wwPDB X-ray Structure Validation Summary Report (i)

Sep 3, 2023 – 10:56 PM EDT

PDB ID	:	3SHJ
Title	:	Proteasome in complex with hydroxyurea derivative HU10
Authors	:	Gallastegui, N.; Beck, P.; Arciniega, M.; Hillebrand, S.; Huber, R.; Groll, M.
Deposited on	:	2011-06-16
Resolution	:	2.80 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

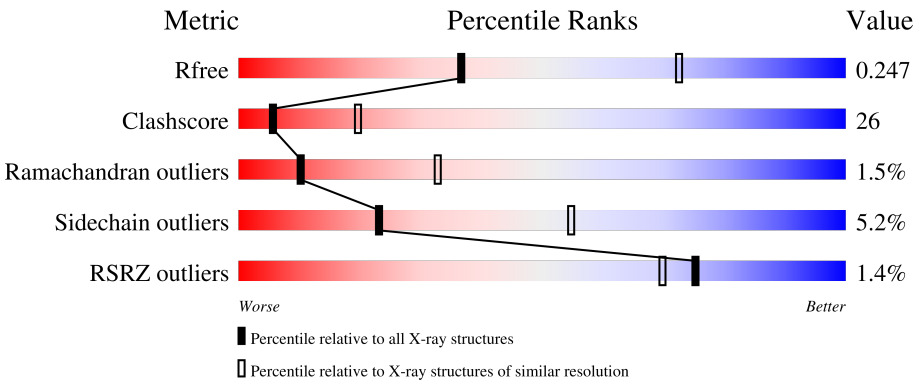
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	250	62% 36% •
1	O	250	2% 64% 34% •
2	B	244	2% 53% 41% 6%
2	P	244	2% 55% 39% 6%
3	C	241	2% 56% 41% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	241	7% 54% 42% •
4	D	242	2% 60% 36% •
4	R	242	3% 55% 40% •
5	E	233	% 49% 45% 6%
5	S	233	4% 51% 43% 6%
6	F	244	% 56% 41% •
6	T	244	59% 38% •
7	G	243	2% 56% 41% •
7	U	243	% 52% 44% •
8	H	222	55% 41% •
8	V	222	% 54% 41% 5%
9	I	204	61% 35% •
9	W	204	61% 36% •
10	J	198	% 58% 39% •
10	X	198	2% 56% 41% •
11	K	212	59% 39% •
11	Y	212	61% 36% •
12	L	222	54% 41% 5%
12	Z	222	% 55% 41% 5%
13	1	233	63% 34% •
13	M	233	59% 38% •
14	2	196	59% 38% •
14	N	196	67% 31% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
15	MES	K	1(L)	-	X	-	-

2 Entry composition (i)

There are 17 unique types of molecules in this entry. The entry contains 50905 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 Proteasome component Y7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	250	1915	1219	315	377	4	0	0	0
1	O	250	1915	1219	315	377	4	0	0	0

- Molecule 2 is a protein called Proteasome component Y13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	244	1905	1201	321	380	3	0	0	0
2	P	244	1905	1201	321	380	3	0	0	0

- Molecule 3 is a protein called Proteasome component PRE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	241	1891	1181	331	375	4	0	0	0
3	Q	241	1891	1181	331	375	4	0	0	0

- Molecule 4 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	242	1862	1162	314	379	7	0	0	0
4	R	242	1862	1162	314	379	7	0	0	0

- Molecule 5 is a protein called Proteasome component PRE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	233	1795	1129	312	350	4	0	0	0
5	S	233	1795	1129	312	350	4	0	0	0

- Molecule 6 is a protein called Proteasome component C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	244	1897	1205	330	358	4	0	0	0
6	T	244	1897	1205	330	358	4	0	0	0

- Molecule 7 is a protein called Proteasome component C7-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	243	1921	1221	322	370	8	0	0	0
7	U	243	1921	1221	322	370	8	0	0	0

- Molecule 8 is a protein called Proteasome component PUP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	222	1685	1061	293	324	7	0	0	0
8	V	222	1685	1061	293	324	7	0	0	0

- Molecule 9 is a protein called Proteasome component PUP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	204	1581	1010	258	305	8	0	0	0
9	W	204	1581	1010	258	305	8	0	0	0

- Molecule 10 is a protein called Proteasome component C11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	198	1585	1005	269	305	6	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	X	198	1585	1005	269	305	6	0	0	0

- Molecule 11 is a protein called Proteasome component PRE2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	212	1644	1045	280	312	7	0	0	0
11	Y	212	1644	1045	280	312	7	0	0	0

- Molecule 12 is a protein called Proteasome component C5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	222	1757	1115	303	335	4	0	0	0
12	Z	222	1757	1115	303	335	4	0	0	0

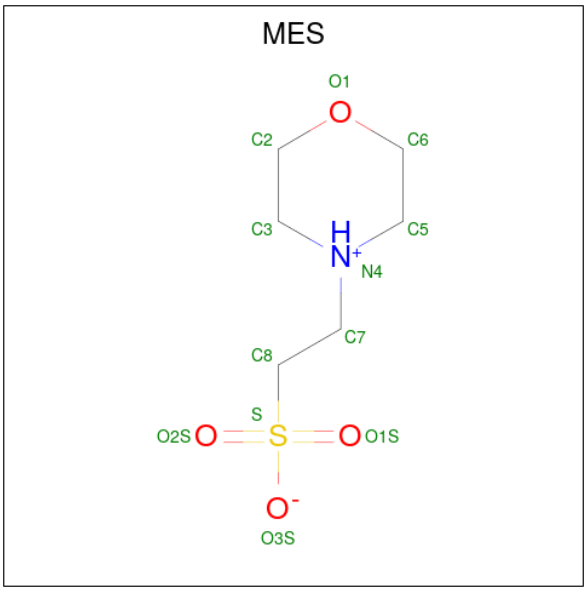
- Molecule 13 is a protein called Proteasome component PRE4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	233	1824	1154	312	351	7	0	0	0
13	1	233	1824	1154	312	351	7	0	0	0

- Molecule 14 is a protein called Proteasome component PRE3.

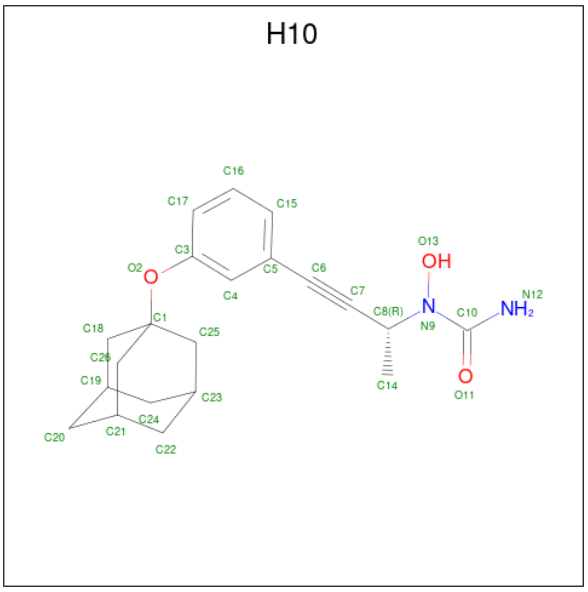
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
14	N	196	1512	955	250	300	7	0	0	0
14	2	196	1512	955	250	300	7	0	0	0

- Molecule 15 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
15	K	1	12	6	1	4	1	0	0
			Total	C	N	O	S		
15	Y	1	12	6	1	4	1	0	0

- Molecule 16 is 1-hydroxy-1-[(2R)-4-{3-[(3S,5S,7S)-tricyclo[3.3.1.1 3,7]dec-1-yloxy]phenyl}but-3-yn-2-yl]urea (three-letter code: H10) (formula: $C_{21}H_{26}N_2O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
16	K	1	26	21	2	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
16	Y	1	26	21	2	3	0	0

- Molecule 17 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
17	A	49	49	49	0	0
17	B	36	36	36	0	0
17	C	37	37	37	0	0
17	D	36	36	36	0	0
17	E	23	23	23	0	0
17	F	47	47	47	0	0
17	G	58	58	58	0	0
17	H	46	46	46	0	0
17	I	65	65	65	0	0
17	J	48	48	48	0	0
17	K	49	49	49	0	0
17	L	58	58	58	0	0
17	M	66	66	66	0	0
17	N	56	56	56	0	0
17	O	30	30	30	0	0
17	P	29	29	29	0	0
17	Q	24	24	24	0	0
17	R	30	30	30	0	0

Continued on next page...

Continued from previous page...

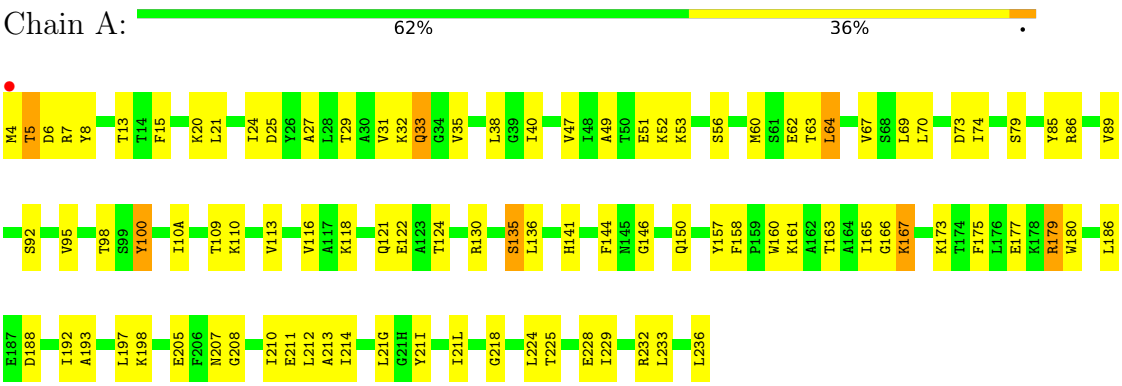
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
17	S	19	19	19	0	0
			Total	O		
17	T	38	38	38	0	0
			Total	O		
17	U	60	60	60	0	0
			Total	O		
17	V	42	42	42	0	0
			Total	O		
17	W	57	57	57	0	0
			Total	O		
17	X	48	48	48	0	0
			Total	O		
17	Y	50	50	50	0	0
			Total	O		
17	Z	50	50	50	0	0
			Total	O		
17	1	69	69	69	0	0
			Total	O		
17	2	61	61	61	0	0

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.

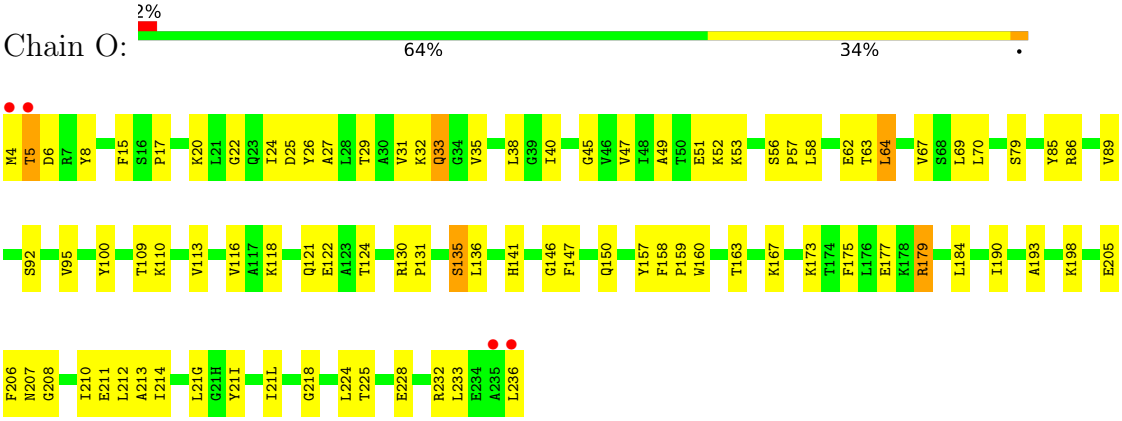
• Molecule 1: Proteasome component Y7

Chain A:



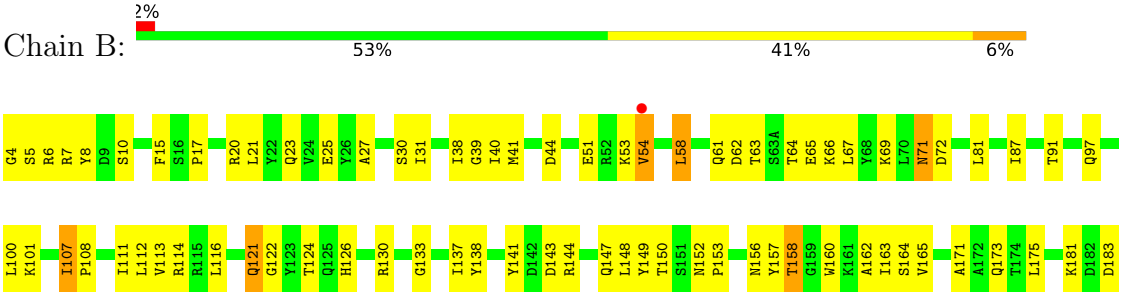
• Molecule 1: Proteasome component Y7

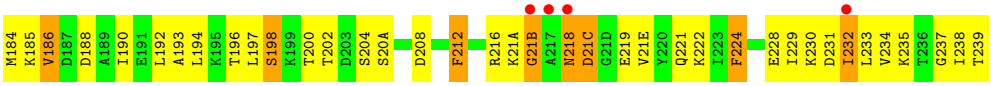
Chain O:



• Molecule 2: Proteasome component Y13

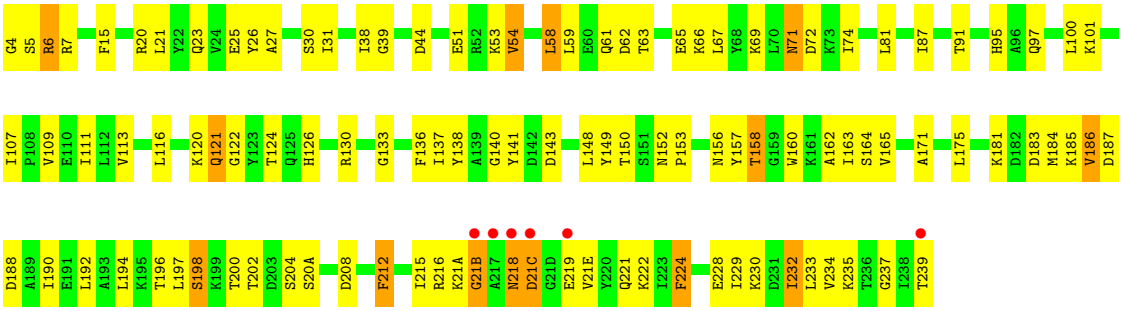
Chain B:





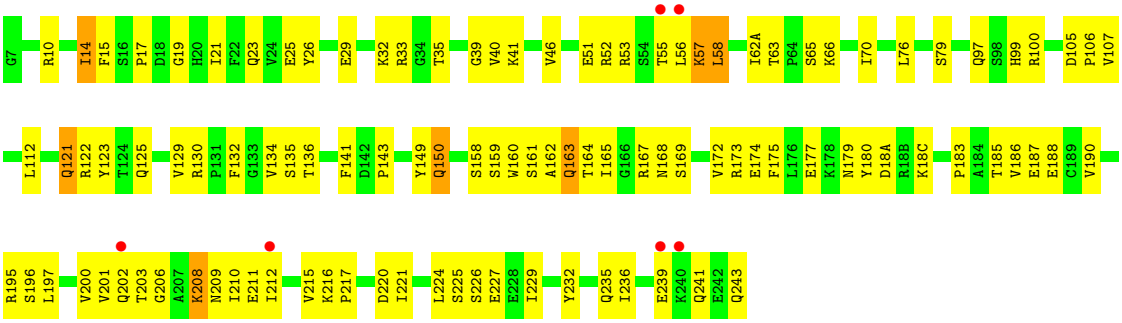
• Molecule 2: Proteasome component Y13

Chain P:



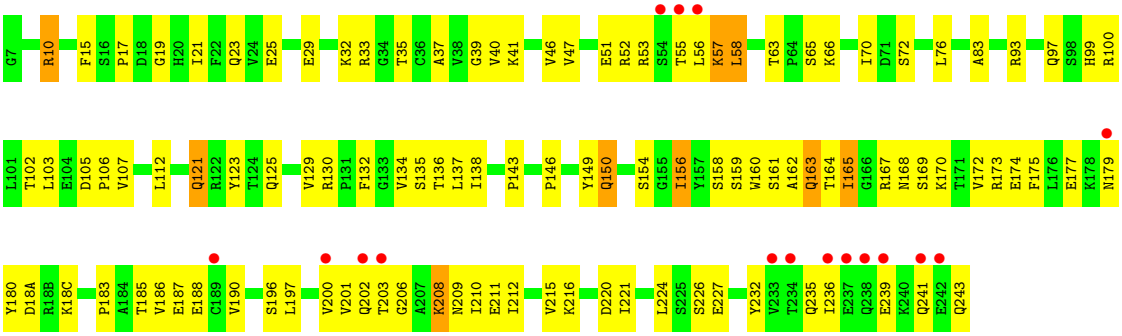
• Molecule 3: Proteasome component PRE6

Chain C:



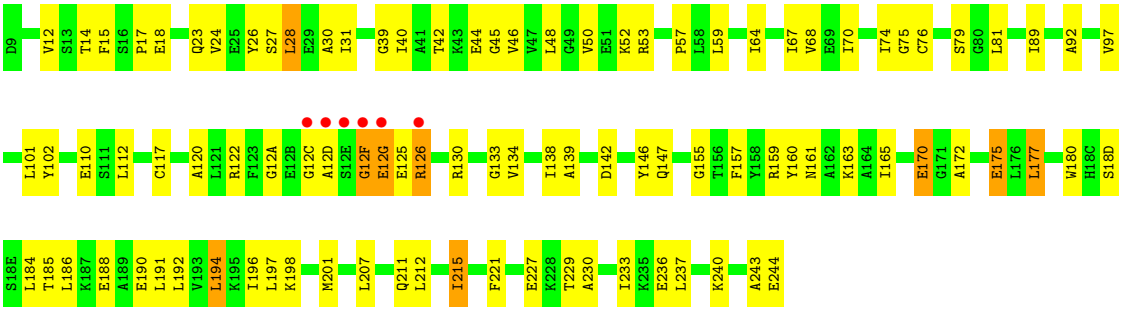
• Molecule 3: Proteasome component PRE6

Chain Q:



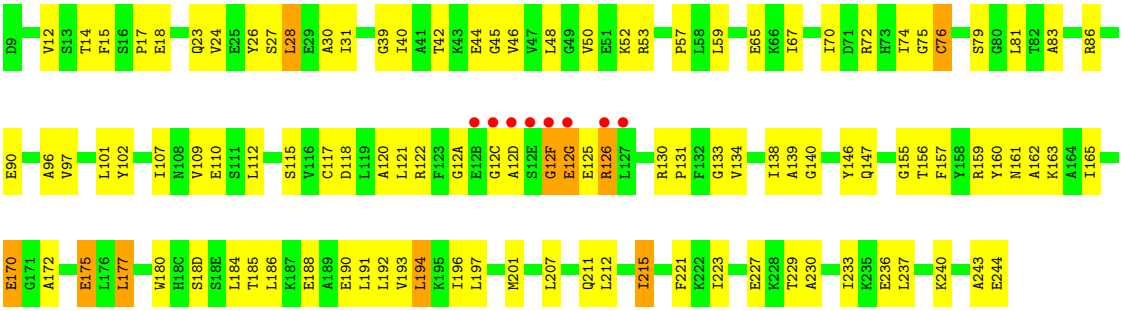
• Molecule 4: Proteasome component PUP2

Chain D:



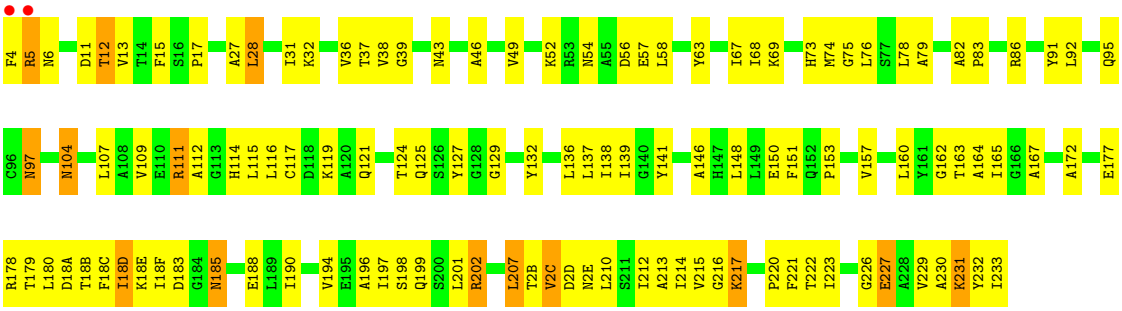
- Molecule 4: Proteasome component PUP2

Chain R:



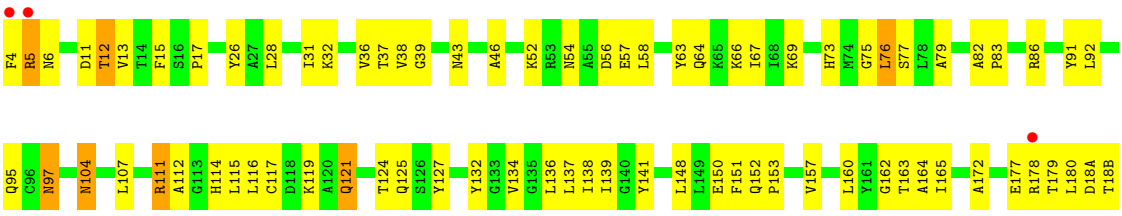
- Molecule 5: Proteasome component PRE5

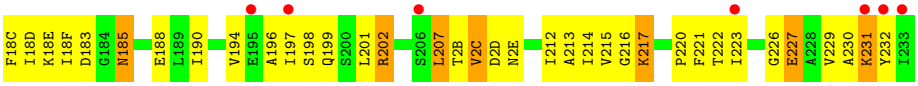
Chain E:



- Molecule 5: Proteasome component PRE5

Chain S:



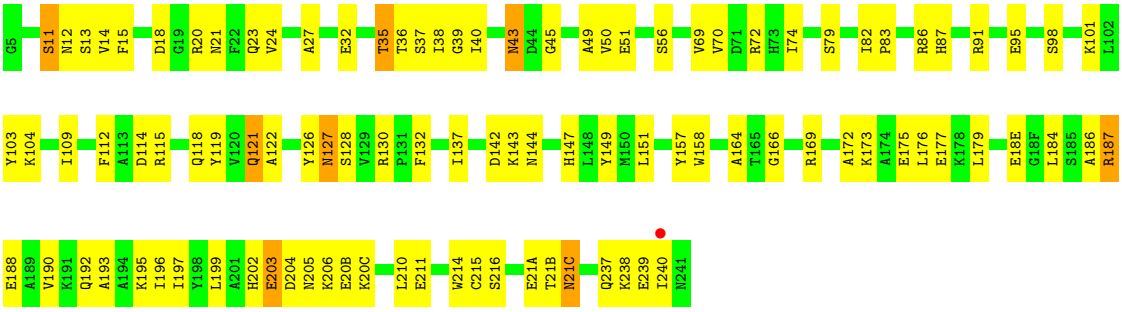


- Molecule 6: Proteasome component C1

Chain F:

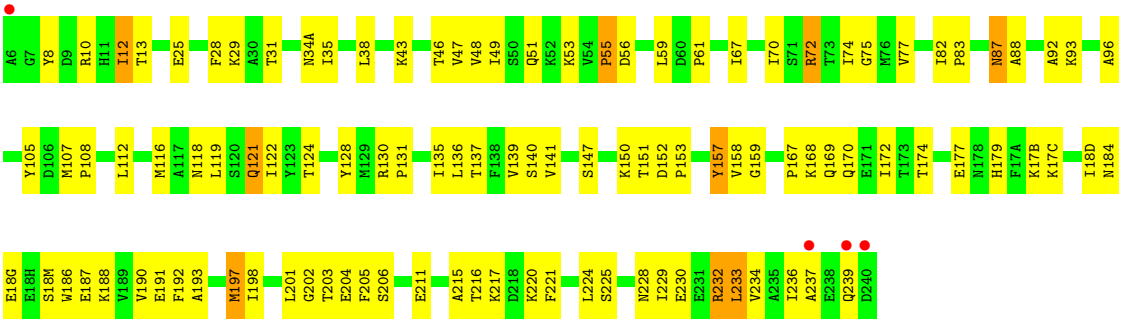
- Molecule 6: Proteasome component C1

Chain T:



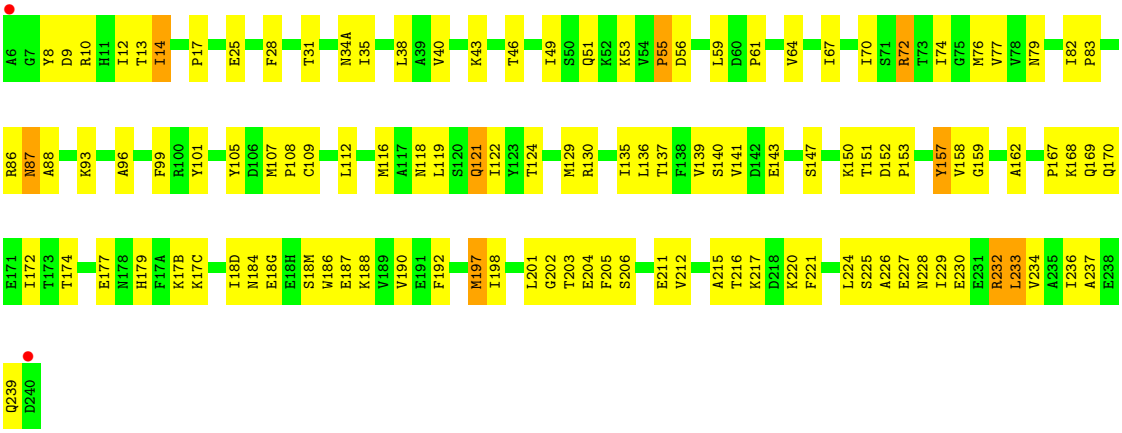
- Molecule 7: Proteasome component C7-alpha

Chain G:



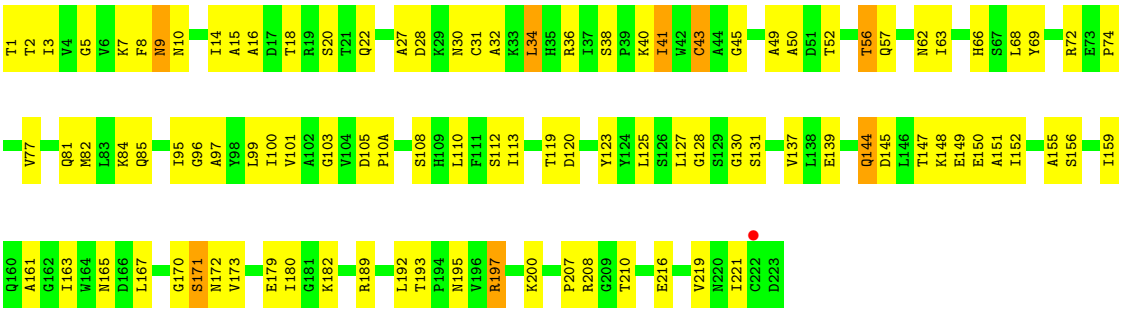
- Molecule 7: Proteasome component C7-alpha

Chain U:



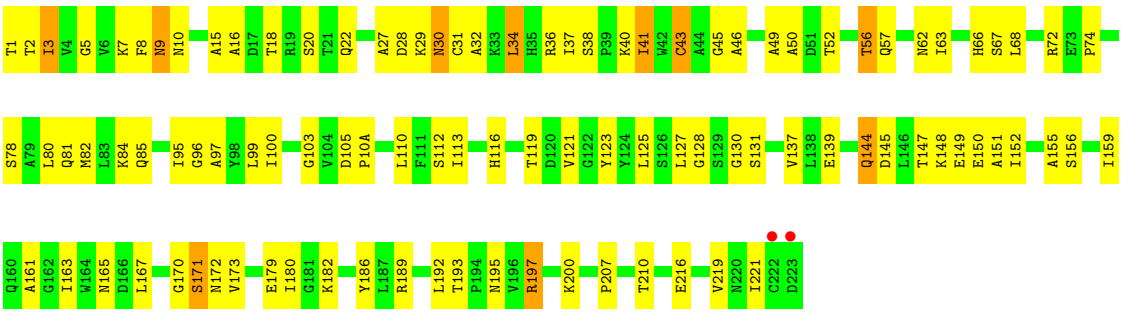
• Molecule 8: Proteasome component PUP1

Chain H:



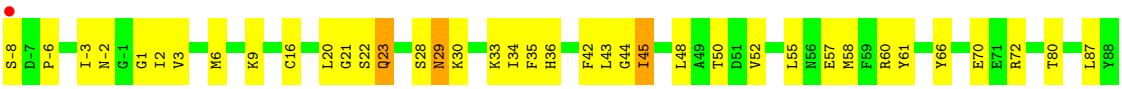
• Molecule 8: Proteasome component PUP1

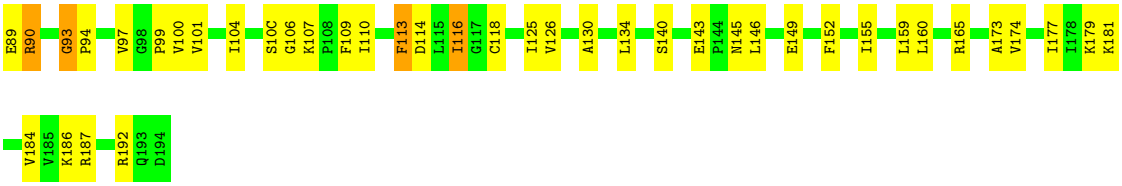
Chain V:



• Molecule 9: Proteasome component PUP3

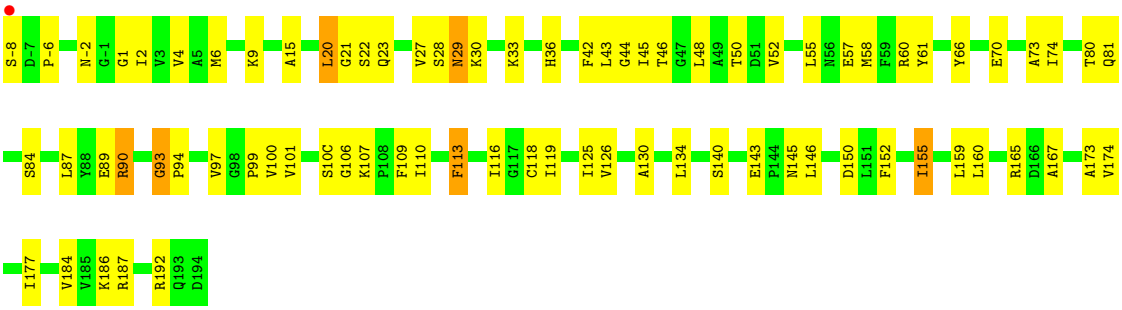
Chain I:





• Molecule 9: Proteasome component PUP3

Chain W:

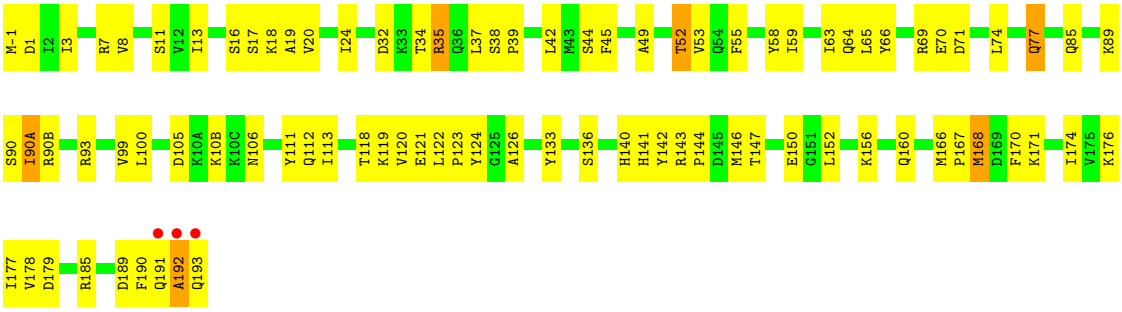


• Molecule 10: Proteasome component C11

Chain J:

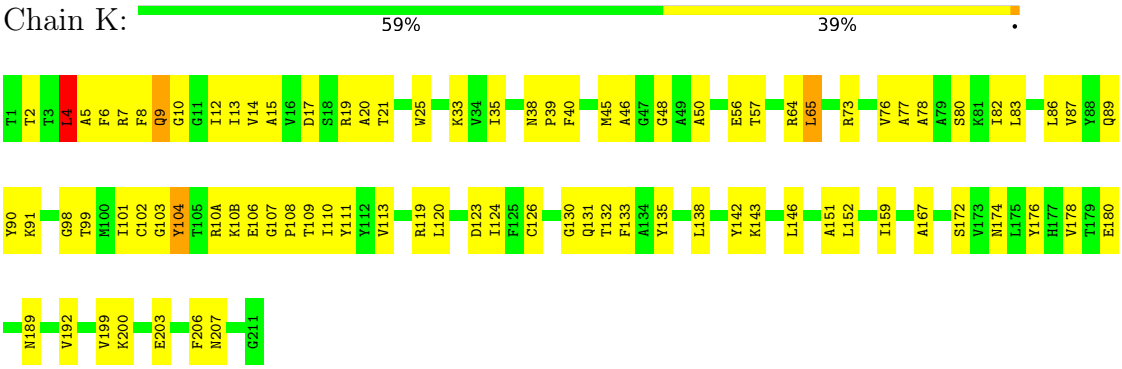
• Molecule 10: Proteasome component C11

Chain X:



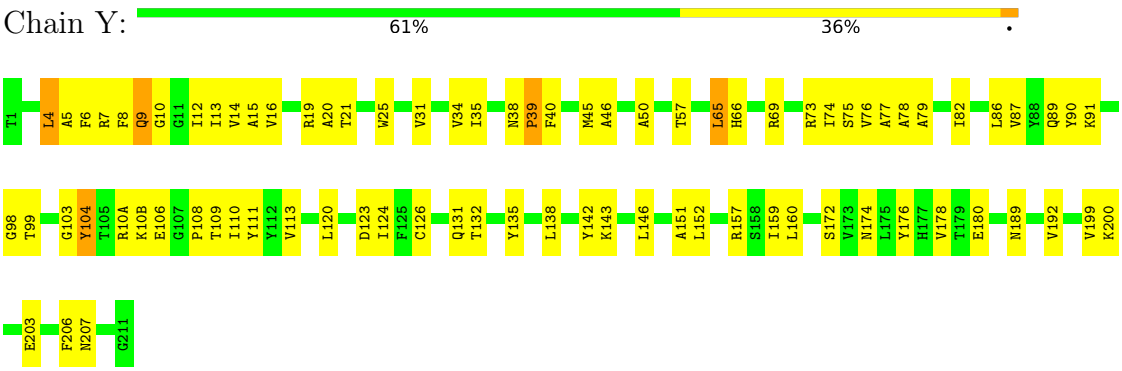
• Molecule 11: Proteasome component PRE2

Chain K:



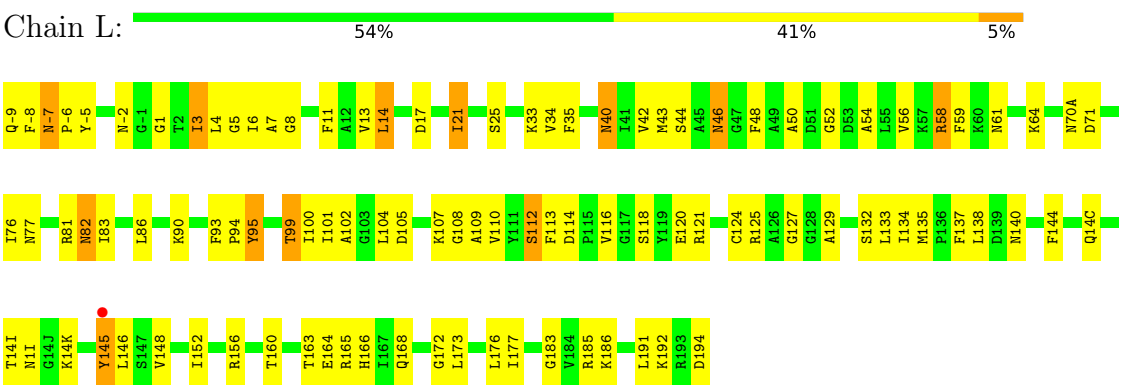
- Molecule 11: Proteasome component PRE2

Chain Y:



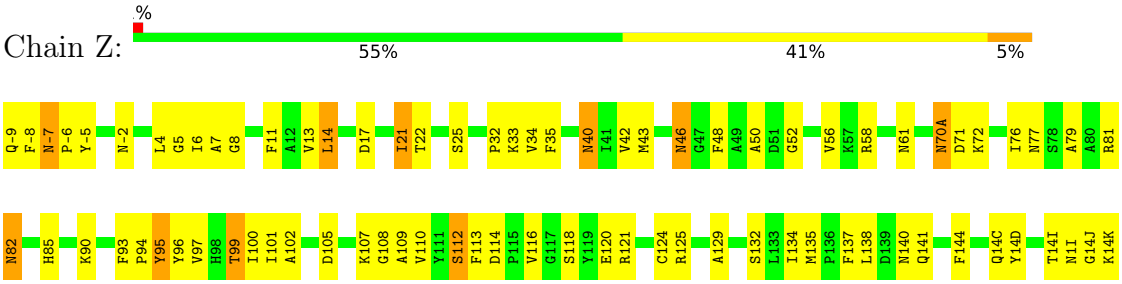
- Molecule 12: Proteasome component C5

Chain L:



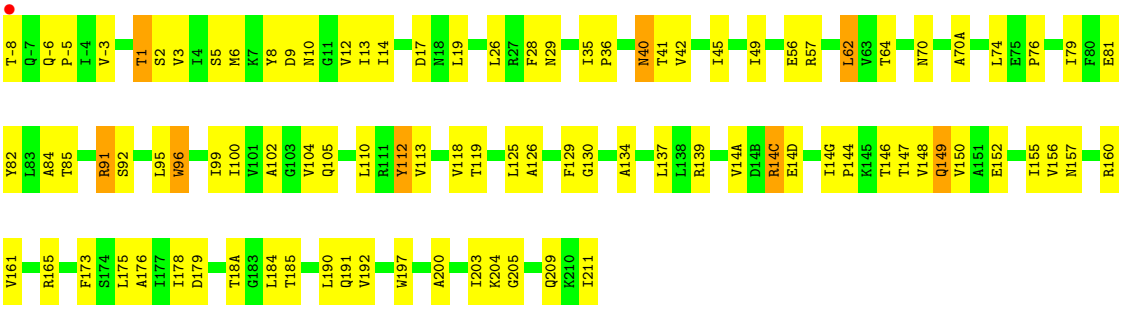
- Molecule 12: Proteasome component C5

Chain Z:



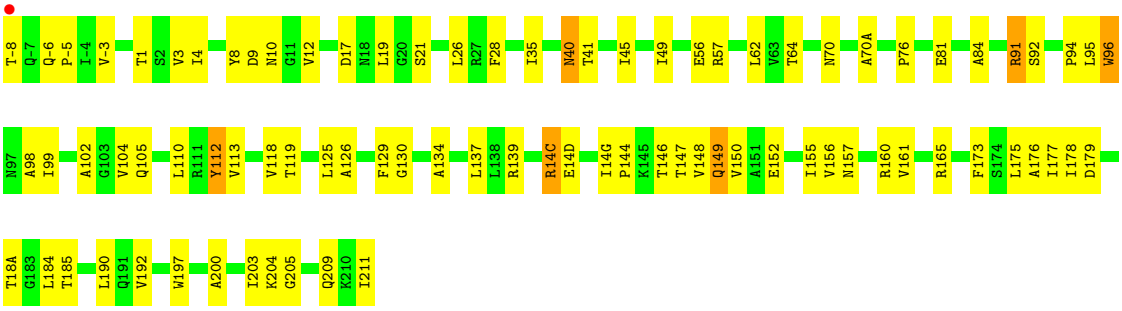
• Molecule 13: Proteasome component PRE4

Chain M:



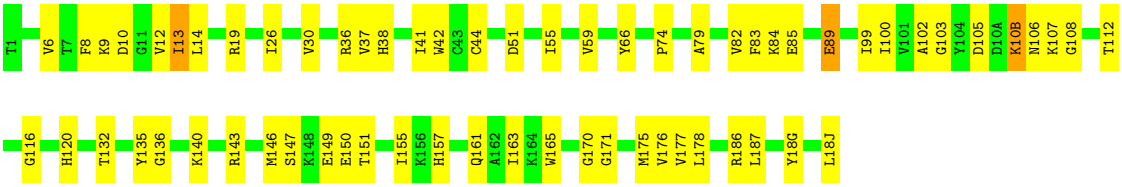
• Molecule 13: Proteasome component PRE4

Chain 1:



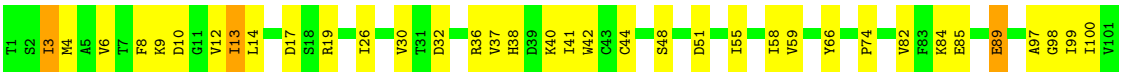
• Molecule 14: Proteasome component PRE3

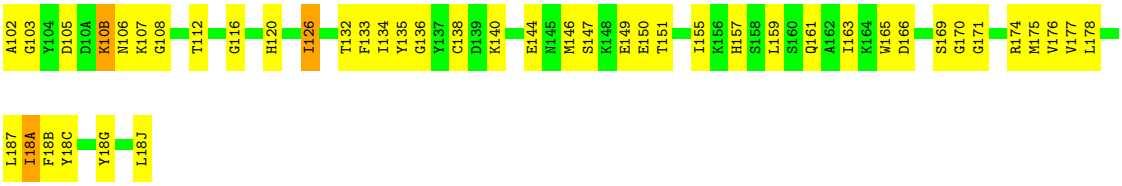
Chain N:



• Molecule 14: Proteasome component PRE3

Chain 2:




A102
G103
Y104
D105
D10A
K10B
N106
K107
G108
T112
G116
H120
I126
T132
F133
I134
Y135
G136
Y137
C138
D139
K140
E144
N145
M146
S147
K148
E149
E150
T151
I155
K156
H157
S158
L159
S160
Q161
A162
I163
K164
W165
D166
S169
G170
G171
R174
M175
V176
V177
L178
L187
I18A
F18B
Y18C
Y18G
L18J

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	134.12Å 300.56Å 144.06Å 90.00° 112.76° 90.00°	Depositor
Resolution (Å)	49.77 – 2.80 49.77 – 2.80	Depositor EDS
% Data completeness (in resolution range)	92.3 (49.77-2.80) 92.3 (49.77-2.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.49 (at 2.81Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.238 , 0.266 0.222 , 0.247	Depositor DCC
R_{free} test set	11760 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	48.1	Xtriage
Anisotropy	0.853	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 59.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	50905	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.41% 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: H10, MES

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	0.37	0/1952	0.63	0/2642
1	O	0.38	0/1952	0.64	0/2642
2	B	0.37	0/1935	0.63	0/2618
2	P	0.38	0/1935	0.64	0/2618
3	C	0.36	0/1920	0.62	0/2598
3	Q	0.36	0/1920	0.62	0/2598
4	D	0.36	0/1887	0.64	0/2541
4	R	0.36	0/1887	0.64	0/2541
5	E	0.35	0/1823	0.60	0/2463
5	S	0.36	0/1823	0.61	0/2463
6	F	0.37	0/1937	0.61	0/2614
6	T	0.37	0/1937	0.62	0/2614
7	G	0.40	0/1959	0.63	0/2652
7	U	0.40	0/1959	0.63	0/2652
8	H	0.38	0/1716	0.66	0/2326
8	V	0.38	0/1716	0.66	0/2326
9	I	0.38	0/1611	0.67	0/2174
9	W	0.40	0/1611	0.68	0/2174
10	J	0.39	0/1613	0.66	0/2173
10	X	0.40	0/1613	0.66	0/2173
11	K	0.40	0/1681	0.66	1/2274 (0.0%)
11	Y	0.40	0/1681	0.65	0/2274
12	L	0.39	0/1795	0.67	1/2420 (0.0%)
12	Z	0.38	0/1795	0.67	1/2420 (0.0%)
13	1	0.39	0/1855	0.67	2/2514 (0.1%)
13	M	0.38	0/1855	0.66	1/2514 (0.0%)
14	2	0.40	0/1541	0.65	0/2087
14	N	0.40	0/1541	0.66	0/2087
All	All	0.38	0/50450	0.64	6/68192 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	M	95	LEU	N-CA-C	-5.85	95.22	111.00
13	1	95	LEU	N-CA-C	-5.76	95.44	111.00
12	L	95	TYR	N-CA-C	-5.49	96.17	111.00
12	Z	95	TYR	N-CA-C	-5.22	96.90	111.00
11	K	4	LEU	CA-CB-CG	5.07	126.95	115.30

There are no chirality outliers.

There are no planarity outliers.

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	1915	0	1926	99	0
1	O	1915	0	1926	89	0
2	B	1905	0	1901	117	0
2	P	1905	0	1901	116	0
3	C	1891	0	1900	121	0
3	Q	1891	0	1900	128	0
4	D	1862	0	1836	89	0
4	R	1862	0	1836	93	0
5	E	1795	0	1797	127	0
5	S	1795	0	1797	119	0
6	F	1897	0	1886	113	0
6	T	1897	0	1886	103	0
7	G	1921	0	1910	109	0
7	U	1921	0	1910	116	0
8	H	1685	0	1688	94	0
8	V	1685	0	1688	103	0
9	I	1581	0	1574	86	0
9	W	1581	0	1574	80	0
10	J	1585	0	1590	91	0
10	X	1585	0	1590	98	0
11	K	1644	0	1595	85	0
11	Y	1644	0	1595	92	0
12	L	1757	0	1711	105	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	Z	1757	0	1711	102	0
13	1	1824	0	1832	81	0
13	M	1824	0	1832	87	0
14	2	1512	0	1481	80	0
14	N	1512	0	1481	61	0
15	K	12	0	13	0	0
15	Y	12	0	13	0	0
16	K	26	0	26	4	0
16	Y	26	0	26	7	0
17	1	69	0	0	9	0
17	2	61	0	0	5	0
17	A	49	0	0	3	0
17	B	36	0	0	3	0
17	C	37	0	0	6	0
17	D	36	0	0	4	0
17	E	23	0	0	5	0
17	F	47	0	0	4	0
17	G	58	0	0	5	0
17	H	46	0	0	2	0
17	I	65	0	0	3	0
17	J	48	0	0	4	0
17	K	49	0	0	4	0
17	L	58	0	0	5	0
17	M	66	0	0	5	0
17	N	56	0	0	4	0
17	O	30	0	0	2	0
17	P	29	0	0	4	0
17	Q	24	0	0	5	0
17	R	30	0	0	4	0
17	S	19	0	0	1	0
17	T	38	0	0	4	0
17	U	60	0	0	7	0
17	V	42	0	0	4	0
17	W	57	0	0	6	0
17	X	48	0	0	8	0
17	Y	50	0	0	3	0
17	Z	50	0	0	11	0
All	All	50905	0	49332	2529	0

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 26.

The worst 5 of 2529 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)
8:H:15:ALA:HB3	8:H:159:ILE:HD11	1.25	1.17
2:P:74:ILE:HD11	2:P:109:VAL:HG22	1.12	1.10
7:U:9:ASP:HA	7:U:14:ILE:HD11	1.25	1.10
2:B:202:THR:HG22	2:B:204:SER:H	1.13	1.09
8:H:50:ALA:HB2	9:I:116:ILE:HD13	1.33	1.09

There are no symmetry-related clashes.

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	248/250 (99%)	226 (91%)	15 (6%)	7 (3%)	5	17
1	O	248/250 (99%)	226 (91%)	14 (6%)	8 (3%)	4	13
2	B	242/244 (99%)	217 (90%)	18 (7%)	7 (3%)	4	15
2	P	242/244 (99%)	217 (90%)	19 (8%)	6 (2%)	5	19
3	C	239/241 (99%)	219 (92%)	15 (6%)	5 (2%)	7	23
3	Q	239/241 (99%)	217 (91%)	17 (7%)	5 (2%)	7	23
4	D	240/242 (99%)	217 (90%)	18 (8%)	5 (2%)	7	23
4	R	240/242 (99%)	216 (90%)	19 (8%)	5 (2%)	7	23
5	E	231/233 (99%)	201 (87%)	24 (10%)	6 (3%)	5	18
5	S	231/233 (99%)	203 (88%)	22 (10%)	6 (3%)	5	18
6	F	242/244 (99%)	224 (93%)	17 (7%)	1 (0%)	34	66
6	T	242/244 (99%)	223 (92%)	18 (7%)	1 (0%)	34	66
7	G	241/243 (99%)	223 (92%)	15 (6%)	3 (1%)	13	39
7	U	241/243 (99%)	222 (92%)	16 (7%)	3 (1%)	13	39
8	H	220/222 (99%)	207 (94%)	10 (4%)	3 (1%)	11	34
8	V	220/222 (99%)	205 (93%)	12 (6%)	3 (1%)	11	34

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	I	202/204 (99%)	189 (94%)	10 (5%)	3 (2%)	10	33
9	W	202/204 (99%)	189 (94%)	11 (5%)	2 (1%)	15	44
10	J	196/198 (99%)	184 (94%)	9 (5%)	3 (2%)	10	33
10	X	196/198 (99%)	184 (94%)	9 (5%)	3 (2%)	10	33
11	K	210/212 (99%)	197 (94%)	13 (6%)	0	100	100
11	Y	210/212 (99%)	200 (95%)	9 (4%)	1 (0%)	29	61
12	L	220/222 (99%)	201 (91%)	18 (8%)	1 (0%)	29	61
12	Z	220/222 (99%)	201 (91%)	18 (8%)	1 (0%)	29	61
13	1	231/233 (99%)	208 (90%)	21 (9%)	2 (1%)	17	46
13	M	231/233 (99%)	210 (91%)	19 (8%)	2 (1%)	17	46
14	2	194/196 (99%)	183 (94%)	11 (6%)	0	100	100
14	N	194/196 (99%)	180 (93%)	14 (7%)	0	100	100
All	All	6312/6368 (99%)	5789 (92%)	431 (7%)	92 (2%)	10	33

5 of 92 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	THR
1	A	167	LYS
2	B	54	VAL
3	C	58	LEU
3	C	203	THR

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	209/209 (100%)	204 (98%)	5 (2%)	49	81
1	O	209/209 (100%)	204 (98%)	5 (2%)	49	81
2	B	203/203 (100%)	188 (93%)	15 (7%)	13	37
2	P	203/203 (100%)	188 (93%)	15 (7%)	13	37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	213/213 (100%)	204 (96%)	9 (4%)	30	63
3	Q	213/213 (100%)	203 (95%)	10 (5%)	26	59
4	D	198/198 (100%)	186 (94%)	12 (6%)	18	48
4	R	198/198 (100%)	186 (94%)	12 (6%)	18	48
5	E	192/192 (100%)	172 (90%)	20 (10%)	7	21
5	S	192/192 (100%)	174 (91%)	18 (9%)	8	26
6	F	201/201 (100%)	188 (94%)	13 (6%)	17	44
6	T	201/201 (100%)	188 (94%)	13 (6%)	17	44
7	G	207/207 (100%)	197 (95%)	10 (5%)	25	58
7	U	207/207 (100%)	197 (95%)	10 (5%)	25	58
8	H	181/181 (100%)	173 (96%)	8 (4%)	28	61
8	V	181/181 (100%)	170 (94%)	11 (6%)	18	48
9	I	172/172 (100%)	165 (96%)	7 (4%)	30	64
9	W	172/172 (100%)	166 (96%)	6 (4%)	36	70
10	J	175/175 (100%)	169 (97%)	6 (3%)	37	71
10	X	175/175 (100%)	169 (97%)	6 (3%)	37	71
11	K	169/169 (100%)	164 (97%)	5 (3%)	41	75
11	Y	169/169 (100%)	164 (97%)	5 (3%)	41	75
12	L	185/185 (100%)	172 (93%)	13 (7%)	15	40
12	Z	185/185 (100%)	173 (94%)	12 (6%)	17	44
13	1	199/199 (100%)	191 (96%)	8 (4%)	31	65
13	M	199/199 (100%)	191 (96%)	8 (4%)	31	65
14	2	162/162 (100%)	154 (95%)	8 (5%)	25	57
14	N	162/162 (100%)	157 (97%)	5 (3%)	40	74
All	All	5332/5332 (100%)	5057 (95%)	275 (5%)	23	55

5 of 275 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	W	29	ASN
10	X	77	GLN
13	1	62	LEU
9	I	116	ILE
9	I	45	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 203 such sidechains are listed below:

Mol	Chain	Res	Type
3	Q	121	GLN
6	T	90	ASN
14	2	38	HIS
3	Q	209	ASN
5	S	64	GLN

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)

4 ligands are modelled in this entry.

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	$\#\|Z\| > 2$
15	MES	Y	1(L)	-	12,12,12	2.74	6 (50%)	14,16,16	3.16	5 (35%)
15	MES	K	1(L)	-	12,12,12	2.75	8 (66%)	14,16,16	3.19	6 (42%)
16	H10	K	2(I)	-	27,29,29	3.41	18 (66%)	36,43,43	1.05	3 (8%)
16	H10	Y	2(I)	-	27,29,29	3.32	18 (66%)	36,43,43	1.01	2 (5%)

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
15	MES	Y	1(L)	-	-	3/6/14/14	0/1/1/1
15	MES	K	1(L)	-	-	5/6/14/14	0/1/1/1
16	H10	K	2(I)	-	-	2/14/45/45	0/5/4/4
16	H10	Y	2(I)	-	-	2/14/45/45	0/5/4/4

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	K	2(I)	H10	O13-N9	-11.07	1.25	1.40
16	Y	2(I)	H10	O13-N9	-10.31	1.26	1.40
16	Y	2(I)	H10	C26-C1	5.58	1.62	1.52
15	Y	1(L)	MES	C8-S	5.23	1.84	1.77
16	K	2(I)	H10	C26-C1	5.22	1.61	1.52

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	Y	1(L)	MES	O1S-S-C8	-9.48	95.50	106.92
15	K	1(L)	MES	O1S-S-C8	-9.44	95.55	106.92
15	Y	1(L)	MES	O2S-S-C8	4.18	111.95	106.92
15	Y	1(L)	MES	O3S-S-O2S	3.87	120.72	111.27
15	K	1(L)	MES	O3S-S-O2S	3.84	120.65	111.27

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	K	1(L)	MES	C7-C8-S-O1S
15	K	1(L)	MES	C7-C8-S-O3S
15	Y	1(L)	MES	C7-C8-S-O1S
15	Y	1(L)	MES	C7-C8-S-O2S
15	Y	1(L)	MES	C7-C8-S-O3S

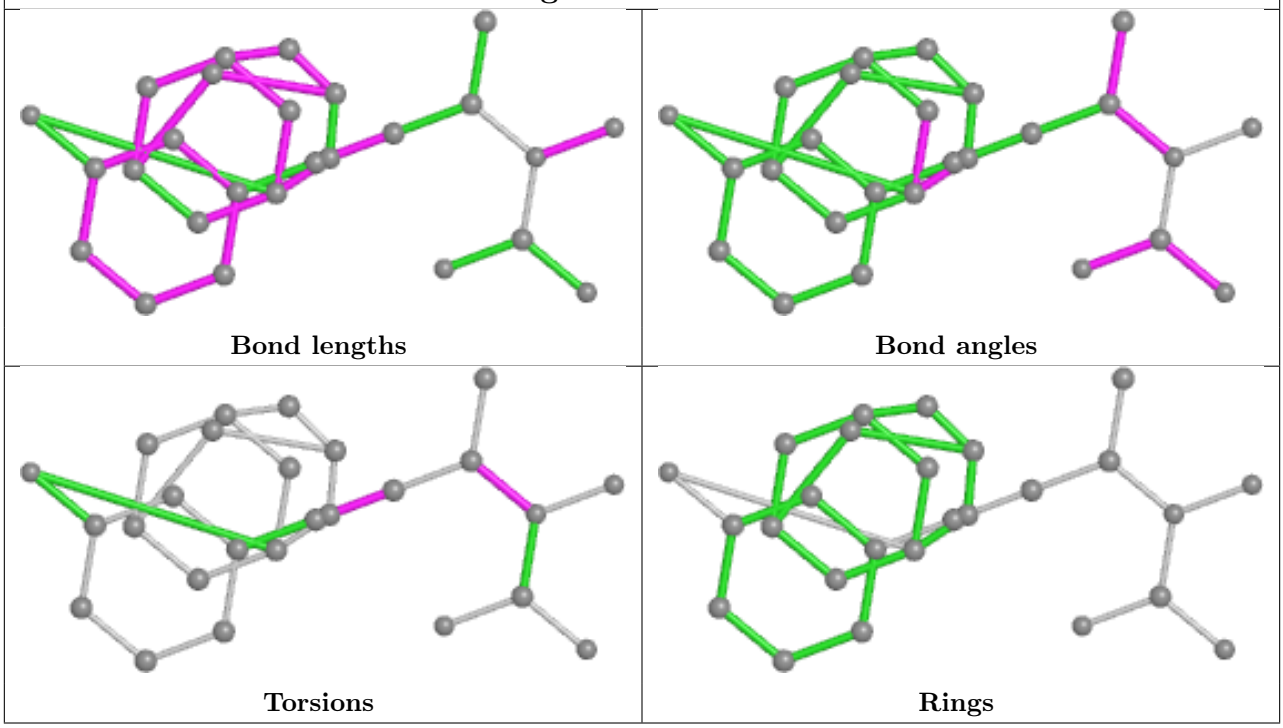
There are no ring outliers.

2 monomers are involved in 11 short contacts:

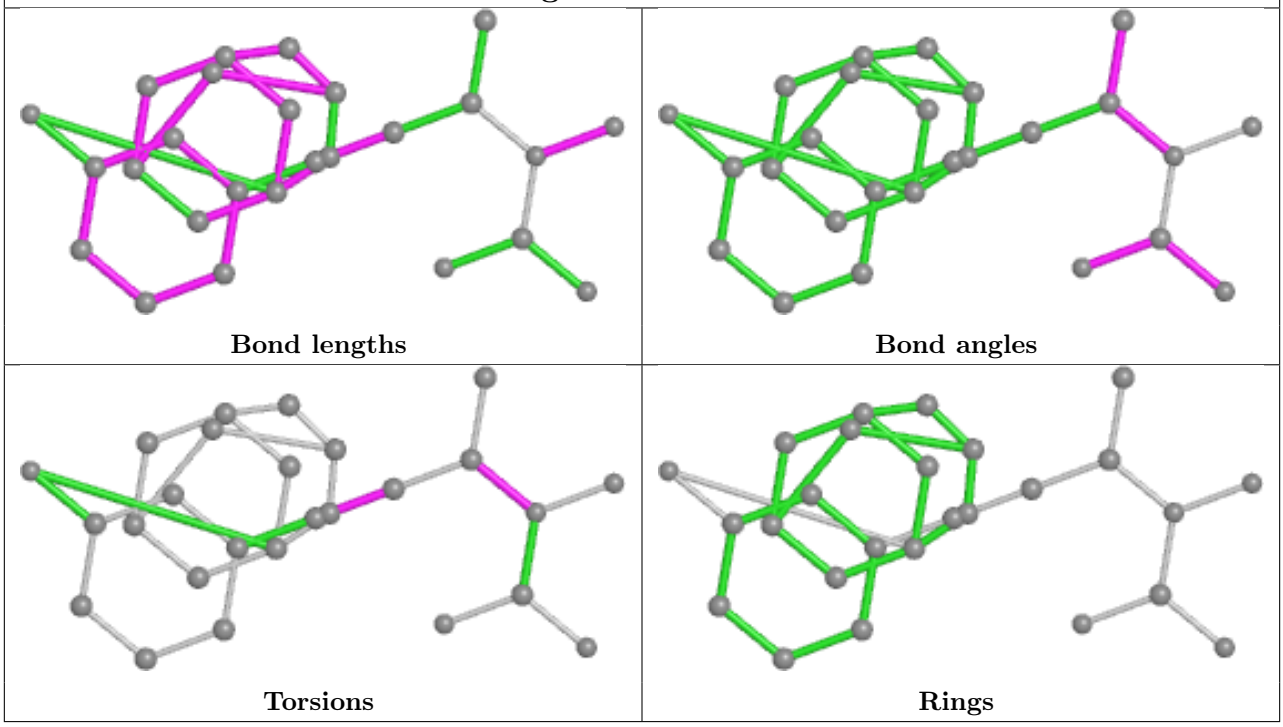
Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	K	2(I)	H10	4	0
16	Y	2(I)	H10	7	0

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 H10 K 2I



Ligand H10 Y 2I



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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	250/250 (100%)	-0.35	1 (0%)	92	91	47, 64, 92, 112	0
1	O	250/250 (100%)	-0.28	4 (1%)	72	66	48, 65, 94, 112	0
2	B	244/244 (100%)	-0.19	5 (2%)	65	56	47, 67, 103, 129	0
2	P	244/244 (100%)	-0.06	6 (2%)	57	47	46, 67, 103, 129	0
3	C	241/241 (100%)	-0.06	6 (2%)	57	47	50, 69, 118, 136	0
3	Q	241/241 (100%)	0.06	16 (6%)	18	11	51, 70, 118, 136	0
4	D	242/242 (100%)	-0.12	6 (2%)	57	47	49, 70, 102, 133	0
4	R	242/242 (100%)	-0.13	8 (3%)	46	36	48, 70, 103, 133	0
5	E	233/233 (100%)	-0.21	2 (0%)	84	80	52, 74, 98, 121	0
5	S	233/233 (100%)	0.02	10 (4%)	35	25	52, 75, 99, 121	0
6	F	244/244 (100%)	-0.24	2 (0%)	86	81	48, 65, 100, 112	0
6	T	244/244 (100%)	-0.24	1 (0%)	92	91	48, 65, 101, 113	0
7	G	243/243 (100%)	-0.37	4 (1%)	72	66	45, 61, 88, 120	0
7	U	243/243 (100%)	-0.32	2 (0%)	86	81	46, 62, 87, 119	0
8	H	222/222 (100%)	-0.39	1 (0%)	91	88	46, 59, 78, 110	0
8	V	222/222 (100%)	-0.41	2 (0%)	84	80	48, 61, 78, 109	0
9	I	204/204 (100%)	-0.36	1 (0%)	91	88	42, 58, 81, 95	0
9	W	204/204 (100%)	-0.21	1 (0%)	91	88	44, 58, 80, 96	0
10	J	198/198 (100%)	-0.46	2 (1%)	82	77	44, 58, 75, 128	0
10	X	198/198 (100%)	-0.43	3 (1%)	73	68	46, 58, 74, 128	0
11	K	212/212 (100%)	-0.43	0	100	100	42, 57, 73, 84	0
11	Y	212/212 (100%)	-0.45	0	100	100	44, 58, 73, 84	0
12	L	222/222 (100%)	-0.37	1 (0%)	91	88	42, 59, 81, 100	0
12	Z	222/222 (100%)	-0.37	2 (0%)	84	80	42, 60, 81, 99	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
13	1	233/233 (100%)	-0.53	1 (0%) 92 91	45, 57, 71, 80	0
13	M	233/233 (100%)	-0.47	1 (0%) 92 91	44, 58, 72, 79	0
14	2	196/196 (100%)	-0.48	0 100 100	44, 54, 75, 87	0
14	N	196/196 (100%)	-0.56	0 100 100	42, 53, 74, 88	0
All	All	6368/6368 (100%)	-0.29	88 (1%) 75 70	42, 62, 95, 136	0

The worst 5 of 88 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	12(D)	ALA	8.6
4	D	12(E)	SER	7.9
4	D	12(F)	GLY	7.6
4	D	12(C)	GLY	7.3
4	R	12(F)	GLY	6.5

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.

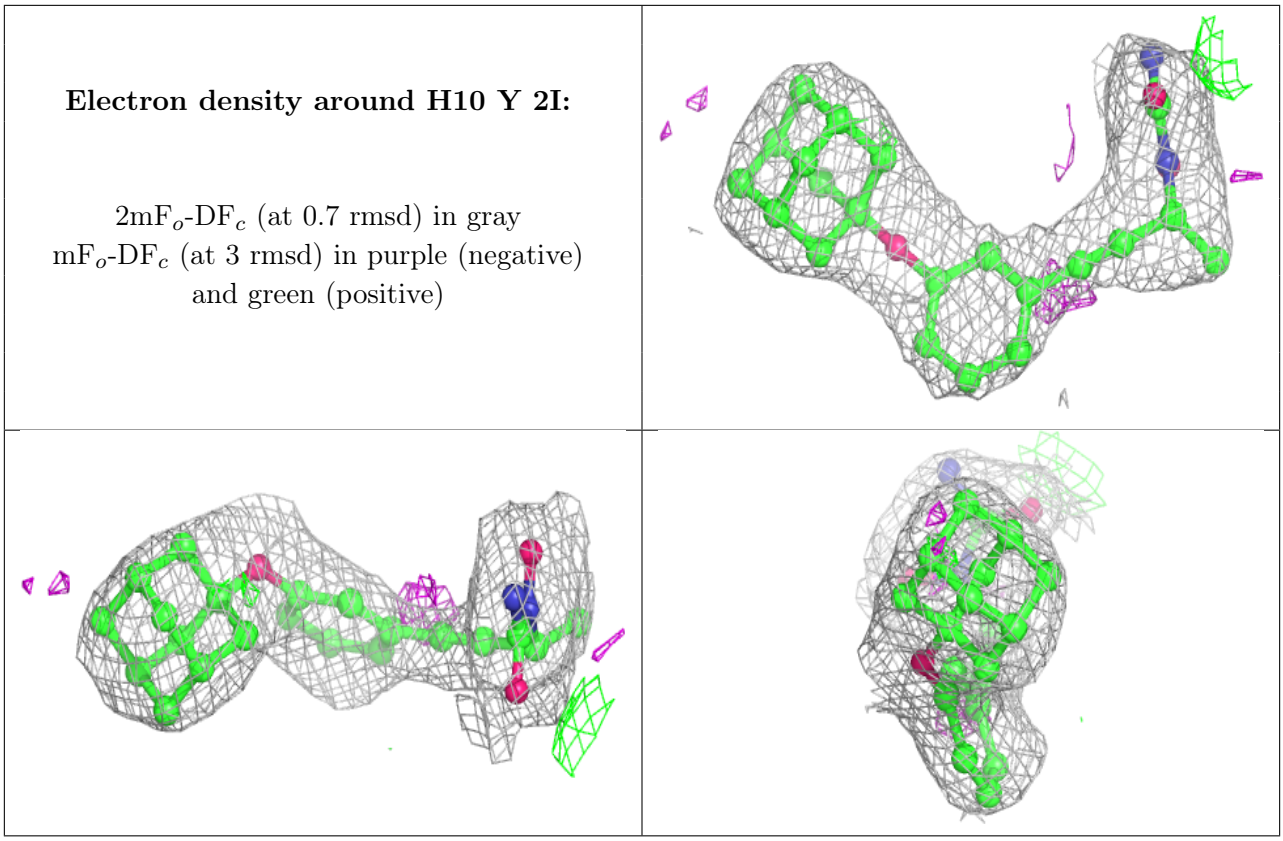
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
16	H10	Y	2(I)	26/26	0.94	0.18	46,51,60,62	0
16	H10	K	2(I)	26/26	0.96	0.20	46,50,57,58	0
15	MES	Y	1(L)	12/12	0.96	0.18	63,70,75,76	0
15	MES	K	1(L)	12/12	0.97	0.16	62,70,76,77	0

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.

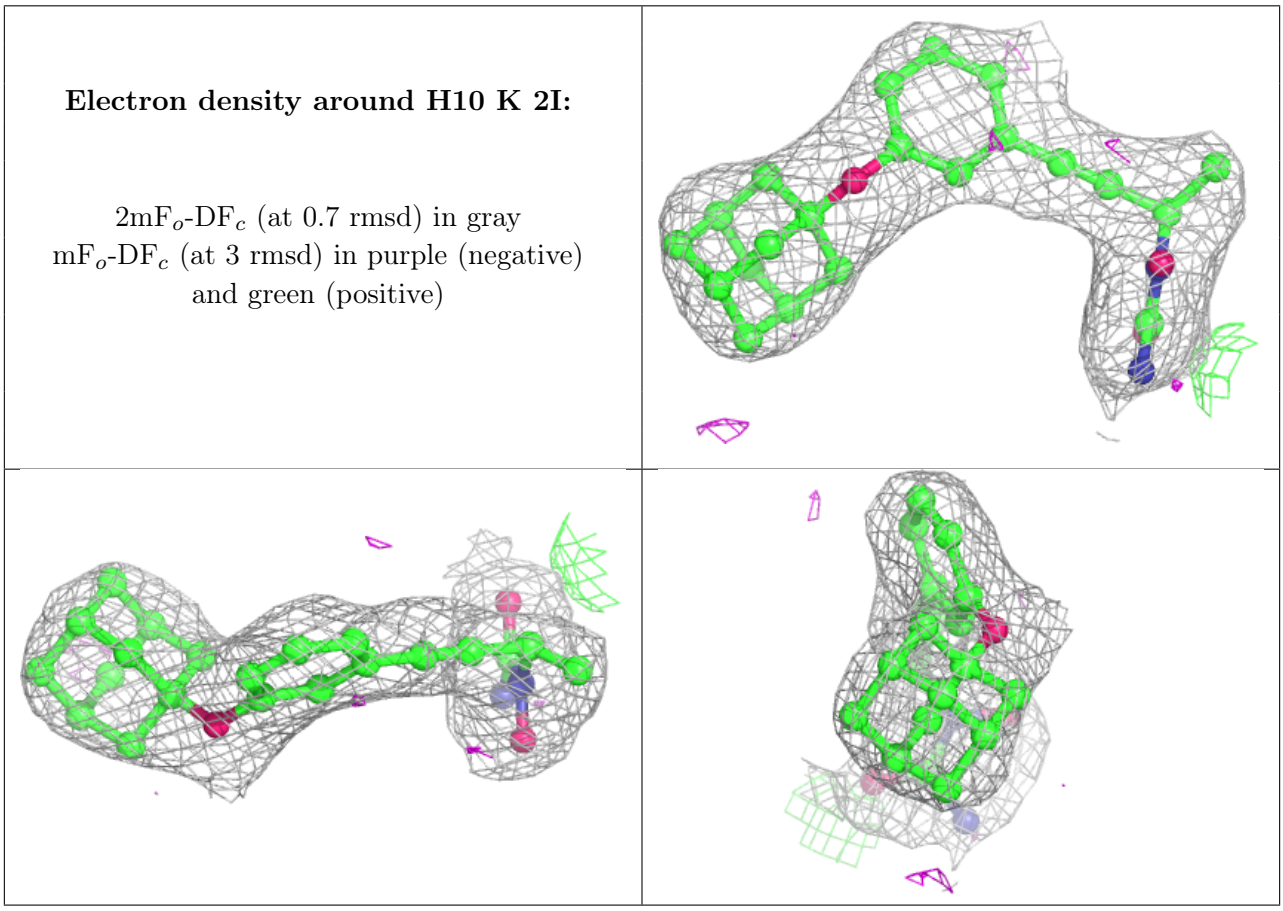
Electron density around H10 Y 2I:

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



Electron density around H10 K 2I:

$2mF_o$-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.