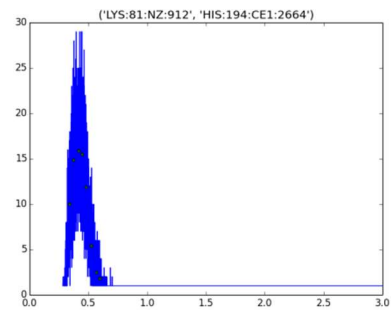
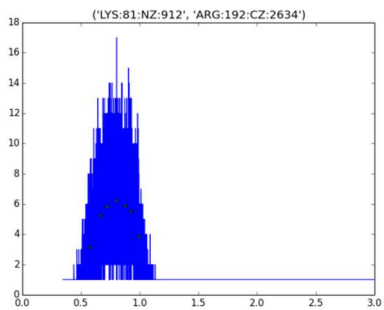
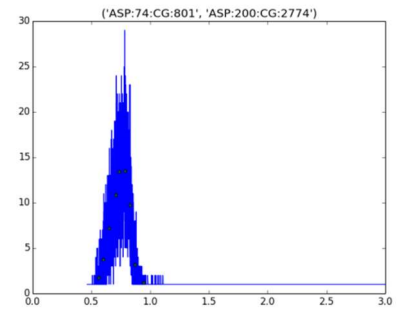
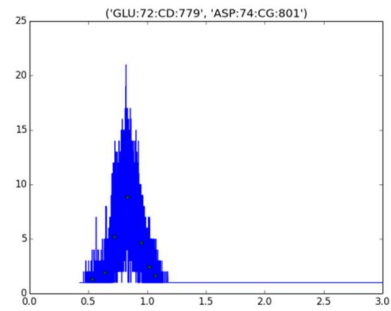
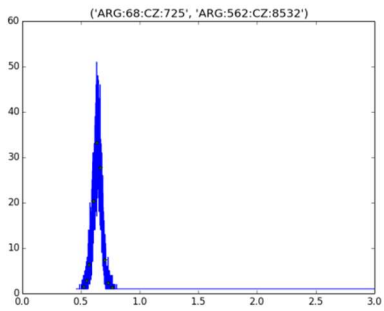
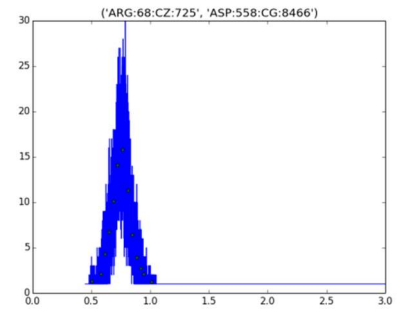
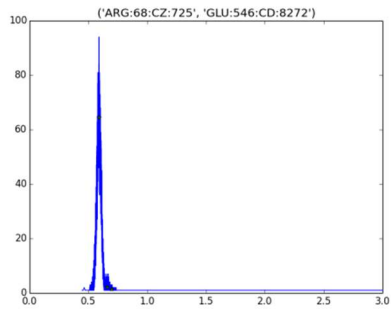
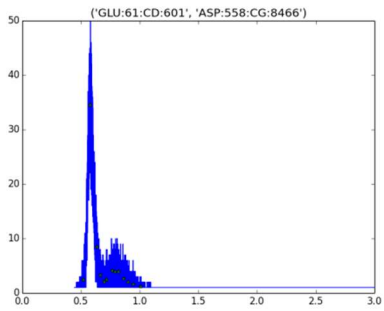
32. K63 NZ – E64 CD

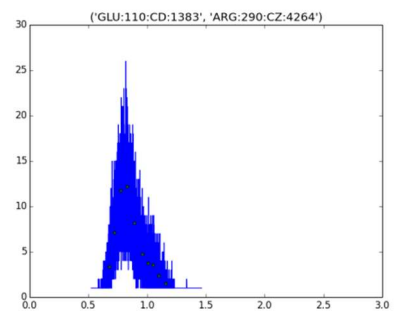
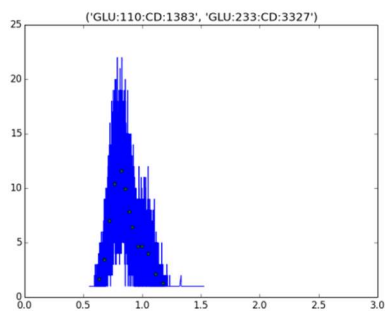
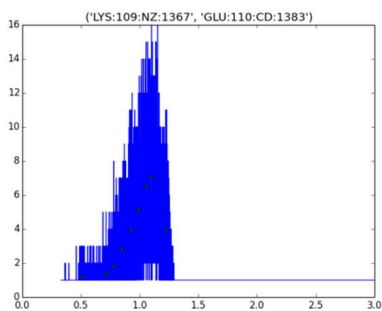
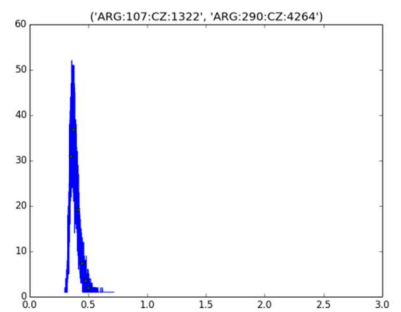
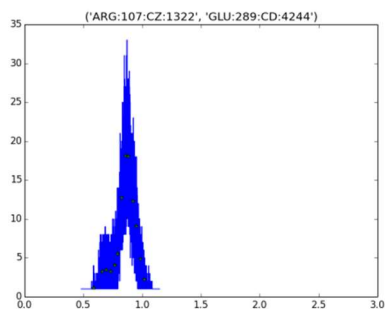
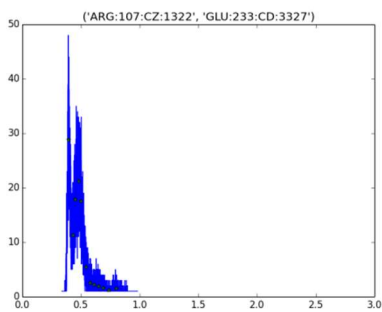
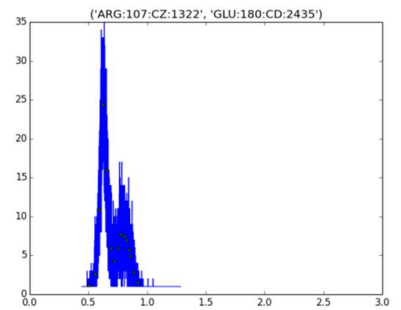
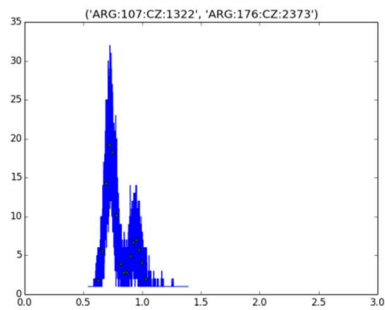
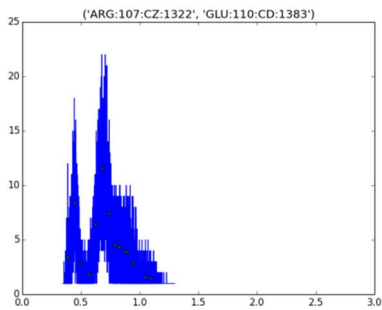
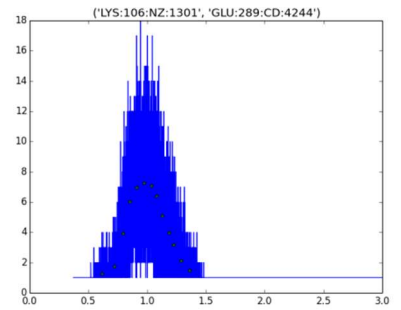
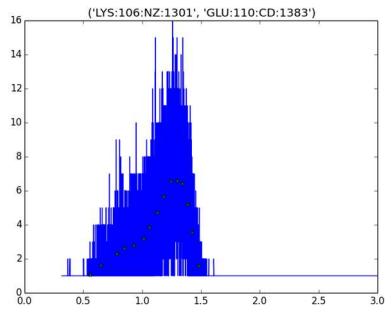
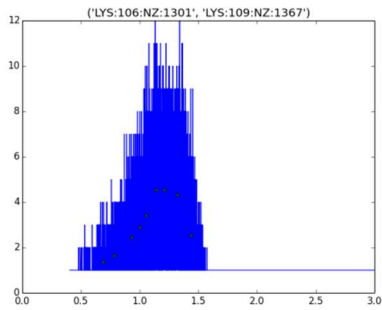
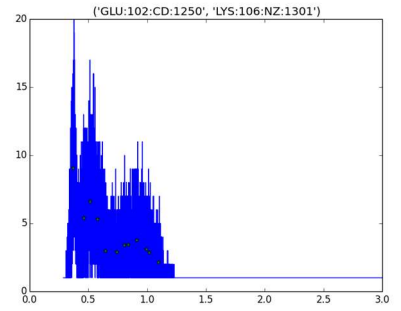
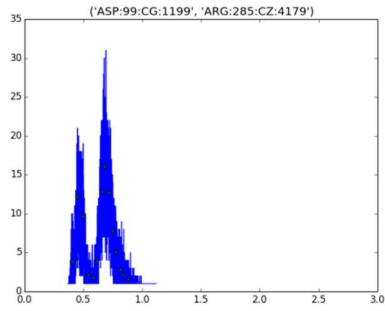
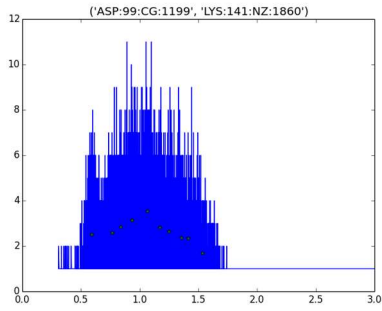
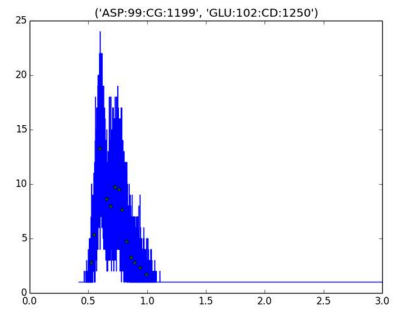
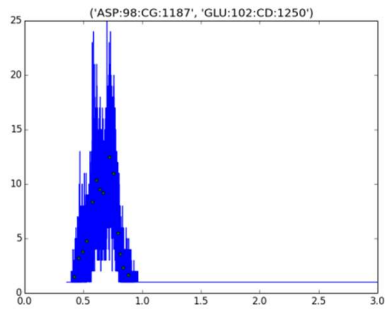
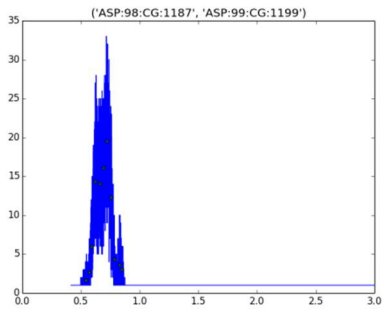
Figure S5: Binned distributions of close distance interactions between D, E, H, K and R side chains in oCE as calculated along the 30 ns MD trajectory with `g_distmap` tool. Only side chain-side chain approaches exhibiting more than 5 % occurrences at  $r_{IC} < 0.6$  nm have been considered.

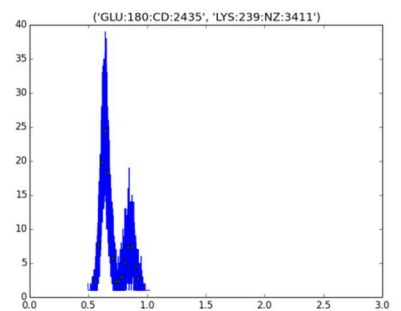
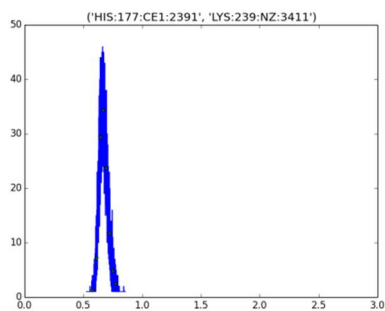
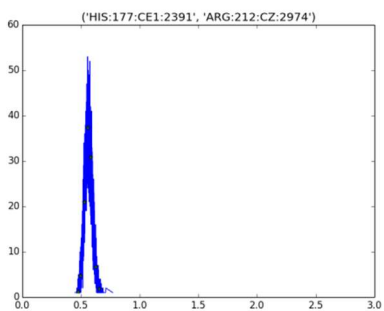
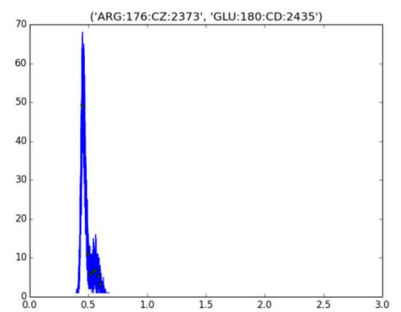
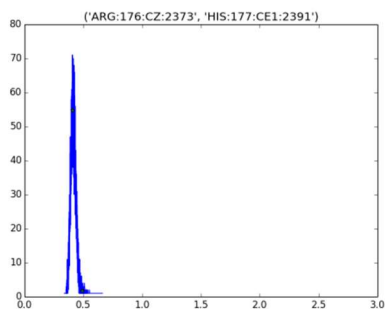
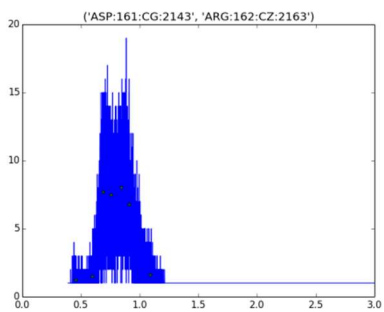
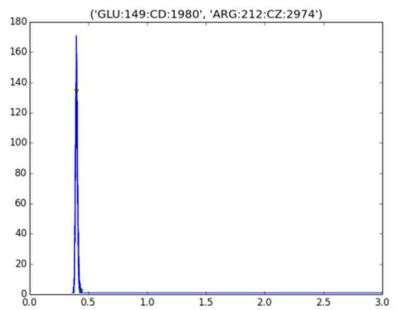
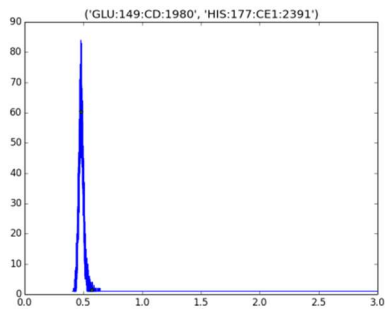
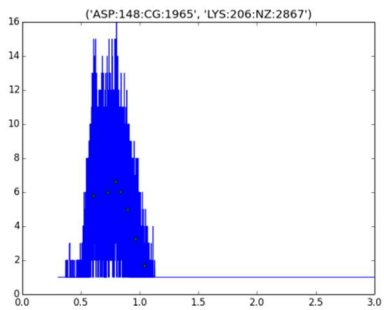
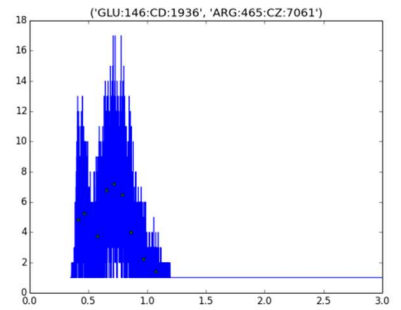
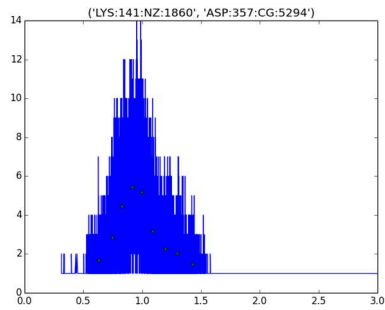
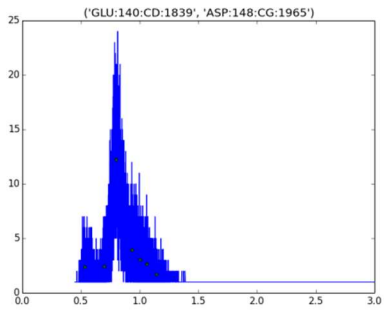
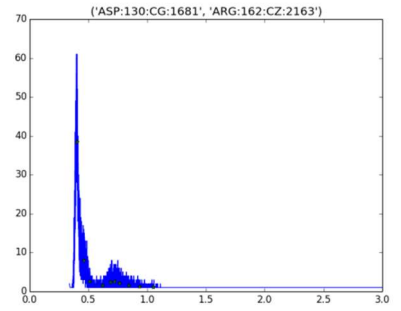
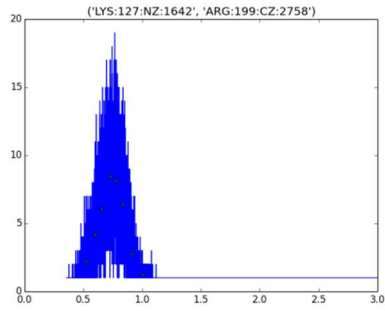
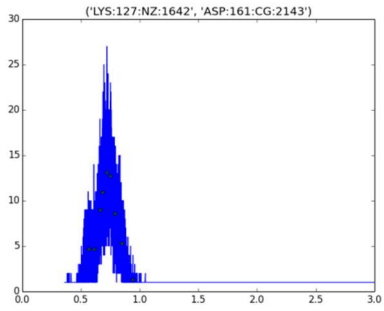
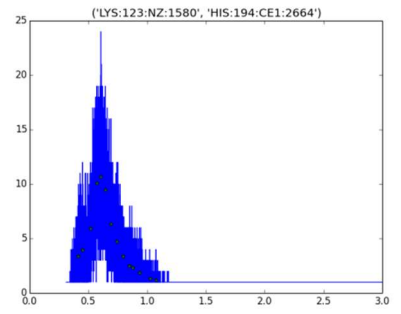
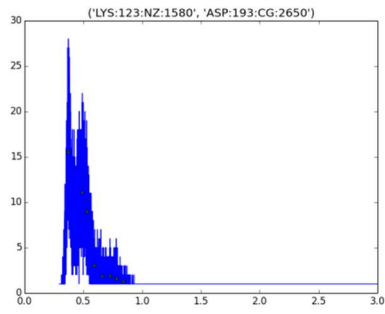
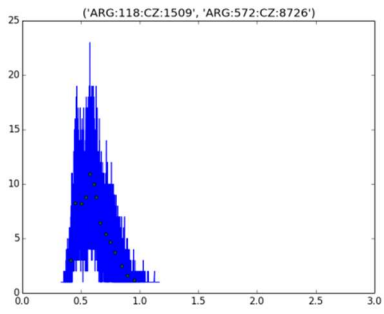


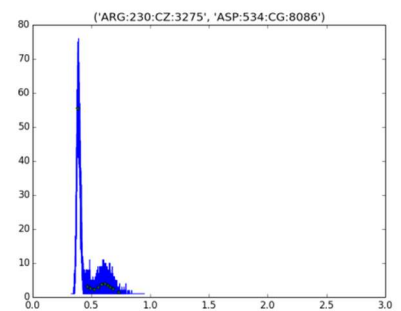
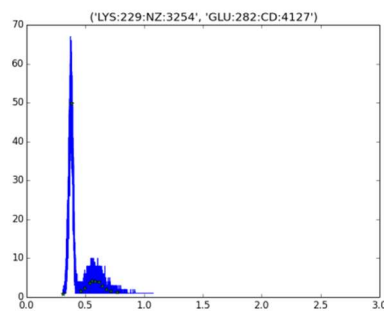
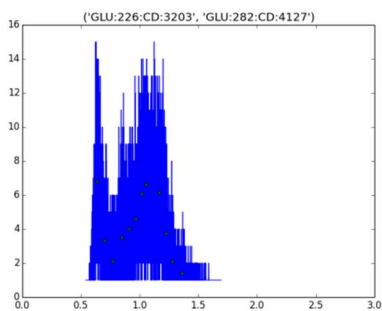
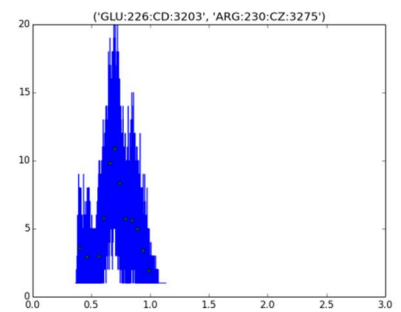
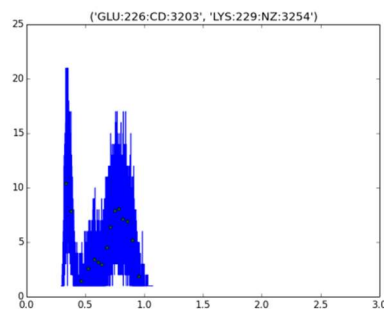
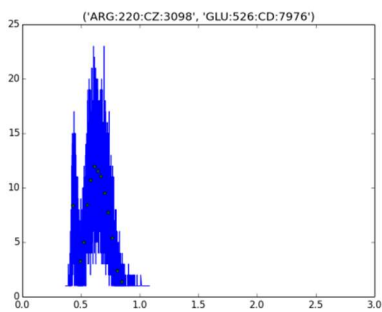
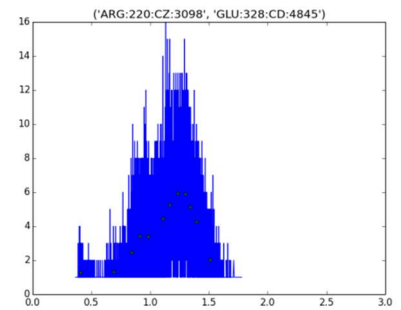
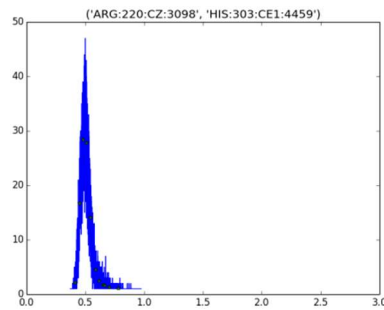
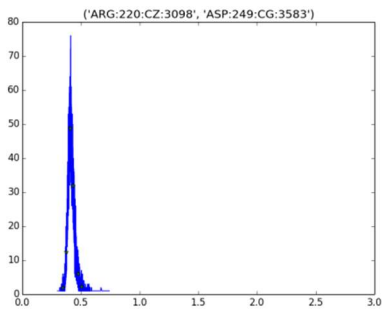
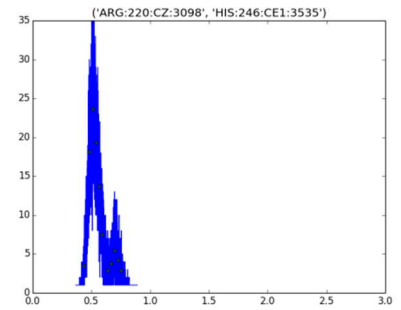
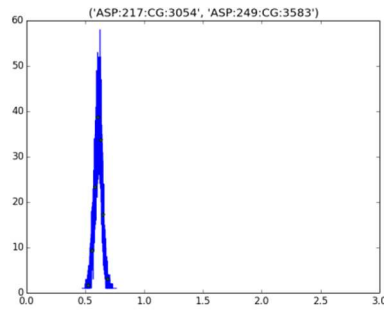
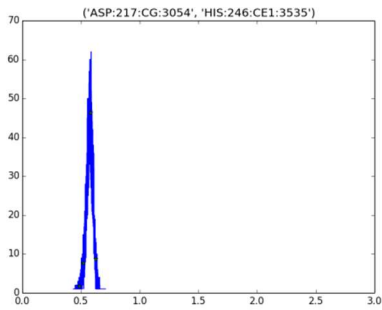
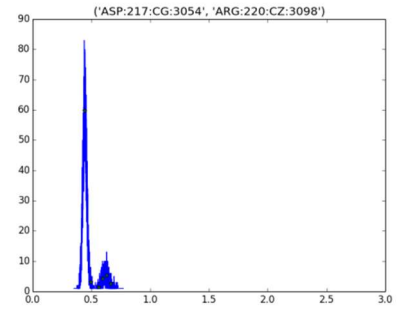
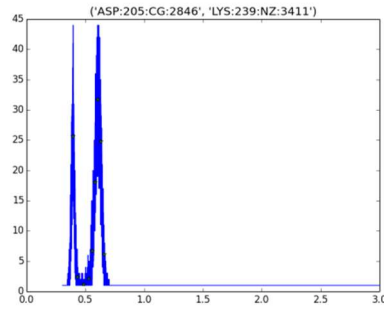
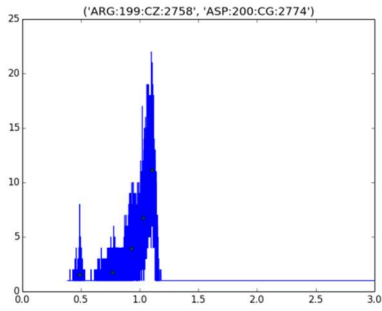
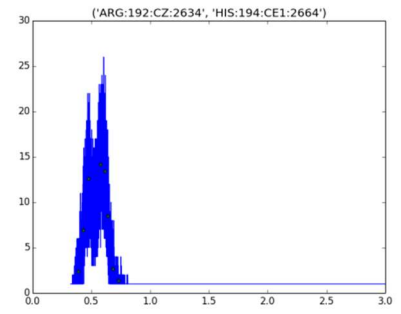
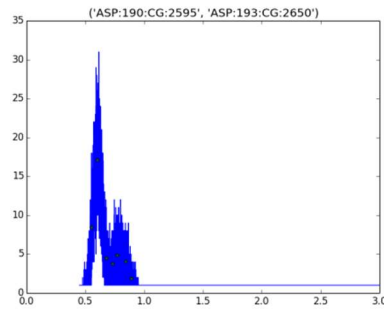
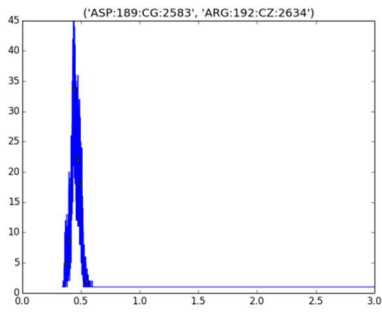


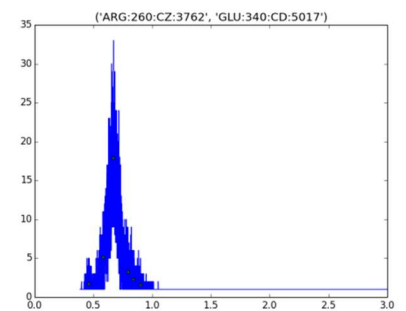
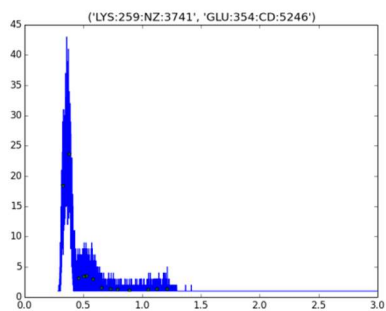
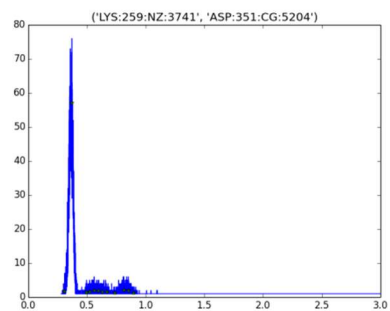
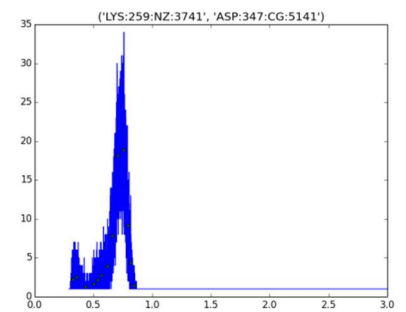
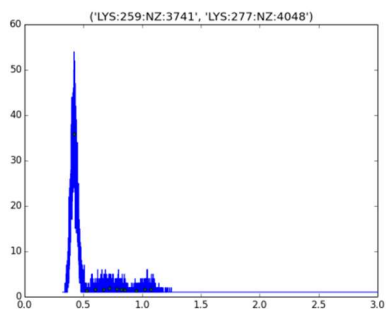
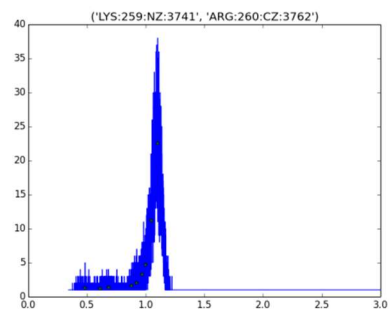
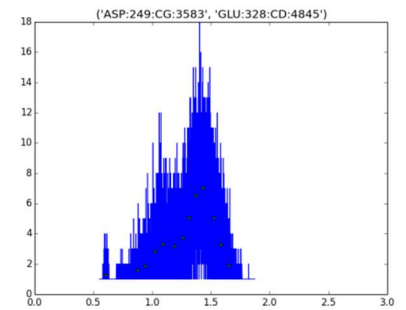
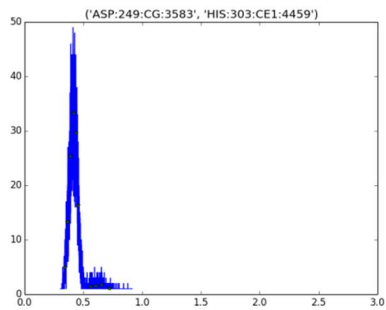
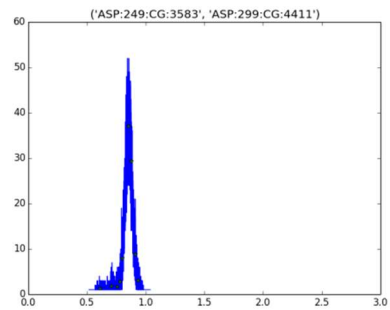
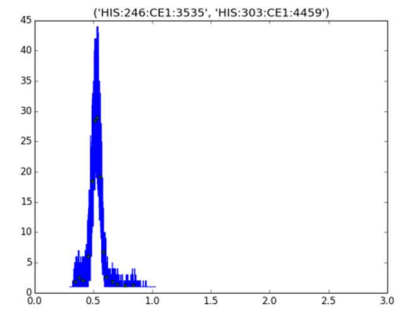
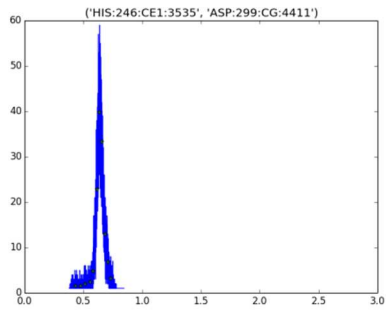
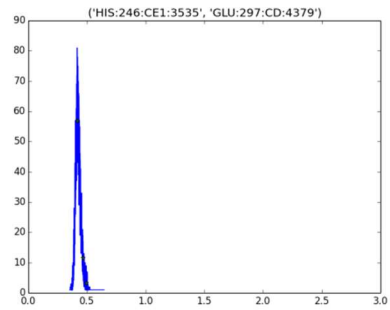
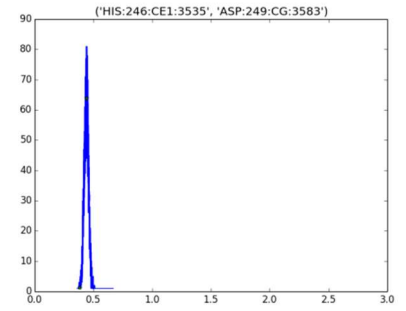
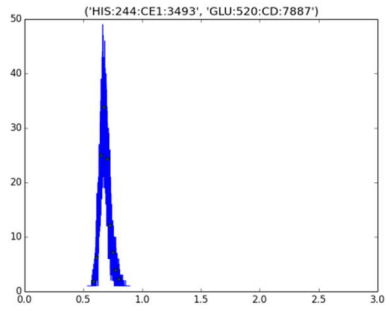
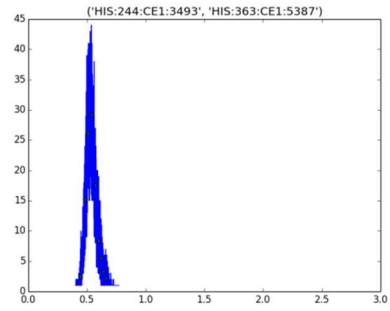
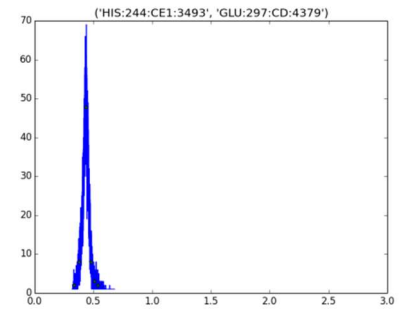
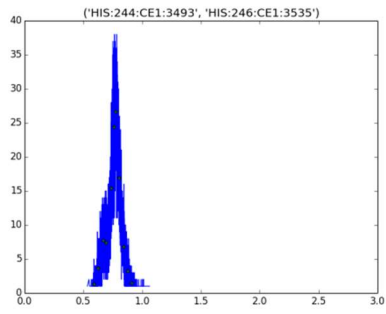
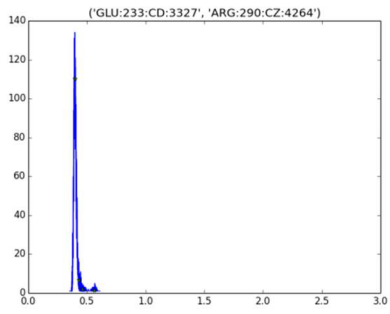


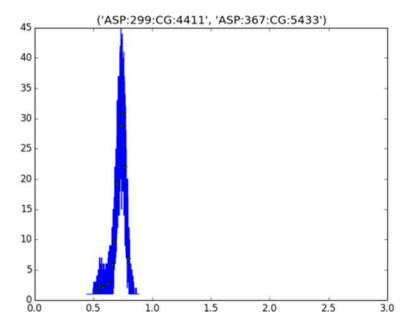
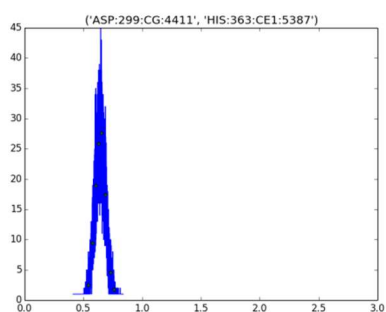
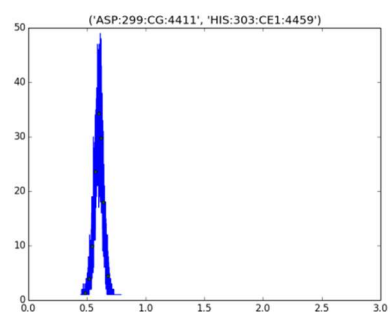
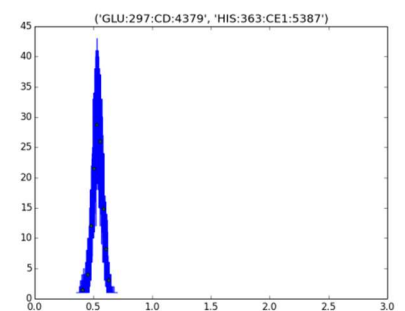
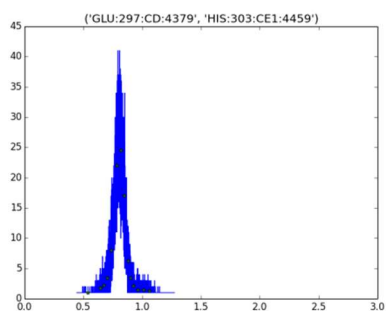
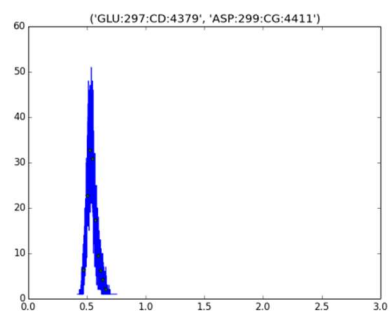
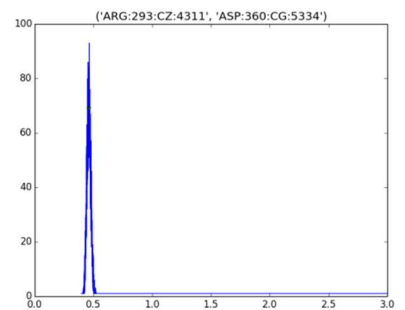
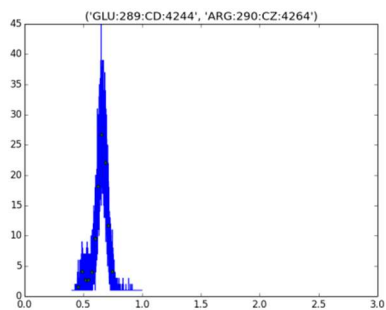
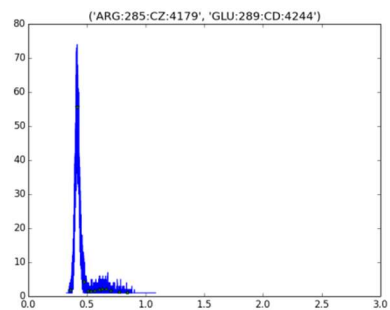
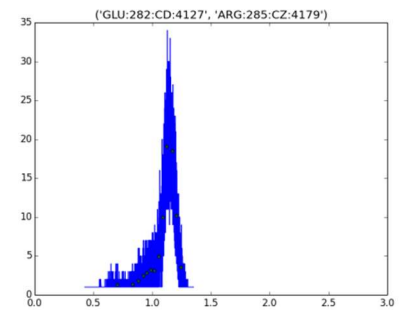
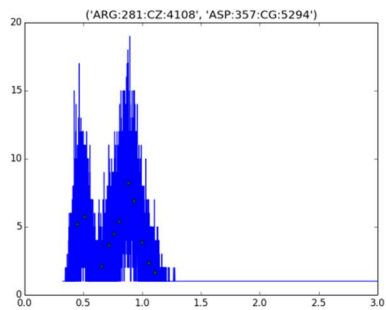
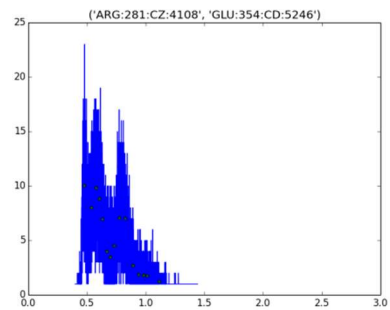
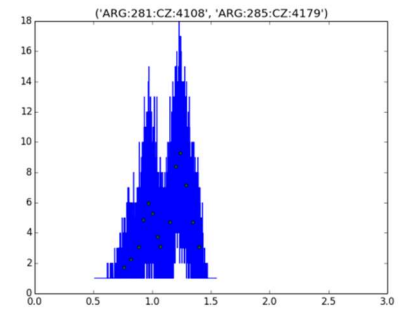
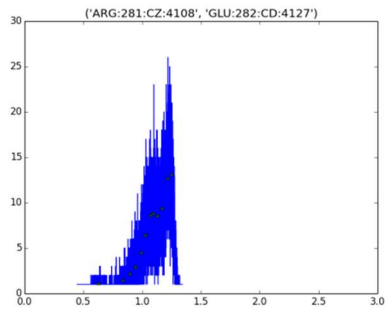
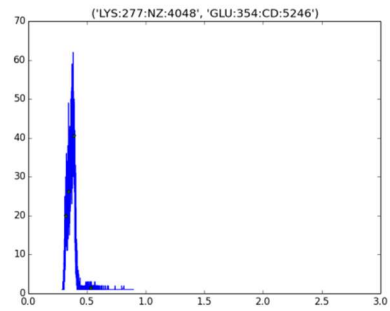
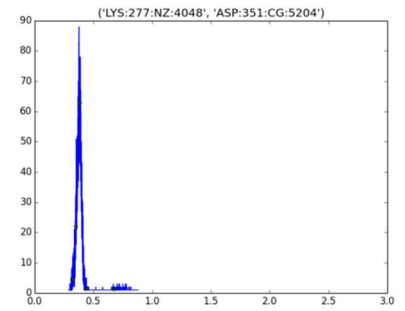
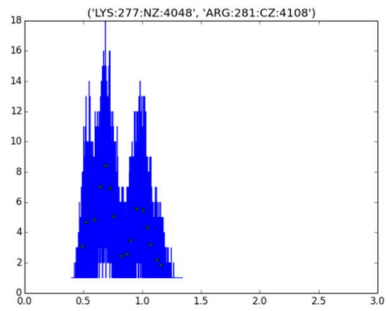
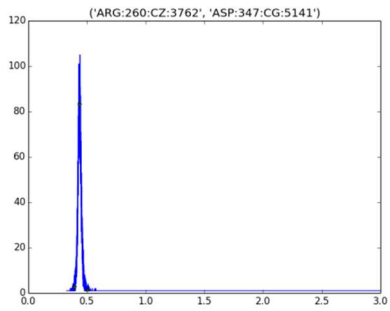


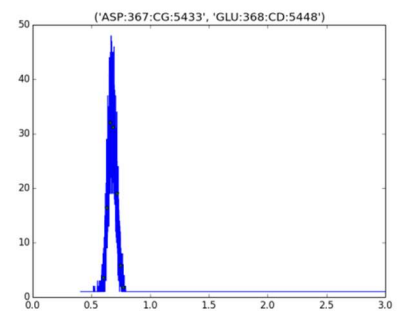
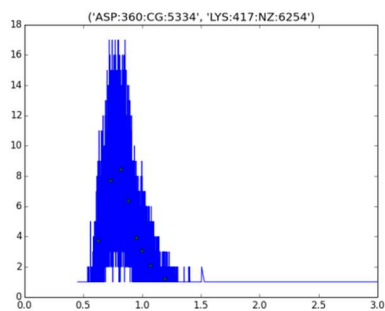
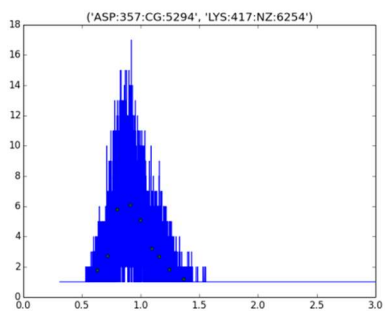
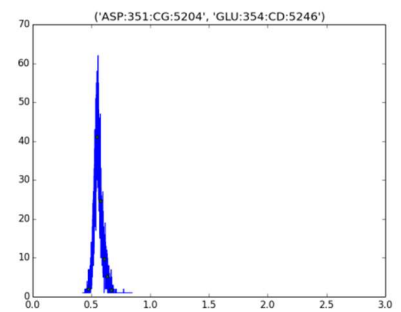
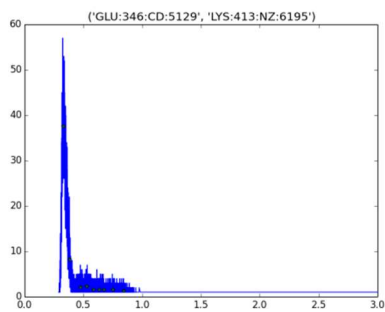
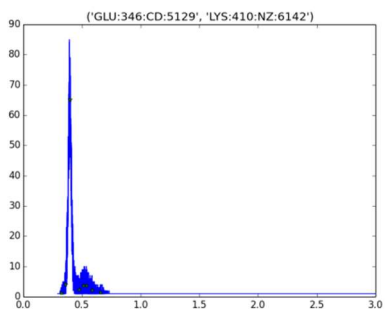
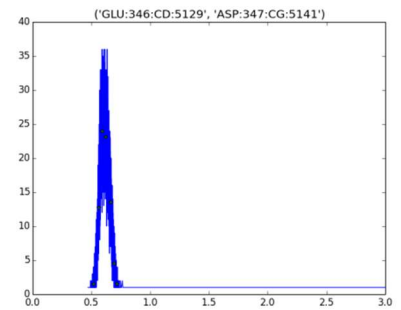
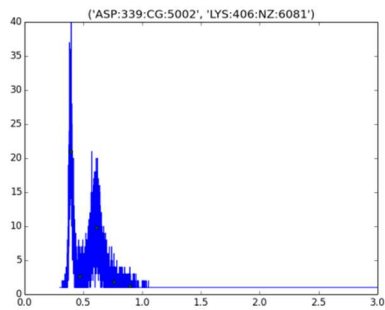
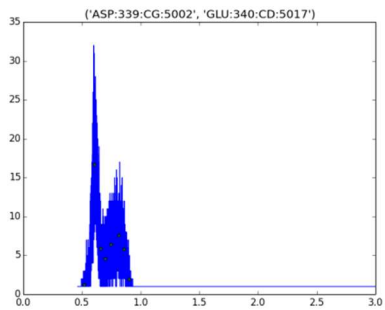
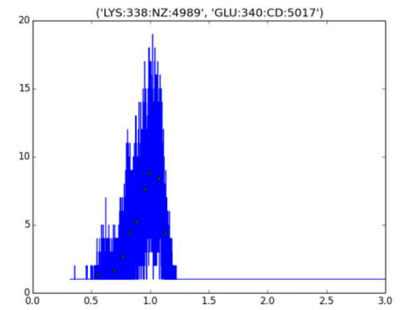
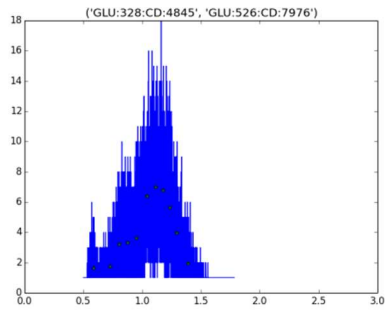
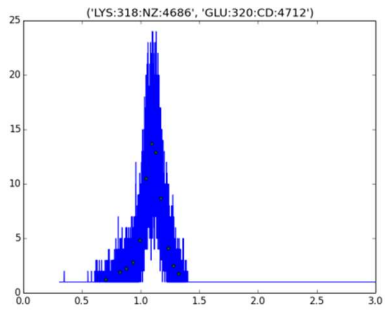
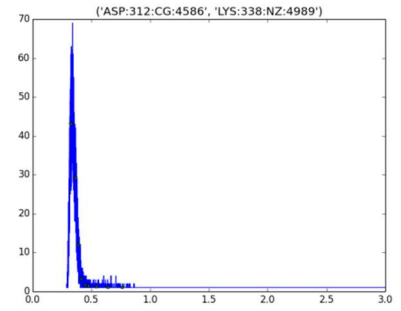
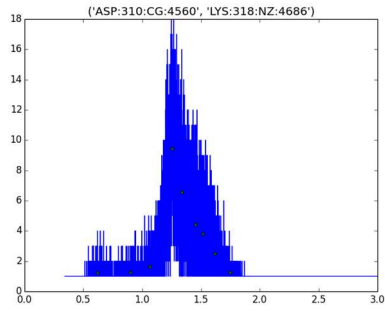
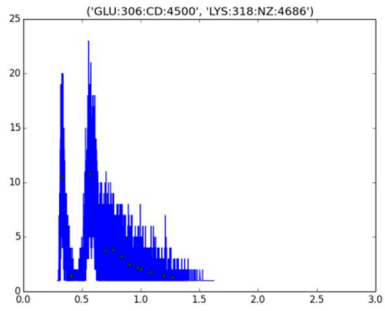
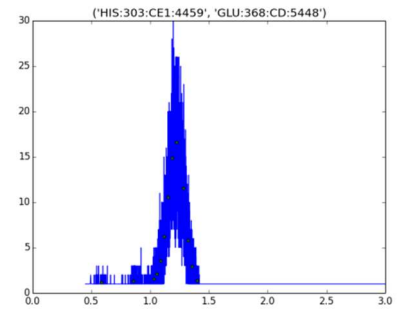
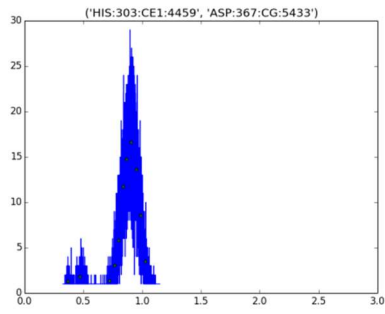
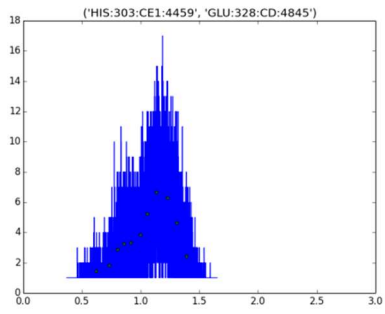




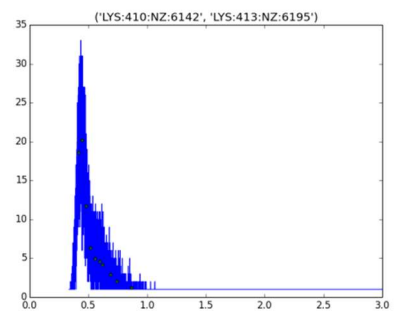
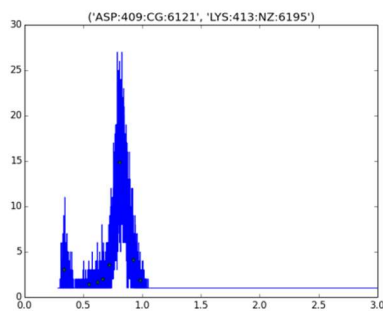
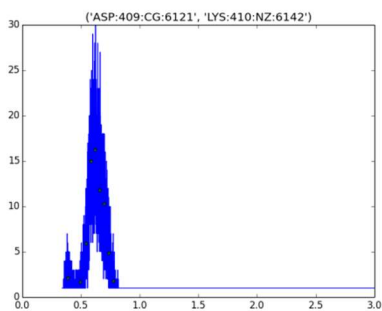
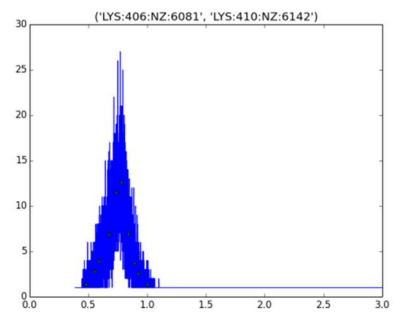
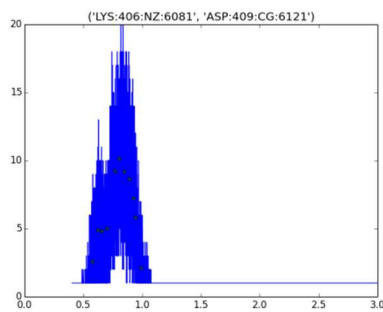
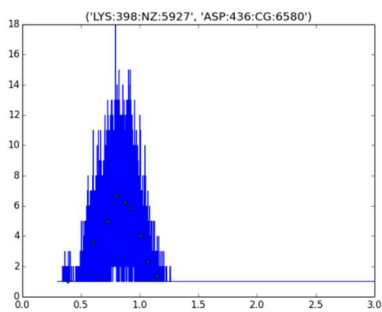
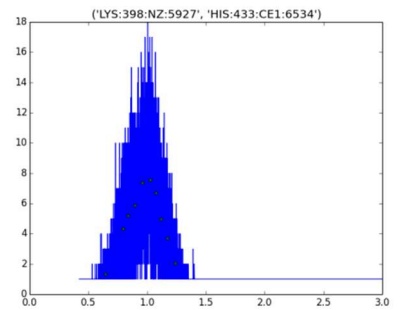
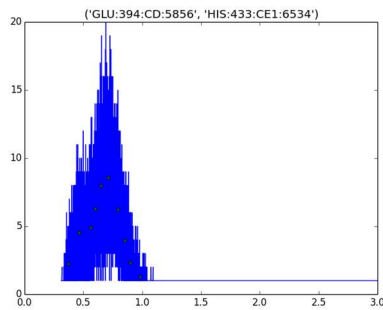
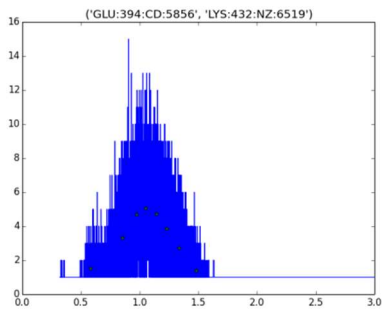
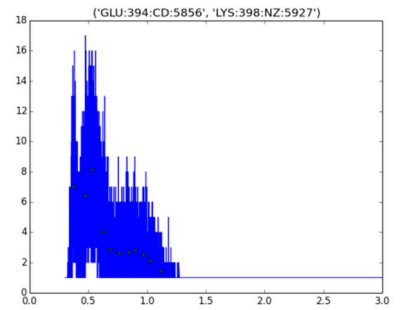
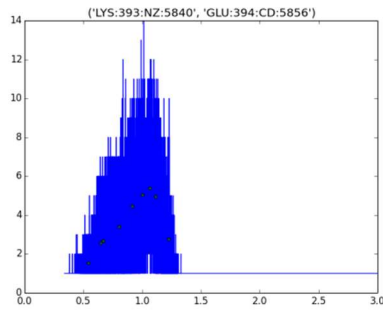
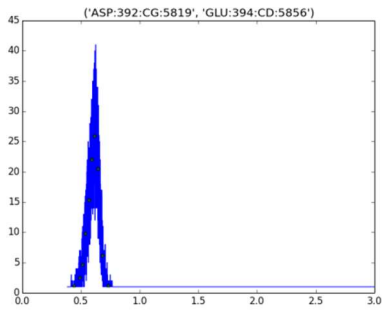
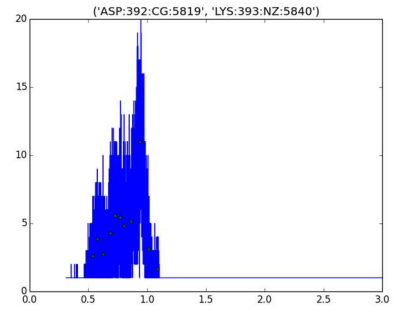
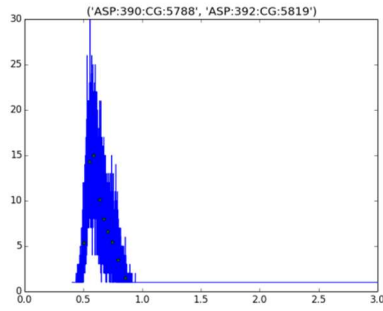
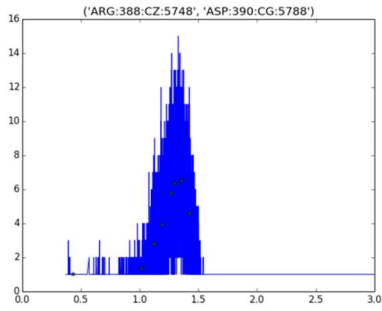
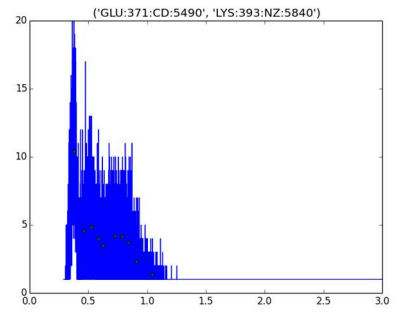
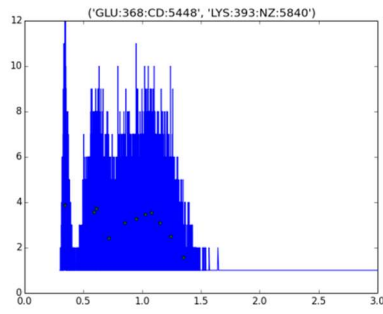
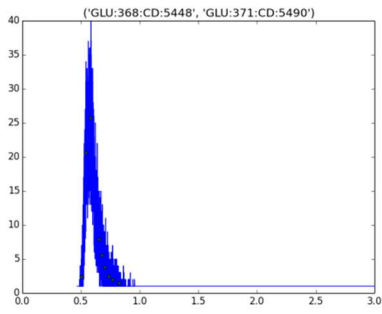




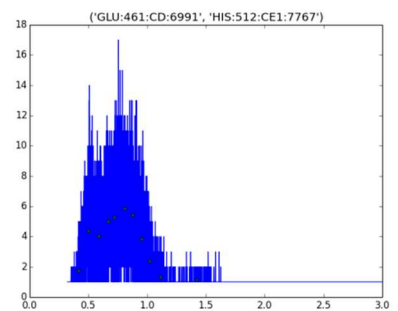
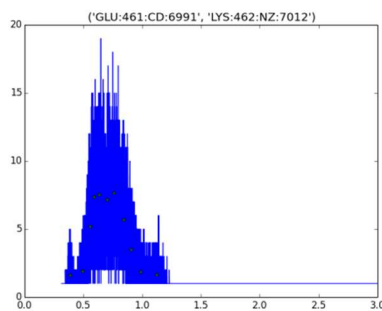
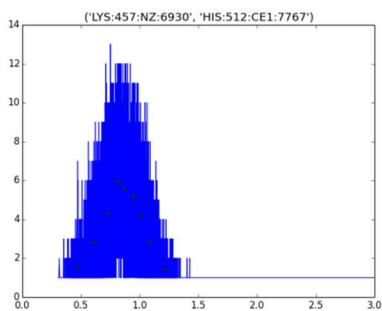
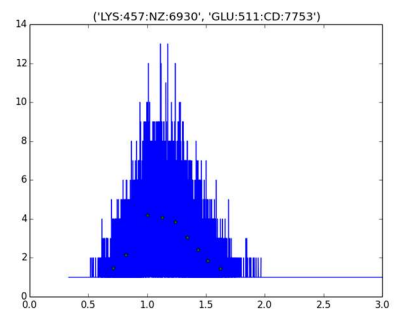
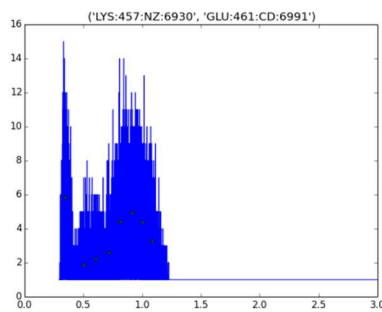
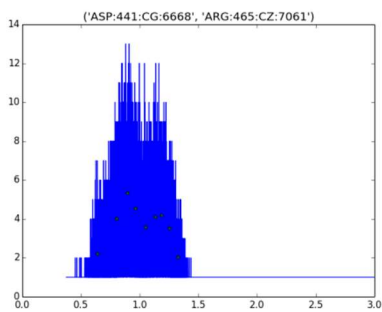
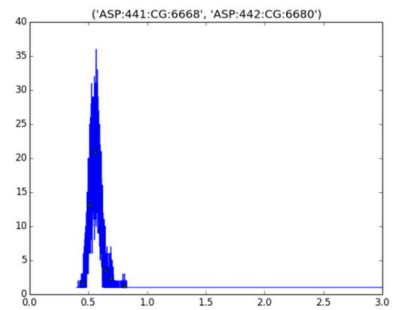
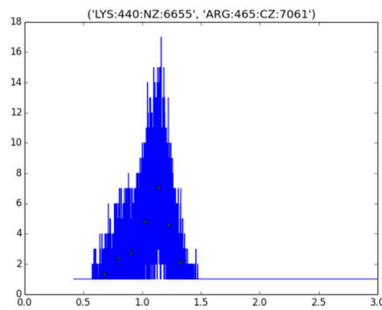
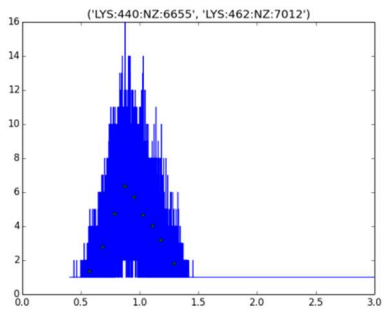
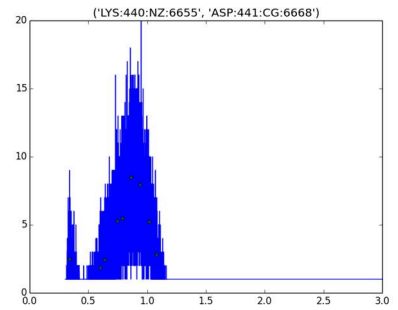
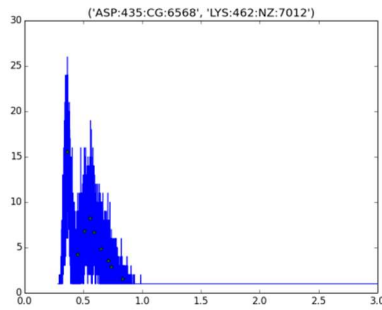
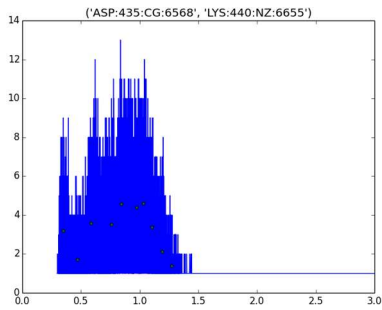
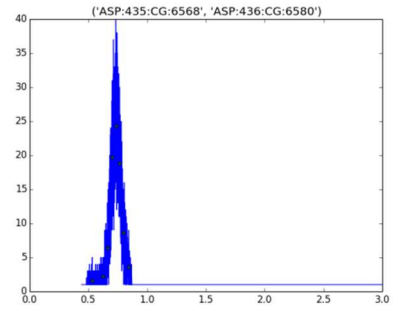
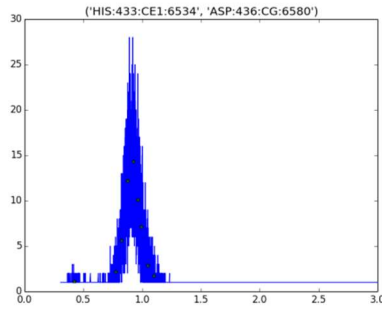
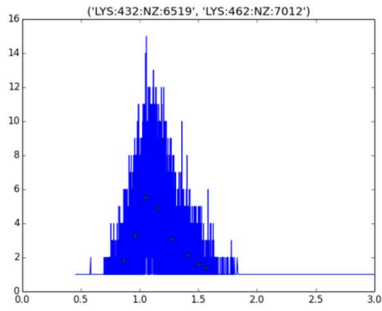
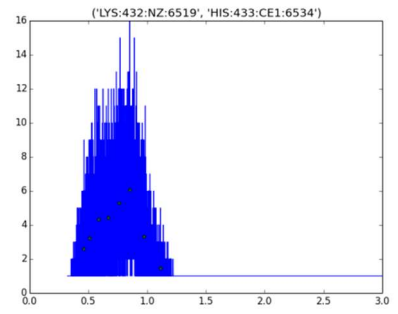
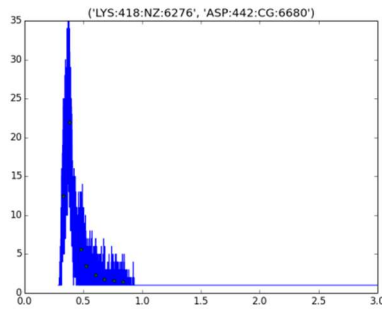
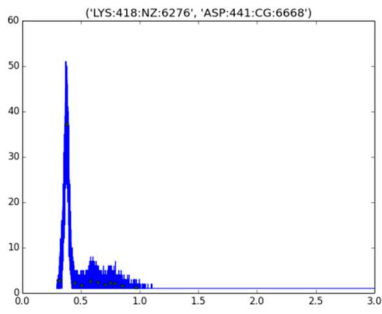


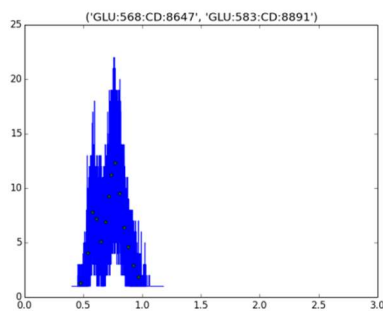
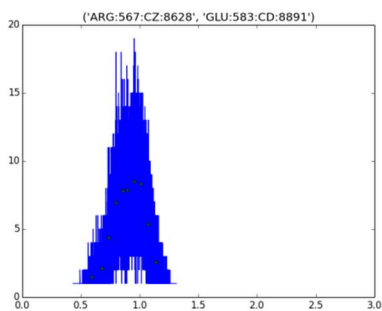
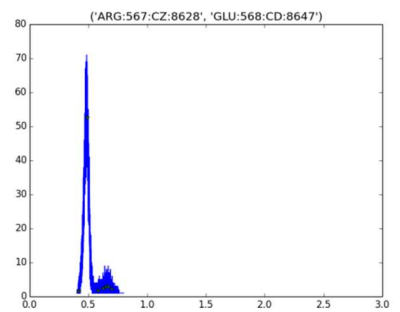
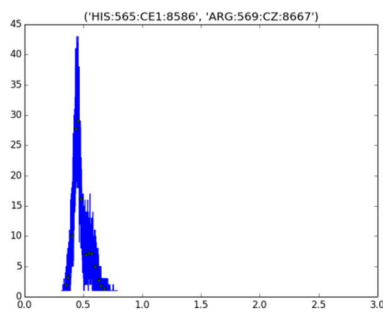
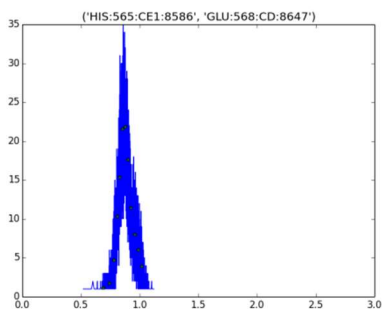
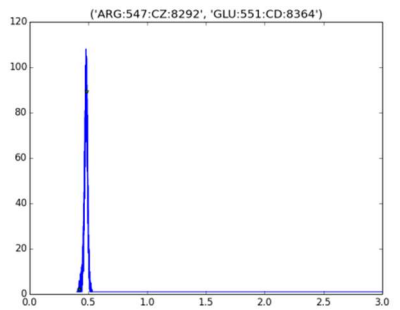
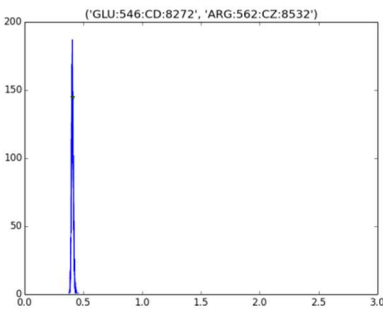
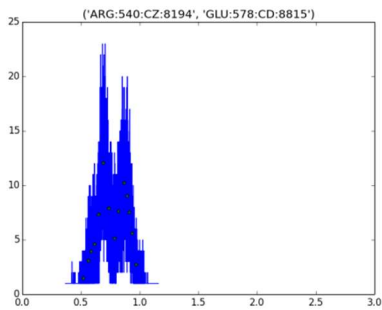
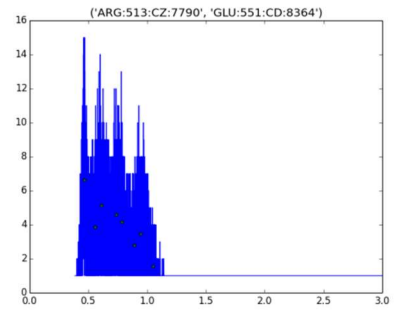
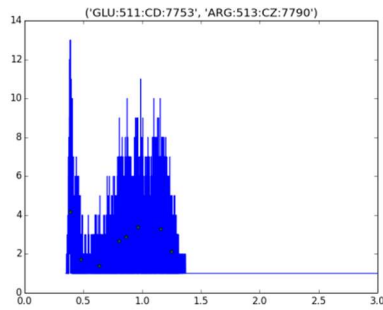
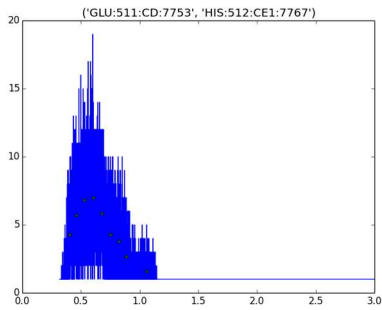
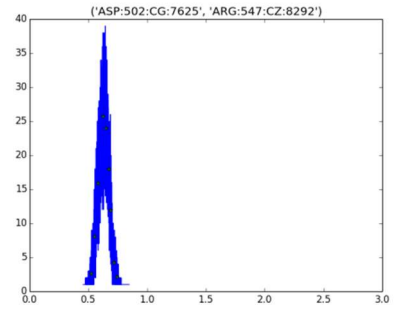
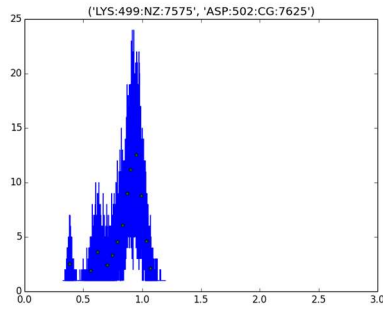
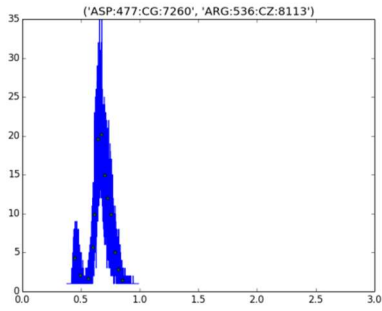
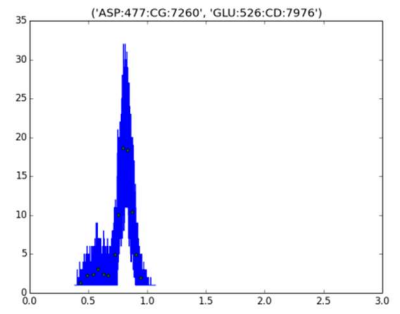
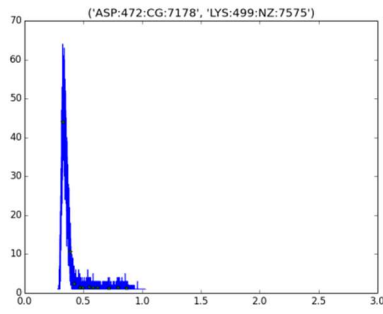
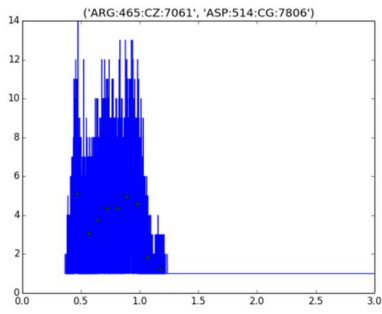


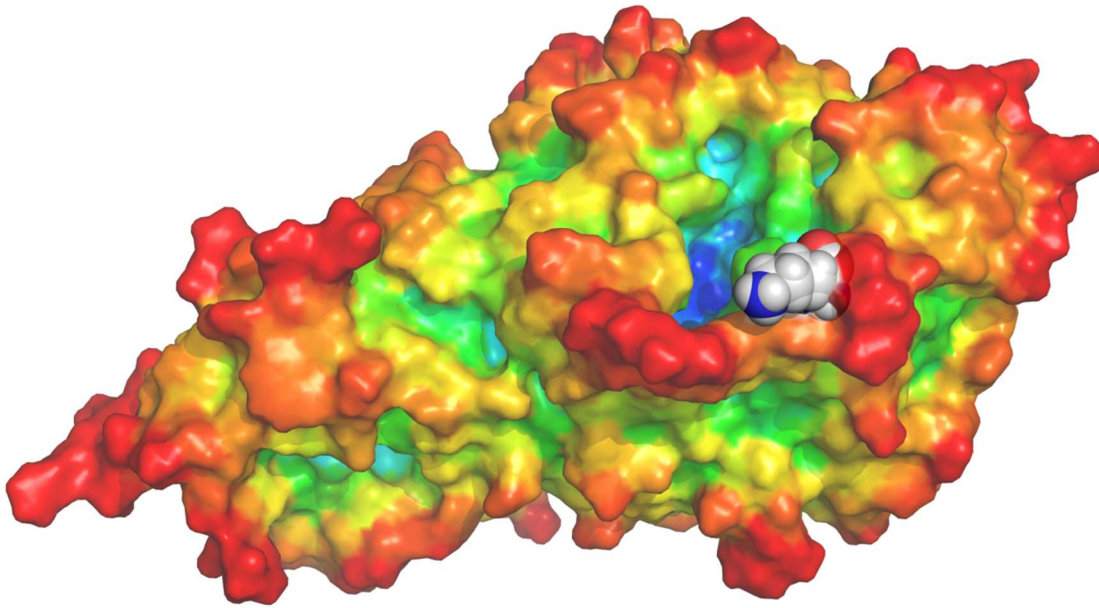












**Figure S6:** The structure of *Ostrinia furnacalis* chitinolytic enzyme: protein surface is colored according to 3D atom depth values, to highlight the catalytic site; K323 side chain is shown with space-fill representation.