



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 24, 2023 – 01:59 PM EDT

PDB ID : 3E47  
Title : Crystal Structure of the Yeast 20S Proteasome in Complex with Homobelactosin C  
Authors : Groll, M.; Larionov, O.V.; de Meijere, A.  
Deposited on : 2008-08-10  
Resolution : 3.00 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](mailto: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 ⓘ 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 ⓘ](#)) 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 : 2.36  
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) : 2.36

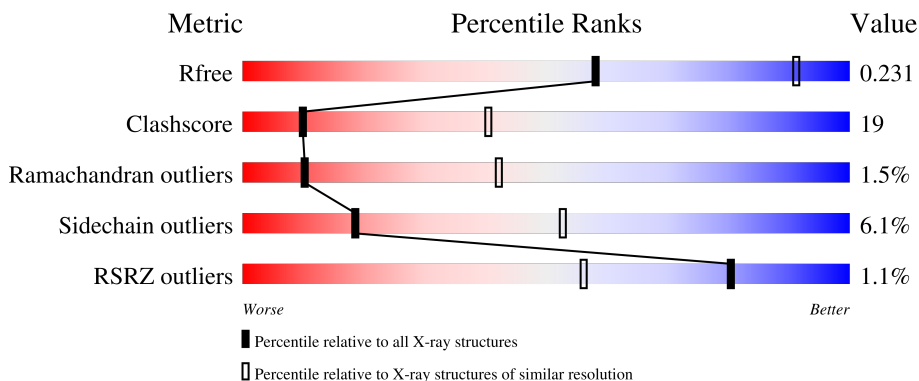
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	
1	O	250	
2	B	244	
2	P	244	
3	C	241	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	241	<p>3% 59% 39%</p>
4	D	242	<p>2% 69% 27%</p>
4	R	242	<p>3% 65% 31%</p>
5	E	233	<p>2% 48% 44% 7%</p>
5	S	233	<p>2% 45% 46% 7%</p>
6	F	244	<p>0% 63% 33% 5%</p>
6	T	244	<p>0% 61% 34% 5%</p>
7	G	243	<p>0% 62% 34% 5%</p>
7	U	243	<p>0% 60% 36%</p>
8	H	222	<p>0% 70% 26%</p>
8	V	222	<p>0% 71% 26%</p>
9	I	204	<p>0% 68% 28%</p>
9	W	204	<p>0% 67% 30%</p>
10	J	198	<p>0% 61% 37%</p>
10	X	198	<p>2% 58% 40%</p>
11	K	212	<p>0% 63% 34%</p>
11	Y	212	<p>0% 63% 35%</p>
12	L	222	<p>0% 66% 30%</p>
12	Z	222	<p>0% 65% 30% 5%</p>
13	1	233	<p>0% 64% 33%</p>
13	M	233	<p>0% 64% 33%</p>
14	2	196	<p>0% 62% 35%</p>
14	N	196	<p>0% 62% 35%</p>

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 50490 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
5	E	233	Total	C	N	O	S	0	0	0
			1795	1129	312	350	4			
5	S	233	Total	C	N	O	S	0	0	0
			1795	1129	312	350	4			

- Molecule 6 is a protein called Proteasome component C1.

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

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

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

- Molecule 8 is a protein called Proteasome component PUP1.

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

- Molecule 9 is a protein called Proteasome component PUP3.

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

- Molecule 10 is a protein called Proteasome component C11.

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

*Continued on next page...*

Continued from previous page...

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

- Molecule 11 is a protein called Proteasome component PRE2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	212	Total	C	N	O	S	0	0	0
			1643	1045	280	311	7			
11	Y	212	Total	C	N	O	S	0	0	0
			1643	1045	280	311	7			

- Molecule 12 is a protein called Proteasome component C5.

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

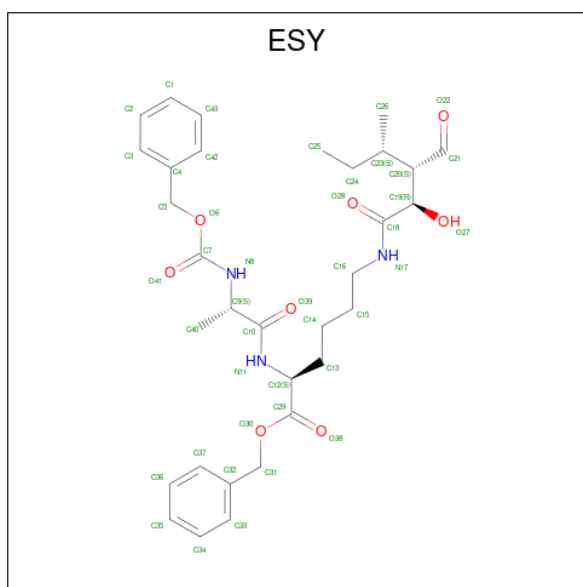
- Molecule 13 is a protein called Proteasome component PRE4.

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

- Molecule 14 is a protein called Proteasome component PRE3.

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

- Molecule 15 is benzyl N-[(benzyloxy)carbonyl]-L-alanyl-N 6 -[(2R,3S,4S)-3-formyl-2-hydroxy-y-4-methylhexanoyl]-L-lysinate (three-letter code: ESY) (formula: C<sub>32</sub>H<sub>43</sub>N<sub>3</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
15	K	1	Total	C	N	O	0	0
			43	32	3	8		
15	Y	1	Total	C	N	O	0	0
			43	32	3	8		

- Molecule 16 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	38	Total	O	0	0
			38	38		
16	B	31	Total	O	0	0
			31	31		
16	C	24	Total	O	0	0
			24	24		
16	D	25	Total	O	0	0
			25	25		
16	E	15	Total	O	0	0
			15	15		
16	F	27	Total	O	0	0
			27	27		
16	G	41	Total	O	0	0
			41	41		
16	H	28	Total	O	0	0
			28	28		
16	I	42	Total	O	0	0
			42	42		
16	J	32	Total	O	0	0
			32	32		

*Continued on next page...*

*Continued from previous page...*

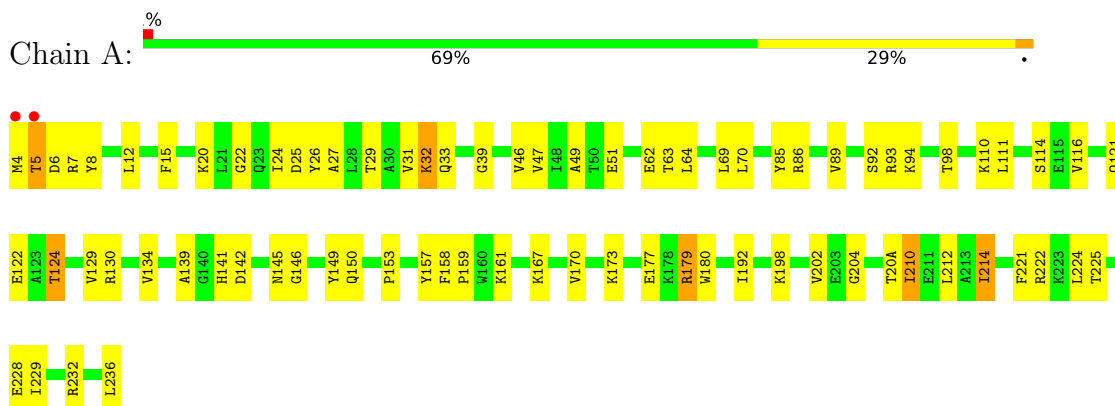
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	K	25	Total 25	O 25	0	0
16	L	35	Total 35	O 35	0	0
16	M	43	Total 43	O 43	0	0
16	N	38	Total 38	O 38	0	0
16	O	18	Total 18	O 18	0	0
16	P	20	Total 20	O 20	0	0
16	Q	17	Total 17	O 17	0	0
16	R	23	Total 23	O 23	0	0
16	S	16	Total 16	O 16	0	0
16	T	26	Total 26	O 26	0	0
16	U	38	Total 38	O 38	0	0
16	V	33	Total 33	O 33	0	0
16	W	39	Total 39	O 39	0	0
16	X	38	Total 38	O 38	0	0
16	Y	24	Total 24	O 24	0	0
16	Z	36	Total 36	O 36	0	0
16	1	45	Total 45	O 45	0	0
16	2	41	Total 41	O 41	0	0



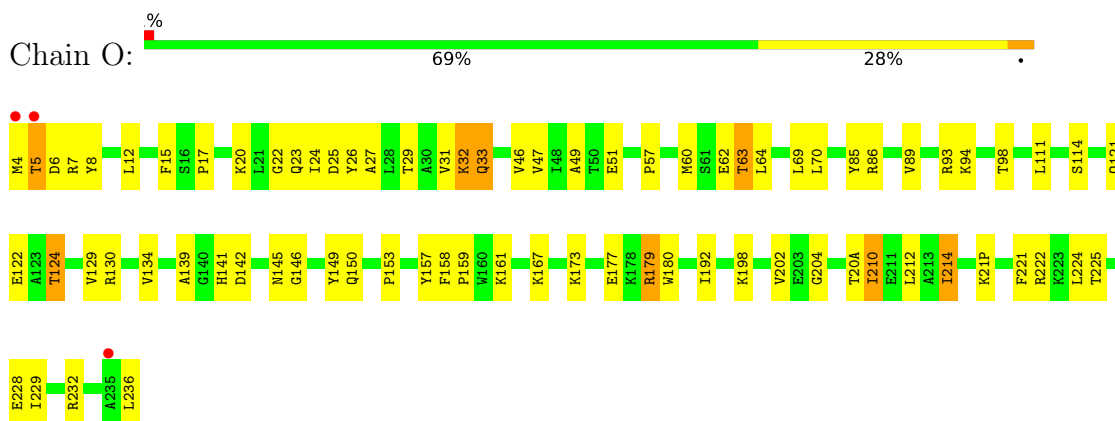
### 3 Residue-property plots [i](#)

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.

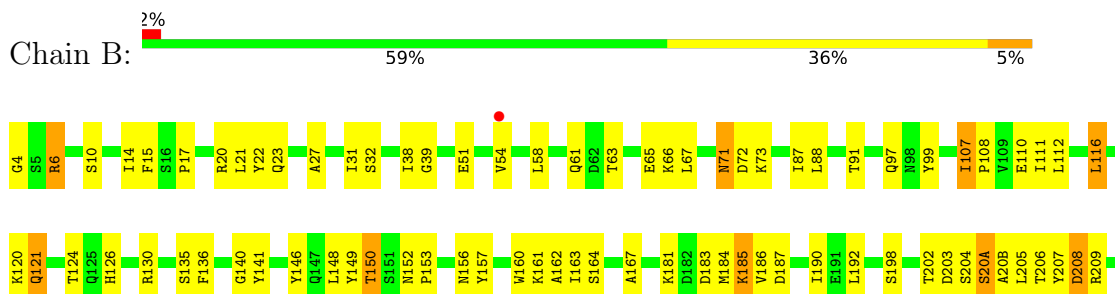
- Molecule 1: Proteasome component Y7



- Molecule 1: Proteasome component Y7



- Molecule 2: Proteasome component Y13

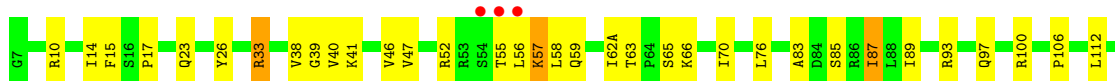




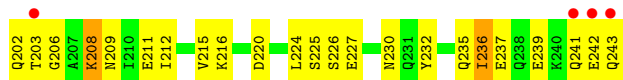
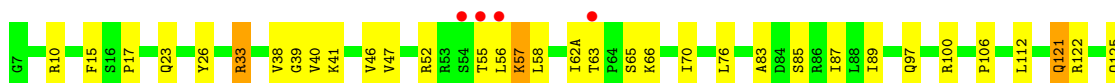
- Molecule 2: Proteasome component Y13



- Molecule 3: Proteasome component PRE6

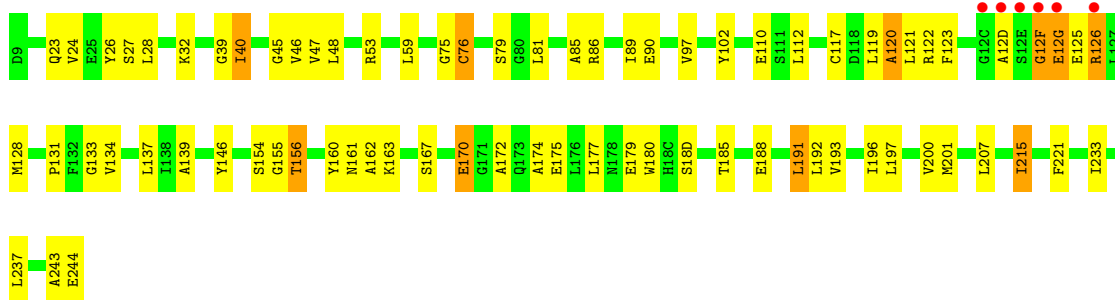


- Molecule 3: Proteasome component PRE6

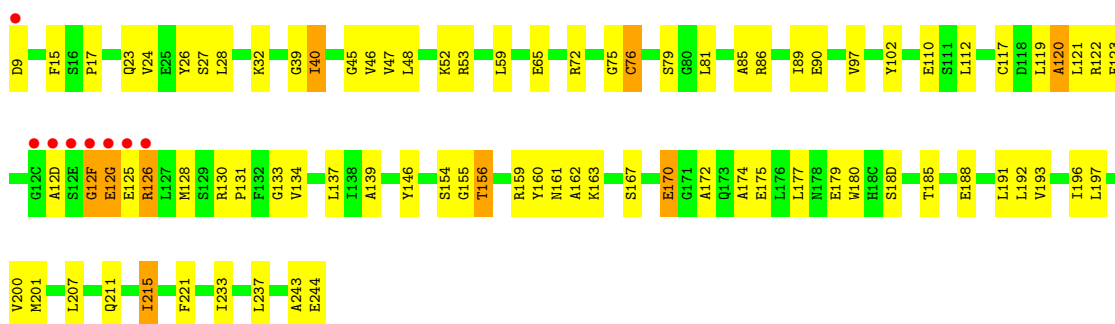


- Molecule 4: Proteasome component PUP2

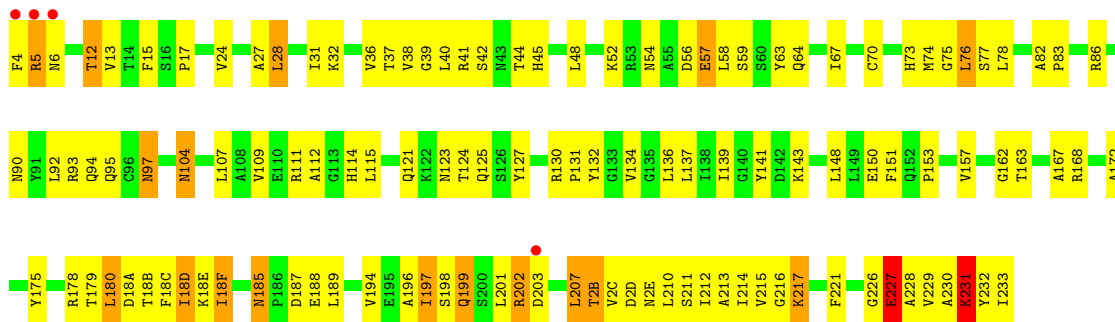




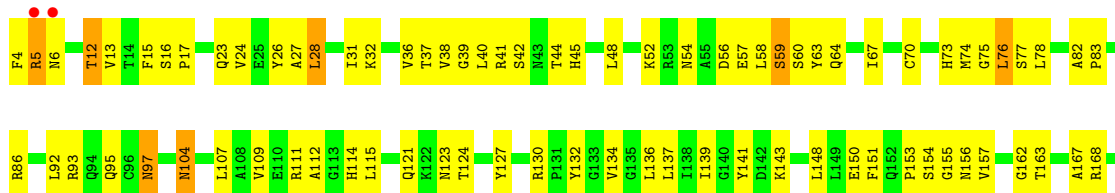
- Molecule 4: Proteasome component PUP2

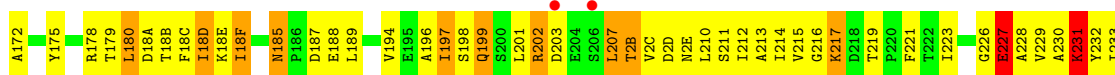


- Molecule 5: Proteasome component PRE5

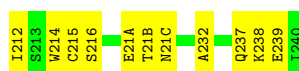
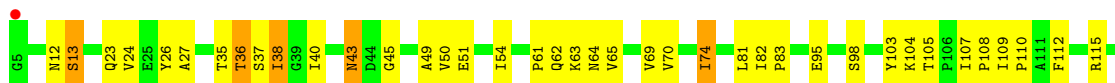


- Molecule 5: Proteasome component PRE5

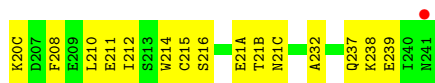
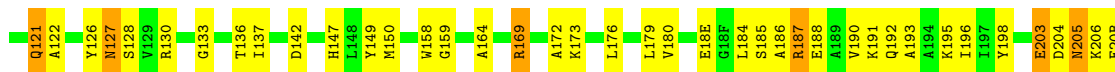
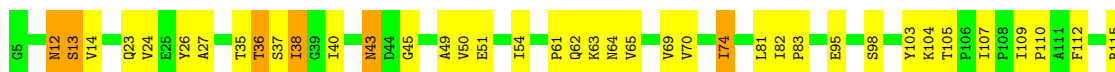




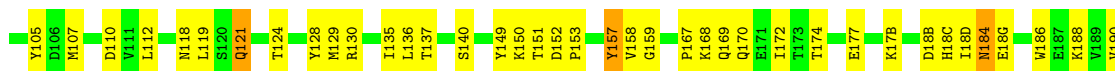
● Molecule 6: Proteasome component C1



● Molecule 6: Proteasome component C1

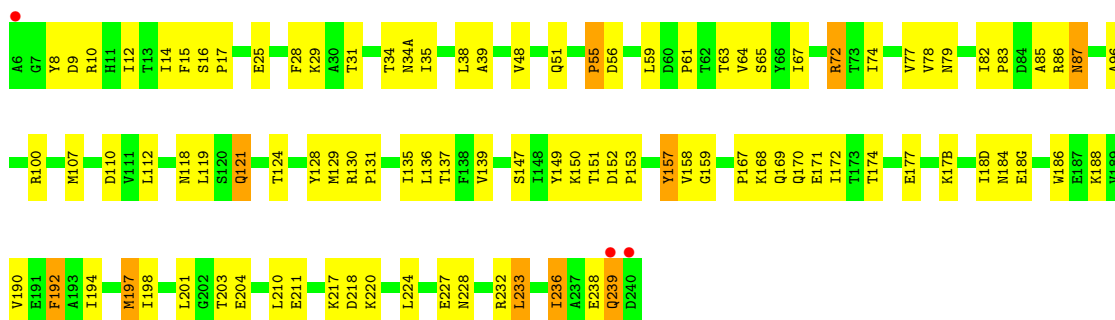


● Molecule 7: Proteasome component C7-alpha

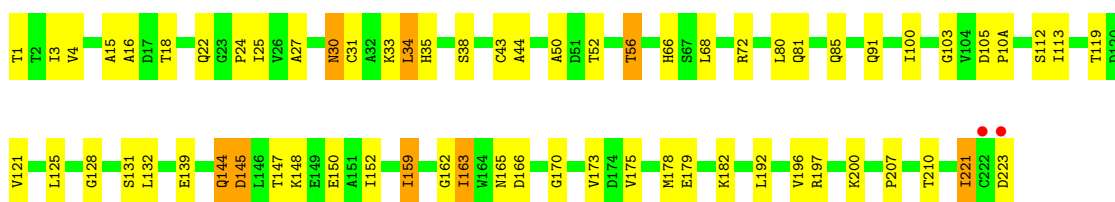


● Molecule 7: Proteasome component C7-alpha

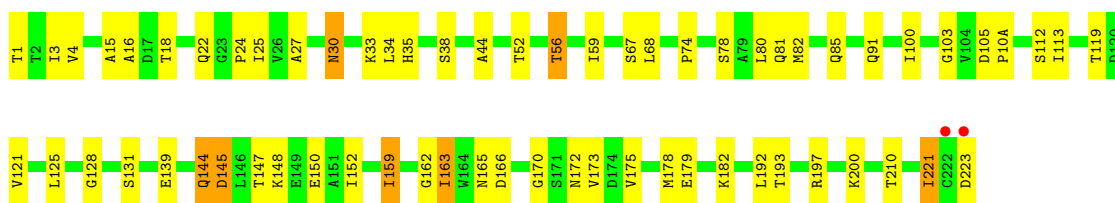




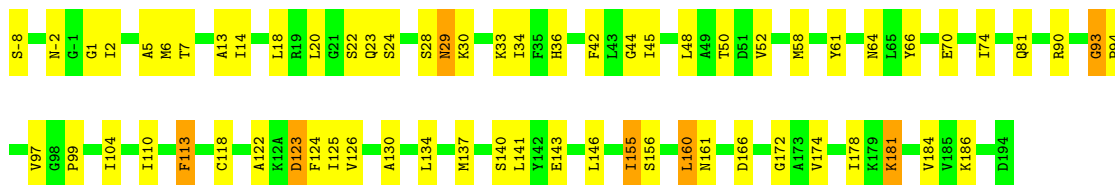
- Molecule 8: Proteasome component PUP1



- Molecule 8: Proteasome component PUP1

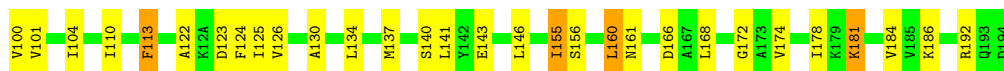


- Molecule 9: Proteasome component PUP3

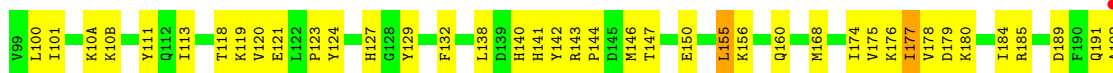


- Molecule 9: Proteasome component PUP3

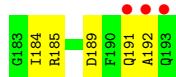
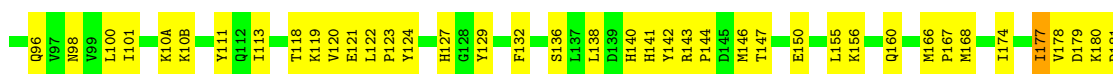
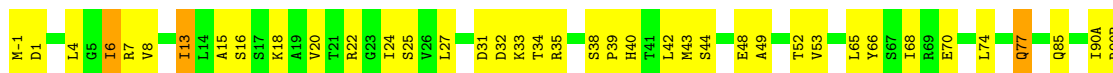




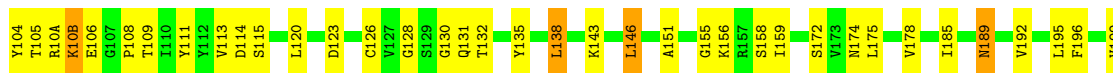
- Molecule 10: Proteasome component C11



- Molecule 10: Proteasome component C11

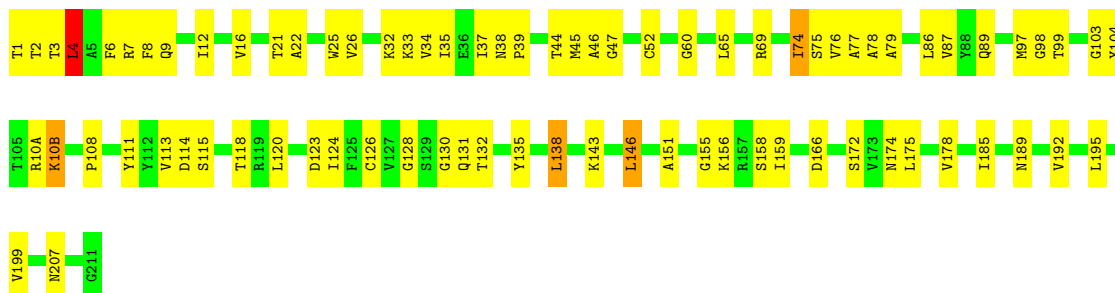


- Molecule 11: Proteasome component PRE2

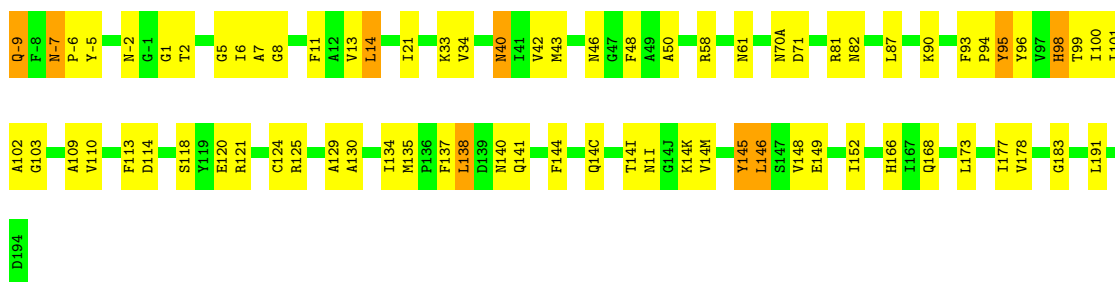


- Molecule 11: Proteasome component PRE2

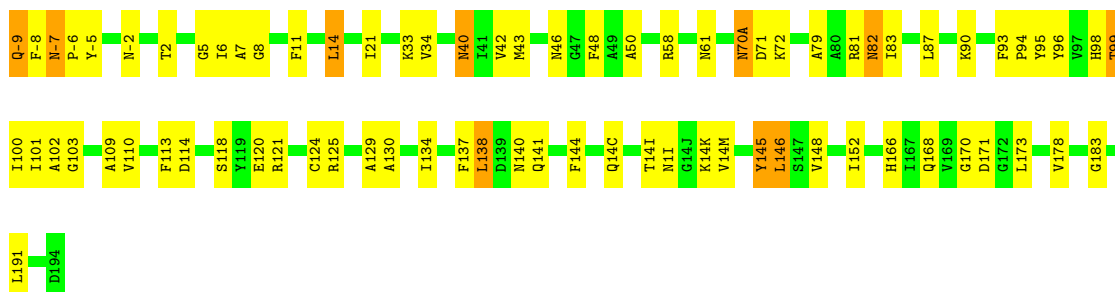




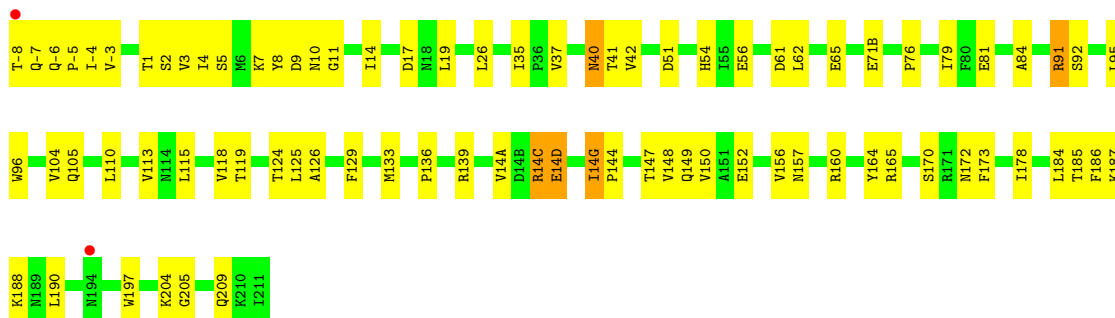
• Molecule 12: Proteasome component C5



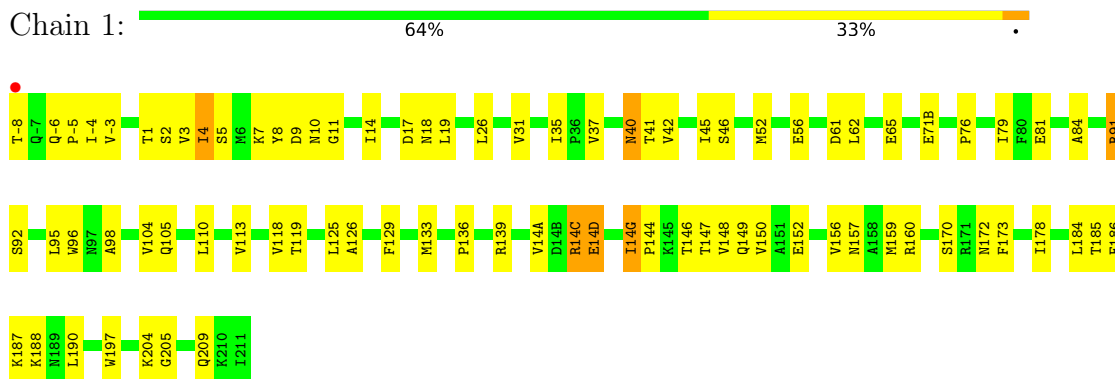
• Molecule 12: Proteasome component C5



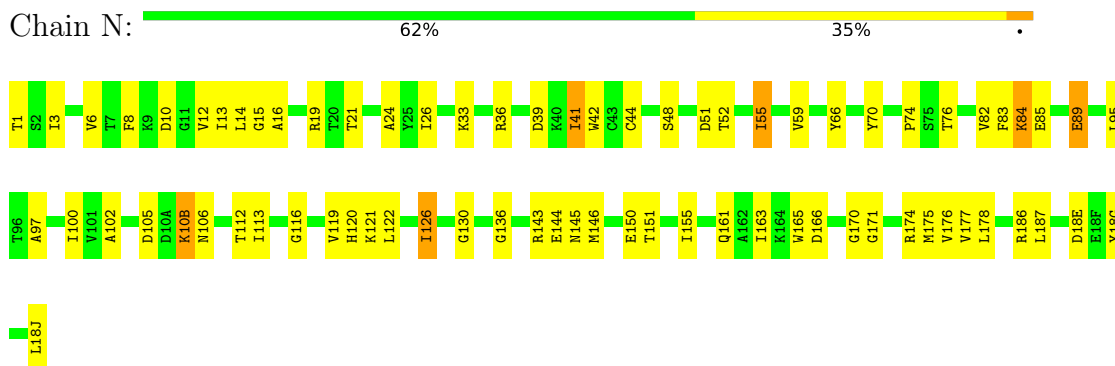
• Molecule 13: Proteasome component PRE4



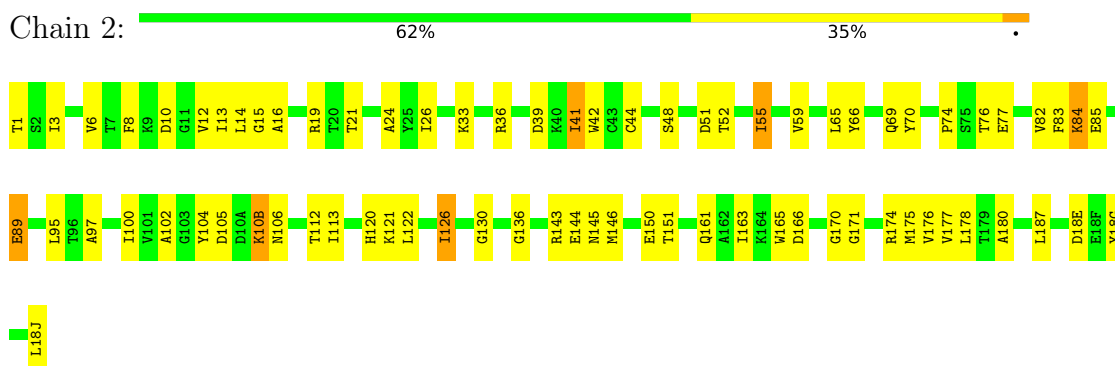
- Molecule 13: Proteasome component PRE4



- Molecule 14: Proteasome component PRE3



- Molecule 14: Proteasome component PRE3





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.49Å 301.06Å 144.45Å 90.00° 112.78° 90.00°	Depositor
Resolution (Å)	15.00 – 3.00 20.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.1 (15.00-3.00) 99.2 (20.00-3.00)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.38 (at 2.98Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.218 , 0.245 0.204 , 0.231	Depositor DCC
$R_{free}$ test set	10446 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.5	Xtrriage
Anisotropy	0.774	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 41.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	50490	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage'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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: ESY

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	0.41	0/1952	0.65	0/2642
1	O	0.40	0/1952	0.65	0/2642
2	B	0.39	0/1935	0.64	0/2618
2	P	0.40	0/1935	0.65	0/2618
3	C	0.37	0/1920	0.64	0/2598
3	Q	0.37	0/1920	0.64	0/2598
4	D	0.38	0/1887	0.64	0/2541
4	R	0.39	0/1887	0.64	0/2541
5	E	0.38	0/1823	0.61	0/2463
5	S	0.38	0/1823	0.61	0/2463
6	F	0.39	0/1937	0.61	0/2614
6	T	0.41	0/1937	0.63	0/2614
7	G	0.42	0/1959	0.64	0/2652
7	U	0.43	0/1959	0.64	0/2652
8	H	0.41	0/1716	0.68	0/2326
8	V	0.40	0/1716	0.68	0/2326
9	I	0.42	0/1611	0.66	0/2174
9	W	0.42	0/1611	0.68	0/2174
10	J	0.41	0/1613	0.64	0/2173
10	X	0.41	0/1613	0.65	0/2173
11	K	0.43	0/1680	0.66	1/2274 (0.0%)
11	Y	0.40	0/1680	0.66	1/2274 (0.0%)
12	L	0.42	0/1795	0.68	1/2420 (0.0%)
12	Z	0.40	0/1795	0.68	0/2420
13	1	0.42	0/1855	0.68	1/2514 (0.0%)
13	M	0.41	0/1855	0.67	1/2514 (0.0%)
14	2	0.41	0/1541	0.65	0/2087
14	N	0.42	0/1541	0.65	0/2087
All	All	0.40	0/50448	0.65	5/68192 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	1	95	LEU	N-CA-C	-5.62	95.83	111.00
13	M	95	LEU	N-CA-C	-5.49	96.18	111.00
12	L	95	TYR	N-CA-C	-5.31	96.67	111.00
11	K	4	LEU	CA-CB-CG	5.14	127.13	115.30
11	Y	4	LEU	CA-CB-CG	5.05	126.93	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	60	0
1	O	1915	0	1926	66	0
2	B	1905	0	1901	85	0
2	P	1905	0	1901	89	0
3	C	1891	0	1900	100	0
3	Q	1891	0	1900	98	0
4	D	1862	0	1836	63	0
4	R	1862	0	1836	77	0
5	E	1795	0	1797	116	0
5	S	1795	0	1797	121	0
6	F	1897	0	1886	81	0
6	T	1897	0	1886	86	0
7	G	1921	0	1910	85	0
7	U	1921	0	1910	94	0
8	H	1685	0	1688	45	0
8	V	1685	0	1688	44	0
9	I	1581	0	1574	55	0
9	W	1581	0	1574	55	0
10	J	1585	0	1590	67	0
10	X	1585	0	1590	73	0
11	K	1643	0	1594	66	0
11	Y	1643	0	1594	68	0
12	L	1757	0	1711	66	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	67	0
13	1	1824	0	1832	68	0
13	M	1824	0	1832	64	0
14	2	1512	0	1481	58	0
14	N	1512	0	1481	57	0
15	K	43	0	42	5	0
15	Y	43	0	42	5	0
16	1	45	0	0	7	0
16	2	41	0	0	2	0
16	A	38	0	0	2	0
16	B	31	0	0	3	0
16	C	24	0	0	2	0
16	D	25	0	0	0	0
16	E	15	0	0	3	0
16	F	27	0	0	1	0
16	G	41	0	0	1	0
16	H	28	0	0	3	0
16	I	42	0	0	2	0
16	J	32	0	0	0	0
16	K	25	0	0	3	0
16	L	35	0	0	4	0
16	M	43	0	0	4	0
16	N	38	0	0	2	0
16	O	18	0	0	3	0
16	P	20	0	0	1	0
16	Q	17	0	0	2	0
16	R	23	0	0	3	0
16	S	16	0	0	1	0
16	T	26	0	0	2	0
16	U	38	0	0	4	0
16	V	33	0	0	2	0
16	W	39	0	0	2	0
16	X	38	0	0	3	0
16	Y	24	0	0	2	0
16	Z	36	0	0	3	0
All	All	50490	0	49336	1891	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 19.

The worst 5 of 1891 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)
11:K:10(B):LYS:H	11:K:10(B):LYS:HD2	1.07	1.11
11:Y:10(B):LYS:H	11:Y:10(B):LYS:HD2	1.07	1.11
2:P:202:THR:HG22	2:P:204:SER:H	1.16	1.05
2:B:202:THR:HG22	2:B:204:SER:H	1.16	1.03
11:Y:35:ILE:HD11	11:Y:45:MET:HE2	1.40	1.03

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%)	228 (92%)	16 (6%)	4 (2%)	9	40
1	O	248/250 (99%)	229 (92%)	15 (6%)	4 (2%)	9	40
2	B	242/244 (99%)	218 (90%)	15 (6%)	9 (4%)	3	19
2	P	242/244 (99%)	216 (89%)	17 (7%)	9 (4%)	3	19
3	C	239/241 (99%)	219 (92%)	15 (6%)	5 (2%)	7	33
3	Q	239/241 (99%)	219 (92%)	15 (6%)	5 (2%)	7	33
4	D	240/242 (99%)	225 (94%)	10 (4%)	5 (2%)	7	33
4	R	240/242 (99%)	224 (93%)	11 (5%)	5 (2%)	7	33
5	E	231/233 (99%)	202 (87%)	20 (9%)	9 (4%)	3	17
5	S	231/233 (99%)	201 (87%)	21 (9%)	9 (4%)	3	17
6	F	242/244 (99%)	220 (91%)	18 (7%)	4 (2%)	9	39
6	T	242/244 (99%)	217 (90%)	21 (9%)	4 (2%)	9	39
7	G	241/243 (99%)	224 (93%)	15 (6%)	2 (1%)	19	57
7	U	241/243 (99%)	224 (93%)	15 (6%)	2 (1%)	19	57
8	H	220/222 (99%)	206 (94%)	12 (6%)	2 (1%)	17	55
8	V	220/222 (99%)	204 (93%)	14 (6%)	2 (1%)	17	55

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	I	202/204 (99%)	196 (97%)	5 (2%)	1 (0%)	29	68
9	W	202/204 (99%)	195 (96%)	6 (3%)	1 (0%)	29	68
10	J	196/198 (99%)	180 (92%)	12 (6%)	4 (2%)	7	34
10	X	196/198 (99%)	178 (91%)	14 (7%)	4 (2%)	7	34
11	K	210/212 (99%)	199 (95%)	11 (5%)	0	100	100
11	Y	210/212 (99%)	200 (95%)	10 (5%)	0	100	100
12	L	220/222 (99%)	206 (94%)	13 (6%)	1 (0%)	29	68
12	Z	220/222 (99%)	206 (94%)	13 (6%)	1 (0%)	29	68
13	1	231/233 (99%)	211 (91%)	20 (9%)	0	100	100
13	M	231/233 (99%)	211 (91%)	20 (9%)	0	100	100
14	2	194/196 (99%)	186 (96%)	7 (4%)	1 (0%)	29	68
14	N	194/196 (99%)	186 (96%)	7 (4%)	1 (0%)	29	68
All	All	6312/6368 (99%)	5830 (92%)	388 (6%)	94 (2%)	10	42

5 of 94 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	167	LYS
2	B	20(A)	SER
3	C	58	LEU
4	D	12(G)	GLU
5	E	5	ARG

### 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%)	200 (96%)	9 (4%)	29	66
1	O	209/209 (100%)	200 (96%)	9 (4%)	29	66
2	B	203/203 (100%)	186 (92%)	17 (8%)	11	38
2	P	203/203 (100%)	185 (91%)	18 (9%)	9	35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	213/213 (100%)	201 (94%)	12 (6%)	21	56
3	Q	213/213 (100%)	201 (94%)	12 (6%)	21	56
4	D	198/198 (100%)	188 (95%)	10 (5%)	24	60
4	R	198/198 (100%)	188 (95%)	10 (5%)	24	60
5	E	192/192 (100%)	171 (89%)	21 (11%)	6	25
5	S	192/192 (100%)	170 (88%)	22 (12%)	5	24
6	F	201/201 (100%)	186 (92%)	15 (8%)	13	43
6	T	201/201 (100%)	186 (92%)	15 (8%)	13	43
7	G	207/207 (100%)	195 (94%)	12 (6%)	20	55
7	U	207/207 (100%)	195 (94%)	12 (6%)	20	55
8	H	181/181 (100%)	168 (93%)	13 (7%)	14	45
8	V	181/181 (100%)	170 (94%)	11 (6%)	18	53
9	I	172/172 (100%)	165 (96%)	7 (4%)	30	67
9	W	172/172 (100%)	166 (96%)	6 (4%)	36	71
10	J	175/175 (100%)	167 (95%)	8 (5%)	27	64
10	X	175/175 (100%)	167 (95%)	8 (5%)	27	64
11	K	169/169 (100%)	160 (95%)	9 (5%)	22	58
11	Y	169/169 (100%)	160 (95%)	9 (5%)	22	58
12	L	185/185 (100%)	172 (93%)	13 (7%)	15	47
12	Z	185/185 (100%)	172 (93%)	13 (7%)	15	47
13	1	199/199 (100%)	189 (95%)	10 (5%)	24	60
13	M	199/199 (100%)	189 (95%)	10 (5%)	24	60
14	2	162/162 (100%)	155 (96%)	7 (4%)	29	66
14	N	162/162 (100%)	154 (95%)	8 (5%)	25	61
All	All	5332/5332 (100%)	5006 (94%)	326 (6%)	18	53

5 of 326 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	S	2(B)	THR
10	X	177	ILE
6	T	74	ILE
7	U	233	LEU
12	Z	58	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 198 such sidechains are listed below:

Mol	Chain	Res	Type
3	Q	243	GLN
7	U	34(A)	ASN
4	R	199	GLN
5	S	156	ASN
7	U	178	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](#)

2 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	ESY	K	2000	11	44,44,44	2.19	15 (34%)	52,56,56	2.32	6 (11%)
15	ESY	Y	2000	11	44,44,44	2.15	11 (25%)	52,56,56	2.21	9 (17%)

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	ESY	K	2000	11	-	11/50/50/50	0/2/2/2
15	ESY	Y	2000	11	-	9/50/50/50	0/2/2/2

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	K	2000	ESY	C10-N11	6.71	1.48	1.34
15	Y	2000	ESY	C10-N11	6.34	1.48	1.34
15	Y	2000	ESY	C20-C19	5.39	1.59	1.54
15	K	2000	ESY	C12-N11	4.77	1.55	1.45
15	Y	2000	ESY	C12-N11	4.55	1.55	1.45

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	K	2000	ESY	C9-N8-C7	10.36	139.92	120.49
15	Y	2000	ESY	C9-N8-C7	9.57	138.43	120.49
15	K	2000	ESY	C10-C9-N8	-8.84	89.70	111.60
15	Y	2000	ESY	C10-C9-N8	-8.08	91.59	111.60
15	K	2000	ESY	C5-O6-C7	-4.71	105.41	115.93

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	K	2000	ESY	C40-C9-N8-C7
15	K	2000	ESY	N11-C12-C13-C14
15	Y	2000	ESY	C40-C9-N8-C7
15	Y	2000	ESY	N11-C12-C13-C14
15	K	2000	ESY	O41-C7-O6-C5

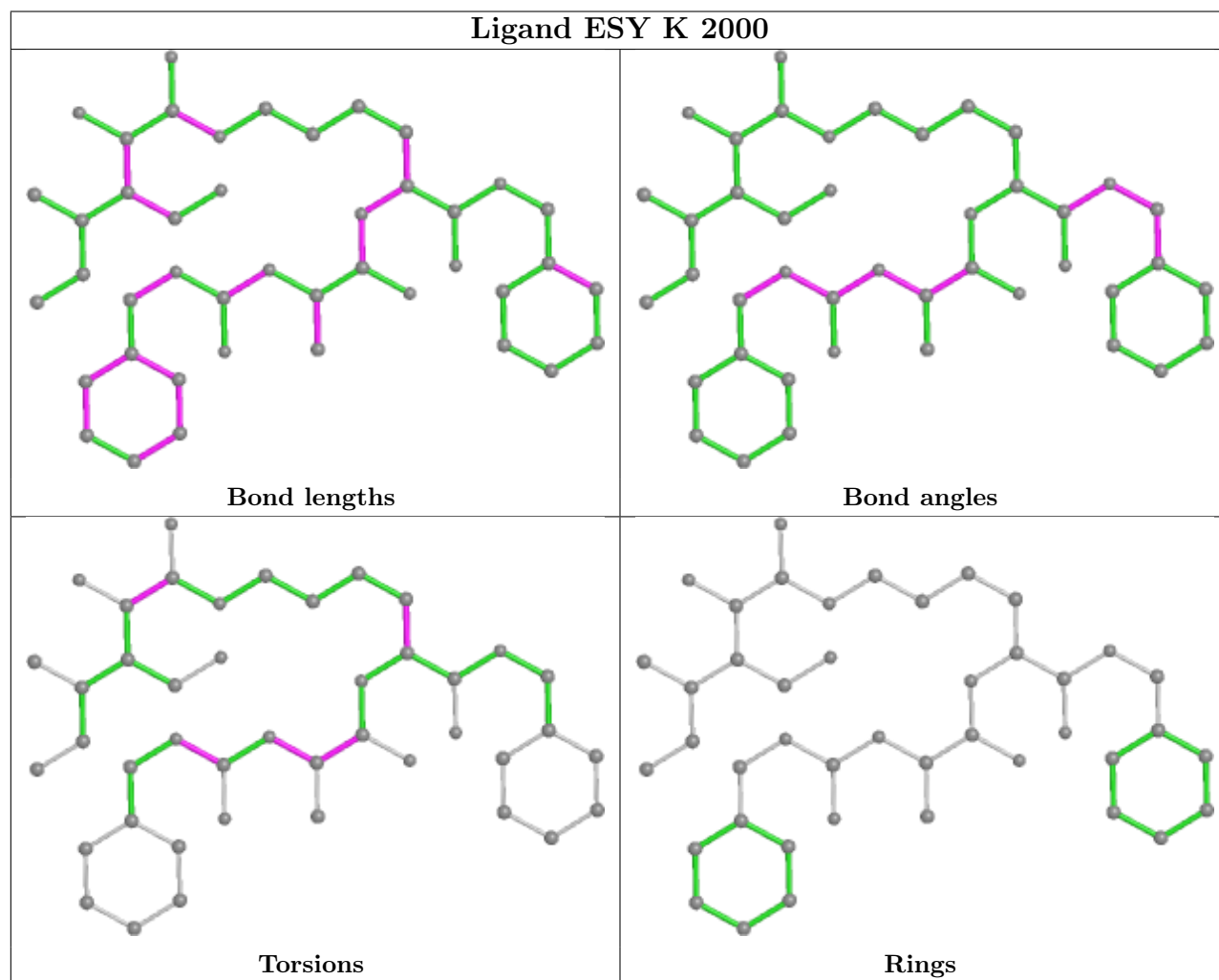
There are no ring outliers.

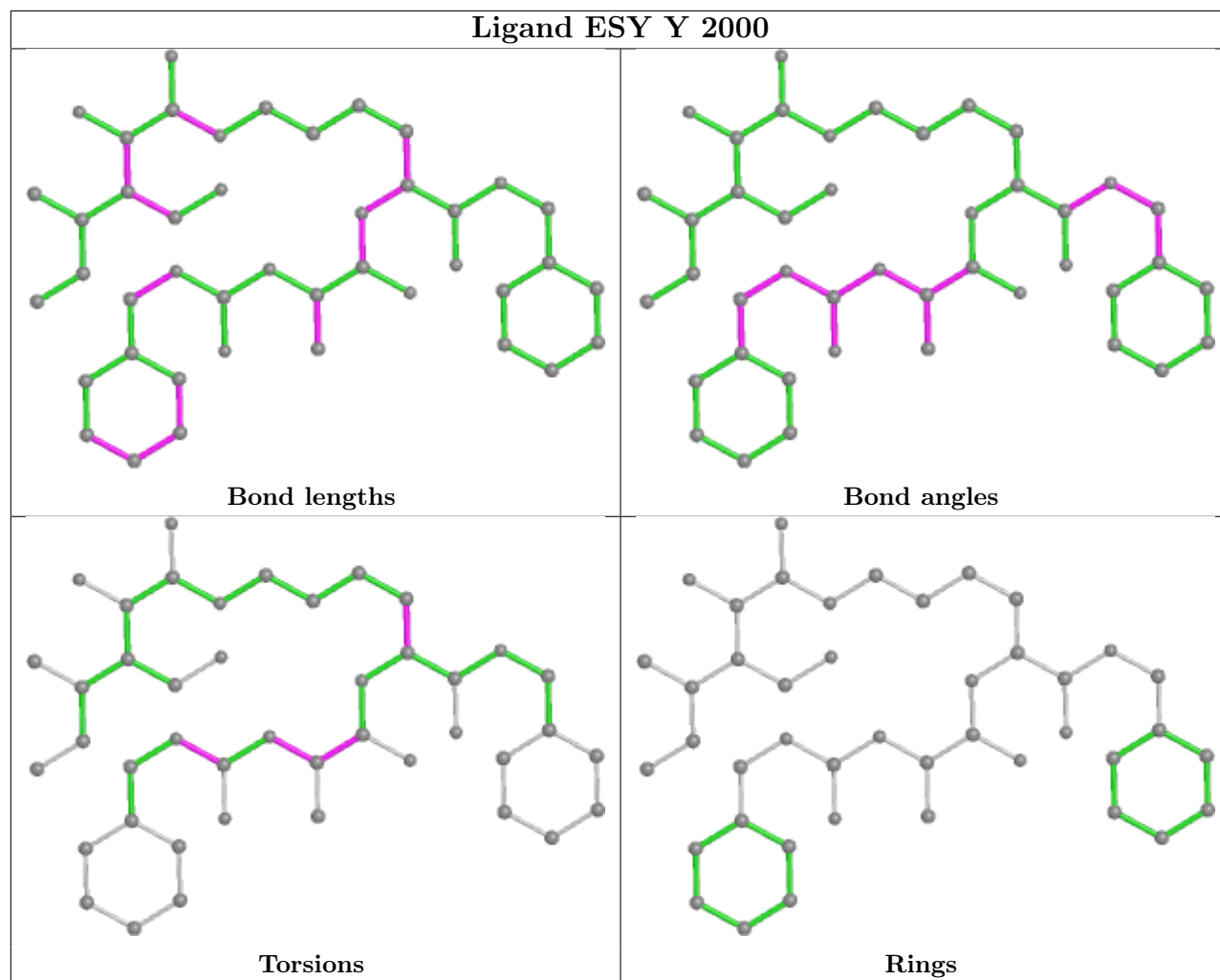
2 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	K	2000	ESY	5	0
15	Y	2000	ESY	5	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 than 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.





## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	250/250 (100%)	-0.59	2 (0%) 86 65	23, 42, 71, 96	0
1	O	250/250 (100%)	-0.53	3 (1%) 79 54	29, 49, 76, 97	0
2	B	244/244 (100%)	-0.48	4 (1%) 72 44	29, 48, 84, 112	0
2	P	244/244 (100%)	-0.39	8 (3%) 46 20	28, 50, 85, 115	0
3	C	241/241 (100%)	-0.39	3 (1%) 79 54	31, 51, 104, 118	0
3	Q	241/241 (100%)	-0.26	8 (3%) 46 20	37, 56, 105, 118	0
4	D	242/242 (100%)	-0.39	6 (2%) 57 29	29, 54, 86, 116	0
4	R	242/242 (100%)	-0.35	8 (3%) 46 20	34, 54, 88, 118	0
5	E	233/233 (100%)	-0.46	4 (1%) 70 41	37, 56, 82, 105	0
5	S	233/233 (100%)	-0.33	4 (1%) 70 41	37, 58, 85, 103	0
6	F	244/244 (100%)	-0.57	2 (0%) 86 65	30, 49, 84, 101	0
6	T	244/244 (100%)	-0.46	1 (0%) 92 79	26, 50, 85, 102	0
7	G	243/243 (100%)	-0.63	2 (0%) 86 65	26, 43, 68, 103	0
7	U	243/243 (100%)	-0.56	3 (1%) 79 54	30, 44, 69, 103	0
8	H	222/222 (100%)	-0.71	2 (0%) 84 63	27, 40, 62, 89	0
8	V	222/222 (100%)	-0.69	2 (0%) 84 63	29, 42, 61, 89	0
9	I	204/204 (100%)	-0.84	0 100 100	26, 40, 57, 73	0
9	W	204/204 (100%)	-0.77	1 (0%) 91 75	27, 40, 59, 74	0
10	J	198/198 (100%)	-0.63	2 (1%) 82 59	25, 41, 60, 114	0
10	X	198/198 (100%)	-0.63	3 (1%) 73 46	27, 43, 62, 116	0
11	K	212/212 (100%)	-0.72	0 100 100	23, 41, 55, 69	0
11	Y	212/212 (100%)	-0.74	0 100 100	27, 43, 58, 68	0
12	L	222/222 (100%)	-0.73	0 100 100	21, 42, 64, 79	0
12	Z	222/222 (100%)	-0.72	0 100 100	28, 42, 65, 82	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	1	233/233 (100%)	-0.69	1 (0%) 92 79	22, 41, 58, 66	0
13	M	233/233 (100%)	-0.66	2 (0%) 84 63	26, 43, 59, 65	0
14	2	196/196 (100%)	-0.71	0 100 100	23, 39, 61, 74	0
14	N	196/196 (100%)	-0.74	0 100 100	26, 38, 59, 72	0
All	All	6368/6368 (100%)	-0.58	71 (1%) 80 56	21, 45, 79, 118	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	R	12(F)	GLY	7.6
4	R	12(E)	SER	7.5
4	D	12(D)	ALA	7.4
7	U	240	ASP	6.5
4	R	12(D)	ALA	6.2

## 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](#)

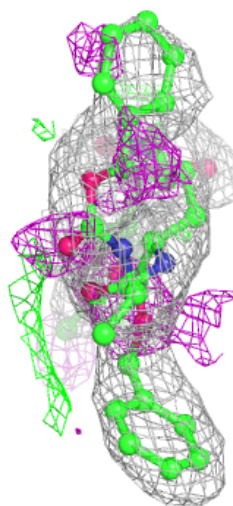
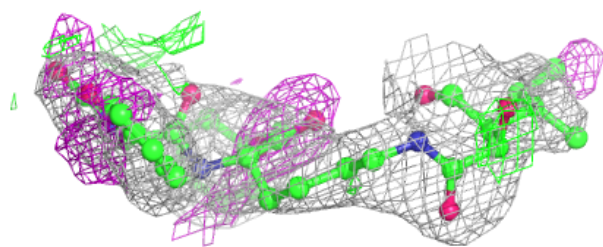
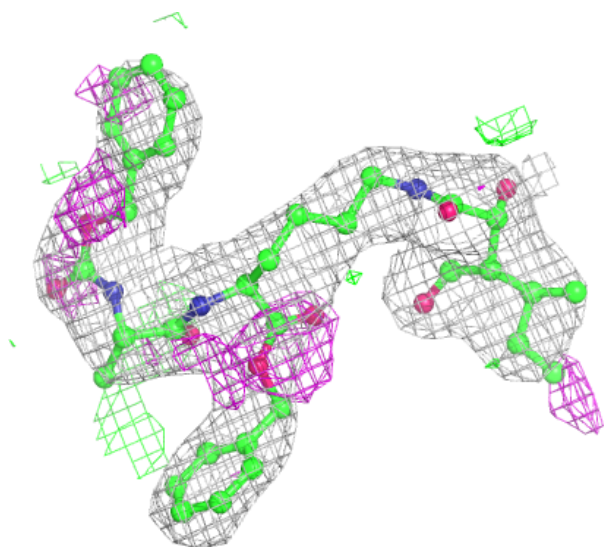
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

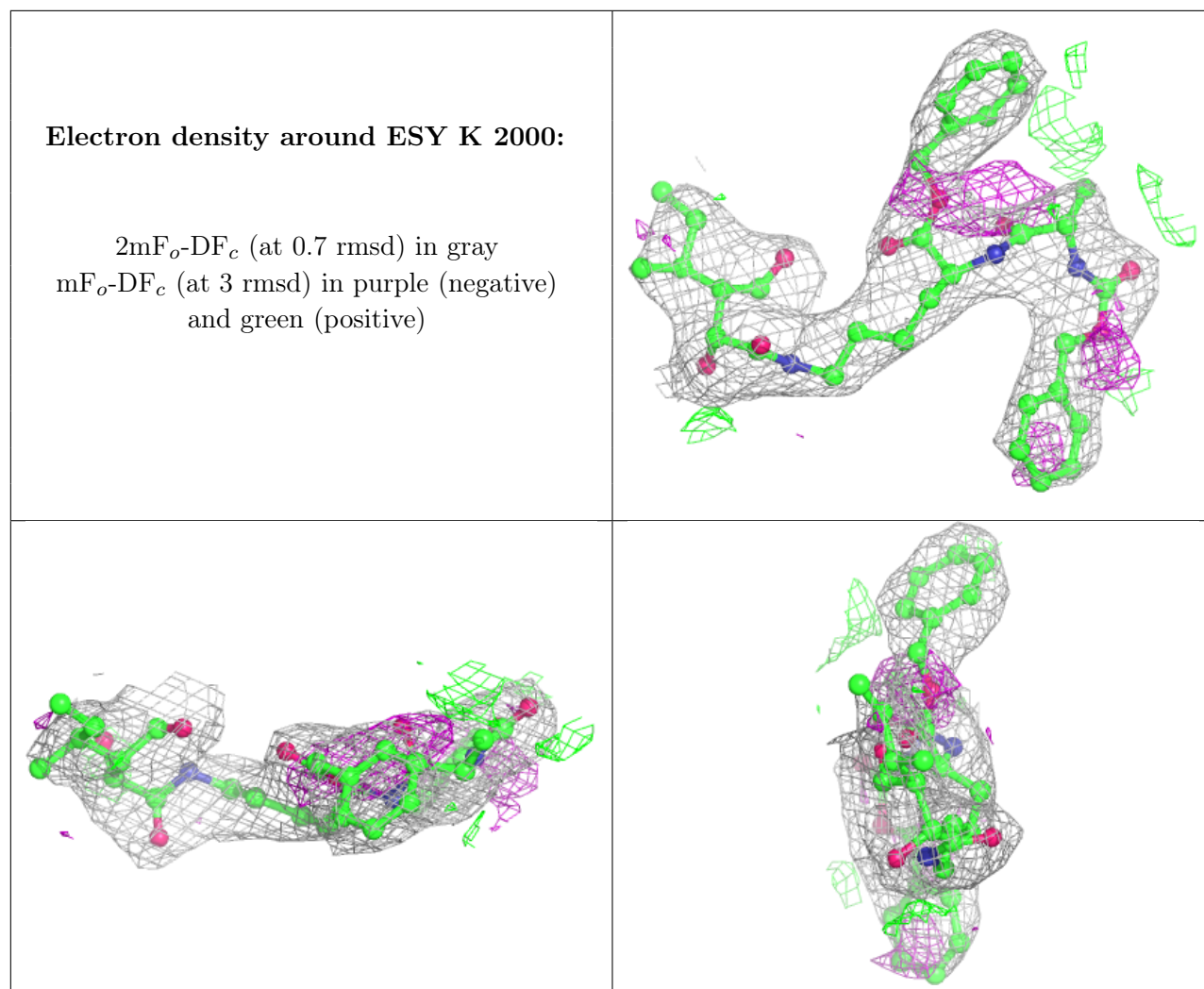
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
15	ESY	Y	2000	43/43	0.92	0.25	20,36,45,46	0
15	ESY	K	2000	43/43	0.93	0.26	20,36,45,46	0

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 ESY Y 2000:**

$2mF_o-DF_c$  (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.