wwPDB X-ray Structure Validation Summary Report (i)

Mar 15, 2022 – 04:20 PM EDT

PDB ID	:	4WYB
Title	:	Structure of the Bud6 flank domain in complex with actin
Authors	:	Eck, M.J.; Park, E.; Zheng, W.
Deposited on	:	2014-11-17
Resolution	:	3.49 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

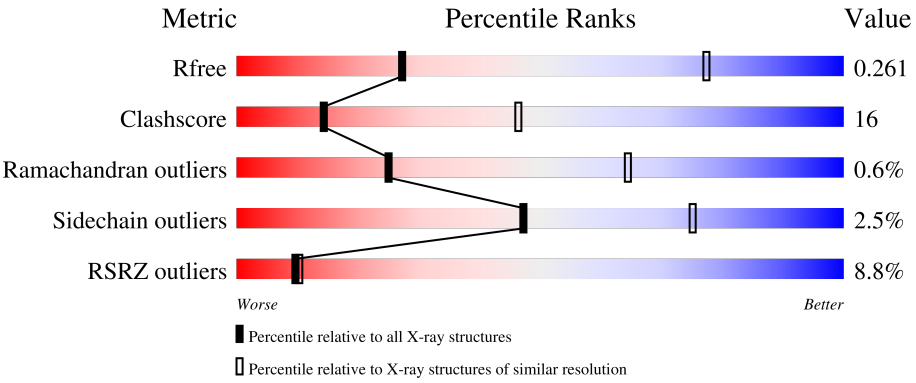
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.27
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.49 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



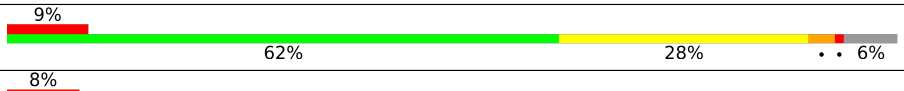
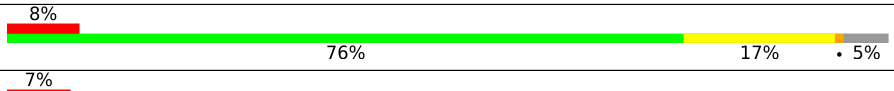
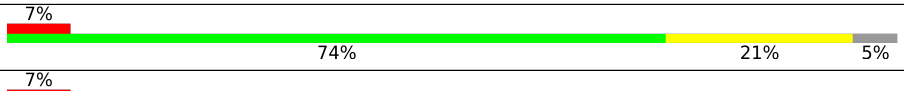
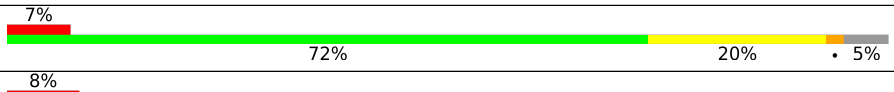
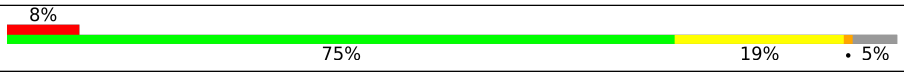
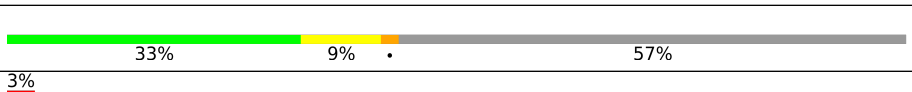
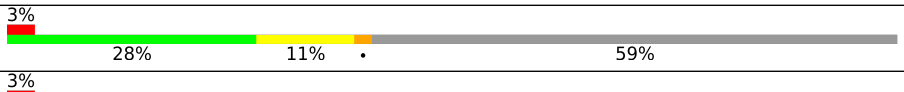
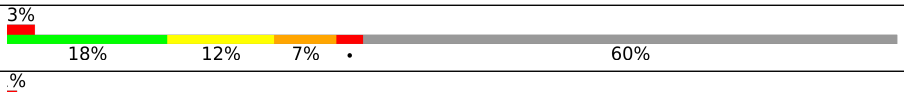
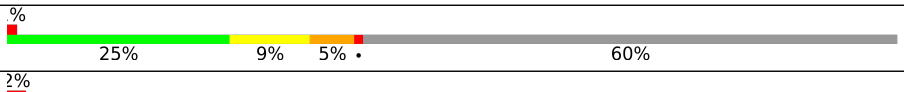
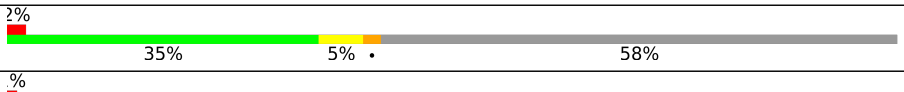
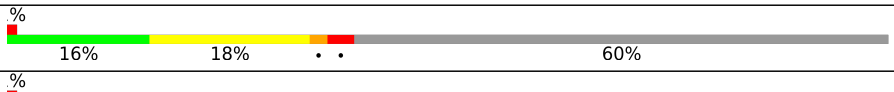
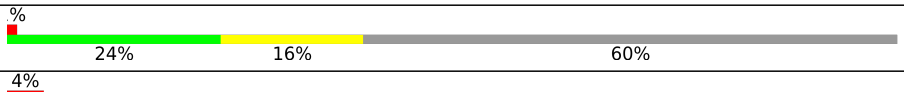
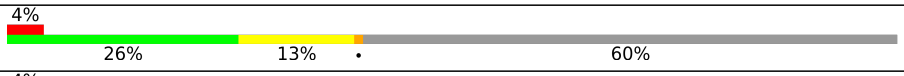
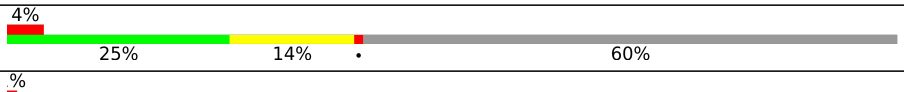
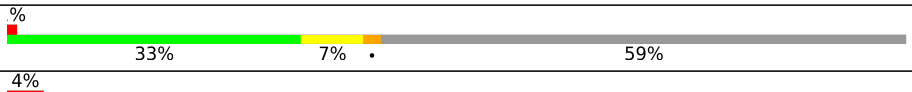
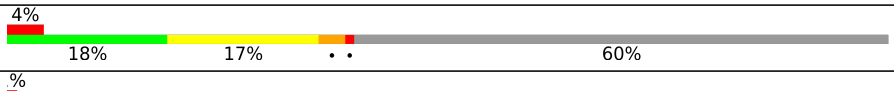
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	377	9% 75% 19% • 5%
1	C	377	9% 75% 20% • 5%
1	E	377	9% 70% 22% • 5%
1	G	377	12% 69% 23% • 6%
1	I	377	8% 72% 22% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	377	
1	M	377	
1	O	377	
1	Q	377	
1	S	377	
1	U	377	
1	X	377	
2	B	92	
2	D	92	
2	F	92	
2	H	92	
2	J	92	
2	L	92	
2	N	92	
2	P	92	
2	R	92	
2	T	92	
2	V	92	
2	Y	92	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 37515 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Actin, alpha skeletal muscle.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	357	Total	C	N	O	S	0	0	0
			2792	1770	468	536	18			
1	C	358	Total	C	N	O	S	0	0	0
			2803	1776	472	537	18			
1	E	358	Total	C	N	O	S	0	0	0
			2797	1773	469	537	18			
1	G	356	Total	C	N	O	S	0	0	0
			2765	1755	461	531	18			
1	I	358	Total	C	N	O	S	0	0	0
			2803	1776	472	537	18			
1	K	357	Total	C	N	O	S	0	0	0
			2782	1764	464	536	18			
1	M	358	Total	C	N	O	S	0	0	0
			2803	1776	472	537	18			
1	O	355	Total	C	N	O	S	0	0	0
			2768	1756	462	532	18			
1	Q	357	Total	C	N	O	S	0	0	0
			2787	1767	466	536	18			
1	S	358	Total	C	N	O	S	0	0	0
			2793	1771	469	535	18			
1	U	358	Total	C	N	O	S	0	0	0
			2787	1767	465	537	18			
1	X	358	Total	C	N	O	S	0	0	0
			2796	1772	471	535	18			

- Molecule 2 is a protein called Bud site selection protein 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	40	Total	C	N	O	S	0	0	0
			331	204	60	64	3			
2	D	38	Total	C	N	O	S	0	0	0
			314	192	57	62	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	37	Total	C	N	O	S	0	0	0
			303	186	53	61	3			
2	H	37	Total	C	N	O	S	0	0	0
			295	182	52	58	3			
2	J	39	Total	C	N	O	S	0	0	0
			309	189	54	63	3			
2	L	37	Total	C	N	O	S	0	0	0
			296	181	52	61	2			
2	N	37	Total	C	N	O	S	0	0	0
			303	186	53	61	3			
2	P	37	Total	C	N	O	S	0	0	0
			303	186	53	61	3			
2	R	37	Total	C	N	O	S	0	0	0
			291	179	51	58	3			
2	T	38	Total	C	N	O	S	0	0	0
			314	192	57	62	3			
2	V	37	Total	C	N	O	S	0	0	0
			293	181	51	59	2			
2	Y	37	Total	C	N	O	S	0	0	0
			303	186	53	61	3			

There are 24 discrepancies between the modelled and reference sequences:

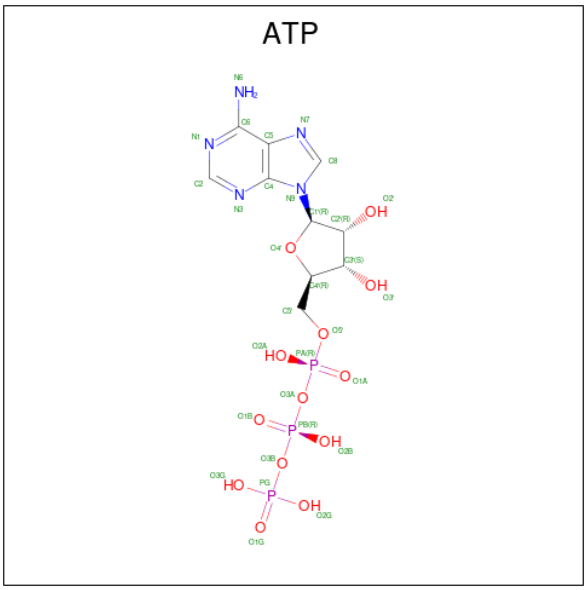
Chain	Residue	Modelled	Actual	Comment	Reference
B	697	GLY	-	expression tag	UNP P41697
B	698	SER	-	expression tag	UNP P41697
D	697	GLY	-	expression tag	UNP P41697
D	698	SER	-	expression tag	UNP P41697
F	697	GLY	-	expression tag	UNP P41697
F	698	SER	-	expression tag	UNP P41697
H	697	GLY	-	expression tag	UNP P41697
H	698	SER	-	expression tag	UNP P41697
J	697	GLY	-	expression tag	UNP P41697
J	698	SER	-	expression tag	UNP P41697
L	697	GLY	-	expression tag	UNP P41697
L	698	SER	-	expression tag	UNP P41697
N	697	GLY	-	expression tag	UNP P41697
N	698	SER	-	expression tag	UNP P41697
P	697	GLY	-	expression tag	UNP P41697
P	698	SER	-	expression tag	UNP P41697
R	697	GLY	-	expression tag	UNP P41697
R	698	SER	-	expression tag	UNP P41697
T	697	GLY	-	expression tag	UNP P41697

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
T	698	SER	-	expression tag	UNP P41697
V	697	GLY	-	expression tag	UNP P41697
V	698	SER	-	expression tag	UNP P41697
Y	697	GLY	-	expression tag	UNP P41697
Y	698	SER	-	expression tag	UNP P41697

- Molecule 3 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	E	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	G	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	I	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	K	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	M	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	O	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	Q	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	S	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	U	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
3	X	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

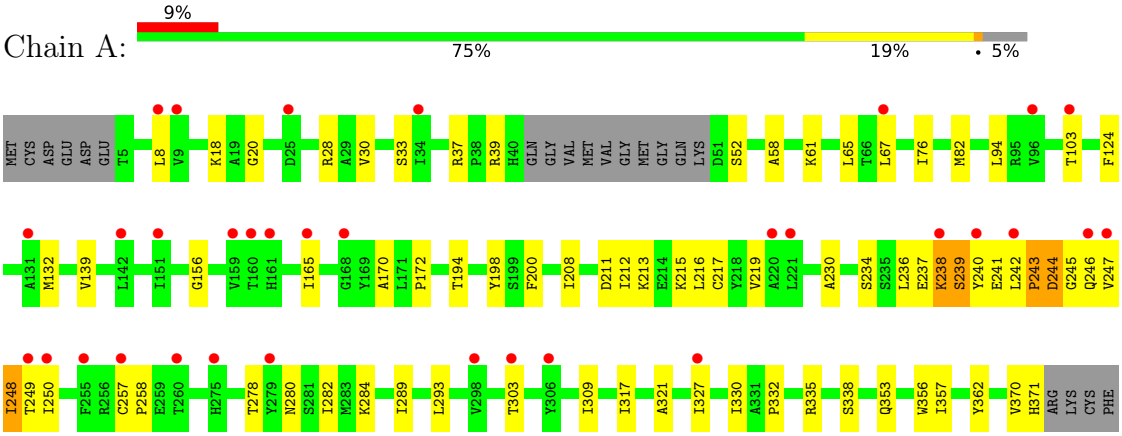
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		
4	C	1	Total	Ca	0	0
			1	1		
4	E	1	Total	Ca	0	0
			1	1		
4	G	1	Total	Ca	0	0
			1	1		
4	I	1	Total	Ca	0	0
			1	1		
4	K	1	Total	Ca	0	0
			1	1		
4	M	1	Total	Ca	0	0
			1	1		
4	O	1	Total	Ca	0	0
			1	1		
4	Q	1	Total	Ca	0	0
			1	1		
4	S	1	Total	Ca	0	0
			1	1		
4	U	1	Total	Ca	0	0
			1	1		
4	X	1	Total	Ca	0	0
			1	1		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

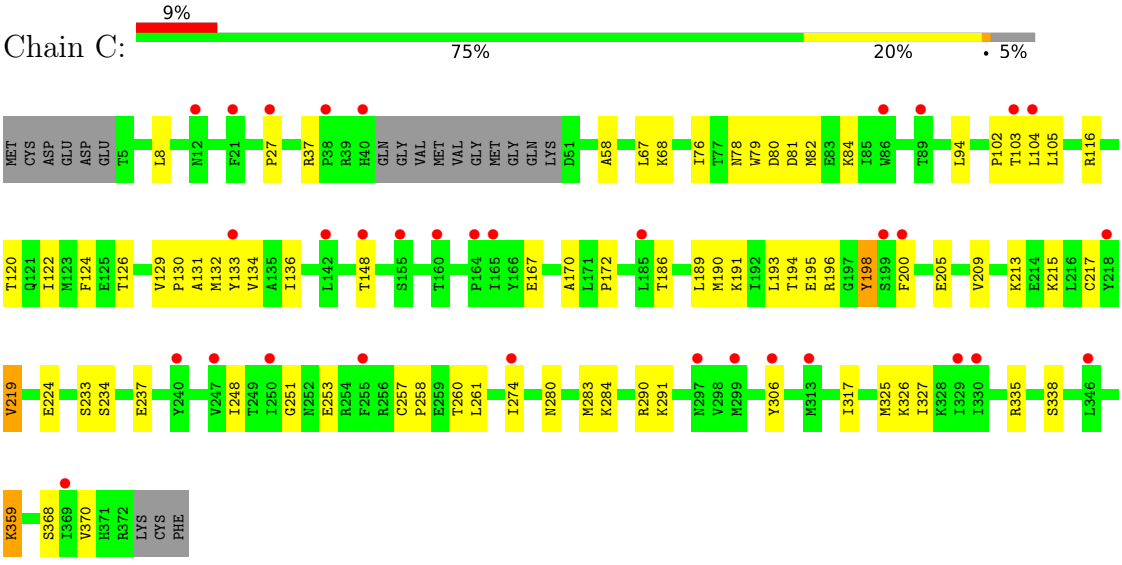
• Molecule 1: Actin, alpha skeletal muscle

Chain A:



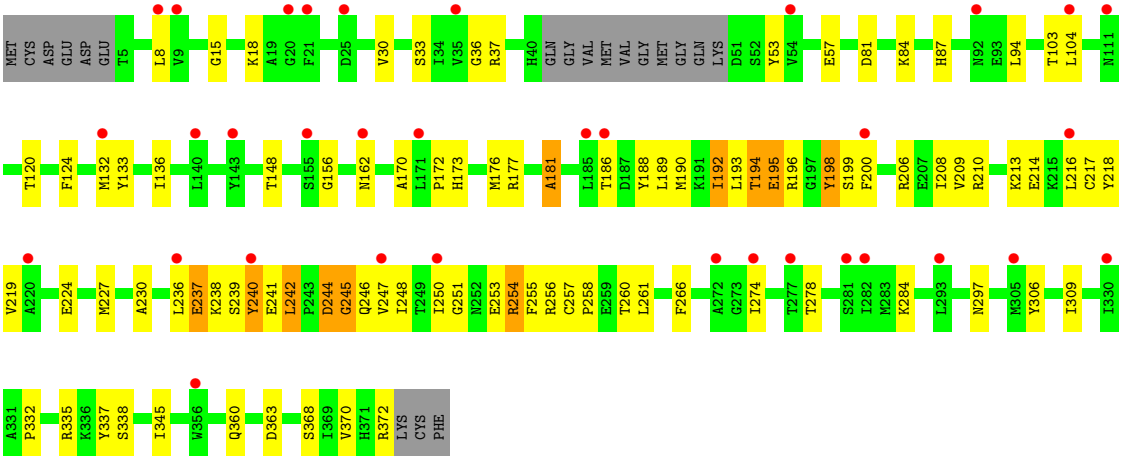
• Molecule 1: Actin, alpha skeletal muscle

Chain C:



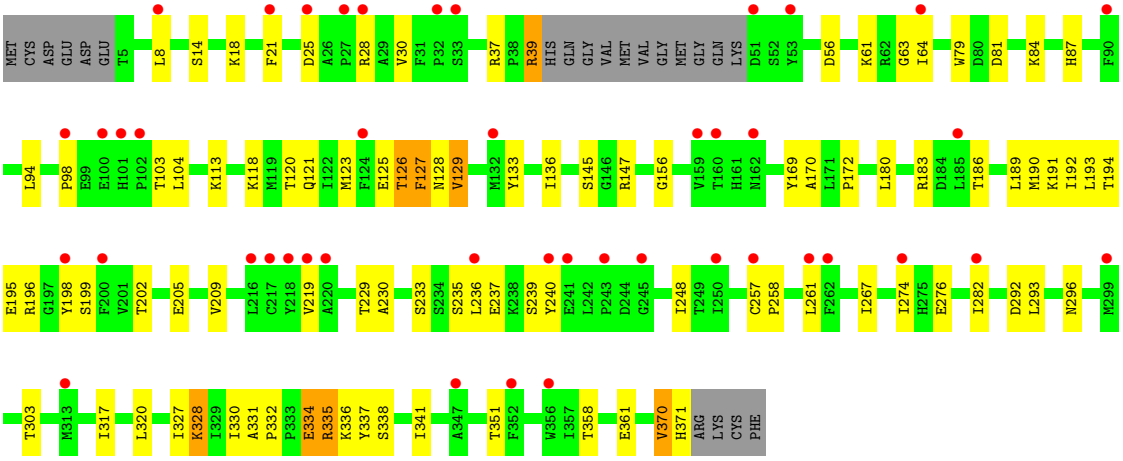
• Molecule 1: Actin, alpha skeletal muscle

Chain E:



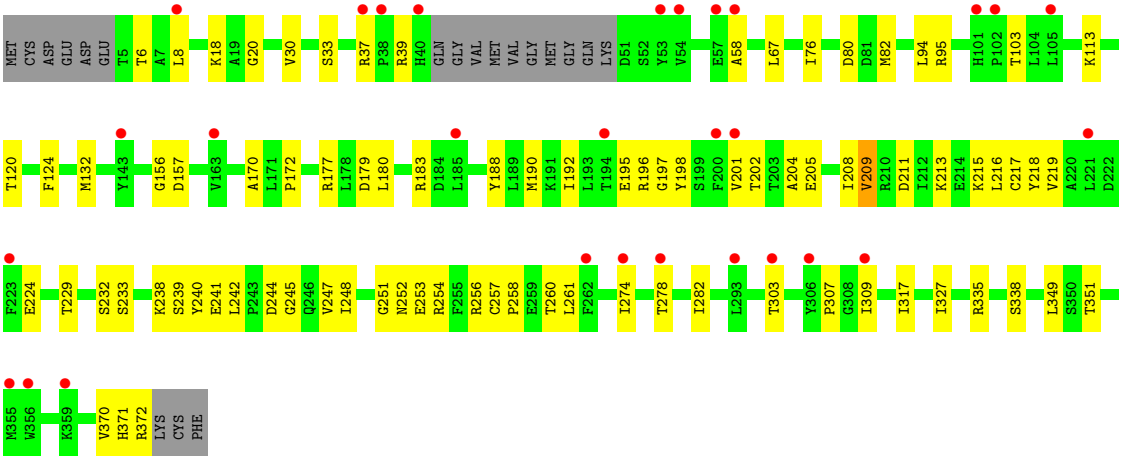
• Molecule 1: Actin, alpha skeletal muscle

Chain G:



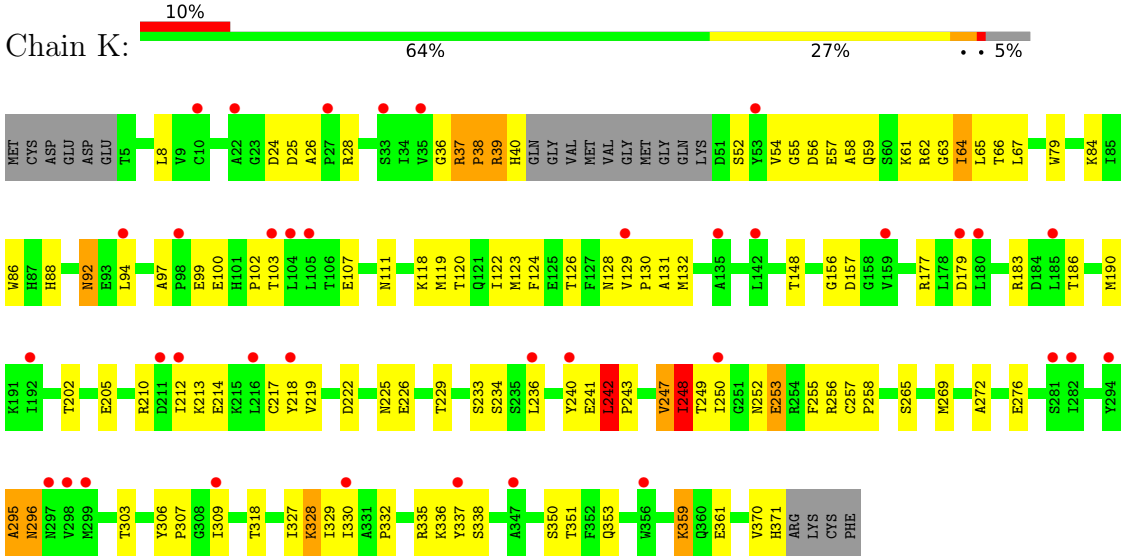
• Molecule 1: Actin, alpha skeletal muscle

Chain I:



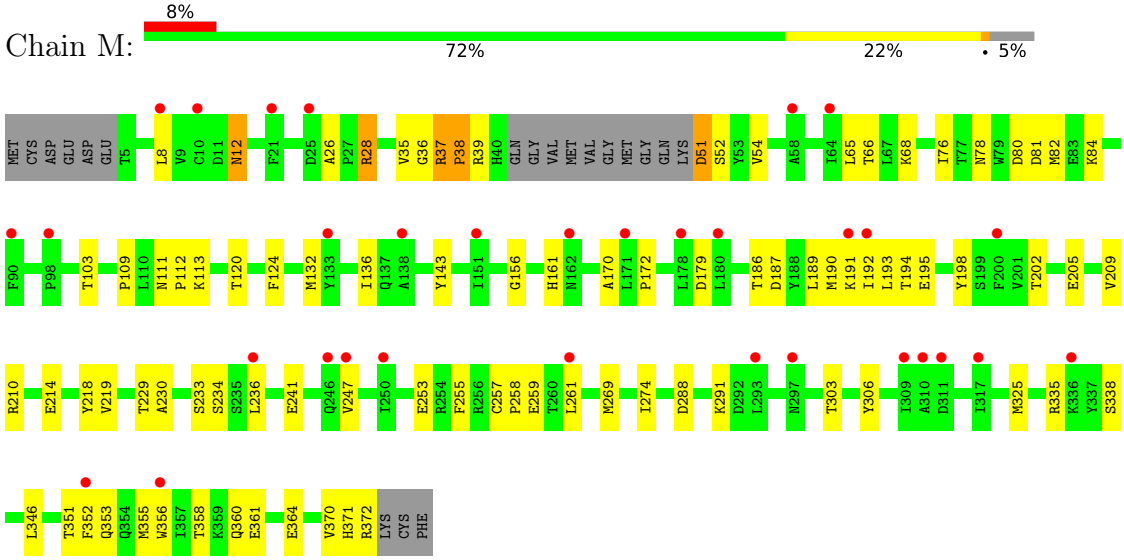
• Molecule 1: Actin, alpha skeletal muscle

Chain K:



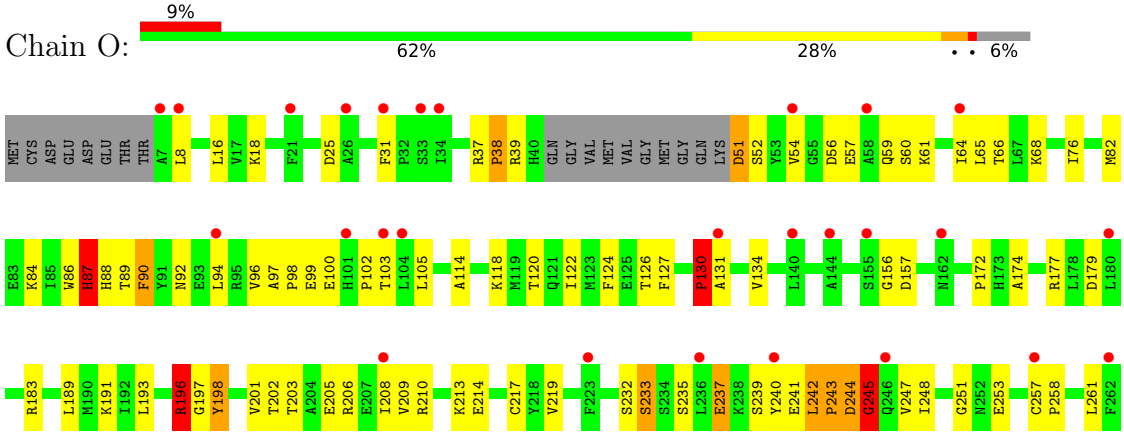
• Molecule 1: Actin, alpha skeletal muscle

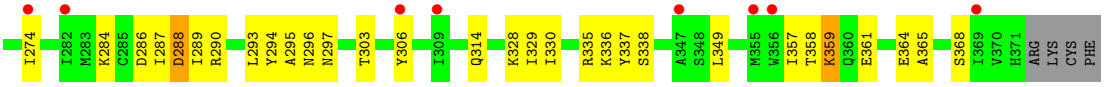
Chain M:



• Molecule 1: Actin, alpha skeletal muscle

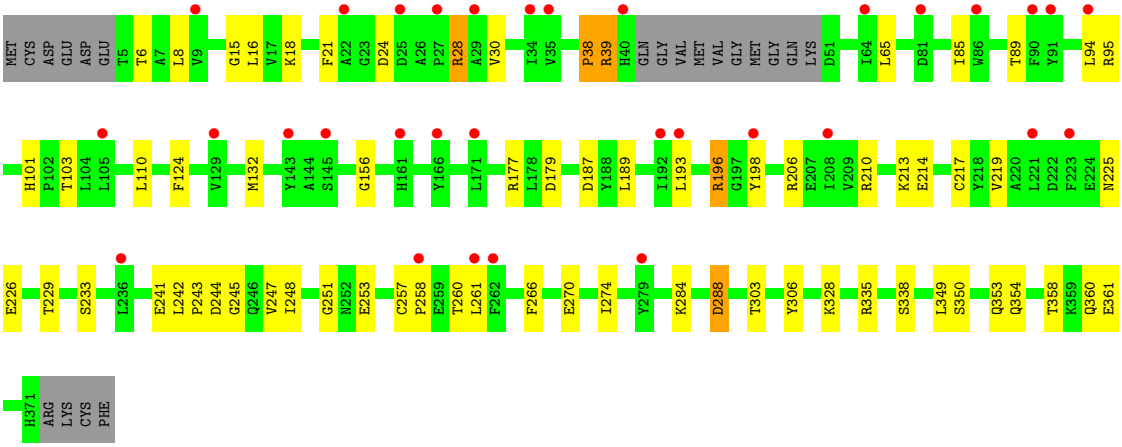
Chain O:





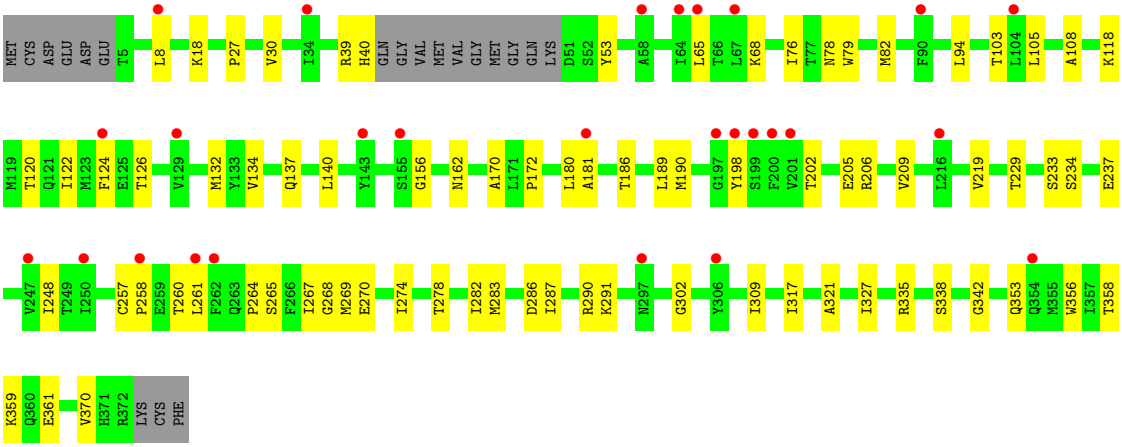
• Molecule 1: Actin, alpha skeletal muscle

Chain Q:



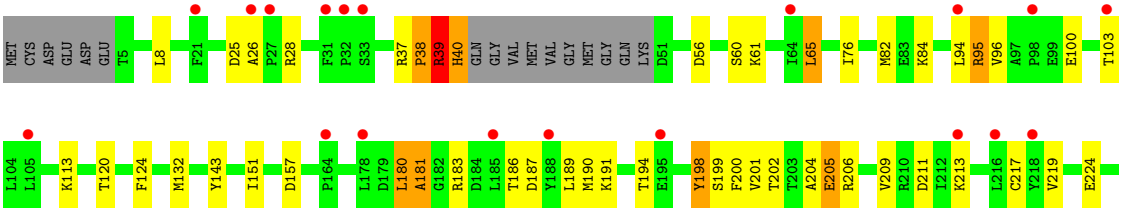
• Molecule 1: Actin, alpha skeletal muscle

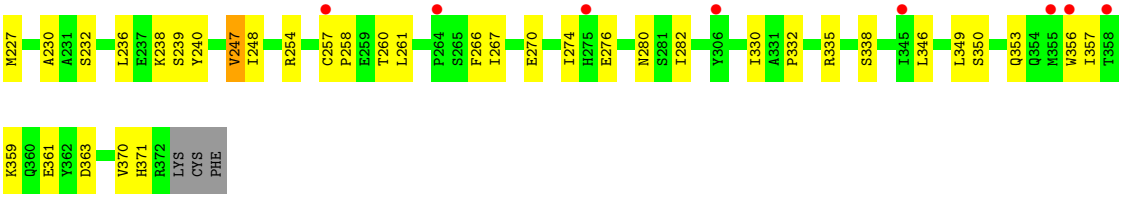
Chain S:



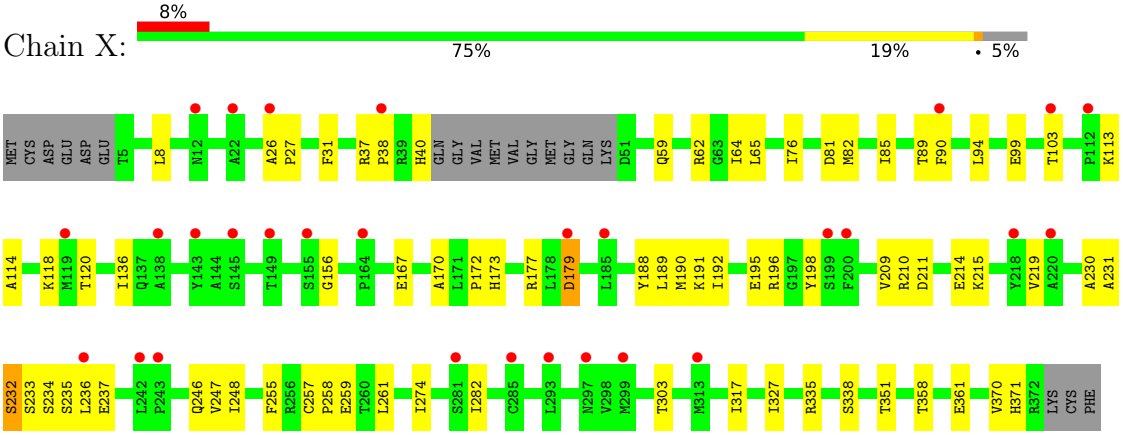
• Molecule 1: Actin, alpha skeletal muscle

Chain U:

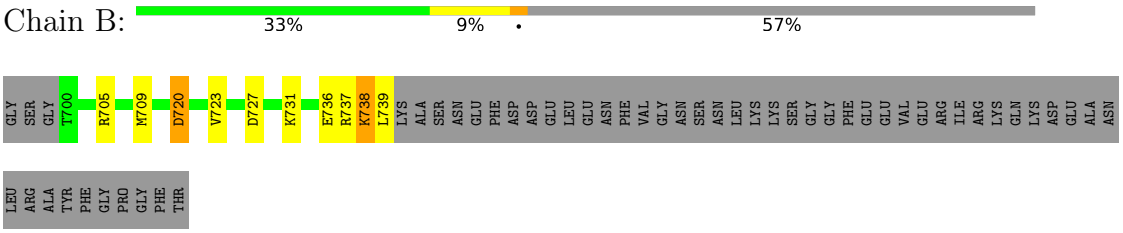




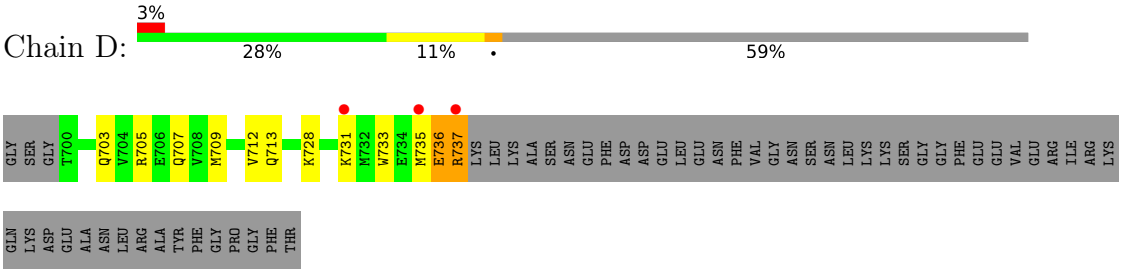
• Molecule 1: Actin, alpha skeletal muscle



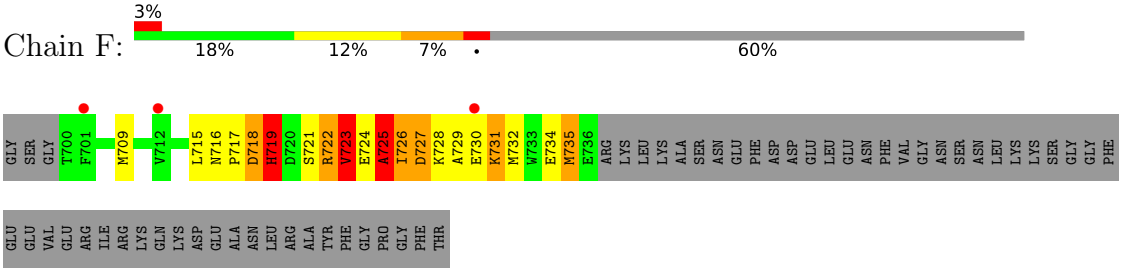
• Molecule 2: Bud site selection protein 6



• Molecule 2: Bud site selection protein 6

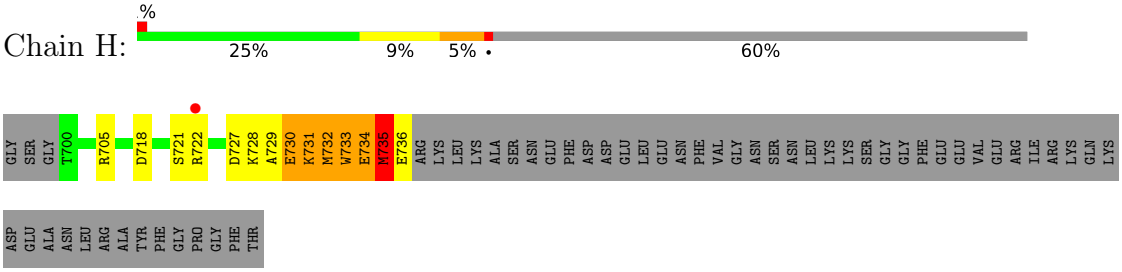


• Molecule 2: Bud site selection protein 6



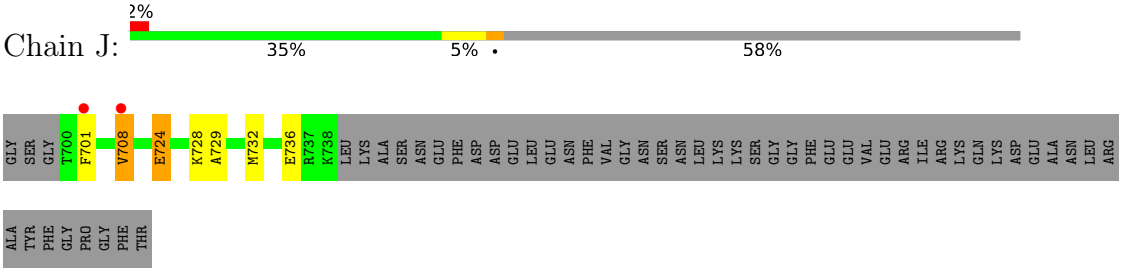
• Molecule 2: Bud site selection protein 6

Chain H:



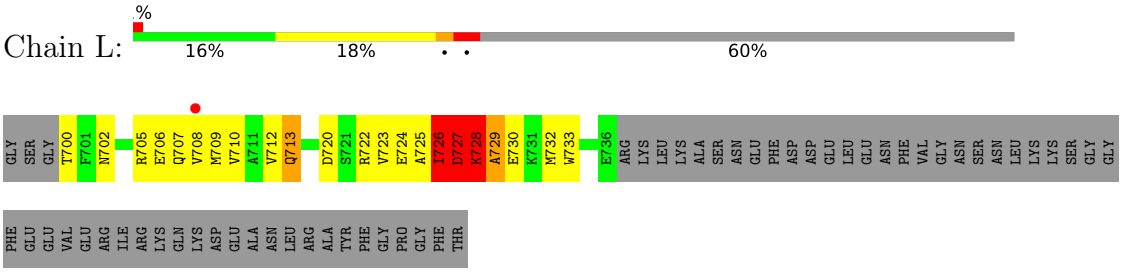
• Molecule 2: Bud site selection protein 6

Chain J:



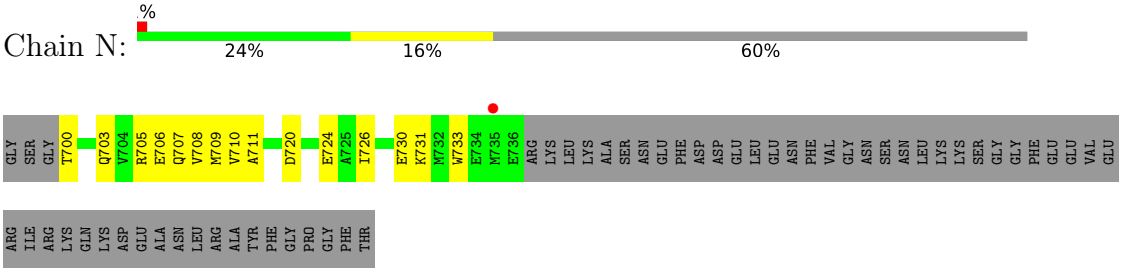
• Molecule 2: Bud site selection protein 6

Chain L:



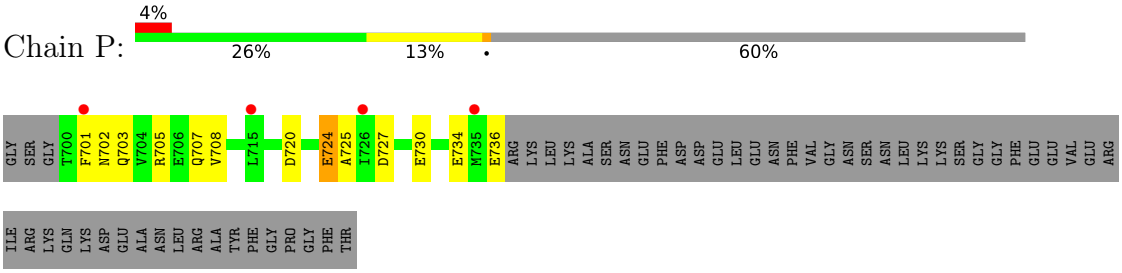
• Molecule 2: Bud site selection protein 6

Chain N:



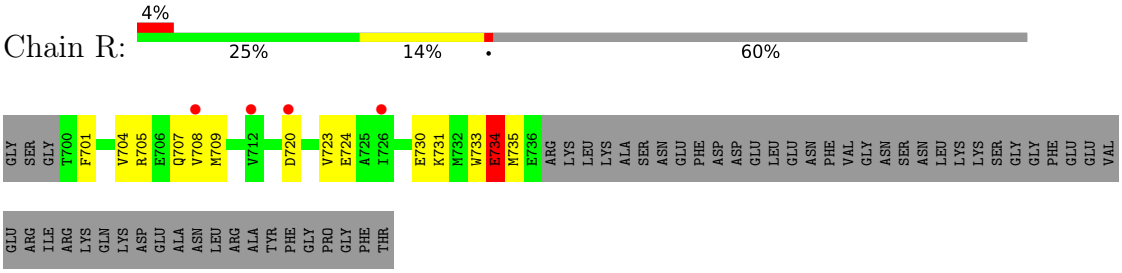
• Molecule 2: Bud site selection protein 6

Chain P:



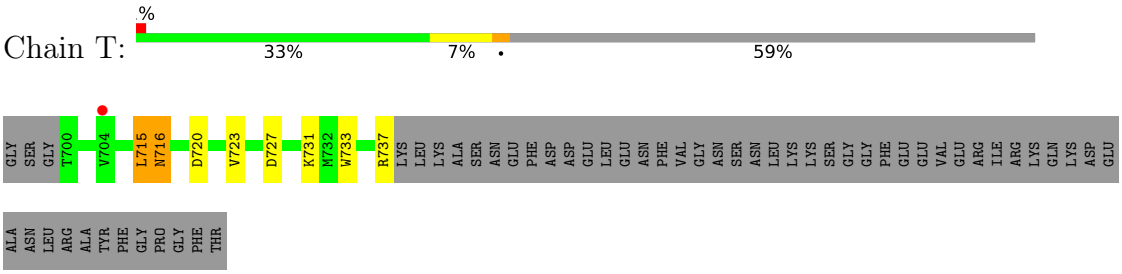
• Molecule 2: Bud site selection protein 6

Chain R:



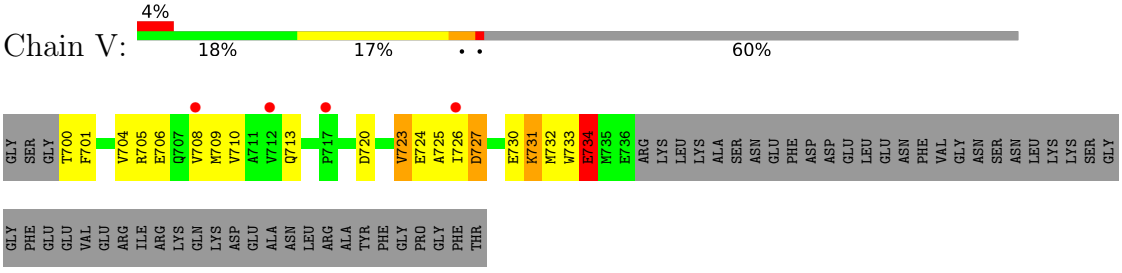
• Molecule 2: Bud site selection protein 6

Chain T:



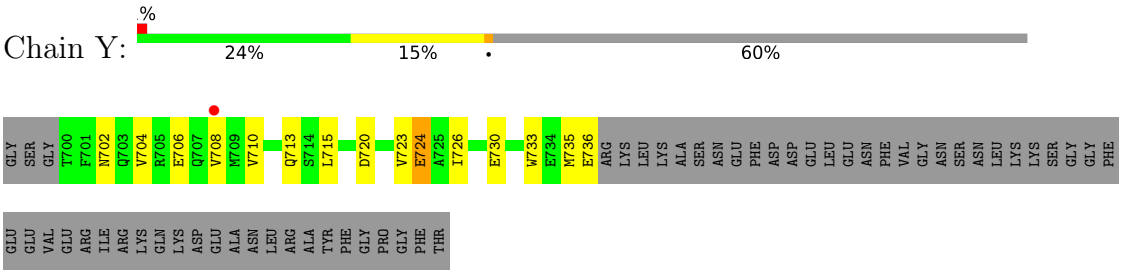
• Molecule 2: Bud site selection protein 6

Chain V:



• Molecule 2: Bud site selection protein 6

Chain Y:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	138.75Å 138.75Å 356.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.42 – 3.49 49.83 – 3.49	Depositor EDS
% Data completeness (in resolution range)	98.5 (45.42-3.49) 93.1 (49.83-3.49)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.212 , 0.258 0.218 , 0.261	Depositor DCC
R_{free} test set	2020 reflections (2.10%)	wwPDB-VP
Wilson B-factor (Å^2)	86.2	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 61.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l 0.429 for h,-h-k,-l 0.028 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	37515	wwPDB-VP
Average B, all atoms (Å^2)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.73% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CA, ATP

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.27	0/2853	0.56	5/3870 (0.1%)
1	C	0.25	0/2864	0.45	0/3884
1	E	0.30	0/2858	0.61	6/3877 (0.2%)
1	G	0.27	0/2825	0.54	5/3835 (0.1%)
1	I	0.25	0/2864	0.46	0/3884
1	K	0.33	0/2843	0.62	7/3859 (0.2%)
1	M	0.27	0/2864	0.47	1/3884 (0.0%)
1	O	0.37	1/2829 (0.0%)	0.64	6/3839 (0.2%)
1	Q	0.25	0/2847	0.47	1/3862 (0.0%)
1	S	0.26	0/2854	0.52	3/3872 (0.1%)
1	U	0.26	0/2848	0.53	4/3866 (0.1%)
1	X	0.25	0/2857	0.54	4/3876 (0.1%)
2	B	0.36	0/335	0.64	0/449
2	D	0.43	0/318	0.61	0/427
2	F	0.49	0/307	1.15	5/413 (1.2%)
2	H	0.36	0/299	0.58	1/403 (0.2%)
2	J	0.27	0/313	0.48	0/423
2	L	0.51	0/300	1.23	4/406 (1.0%)
2	N	0.30	0/307	0.53	0/413
2	P	0.38	0/307	0.88	1/413 (0.2%)
2	R	0.35	0/295	0.66	0/399
2	T	0.36	0/318	0.67	0/427
2	V	0.51	0/297	1.00	3/401 (0.7%)
2	Y	0.28	0/307	0.54	0/413
All	All	0.29	1/37909 (0.0%)	0.57	56/51395 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	O	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	O	359	LYS	CD-CE	-5.10	1.38	1.51

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	O	196	ARG	CB-CA-C	-13.46	83.48	110.40
1	E	240	TYR	CB-CA-C	-11.91	86.59	110.40
2	L	726	ILE	N-CA-C	-11.31	80.45	111.00
1	X	231	ALA	CB-CA-C	-10.64	94.14	110.10
1	A	243	PRO	N-CA-C	-10.58	84.59	112.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	O	245	GLY	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2792	0	2753	93	0
1	C	2803	0	2769	62	0
1	E	2797	0	2758	114	0
1	G	2765	0	2721	96	0
1	I	2803	0	2769	63	0
1	K	2782	0	2732	123	1
1	M	2803	0	2769	64	1
1	O	2768	0	2719	155	0
1	Q	2787	0	2751	57	0
1	S	2793	0	2754	50	0
1	U	2787	0	2736	63	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	X	2796	0	2756	54	0
2	B	331	0	325	7	0
2	D	314	0	301	9	0
2	F	303	0	287	47	0
2	H	295	0	278	21	0
2	J	309	0	281	10	0
2	L	296	0	270	37	0
2	N	303	0	288	8	0
2	P	303	0	288	10	0
2	R	291	0	267	17	0
2	T	314	0	301	12	0
2	V	293	0	271	16	0
2	Y	303	0	288	12	0
3	A	31	0	12	0	0
3	C	31	0	12	1	0
3	E	31	0	12	2	0
3	G	31	0	12	0	0
3	I	31	0	12	0	0
3	K	31	0	12	1	0
3	M	31	0	12	1	0
3	O	31	0	12	3	0
3	Q	31	0	12	1	0
3	S	31	0	12	1	0
3	U	31	0	12	0	0
3	X	31	0	12	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
4	G	1	0	0	0	0
4	I	1	0	0	0	0
4	K	1	0	0	0	0
4	M	1	0	0	0	0
4	O	1	0	0	0	0
4	Q	1	0	0	0	0
4	S	1	0	0	0	0
4	U	1	0	0	0	0
4	X	1	0	0	0	0
All	All	37515	0	36576	1149	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 16.

The worst 5 of 1149 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:240:TYR:CB	1:K:248:ILE:HG21	1.37	1.53
1:E:240:TYR:O	1:E:248:ILE:HG22	1.20	1.34
1:E:198:TYR:CE2	1:E:248:ILE:HG13	1.67	1.29
1:K:212:ILE:CG1	1:K:240:TYR:CE2	2.16	1.28
1:A:239:SER:CB	1:A:248:ILE:O	1.79	1.27

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:39:ARG:NH1	1:M:187:ASP:OD2[3_755]	1.82	0.38

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	353/377 (94%)	339 (96%)	13 (4%)	1 (0%)	41 75
1	C	354/377 (94%)	334 (94%)	19 (5%)	1 (0%)	41 75
1	E	354/377 (94%)	325 (92%)	29 (8%)	0	100 100
1	G	352/377 (93%)	326 (93%)	24 (7%)	2 (1%)	25 64
1	I	354/377 (94%)	333 (94%)	20 (6%)	1 (0%)	41 75
1	K	353/377 (94%)	329 (93%)	20 (6%)	4 (1%)	14 52
1	M	354/377 (94%)	337 (95%)	15 (4%)	2 (1%)	25 64
1	O	351/377 (93%)	322 (92%)	24 (7%)	5 (1%)	11 46
1	Q	353/377 (94%)	340 (96%)	12 (3%)	1 (0%)	41 75
1	S	354/377 (94%)	334 (94%)	20 (6%)	0	100 100
1	U	354/377 (94%)	338 (96%)	14 (4%)	2 (1%)	25 64
1	X	354/377 (94%)	335 (95%)	17 (5%)	2 (1%)	25 64
2	B	38/92 (41%)	35 (92%)	2 (5%)	1 (3%)	5 33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	D	36/92 (39%)	33 (92%)	3 (8%)	0	100 100
2	F	35/92 (38%)	30 (86%)	3 (9%)	2 (6%)	1 16
2	H	35/92 (38%)	29 (83%)	3 (9%)	3 (9%)	1 9
2	J	37/92 (40%)	36 (97%)	1 (3%)	0	100 100
2	L	35/92 (38%)	25 (71%)	9 (26%)	1 (3%)	4 31
2	N	35/92 (38%)	31 (89%)	4 (11%)	0	100 100
2	P	35/92 (38%)	32 (91%)	3 (9%)	0	100 100
2	R	35/92 (38%)	32 (91%)	2 (6%)	1 (3%)	4 31
2	T	36/92 (39%)	34 (94%)	2 (6%)	0	100 100
2	V	35/92 (38%)	31 (89%)	4 (11%)	0	100 100
2	Y	35/92 (38%)	34 (97%)	1 (3%)	0	100 100
All	All	4667/5628 (83%)	4374 (94%)	264 (6%)	29 (1%)	25 64

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	244	ASP
2	B	738	LYS
2	F	723	VAL
2	F	725	ALA
1	G	128	ASN

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	303/320 (95%)	300 (99%)	3 (1%)	76 88
1	C	304/320 (95%)	299 (98%)	5 (2%)	62 83
1	E	303/320 (95%)	296 (98%)	7 (2%)	50 77
1	G	298/320 (93%)	290 (97%)	8 (3%)	44 73
1	I	304/320 (95%)	302 (99%)	2 (1%)	84 93

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	K	301/320 (94%)	290 (96%)	11 (4%)	34 65
1	M	304/320 (95%)	299 (98%)	5 (2%)	62 83
1	O	299/320 (93%)	286 (96%)	13 (4%)	29 62
1	Q	302/320 (94%)	298 (99%)	4 (1%)	69 86
1	S	302/320 (94%)	301 (100%)	1 (0%)	92 97
1	U	301/320 (94%)	291 (97%)	10 (3%)	38 68
1	X	302/320 (94%)	297 (98%)	5 (2%)	60 82
2	B	37/79 (47%)	35 (95%)	2 (5%)	22 55
2	D	35/79 (44%)	33 (94%)	2 (6%)	20 53
2	F	34/79 (43%)	29 (85%)	5 (15%)	3 18
2	H	32/79 (40%)	29 (91%)	3 (9%)	8 35
2	J	33/79 (42%)	31 (94%)	2 (6%)	18 51
2	L	32/79 (40%)	29 (91%)	3 (9%)	8 35
2	N	34/79 (43%)	33 (97%)	1 (3%)	42 71
2	P	34/79 (43%)	33 (97%)	1 (3%)	42 71
2	R	31/79 (39%)	29 (94%)	2 (6%)	17 50
2	T	35/79 (44%)	33 (94%)	2 (6%)	20 53
2	V	31/79 (39%)	29 (94%)	2 (6%)	17 50
2	Y	34/79 (43%)	33 (97%)	1 (3%)	42 71
All	All	4025/4788 (84%)	3925 (98%)	100 (2%)	47 75

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	351	THR
1	O	297	ASN
2	Y	724	GLU
1	O	39	ARG
1	O	198	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	F	719	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	M	12	ASN
1	M	40	HIS
2	P	702	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 24 ligands modelled in this entry, 12 are monoatomic - leaving 12 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	ATP	I	401	4	26,33,33	4.62	9 (34%)	31,52,52	2.88	7 (22%)
3	ATP	A	401	4	26,33,33	4.66	9 (34%)	31,52,52	2.87	7 (22%)
3	ATP	G	401	4	26,33,33	4.64	9 (34%)	31,52,52	2.87	7 (22%)
3	ATP	C	401	4	26,33,33	4.63	9 (34%)	31,52,52	2.89	8 (25%)
3	ATP	E	401	4	26,33,33	4.64	9 (34%)	31,52,52	2.90	7 (22%)
3	ATP	U	401	4	26,33,33	4.64	9 (34%)	31,52,52	2.88	7 (22%)
3	ATP	Q	401	4	26,33,33	4.66	9 (34%)	31,52,52	2.87	7 (22%)
3	ATP	K	401	4	26,33,33	4.62	9 (34%)	31,52,52	2.92	7 (22%)
3	ATP	O	401	4	26,33,33	4.59	9 (34%)	31,52,52	3.00	7 (22%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	ATP	X	401	4	26,33,33	4.61	9 (34%)	31,52,52	2.89	7 (22%)
3	ATP	M	401	4	26,33,33	4.66	9 (34%)	31,52,52	2.87	8 (25%)
3	ATP	S	401	4	26,33,33	4.64	9 (34%)	31,52,52	2.88	6 (19%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ATP	I	401	4	-	4/18/38/38	0/3/3/3
3	ATP	A	401	4	-	1/18/38/38	0/3/3/3
3	ATP	G	401	4	-	5/18/38/38	0/3/3/3
3	ATP	C	401	4	-	5/18/38/38	0/3/3/3
3	ATP	E	401	4	-	4/18/38/38	0/3/3/3
3	ATP	U	401	4	-	2/18/38/38	0/3/3/3
3	ATP	Q	401	4	-	6/18/38/38	0/3/3/3
3	ATP	K	401	4	-	7/18/38/38	0/3/3/3
3	ATP	O	401	4	-	7/18/38/38	0/3/3/3
3	ATP	X	401	4	-	3/18/38/38	0/3/3/3
3	ATP	M	401	4	-	7/18/38/38	0/3/3/3
3	ATP	S	401	4	-	5/18/38/38	0/3/3/3

The worst 5 of 108 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	Q	401	ATP	C2'-C1'	-16.15	1.29	1.53
3	C	401	ATP	C2'-C1'	-16.15	1.29	1.53
3	M	401	ATP	C2'-C1'	-16.11	1.29	1.53
3	U	401	ATP	C2'-C1'	-16.11	1.29	1.53
3	I	401	ATP	C2'-C1'	-16.09	1.29	1.53

The worst 5 of 85 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	401	ATP	C5-C6-N6	9.73	135.13	120.35
3	O	401	ATP	C5-C6-N6	9.69	135.07	120.35
3	K	401	ATP	C5-C6-N6	9.62	134.97	120.35
3	Q	401	ATP	C5-C6-N6	9.58	134.91	120.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	401	ATP	C5-C6-N6	9.58	134.91	120.35

There are no chirality outliers.

5 of 56 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	401	ATP	C3'-C4'-C5'-O5'
3	E	401	ATP	C5'-O5'-PA-O1A
3	E	401	ATP	C3'-C4'-C5'-O5'
3	G	401	ATP	C5'-O5'-PA-O1A
3	I	401	ATP	C5'-O5'-PA-O1A

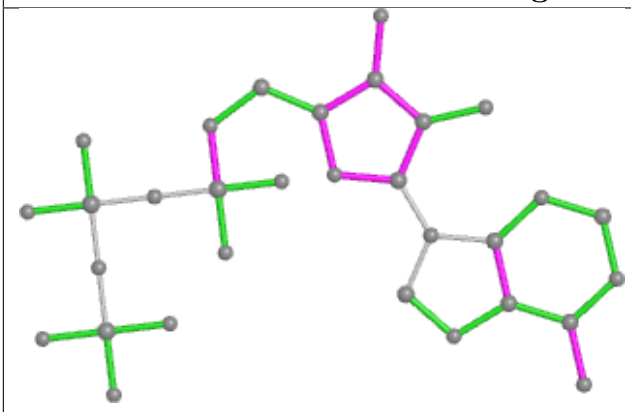
There are no ring outliers.

7 monomers are involved in 10 short contacts:

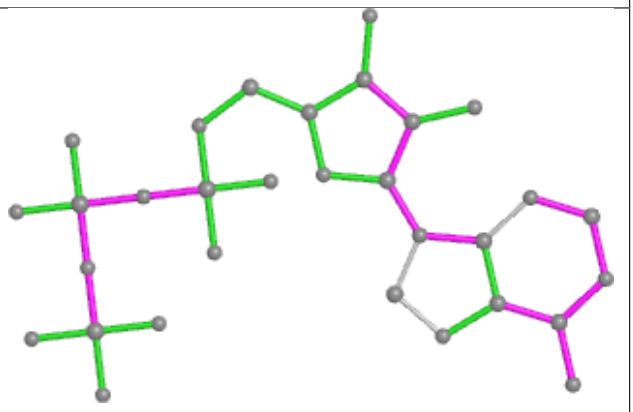
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	401	ATP	1	0
3	E	401	ATP	2	0
3	Q	401	ATP	1	0
3	K	401	ATP	1	0
3	O	401	ATP	3	0
3	M	401	ATP	1	0
3	S	401	ATP	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

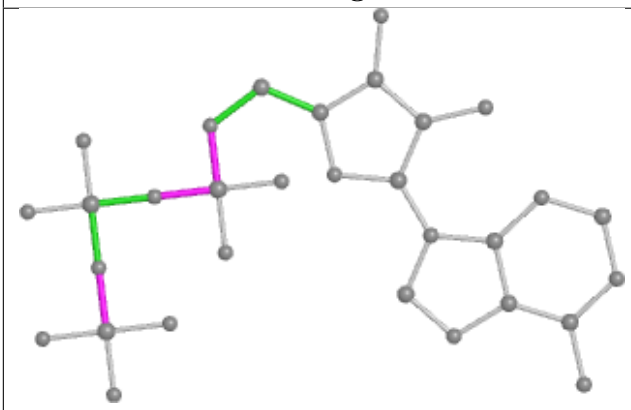
Ligand ATP I 401



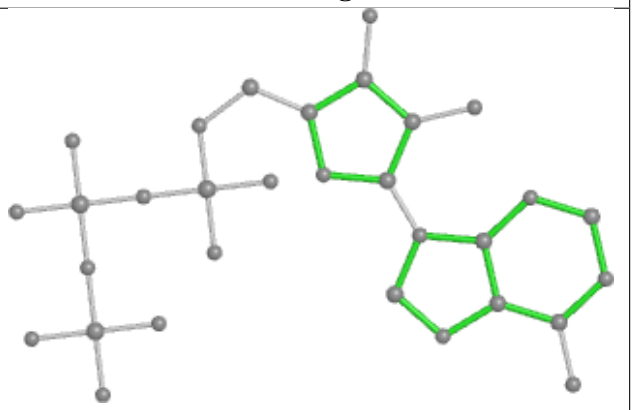
Bond lengths



Bond angles

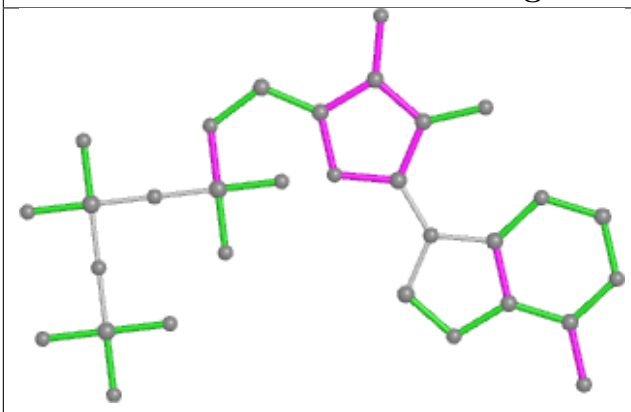


Torsions

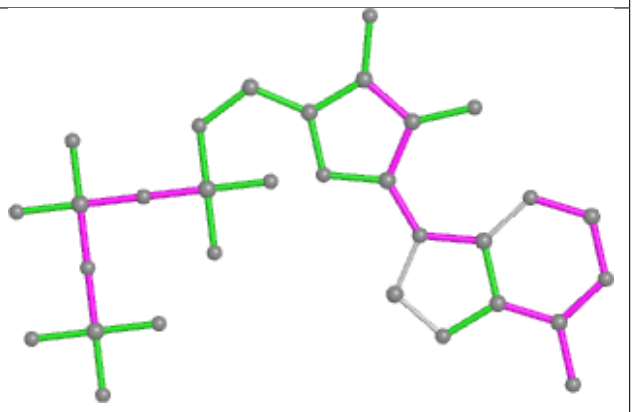


Rings

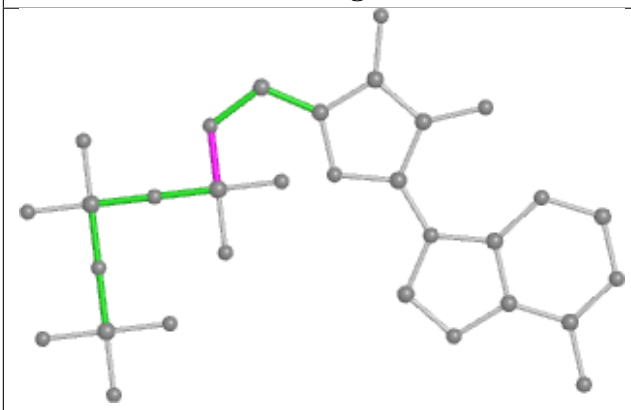
Ligand ATP A 401



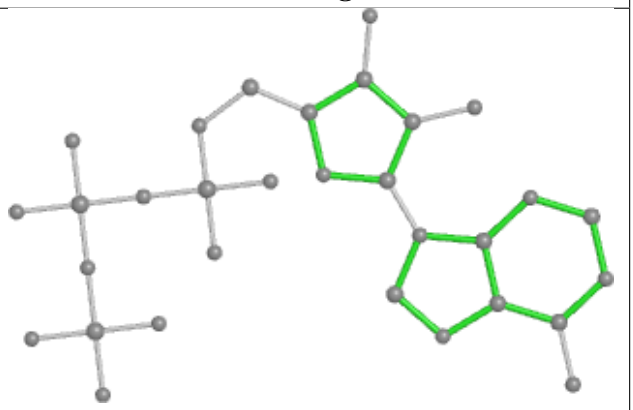
Bond lengths



Bond angles

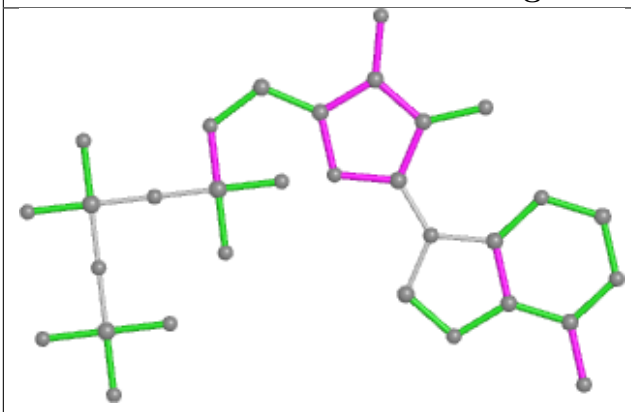


Torsions

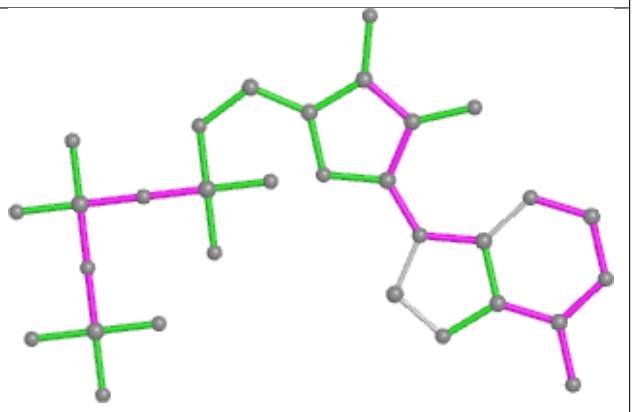


Rings

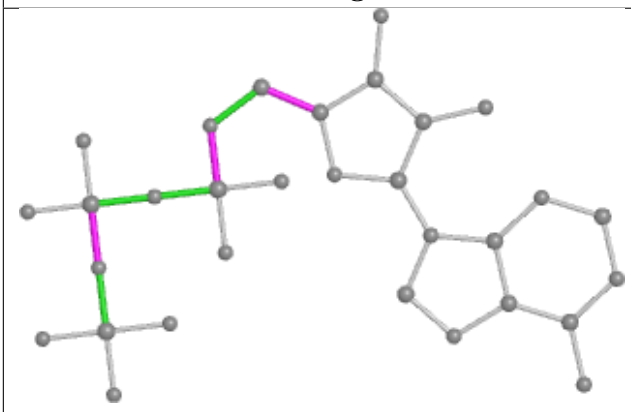
Ligand ATP G 401



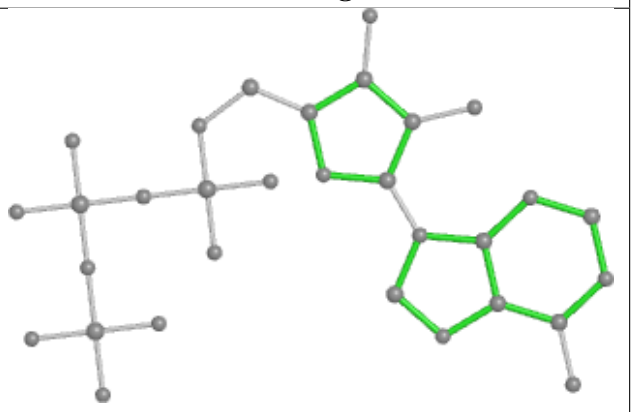
Bond lengths



Bond angles

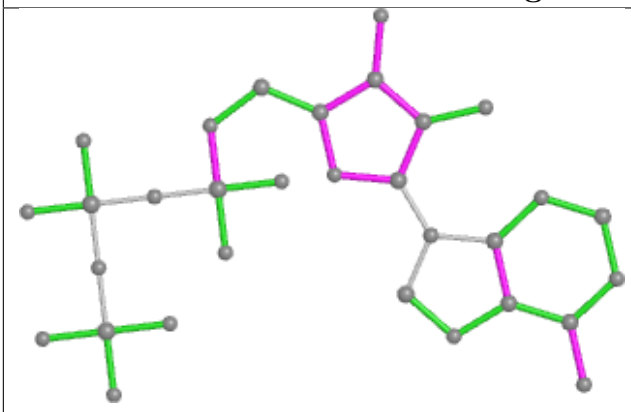


Torsions

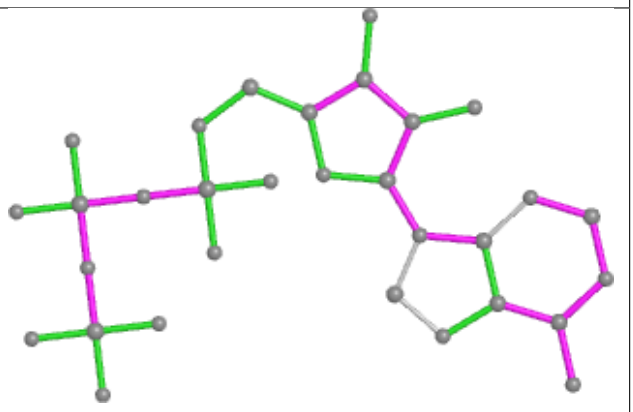


Rings

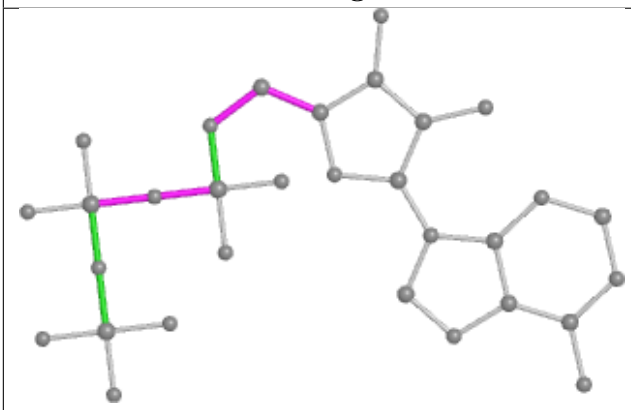
Ligand ATP C 401



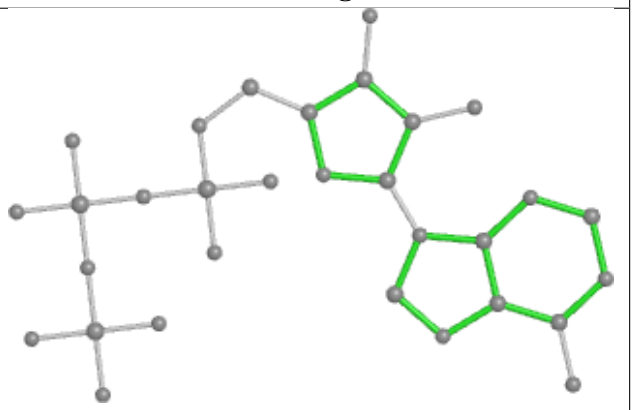
Bond lengths



Bond angles

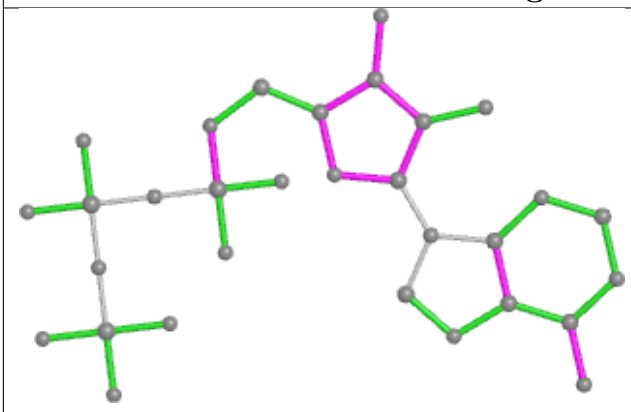


Torsions

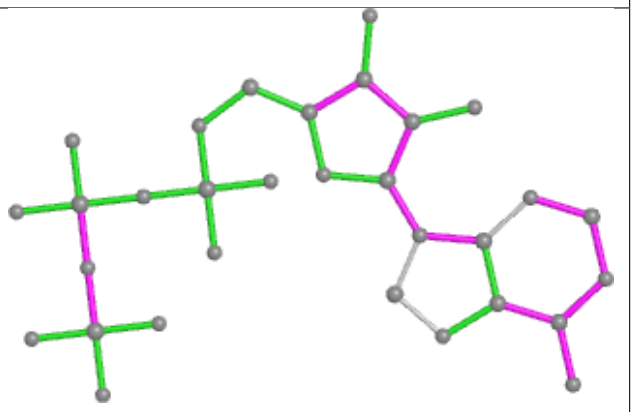


Rings

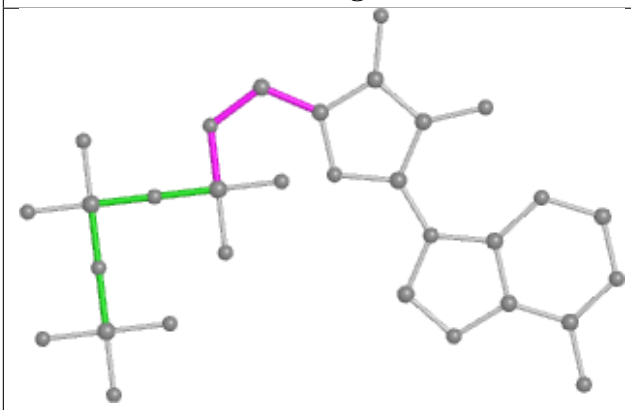
Ligand ATP E 401



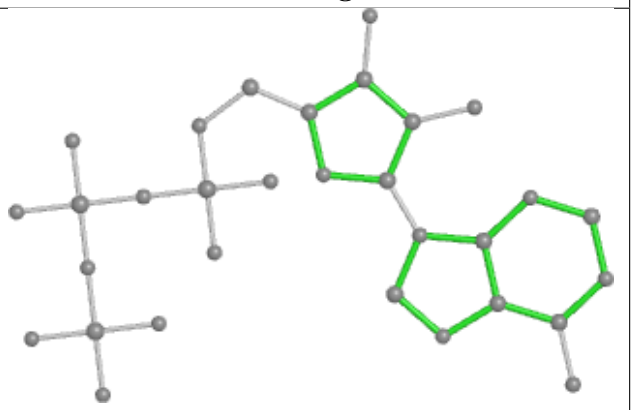
Bond lengths



Bond angles

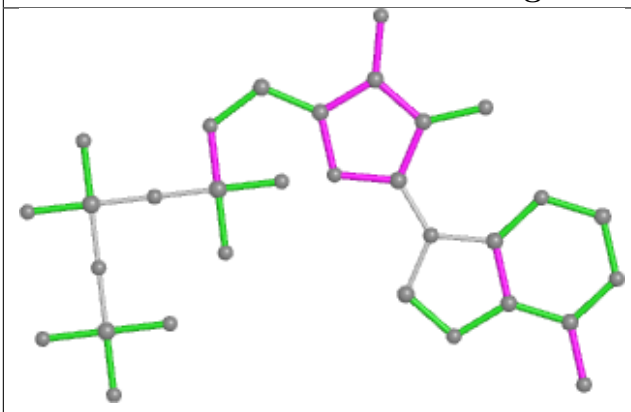


Torsions

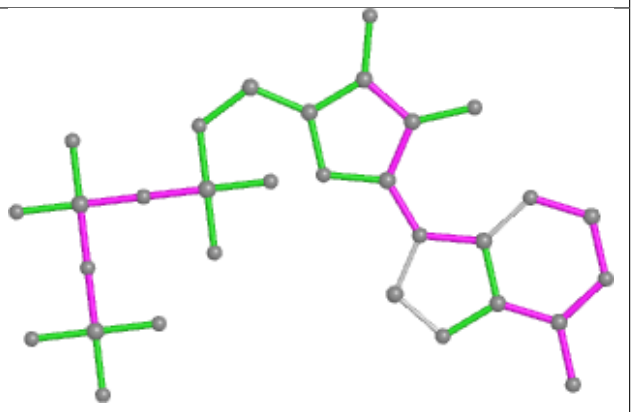


Rings

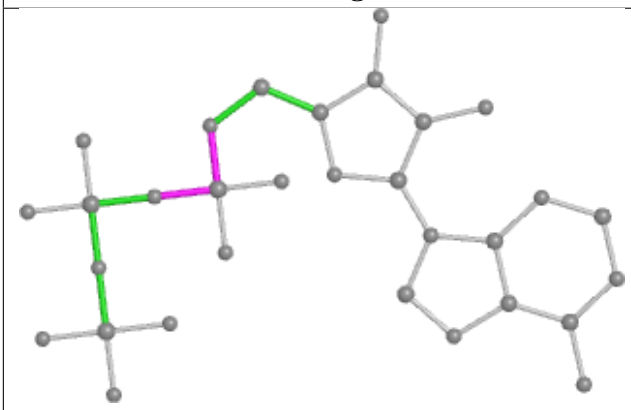
Ligand ATP U 401



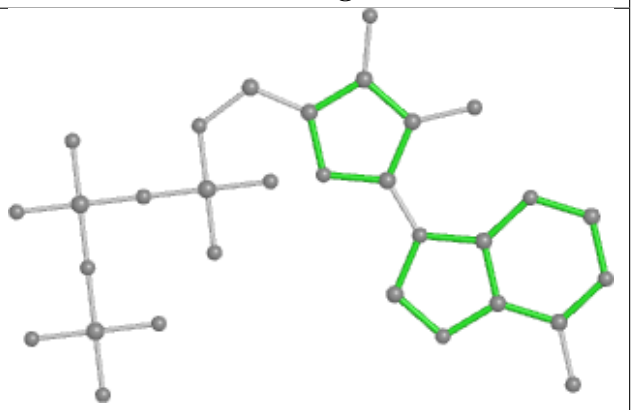
Bond lengths



Bond angles

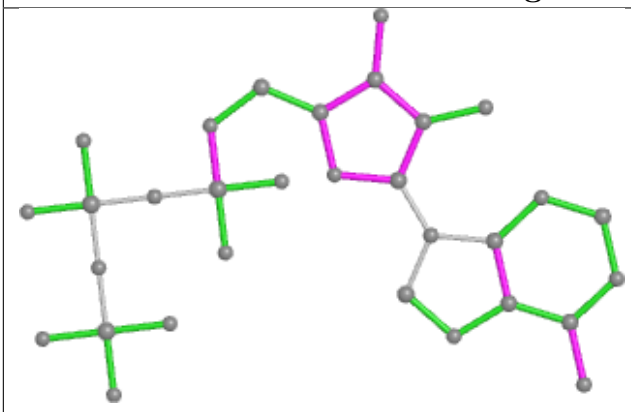


Torsions

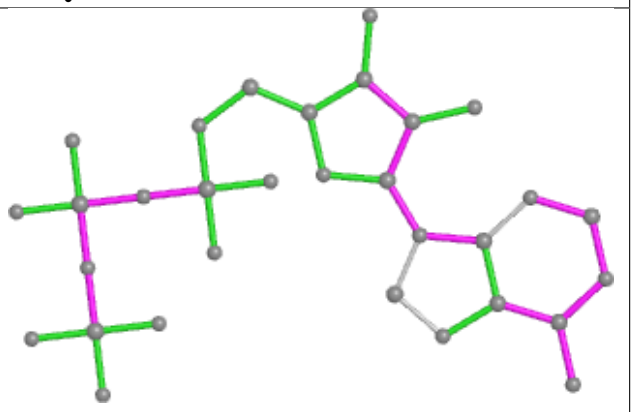


Rings

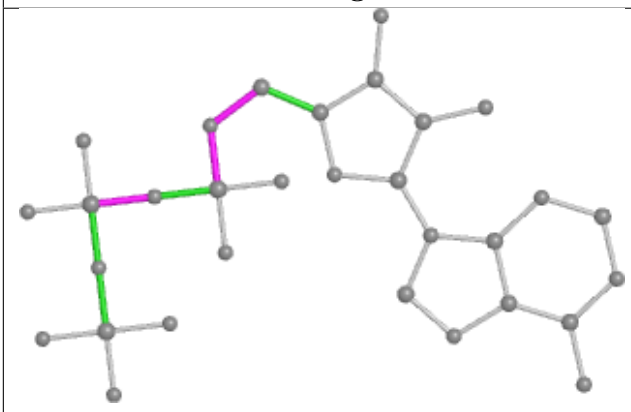
Ligand ATP Q 401



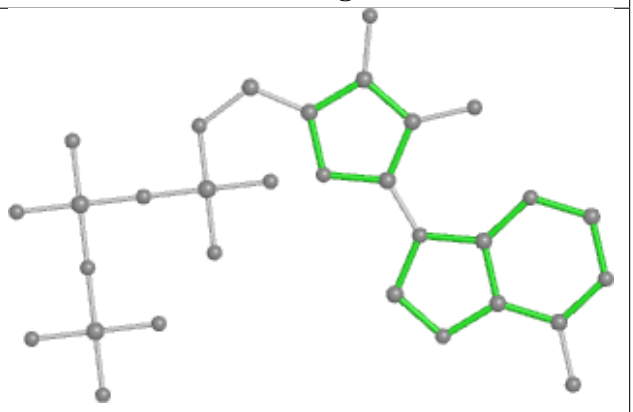
Bond lengths



Bond angles

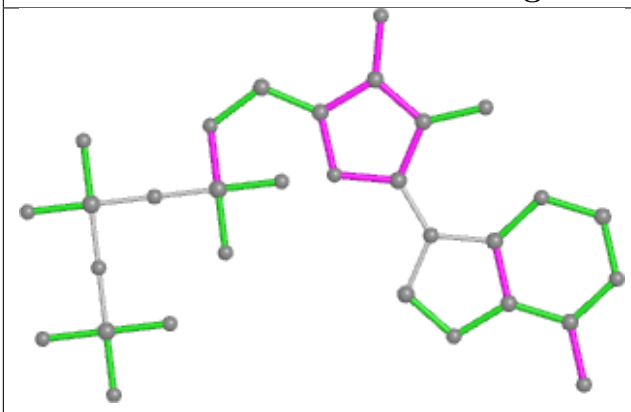


Torsions

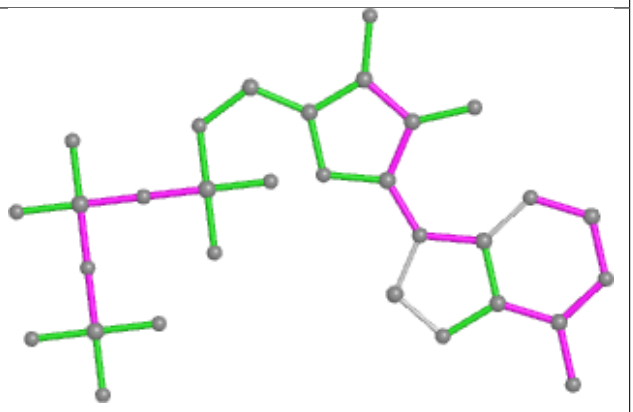


Rings

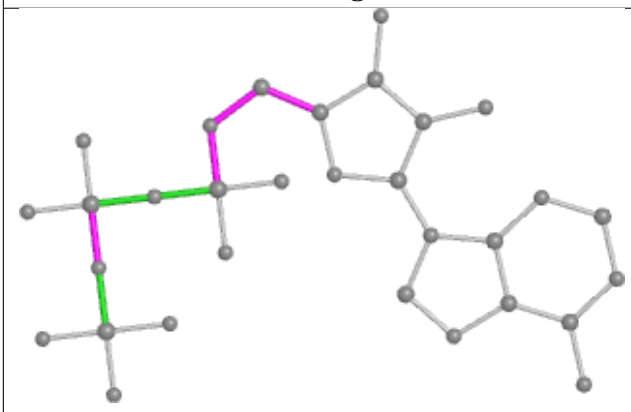
Ligand ATP K 401



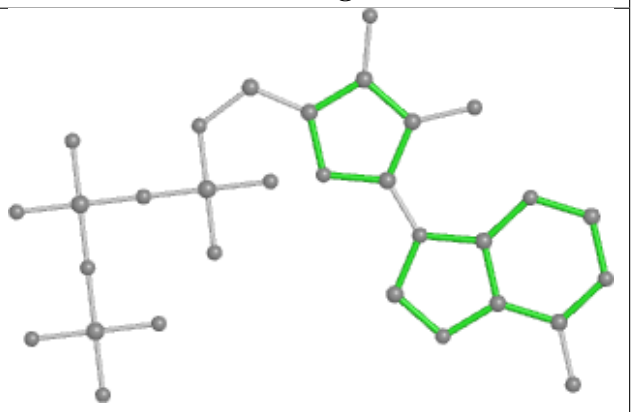
Bond lengths



Bond angles

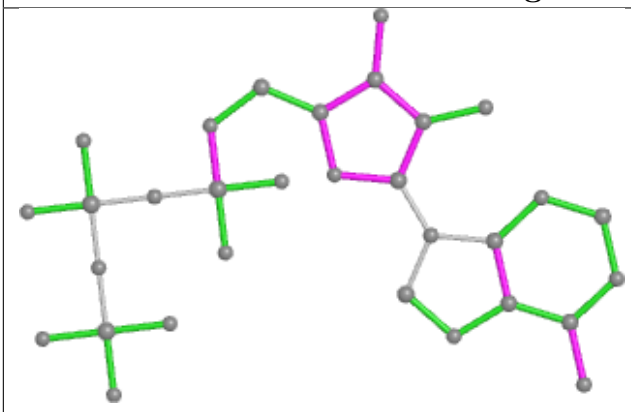


Torsions

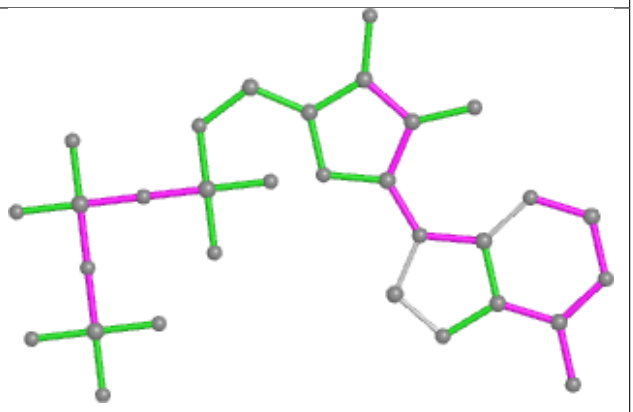


Rings

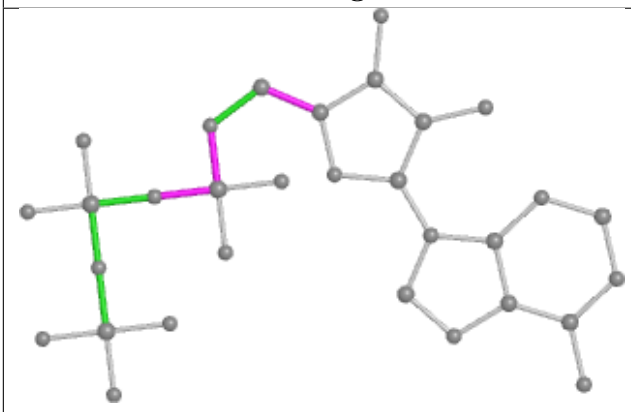
Ligand ATP O 401



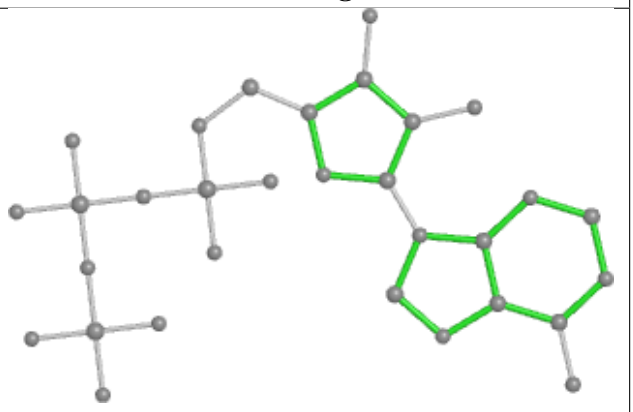
Bond lengths



Bond angles

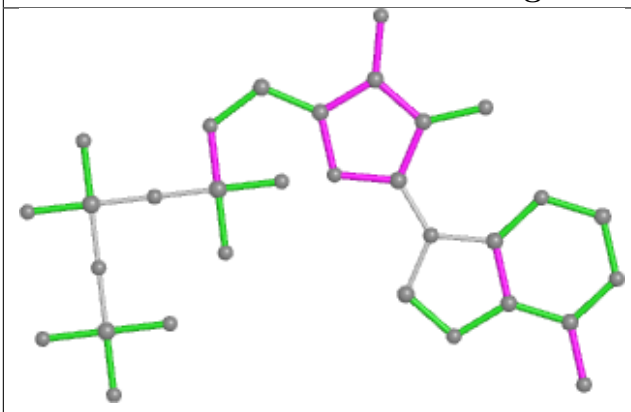


Torsions

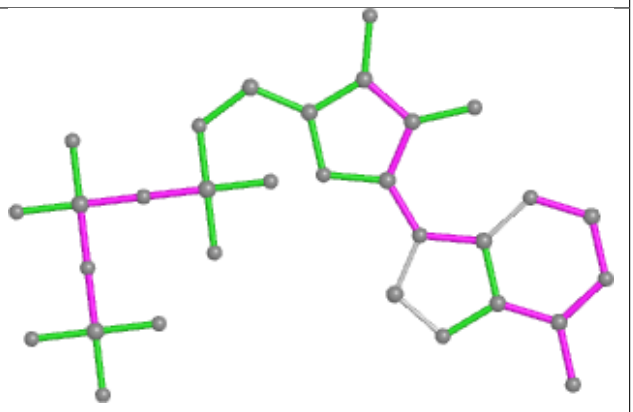


Rings

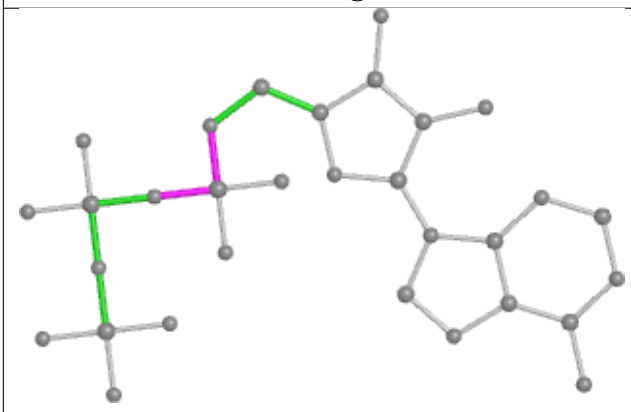
Ligand ATP X 401



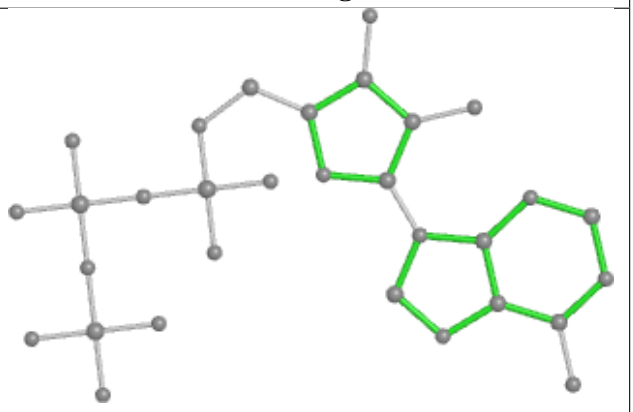
Bond lengths



Bond angles

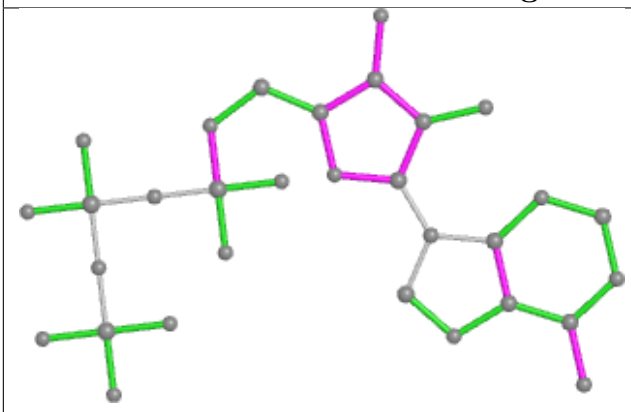


Torsions

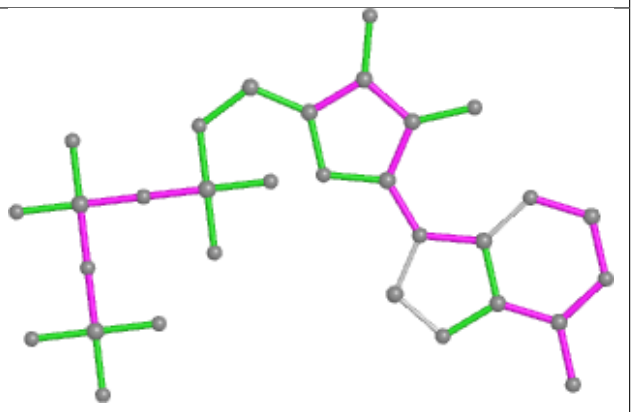


Rings

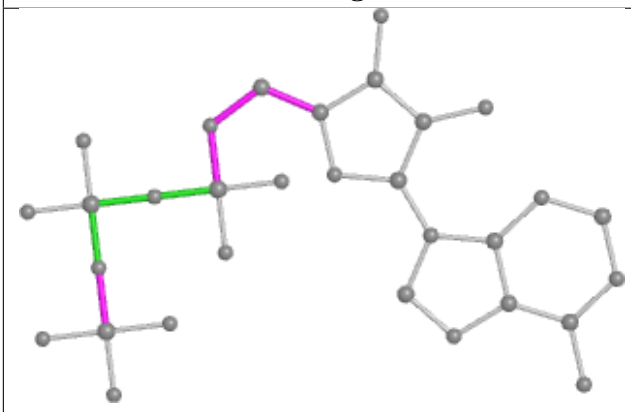
Ligand ATP M 401



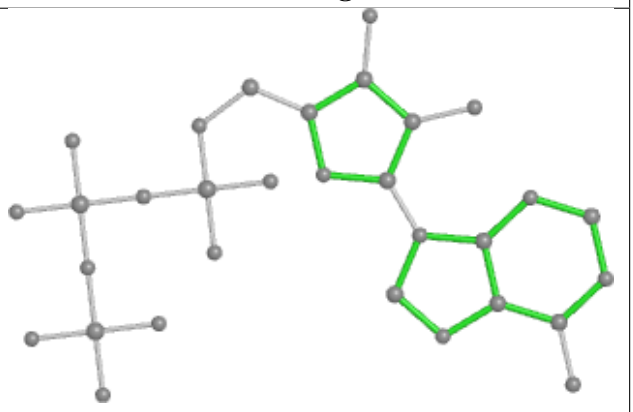
Bond lengths



Bond angles

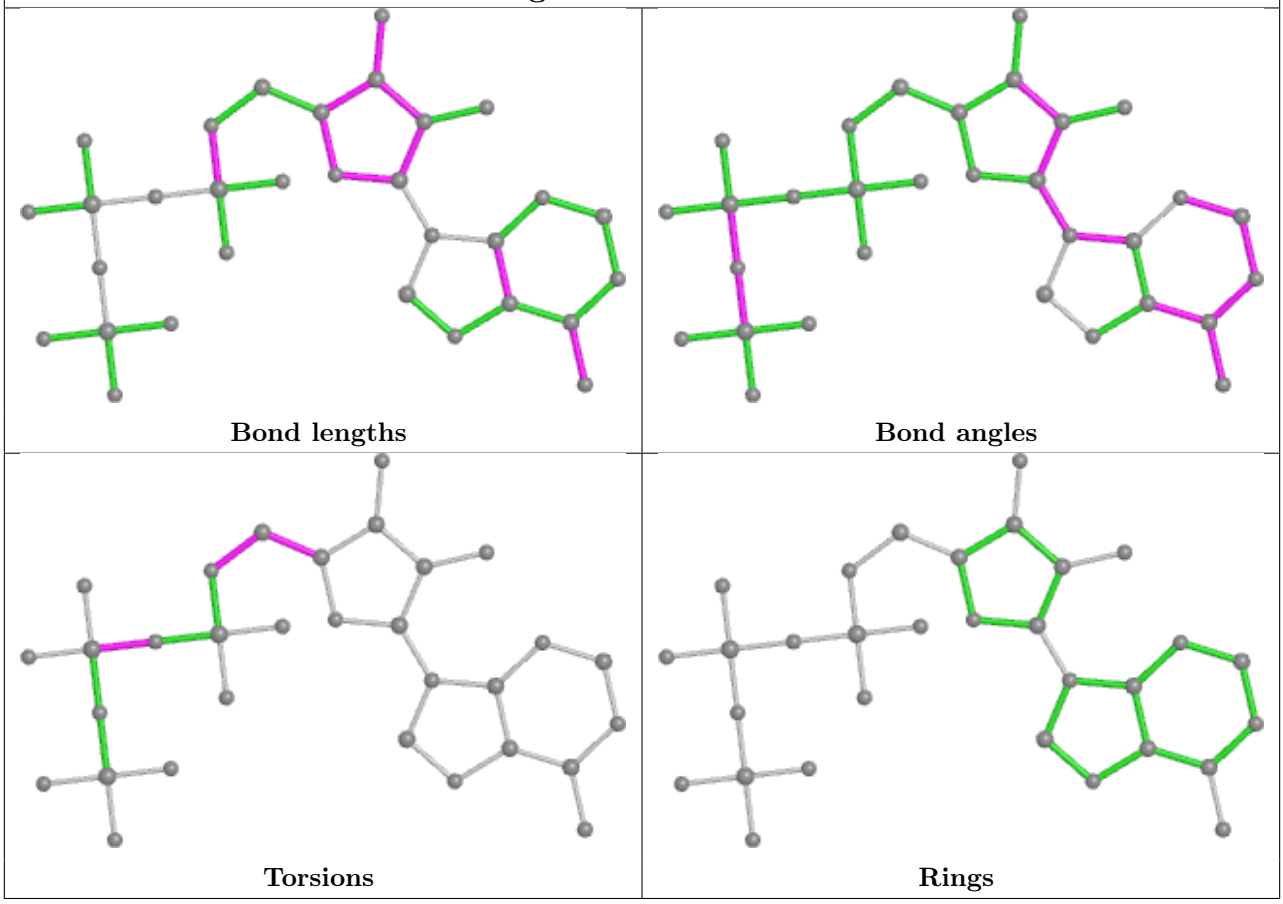


Torsions



Rings

Ligand ATP S 401



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	357/377 (94%)	0.86	33 (9%) 9 9	47, 76, 125, 154	0
1	C	358/377 (94%)	0.80	33 (9%) 9 9	50, 79, 125, 174	0
1	E	358/377 (94%)	0.83	34 (9%) 8 8	57, 89, 133, 161	0
1	G	356/377 (94%)	0.91	44 (12%) 4 5	56, 89, 130, 155	0
1	I	358/377 (94%)	0.82	29 (8%) 12 12	48, 76, 125, 168	0
1	K	357/377 (94%)	0.92	37 (10%) 6 7	62, 94, 144, 169	0
1	M	358/377 (94%)	0.85	32 (8%) 9 10	53, 79, 121, 171	0
1	O	355/377 (94%)	0.85	35 (9%) 7 8	61, 95, 144, 161	0
1	Q	357/377 (94%)	0.84	32 (8%) 9 10	59, 87, 126, 171	0
1	S	358/377 (94%)	0.77	27 (7%) 14 14	53, 78, 123, 172	0
1	U	358/377 (94%)	0.78	27 (7%) 14 14	57, 87, 123, 167	0
1	X	358/377 (94%)	0.84	29 (8%) 12 12	52, 78, 125, 168	0
2	B	40/92 (43%)	0.28	0 100 100	69, 85, 119, 129	0
2	D	38/92 (41%)	0.67	3 (7%) 12 13	71, 92, 129, 151	0
2	F	37/92 (40%)	0.34	3 (8%) 12 12	95, 108, 133, 149	0
2	H	37/92 (40%)	0.34	1 (2%) 54 48	24, 103, 125, 131	0
2	J	39/92 (42%)	0.26	2 (5%) 28 25	68, 83, 102, 113	0
2	L	37/92 (40%)	0.18	1 (2%) 54 48	95, 121, 165, 173	0
2	N	37/92 (40%)	0.46	1 (2%) 54 48	78, 97, 111, 118	0
2	P	37/92 (40%)	0.67	4 (10%) 5 6	98, 115, 158, 172	0
2	R	37/92 (40%)	0.51	4 (10%) 5 6	87, 108, 127, 135	0
2	T	38/92 (41%)	0.44	1 (2%) 56 49	73, 88, 116, 122	0
2	V	37/92 (40%)	0.54	4 (10%) 5 6	92, 113, 126, 130	0
2	Y	37/92 (40%)	0.29	1 (2%) 54 48	79, 93, 107, 110	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
All	All	4739/5628 (84%)	0.80	417 (8%)	10	10	24, 86, 130, 174	0

The worst 5 of 417 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	199	SER	5.6
1	M	236	LEU	5.4
1	S	198	TYR	4.8
2	P	726	ILE	4.6
1	X	218	TYR	4.3

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CA	K	402	1/1	0.92	0.23	83,83,83,83	0
4	CA	G	402	1/1	0.93	0.22	80,80,80,80	0
4	CA	I	402	1/1	0.93	0.25	72,72,72,72	0
4	CA	A	402	1/1	0.93	0.28	62,62,62,62	0
4	CA	U	402	1/1	0.93	0.23	67,67,67,67	0
3	ATP	G	401	31/31	0.94	0.28	75,91,100,102	0
3	ATP	K	401	31/31	0.94	0.23	83,97,101,105	0
3	ATP	O	401	31/31	0.94	0.25	81,94,104,111	0
4	CA	O	402	1/1	0.94	0.20	73,73,73,73	0
3	ATP	E	401	31/31	0.94	0.24	76,91,107,118	0
3	ATP	A	401	31/31	0.95	0.29	66,77,85,89	0
3	ATP	C	401	31/31	0.95	0.22	72,82,91,95	0
3	ATP	S	401	31/31	0.95	0.22	72,78,88,90	0

Continued on next page...

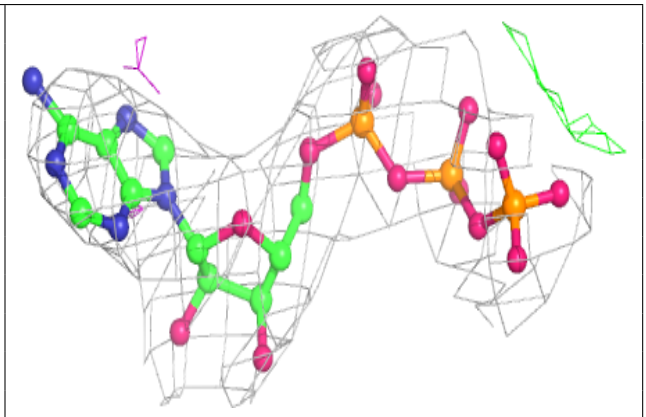
Continued from previous page...

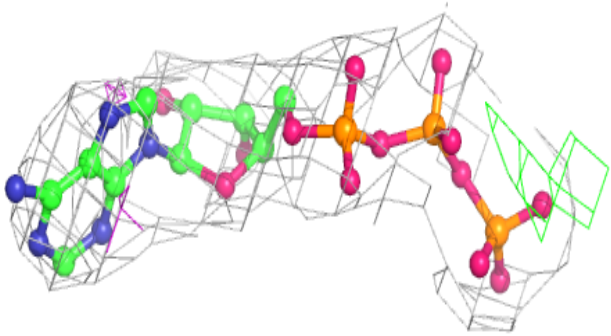
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	ATP	U	401	31/31	0.95	0.25	69,90,98,108	0
3	ATP	I	401	31/31	0.95	0.28	62,81,88,92	0
3	ATP	X	401	31/31	0.96	0.24	73,81,87,89	0
3	ATP	M	401	31/31	0.96	0.21	75,83,94,99	0
4	CA	E	402	1/1	0.96	0.20	73,73,73,73	0
4	CA	Q	402	1/1	0.96	0.27	68,68,68,68	0
3	ATP	Q	401	31/31	0.96	0.26	69,88,94,95	0
4	CA	M	402	1/1	0.97	0.22	67,67,67,67	0
4	CA	S	402	1/1	0.97	0.31	66,66,66,66	0
4	CA	C	402	1/1	0.97	0.21	65,65,65,65	0
4	CA	X	402	1/1	0.98	0.27	70,70,70,70	0

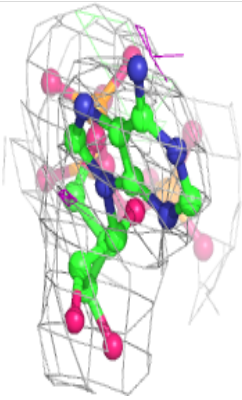
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around ATP G 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

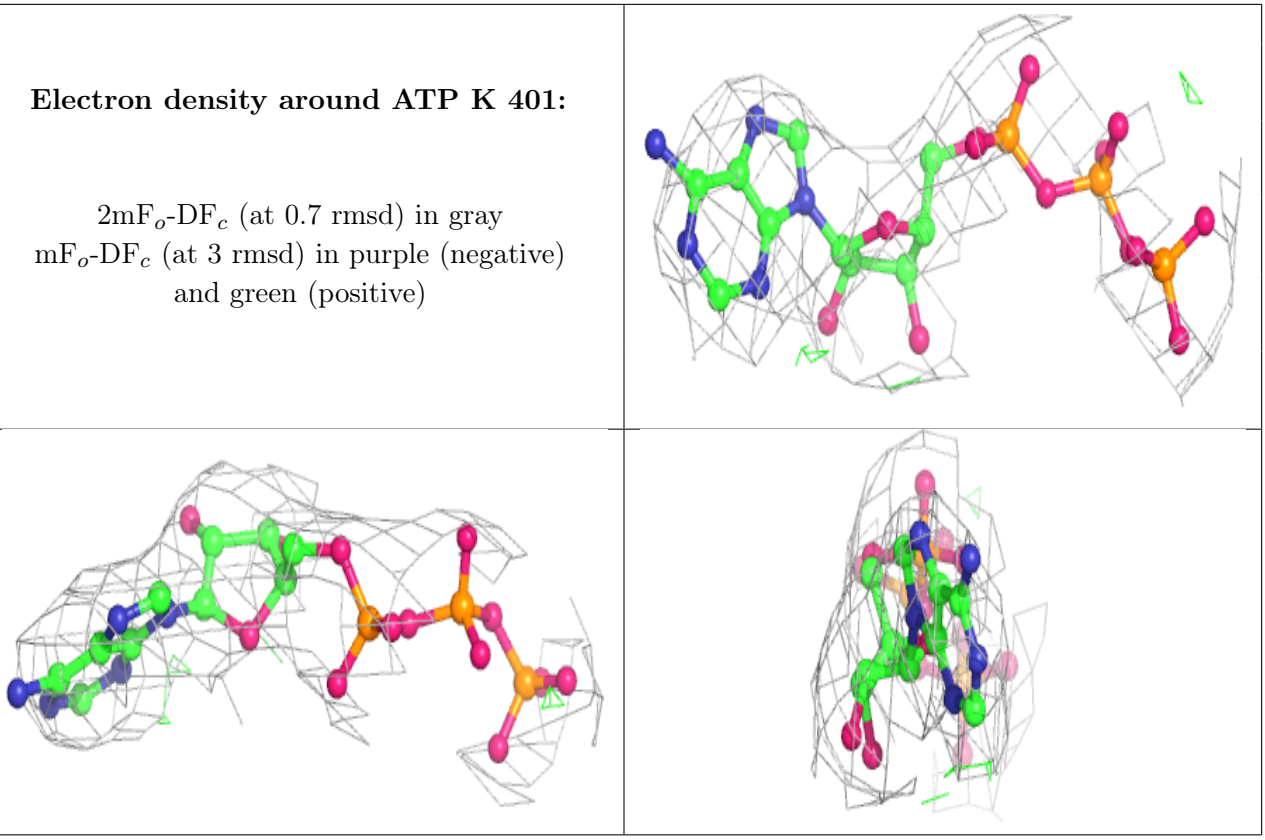






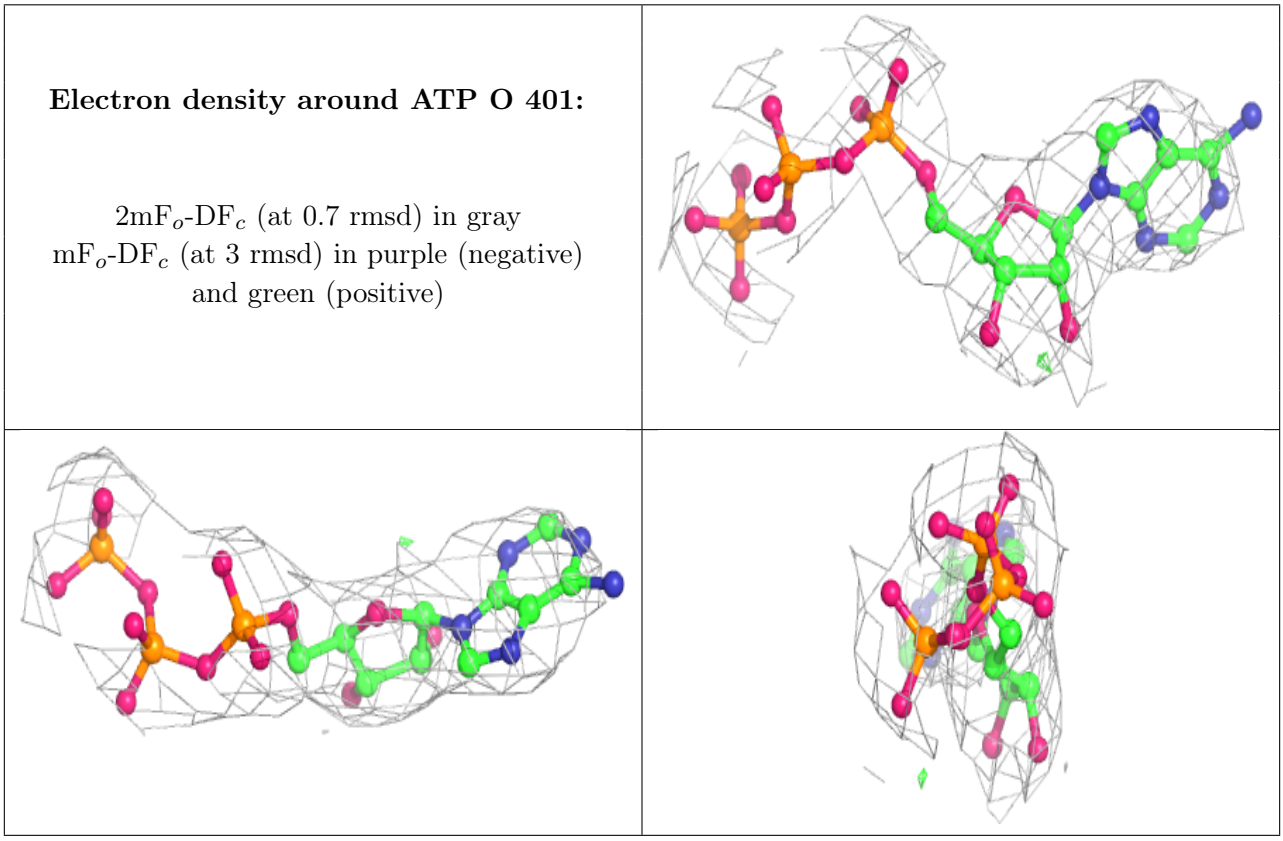
Electron density around ATP K 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



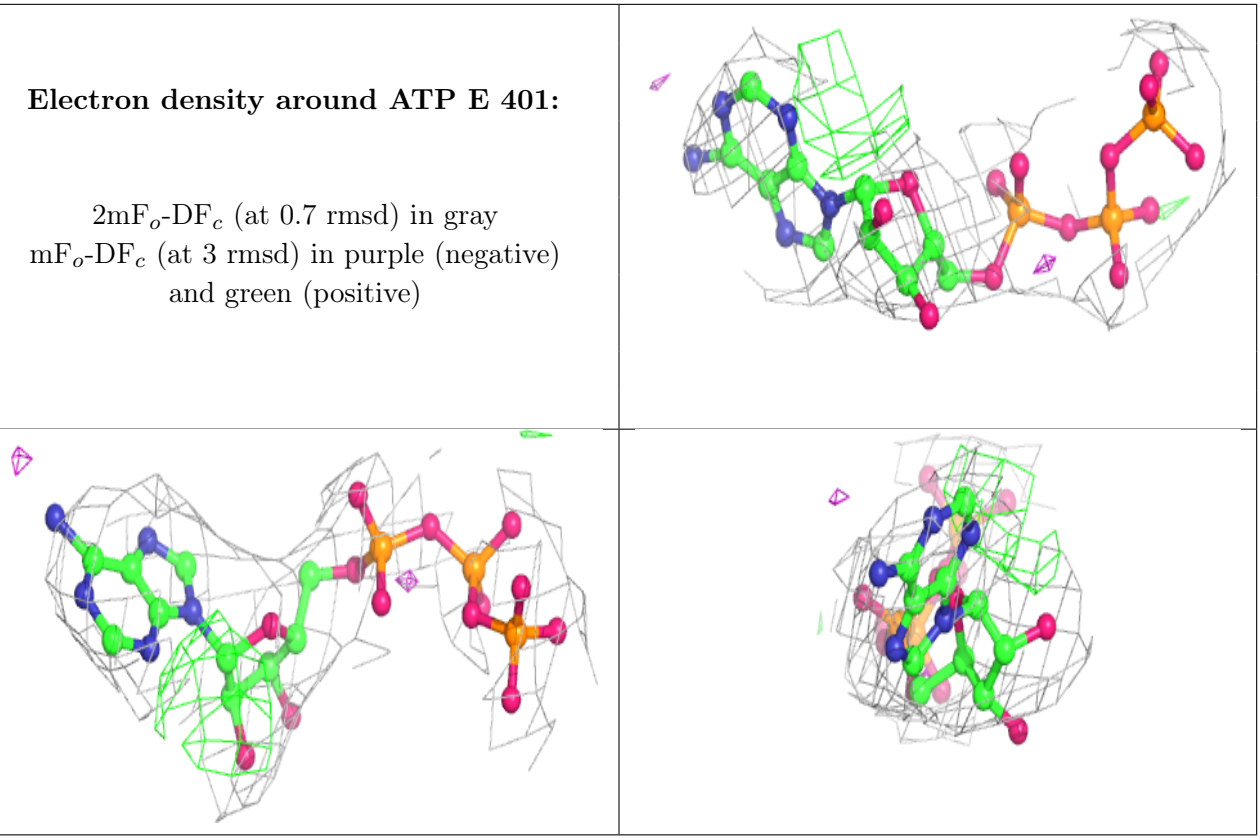
Electron density around ATP O 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



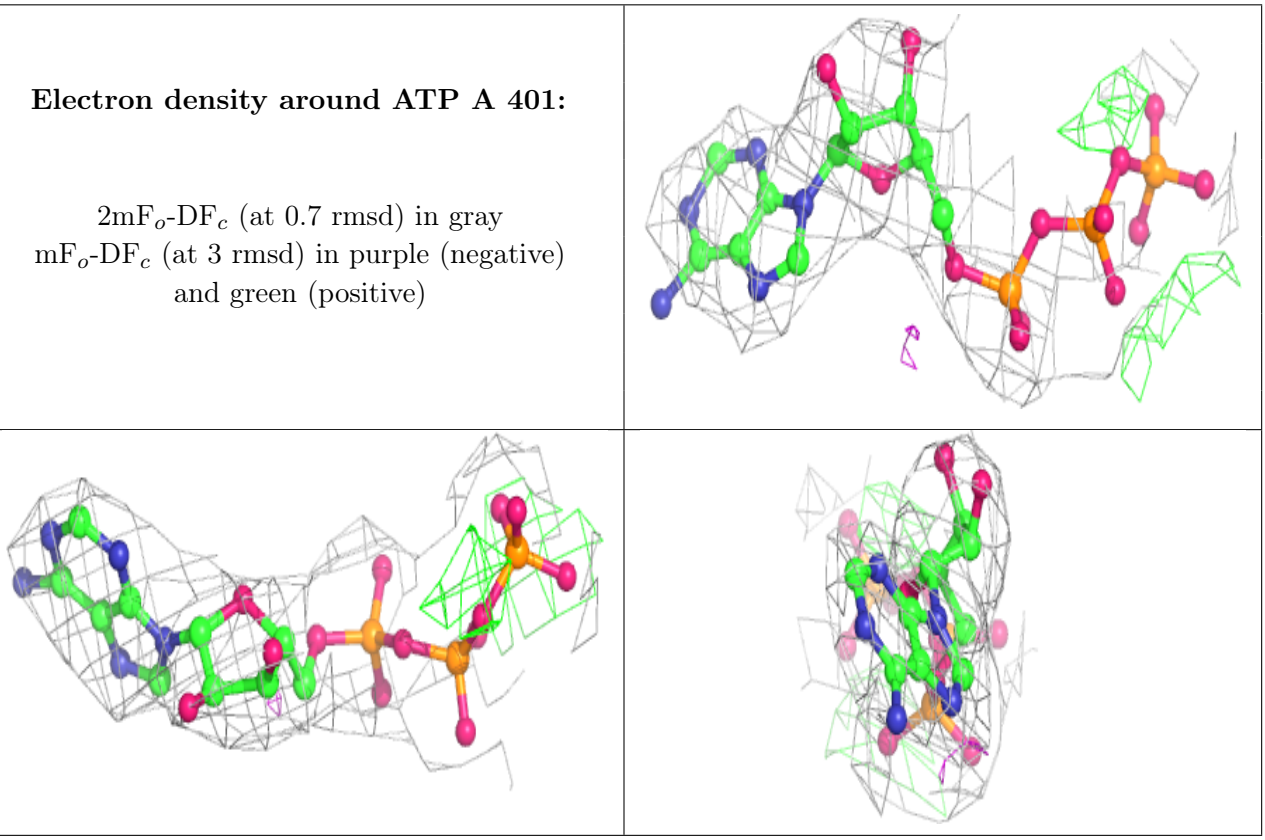
Electron density around ATP E 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



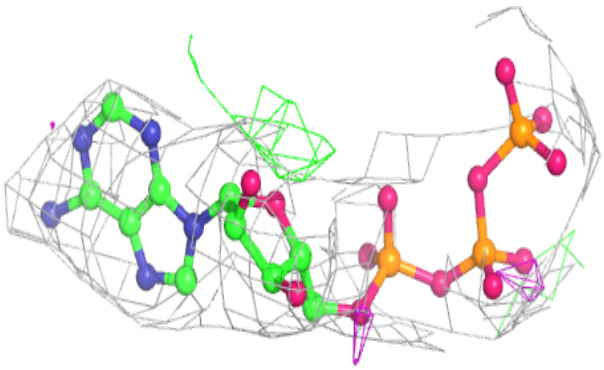
Electron density around ATP A 401:

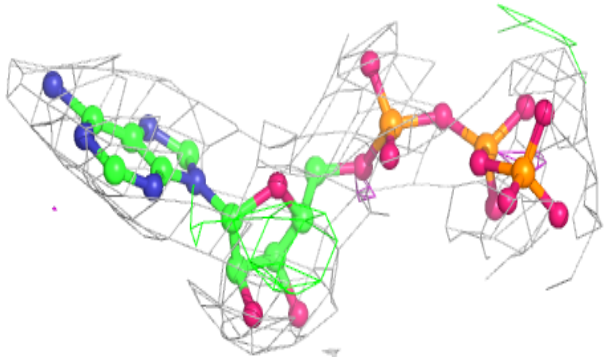
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

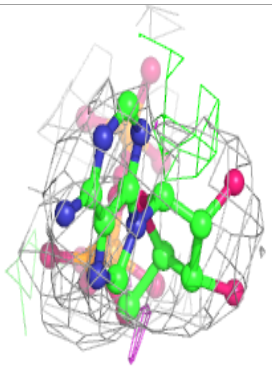


Electron density around ATP C 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

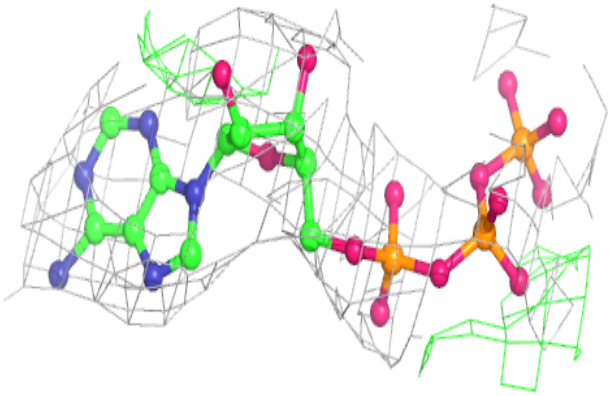


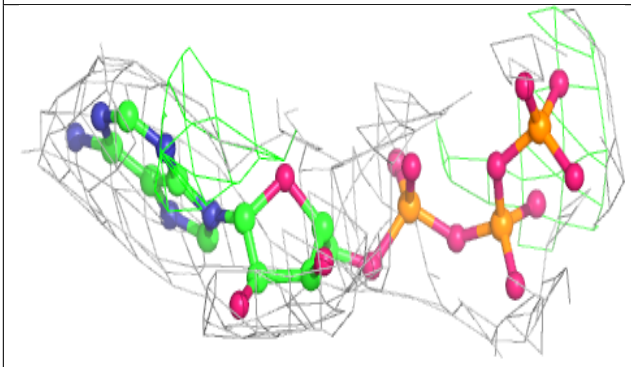


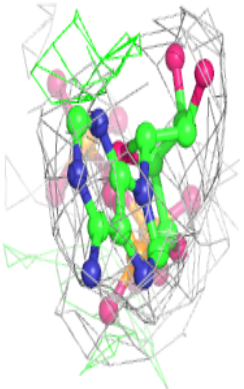


Electron density around ATP S 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

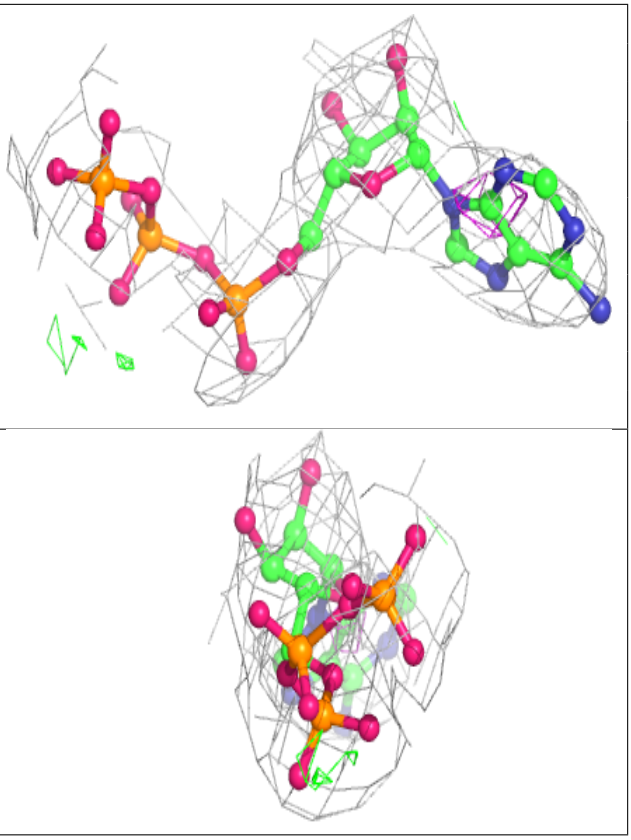






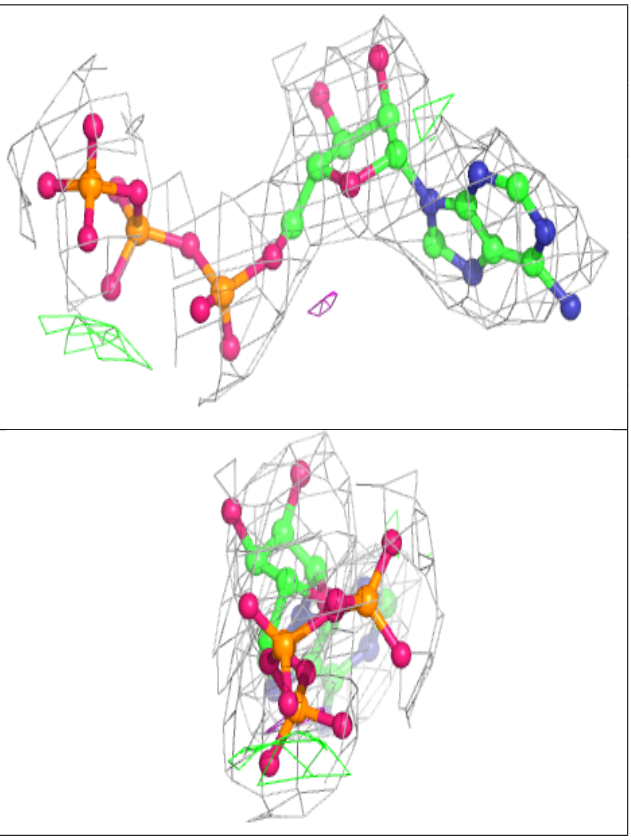
Electron density around ATP U 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



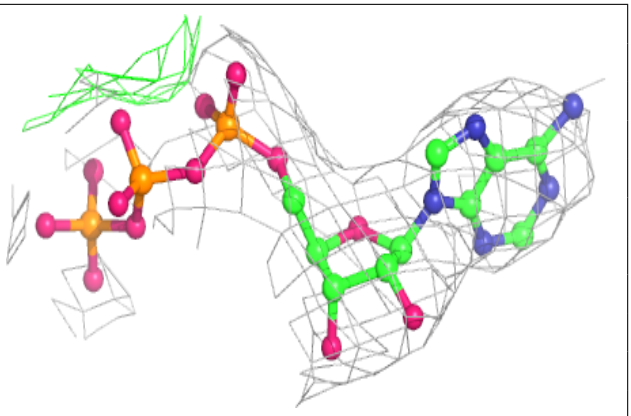
Electron density around ATP I 401:

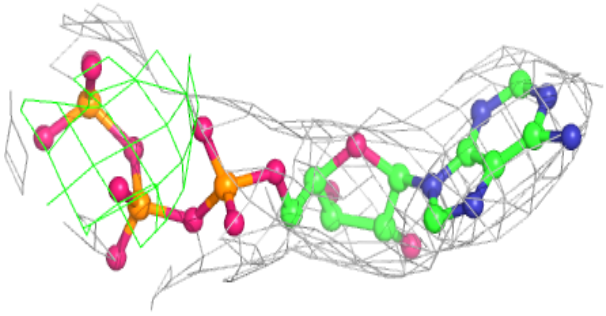
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

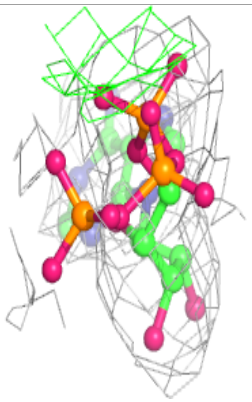


Electron density around ATP X 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

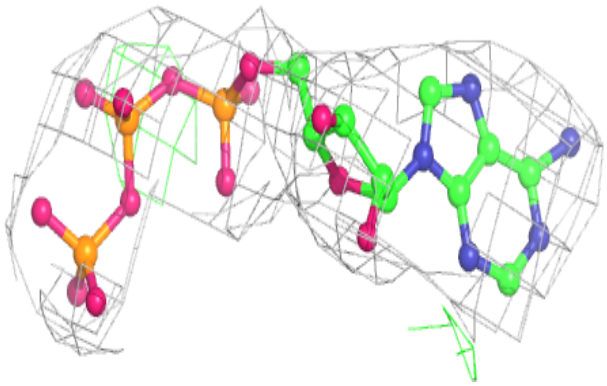


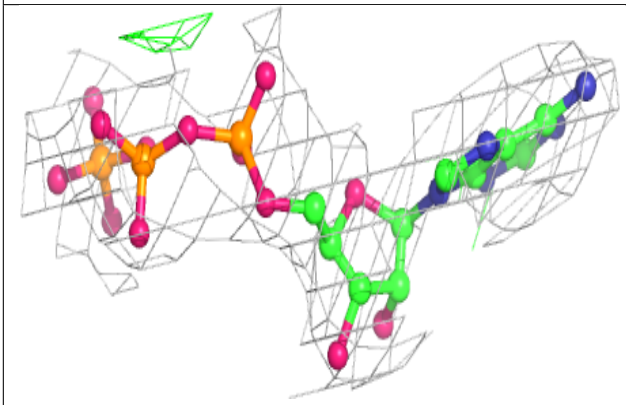


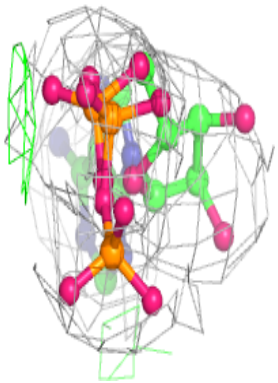


Electron density around ATP M 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

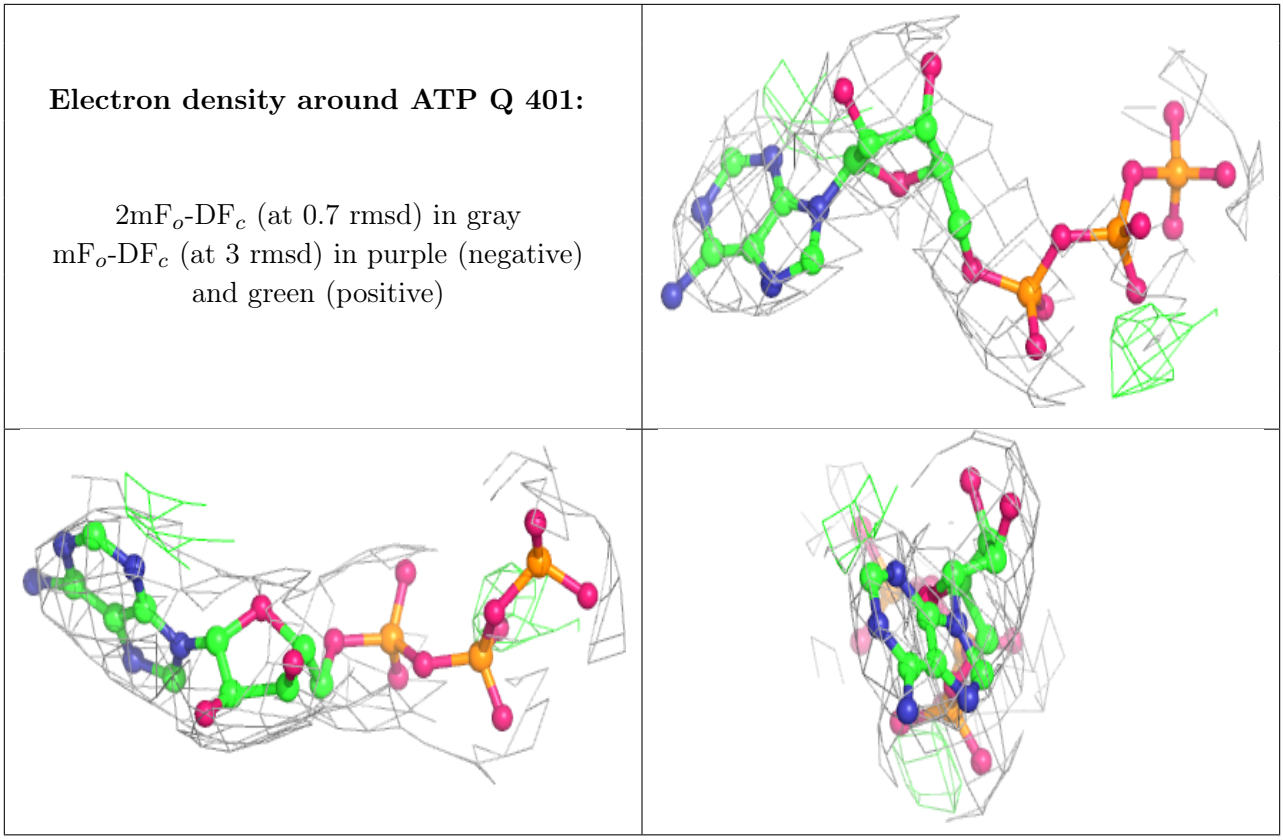






Electron density around ATP Q 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.