



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 08:00 PM UTC

PDB ID : 4INR / pdb_00004inr
Title : Yeast 20S proteasome in complex with the vinyl sulfone LU102
Authors : Geurink, P.P.; van der Linden, W.A.; Mirabella, A.C.; Gallastegui, N.; de Bruin, G.; Blom, A.E.M.; Voges, M.J.; Mock, E.D.; Florea, B.I.; van der Marel, G.A.; Driessen, C.; van der Stelt, M.; Groll, M.; Overkleeft, H.S.; Kisselev, A.F.
Deposited on : 2013-01-06
Resolution : 2.70 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 ⓘ 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](#) ①) were used in the production of this report:

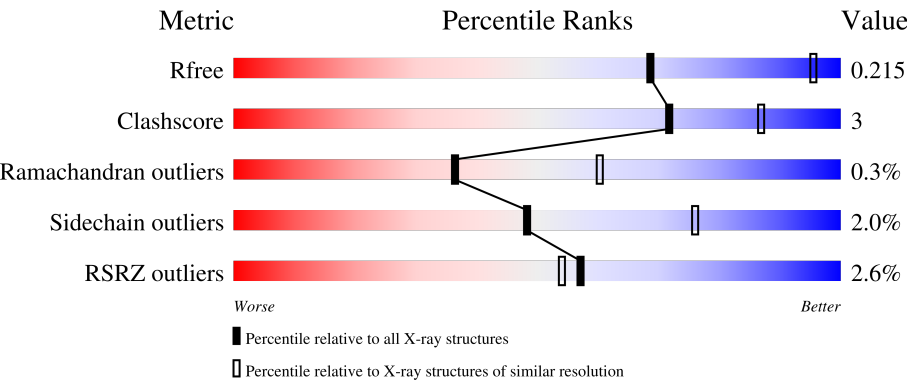
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.70 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}	180053	3538 (2.70-2.70)
Clashscore	190562	3843 (2.70-2.70)
Ramachandran outliers	187476	3778 (2.70-2.70)
Sidechain outliers	187428	3778 (2.70-2.70)
RSRZ outliers	180081	3538 (2.70-2.70)

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 $\leq 5\%$. 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	2% 93% 7%
1	O	250	3% 92% 7%
2	B	258	6% 83% 11% 5%
2	P	258	8% 84% 10% 5%
3	C	254	2% 86% 8% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	
4	D	260	
4	R	260	
5	E	234	
5	S	234	
6	F	288	
6	T	288	
7	G	252	
7	U	252	
8	H	232	
8	V	232	
9	I	205	
9	W	205	
10	J	198	
10	X	198	
11	K	212	
11	Y	212	
12	L	222	
12	Z	222	
13	M	233	
13	a	233	
14	N	196	
14	b	196	

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 51061 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
1	A	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			
1	O	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			

- Molecule 2 is a protein called Proteasome component Y13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	244	Total	C	N	O	S	0	0	0
			1904	1201	321	379	3			
2	P	244	Total	C	N	O	S	0	0	0
			1904	1201	321	379	3			

- Molecule 3 is a protein called Proteasome component PRE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	241	Total	C	N	O	S	0	0	0
			1890	1181	331	374	4			
3	Q	241	Total	C	N	O	S	0	0	0
			1890	1181	331	374	4			

- Molecule 4 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	242	Total	C	N	O	S	0	0	0
			1861	1162	314	378	7			
4	R	242	Total	C	N	O	S	0	0	0
			1861	1162	314	378	7			

- 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
			1896	1205	330	357	4			
6	T	244	Total	C	N	O	S	0	0	0
			1896	1205	330	357	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
			1684	1061	293	323	7			
8	V	222	Total	C	N	O	S	0	0	0
			1684	1061	293	323	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
			1644	1045	280	312	7			
11	Y	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	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			

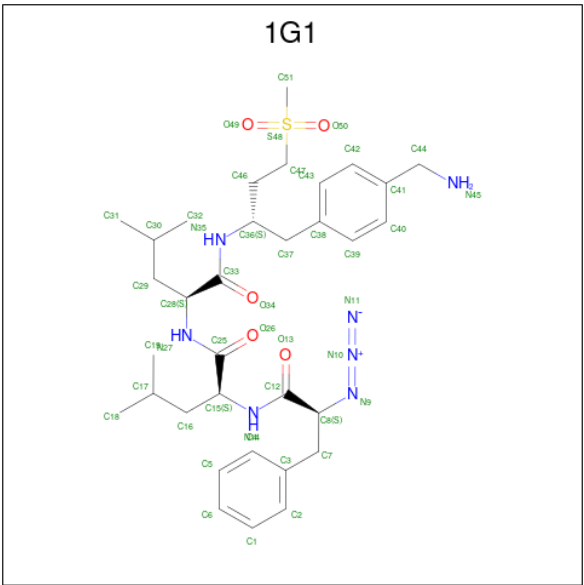
- Molecule 13 is a protein called Proteasome component PRE4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			
13	a	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	b	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			

- Molecule 15 is N3Phe-Leu-Leu-Phe(4-NH₂CH₂)-methyl vinyl sulfone, bound form (CCD ID: 1G1) (formula: C₃₃H₄₉N₇O₅S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
15	H	1	Total	C	N	O	S	0	0
			46	33	7	5	1		
15	K	1	Total	C	N	O	S	0	0
			46	33	7	5	1		
15	V	1	Total	C	N	O	S	0	0
			46	33	7	5	1		
15	Y	1	Total	C	N	O	S	0	0
			46	33	7	5	1		

- Molecule 16 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	58	Total	O	0	0
			58	58		
16	B	37	Total	O	0	0
			37	37		
16	C	43	Total	O	0	0
			43	43		
16	D	37	Total	O	0	0
			37	37		
16	E	21	Total	O	0	0
			21	21		
16	F	47	Total	O	0	0
			47	47		
16	G	60	Total	O	0	0
			60	60		
16	H	54	Total	O	0	0
			54	54		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	I	68	Total 68	O 68	0	0
16	J	53	Total 53	O 53	0	0
16	K	48	Total 48	O 48	0	0
16	L	57	Total 57	O 57	0	0
16	M	75	Total 75	O 75	0	0
16	N	55	Total 55	O 55	0	0
16	O	35	Total 35	O 35	0	0
16	P	28	Total 28	O 28	0	0
16	Q	29	Total 29	O 29	0	0
16	R	27	Total 27	O 27	0	0
16	S	17	Total 17	O 17	0	0
16	T	42	Total 42	O 42	0	0
16	U	60	Total 60	O 60	0	0
16	V	50	Total 50	O 50	0	0
16	W	57	Total 57	O 57	0	0
16	X	46	Total 46	O 46	0	0
16	Y	52	Total 52	O 52	0	0
16	Z	49	Total 49	O 49	0	0
16	a	77	Total 77	O 77	0	0
16	b	57	Total 57	O 57	0	0

3 Residue-property plots

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.

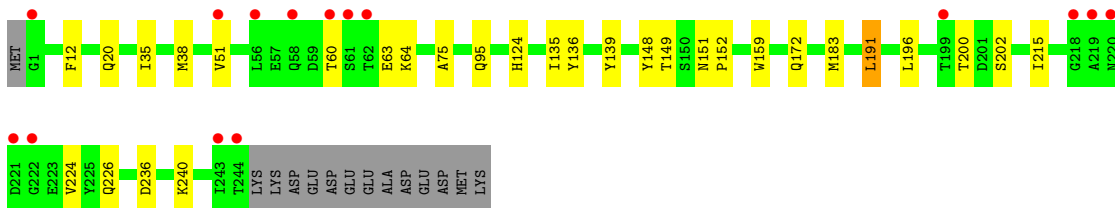
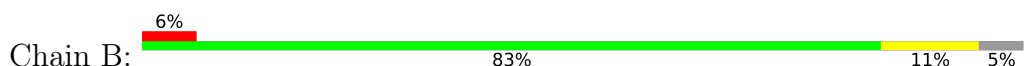
- Molecule 1: Proteasome component Y7



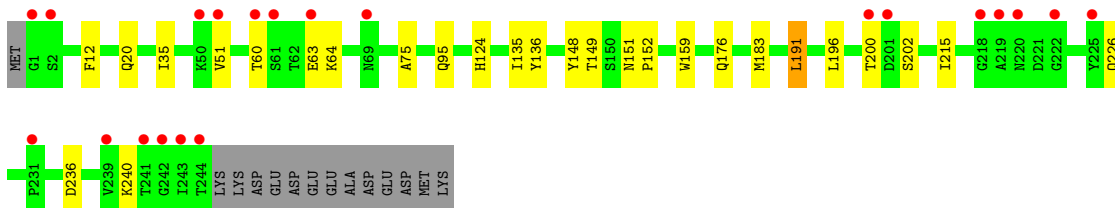
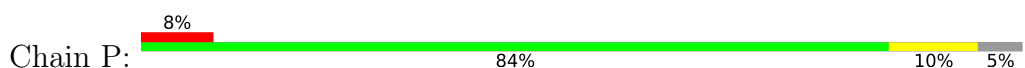
- Molecule 1: Proteasome component Y7



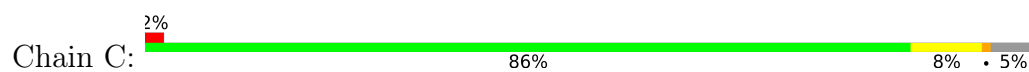
- Molecule 2: Proteasome component Y13



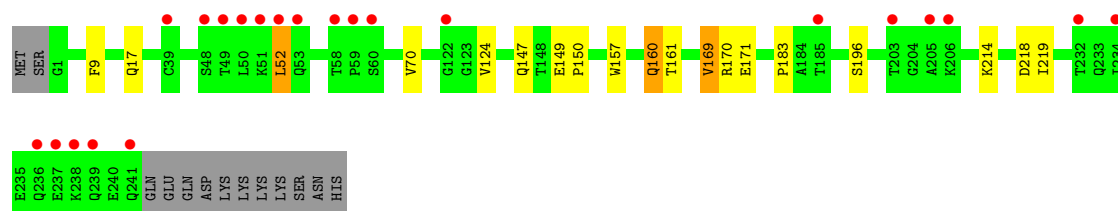
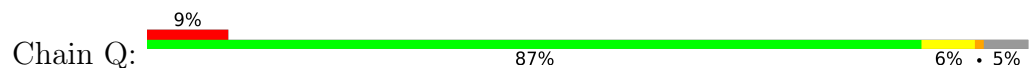
- Molecule 2: Proteasome component Y13



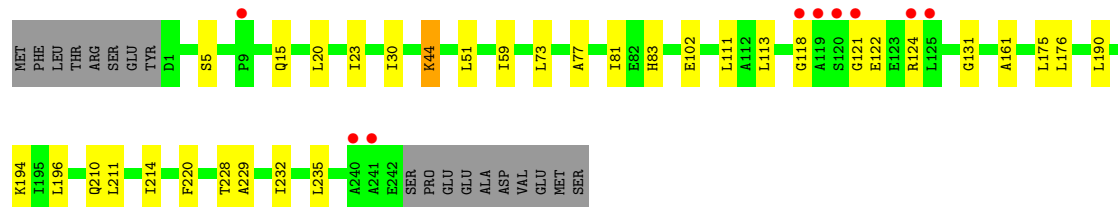
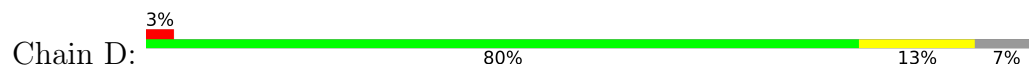
- Molecule 3: Proteasome component PRE6



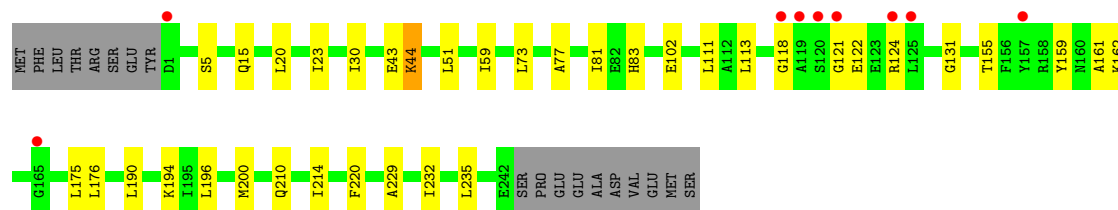
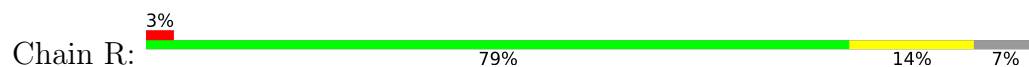
• Molecule 3: Proteasome component PRE6



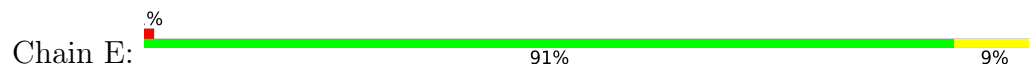
• Molecule 4: Proteasome component PUP2



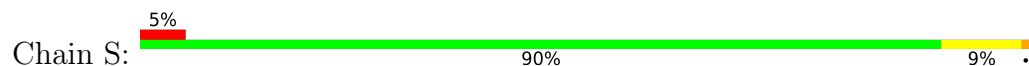
• Molecule 4: Proteasome component PUP2



• Molecule 5: Proteasome component PRE5

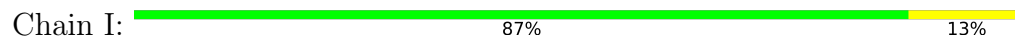


• Molecule 5: Proteasome component PRE5

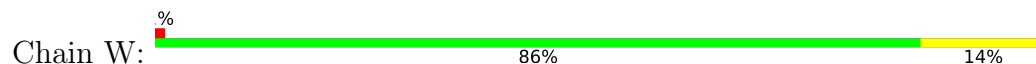




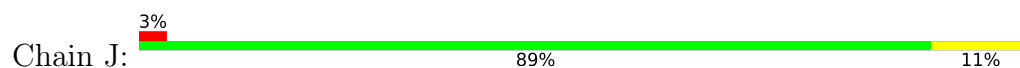
- Molecule 9: Proteasome component PUP3



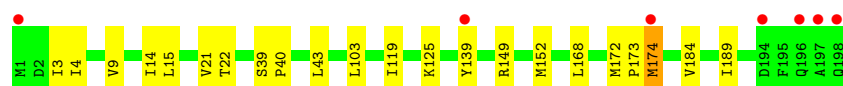
- Molecule 9: Proteasome component PUP3



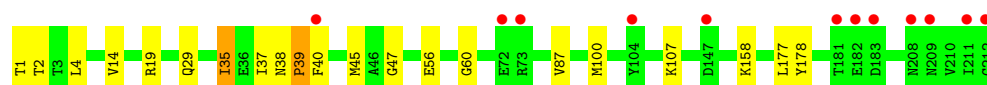
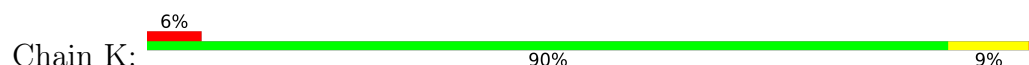
- Molecule 10: Proteasome component C11



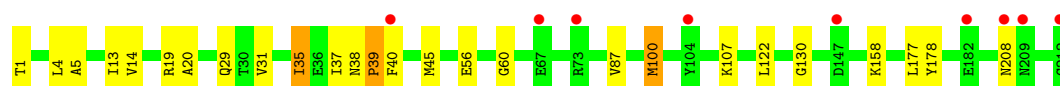
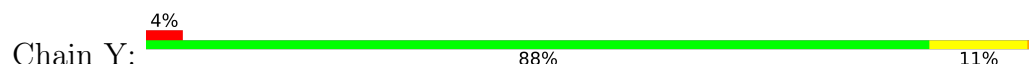
- Molecule 10: Proteasome component C11



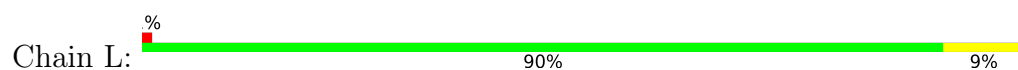
- Molecule 11: Proteasome component PRE2



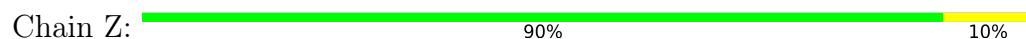
- Molecule 11: Proteasome component PRE2



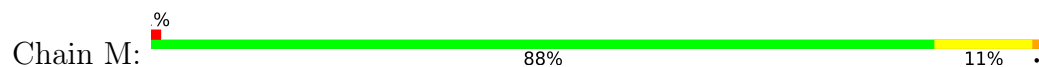
- Molecule 12: Proteasome component C5



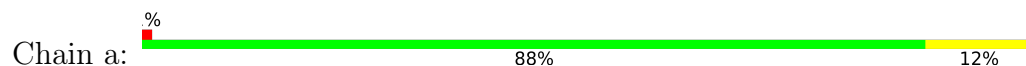
- Molecule 12: Proteasome component C5



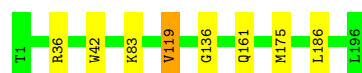
- Molecule 13: Proteasome component PRE4



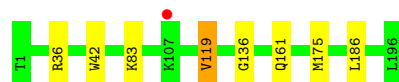
- Molecule 13: Proteasome component PRE4



- 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, α , β , γ	136.32Å 301.05Å 144.32Å 90.00° 113.13° 90.00°	Depositor
Resolution (Å)	15.00 – 2.70 15.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.7 (15.00-2.70) 98.2 (15.00-2.70)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.88 (at 2.69Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.224 , 0.231 (Not available) , 0.215	Depositor DCC
R_{free} test set	14301 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	66.2	Xtriage
Anisotropy	0.085	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 42.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	51061	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 1G1

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	0.48	0/1952	0.77	0/2642
1	O	0.49	0/1952	0.77	0/2642
2	B	0.46	0/1934	0.74	0/2618
2	P	0.46	0/1934	0.74	0/2618
3	C	0.47	0/1919	0.76	0/2598
3	Q	0.47	0/1919	0.77	0/2598
4	D	0.52	0/1886	0.72	0/2541
4	R	0.52	0/1886	0.73	0/2541
5	E	0.47	0/1823	0.73	0/2463
5	S	0.48	0/1823	0.73	2/2463 (0.1%)
6	F	0.58	0/1936	0.77	0/2614
6	T	0.58	0/1936	0.79	0/2614
7	G	0.49	0/1959	0.75	0/2652
7	U	0.49	0/1959	0.75	0/2652
8	H	0.62	0/1715	0.74	0/2326
8	V	0.61	0/1715	0.74	0/2326
9	I	0.45	0/1611	0.73	0/2174
9	W	0.45	0/1611	0.73	0/2174
10	J	0.48	0/1613	0.73	0/2173
10	X	0.48	0/1613	0.73	0/2173
11	K	0.59	0/1681	0.74	0/2274
11	Y	0.58	0/1681	0.74	0/2274
12	L	0.51	0/1795	0.73	0/2420
12	Z	0.51	0/1795	0.72	0/2420
13	M	0.47	0/1855	0.71	0/2514
13	a	0.47	0/1855	0.73	0/2514
14	N	0.52	0/1541	0.75	0/2087
14	b	0.52	0/1541	0.74	0/2087
All	All	0.51	0/50440	0.74	2/68192 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	S	147	GLN	CA-C-N	5.24	125.55	119.47
5	S	147	GLN	C-N-CA	5.24	125.55	119.47

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1929	9	0
1	O	1915	0	1929	9	0
2	B	1904	0	1904	17	0
2	P	1904	0	1904	17	0
3	C	1890	0	1903	15	0
3	Q	1890	0	1903	12	0
4	D	1861	0	1839	16	0
4	R	1861	0	1839	19	0
5	E	1795	0	1800	12	0
5	S	1795	0	1800	13	0
6	F	1896	0	1889	15	0
6	T	1896	0	1889	16	0
7	G	1921	0	1913	8	0
7	U	1921	0	1913	10	0
8	H	1684	0	1687	7	0
8	V	1684	0	1687	13	0
9	I	1581	0	1574	14	0
9	W	1581	0	1574	16	0
10	J	1585	0	1590	14	0
10	X	1585	0	1590	13	0
11	K	1644	0	1594	11	0
11	Y	1644	0	1594	17	0
12	L	1757	0	1711	11	0
12	Z	1757	0	1711	11	0
13	M	1824	0	1832	17	0
13	a	1824	0	1832	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
14	N	1512	0	1481	5	0
14	b	1512	0	1481	6	0
15	H	46	0	48	2	0
15	K	46	0	48	2	0
15	V	46	0	48	5	0
15	Y	46	0	48	6	0
16	A	58	0	0	0	0
16	B	37	0	0	0	0
16	C	43	0	0	0	0
16	D	37	0	0	0	0
16	E	21	0	0	0	0
16	F	47	0	0	0	0
16	G	60	0	0	0	0
16	H	54	0	0	0	0
16	I	68	0	0	0	0
16	J	53	0	0	0	0
16	K	48	0	0	0	0
16	L	57	0	0	0	0
16	M	75	0	0	0	0
16	N	55	0	0	0	0
16	O	35	0	0	0	0
16	P	28	0	0	0	0
16	Q	29	0	0	0	0
16	R	27	0	0	0	0
16	S	17	0	0	0	0
16	T	42	0	0	0	0
16	U	60	0	0	0	0
16	V	50	0	0	0	0
16	W	57	0	0	0	0
16	X	46	0	0	1	0
16	Y	52	0	0	3	0
16	Z	49	0	0	0	0
16	a	77	0	0	0	0
16	b	57	0	0	0	0
All	All	51061	0	49484	323	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 3.

The worst 5 of 323 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)
15:Y:301:1G1:H19	16:Y:442:HOH:O	1.74	0.87
12:L:109:THR:HG23	12:L:125:PHE:HB2	1.66	0.77
10:X:4:ILE:HG22	10:X:103:LEU:HD12	1.66	0.76
5:E:12:PHE:H	6:F:19:GLN:HE22	1.38	0.71
5:S:12:PHE:H	6:T:19:GLN:HE22	1.39	0.71

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%)	241 (97%)	6 (2%)	1 (0%)	30	54
1	O	248/250 (99%)	240 (97%)	7 (3%)	1 (0%)	30	54
2	B	242/258 (94%)	236 (98%)	5 (2%)	1 (0%)	30	54
2	P	242/258 (94%)	236 (98%)	5 (2%)	1 (0%)	30	54
3	C	239/254 (94%)	231 (97%)	6 (2%)	2 (1%)	16	37
3	Q	239/254 (94%)	232 (97%)	5 (2%)	2 (1%)	16	37
4	D	240/260 (92%)	235 (98%)	2 (1%)	3 (1%)	9	25
4	R	240/260 (92%)	235 (98%)	2 (1%)	3 (1%)	9	25
5	E	231/234 (99%)	224 (97%)	7 (3%)	0	100	100
5	S	231/234 (99%)	224 (97%)	7 (3%)	0	100	100
6	F	242/288 (84%)	233 (96%)	9 (4%)	0	100	100
6	T	242/288 (84%)	233 (96%)	9 (4%)	0	100	100
7	G	241/252 (96%)	234 (97%)	7 (3%)	0	100	100
7	U	241/252 (96%)	234 (97%)	7 (3%)	0	100	100
8	H	220/232 (95%)	213 (97%)	7 (3%)	0	100	100
8	V	220/232 (95%)	212 (96%)	8 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	I	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
9	W	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
10	J	196/198 (99%)	189 (96%)	6 (3%)	1 (0%)	24	48
10	X	196/198 (99%)	189 (96%)	6 (3%)	1 (0%)	24	48
11	K	210/212 (99%)	203 (97%)	6 (3%)	1 (0%)	24	48
11	Y	210/212 (99%)	200 (95%)	9 (4%)	1 (0%)	24	48
12	L	220/222 (99%)	214 (97%)	6 (3%)	0	100	100
12	Z	220/222 (99%)	214 (97%)	6 (3%)	0	100	100
13	M	231/233 (99%)	221 (96%)	9 (4%)	1 (0%)	30	54
13	a	231/233 (99%)	222 (96%)	8 (4%)	1 (0%)	30	54
14	N	194/196 (99%)	187 (96%)	7 (4%)	0	100	100
14	b	194/196 (99%)	188 (97%)	6 (3%)	0	100	100
All	All	6312/6588 (96%)	6111 (97%)	181 (3%)	20 (0%)	36	60

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	THR
3	C	52	LEU
4	D	121	GLY
1	O	2	THR
3	Q	52	LEU

5.3.2 Protein sidechains ⓘ

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%)	206 (99%)	3 (1%)	59	82
1	O	209/209 (100%)	206 (99%)	3 (1%)	59	82
2	B	203/216 (94%)	200 (98%)	3 (2%)	57	81
2	P	203/216 (94%)	201 (99%)	2 (1%)	68	86

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	213/226 (94%)	209 (98%)	4 (2%)	50	77
3	Q	213/226 (94%)	209 (98%)	4 (2%)	50	77
4	D	198/215 (92%)	192 (97%)	6 (3%)	36	66
4	R	198/215 (92%)	192 (97%)	6 (3%)	36	66
5	E	192/193 (100%)	187 (97%)	5 (3%)	40	70
5	S	192/193 (100%)	186 (97%)	6 (3%)	35	65
6	F	201/239 (84%)	194 (96%)	7 (4%)	32	61
6	T	201/239 (84%)	194 (96%)	7 (4%)	32	61
7	G	207/210 (99%)	204 (99%)	3 (1%)	59	82
7	U	207/210 (99%)	203 (98%)	4 (2%)	50	77
8	H	181/190 (95%)	179 (99%)	2 (1%)	65	85
8	V	181/190 (95%)	179 (99%)	2 (1%)	65	85
9	I	172/173 (99%)	170 (99%)	2 (1%)	63	84
9	W	172/173 (99%)	170 (99%)	2 (1%)	63	84
10	J	175/175 (100%)	174 (99%)	1 (1%)	78	91
10	X	175/175 (100%)	173 (99%)	2 (1%)	65	85
11	K	169/169 (100%)	165 (98%)	4 (2%)	43	72
11	Y	169/169 (100%)	165 (98%)	4 (2%)	43	72
12	L	185/185 (100%)	182 (98%)	3 (2%)	55	80
12	Z	185/185 (100%)	182 (98%)	3 (2%)	55	80
13	M	199/199 (100%)	192 (96%)	7 (4%)	32	61
13	a	199/199 (100%)	192 (96%)	7 (4%)	32	61
14	N	162/162 (100%)	161 (99%)	1 (1%)	78	91
14	b	162/162 (100%)	161 (99%)	1 (1%)	78	91
All	All	5332/5522 (97%)	5228 (98%)	104 (2%)	48	76

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	Q	160	GLN
5	S	231	LYS
13	a	146	PHE
3	Q	171	GLU
4	R	214	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 177 such sidechains are listed below:

Mol	Chain	Res	Type
4	R	210	GLN
8	V	30	ASN
5	S	68	HIS
6	T	39	ASN
9	W	63	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 oligosaccharides 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	1G1	V	301	8	47,47,47	1.33	3 (6%)	58,63,63	1.58	4 (6%)
15	1G1	Y	301	11	47,47,47	1.23	3 (6%)	58,63,63	1.64	4 (6%)
15	1G1	K	301	11	47,47,47	1.21	3 (6%)	58,63,63	1.52	4 (6%)
15	1G1	H	301	8	47,47,47	1.26	3 (6%)	58,63,63	1.58	5 (8%)

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	1G1	V	301	8	-	9/51/51/51	0/2/2/2
15	1G1	Y	301	11	-	15/51/51/51	0/2/2/2
15	1G1	K	301	11	-	15/51/51/51	0/2/2/2
15	1G1	H	301	8	-	15/51/51/51	0/2/2/2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	V	301	1G1	C46-C47	4.79	1.57	1.52
15	H	301	1G1	C46-C47	4.32	1.56	1.52
15	Y	301	1G1	C46-C47	3.85	1.56	1.52
15	K	301	1G1	C46-C47	3.78	1.56	1.52
15	V	301	1G1	C47-S48	3.21	1.82	1.78

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	Y	301	1G1	C8-N9-N10	9.83	125.06	115.36
15	V	301	1G1	C8-N9-N10	8.64	123.88	115.36
15	K	301	1G1	C8-N9-N10	8.61	123.85	115.36
15	H	301	1G1	C8-N9-N10	8.10	123.35	115.36
15	V	301	1G1	O49-S48-C47	4.60	111.65	108.33

There are no chirality outliers.

5 of 54 torsion outliers are listed below:

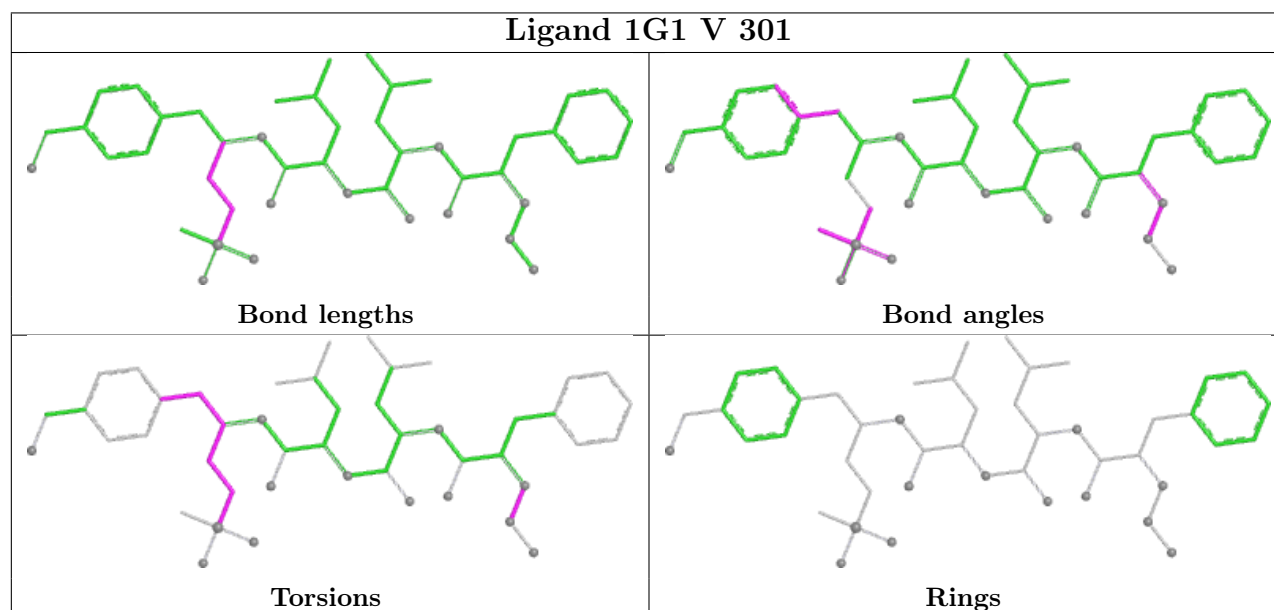
Mol	Chain	Res	Type	Atoms
15	H	301	1G1	C46-C47-S48-O49
15	K	301	1G1	C7-C8-N9-N10
15	K	301	1G1	N11-N10-N9-C8
15	K	301	1G1	C46-C47-S48-O50
15	K	301	1G1	C46-C47-S48-O49

There are no ring outliers.

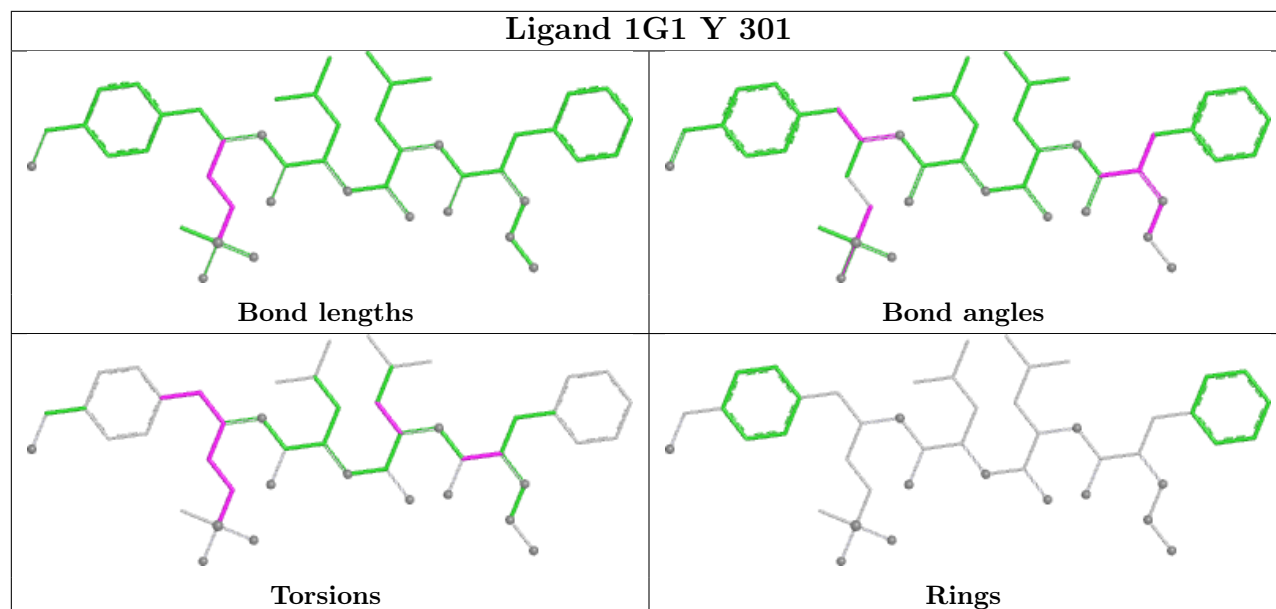
4 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	V	301	1G1	5	0
15	Y	301	1G1	6	0
15	K	301	1G1	2	0
15	H	301	1G1	2	0

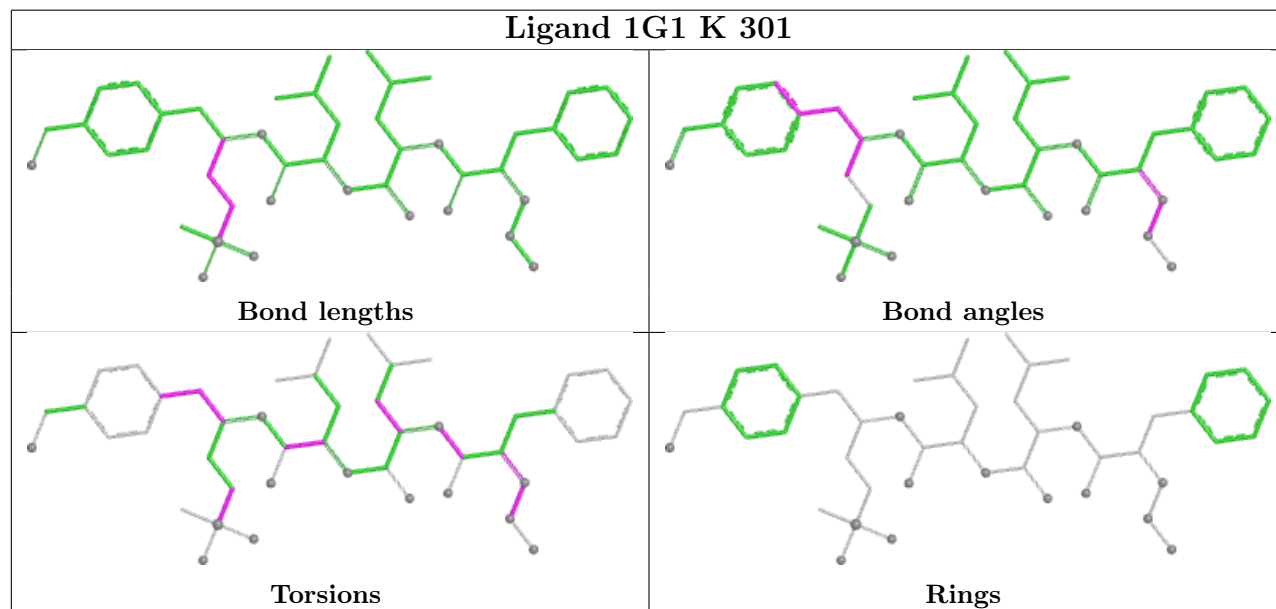
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

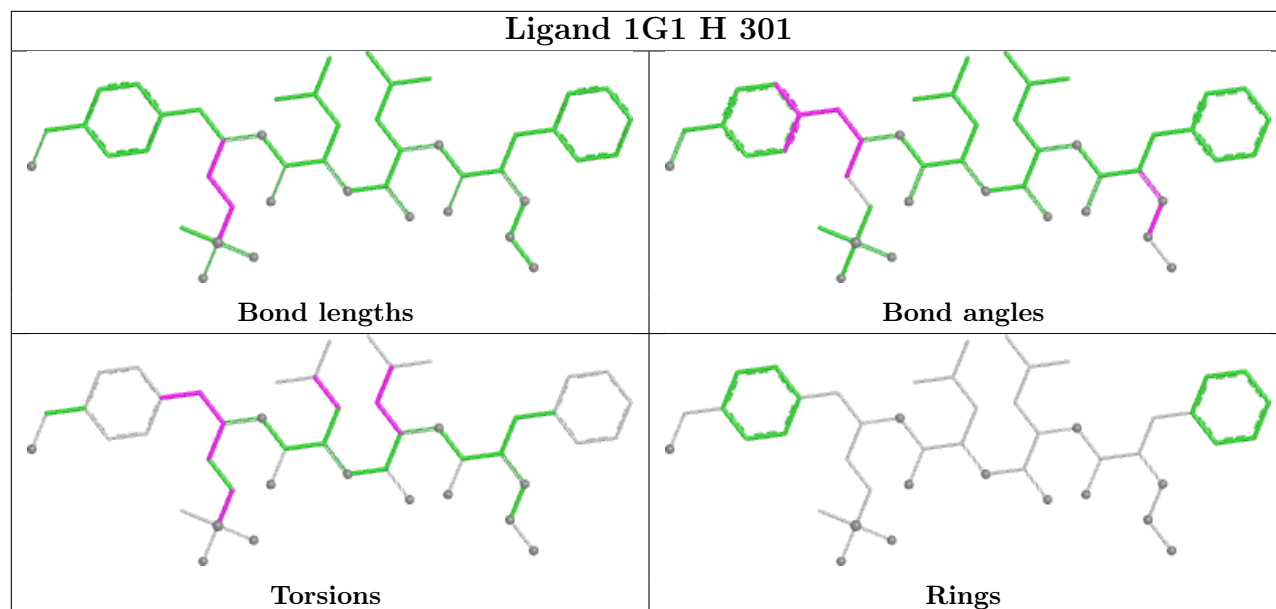


Ligand 1G1 Y 301



Ligand 1G1 K 301





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, 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%)	0.05	4 (1%) 70 68	61, 76, 101, 119	0
1	O	250/250 (100%)	0.07	8 (3%) 50 46	63, 81, 106, 123	0
2	B	244/258 (94%)	0.28	15 (6%) 27 24	58, 80, 121, 136	0
2	P	244/258 (94%)	0.46	21 (8%) 16 13	62, 82, 117, 134	0
3	C	241/254 (94%)	0.19	5 (2%) 63 61	59, 82, 123, 156	0
3	Q	241/254 (94%)	0.50	22 (9%) 15 12	68, 97, 148, 191	0
4	D	242/260 (93%)	0.18	9 (3%) 45 41	61, 82, 108, 127	0
4	R	242/260 (93%)	0.27	9 (3%) 45 41	65, 89, 120, 138	0
5	E	233/234 (99%)	0.19	3 (1%) 75 73	66, 90, 116, 125	0
5	S	233/234 (99%)	0.41	11 (4%) 36 33	68, 94, 125, 135	0
6	F	244/288 (84%)	0.09	4 (1%) 70 68	63, 81, 113, 137	0
6	T	244/288 (84%)	0.13	3 (1%) 76 75	63, 82, 118, 140	0
7	G	243/252 (96%)	-0.01	1 (0%) 88 87	60, 76, 99, 129	0
7	U	243/252 (96%)	0.06	3 (1%) 76 75	61, 75, 97, 129	0
8	H	222/232 (95%)	0.04	1 (0%) 87 86	59, 70, 86, 104	0
8	V	222/232 (95%)	0.02	1 (0%) 87 86	58, 69, 86, 109	0
9	I	204/205 (99%)	-0.20	1 (0%) 87 86	55, 67, 81, 88	0
9	W	204/205 (99%)	-0.08	2 (0%) 79 78	58, 67, 84, 94	0
10	J	198/198 (100%)	0.01	5 (2%) 58 55	54, 67, 87, 123	0
10	X	198/198 (100%)	0.02	7 (3%) 47 43	58, 68, 84, 124	0
11	K	212/212 (100%)	0.16	12 (5%) 29 26	54, 68, 89, 95	0
11	Y	212/212 (100%)	0.15	9 (4%) 40 37	57, 70, 92, 100	0
12	L	222/222 (100%)	-0.06	3 (1%) 73 72	55, 68, 96, 104	0
12	Z	222/222 (100%)	0.01	1 (0%) 87 86	56, 70, 96, 103	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
13	M	233/233 (100%)	-0.13	3 (1%) 75 73	53, 69, 86, 101	0
13	a	233/233 (100%)	-0.14	2 (0%) 81 80	55, 68, 82, 96	0
14	N	196/196 (100%)	-0.14	0 100 100	57, 65, 82, 89	0
14	b	196/196 (100%)	-0.19	1 (0%) 87 86	55, 64, 82, 89	0
All	All	6368/6588 (96%)	0.09	166 (2%) 57 54	53, 74, 113, 191	0

The worst 5 of 166 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	50	LEU	9.6
11	Y	104	TYR	7.7
4	D	119	ALA	7.5
11	K	104	TYR	7.2
4	R	119	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 oligosaccharides 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.

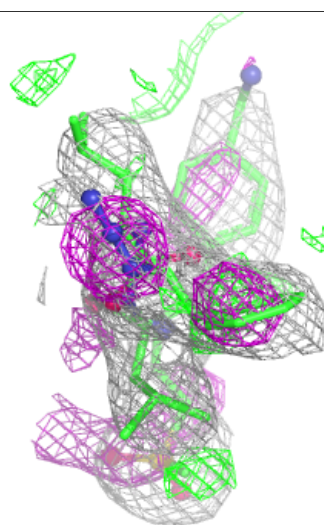
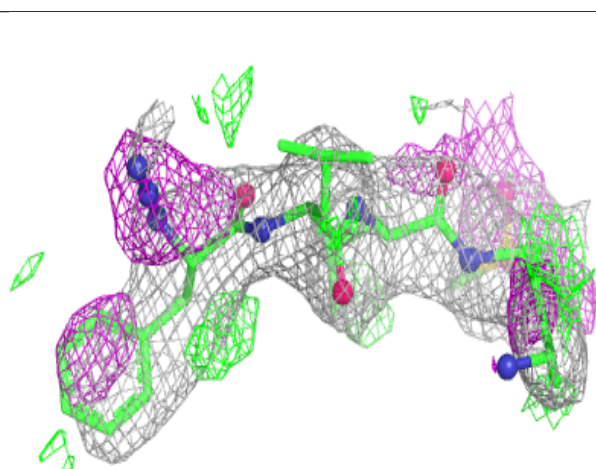
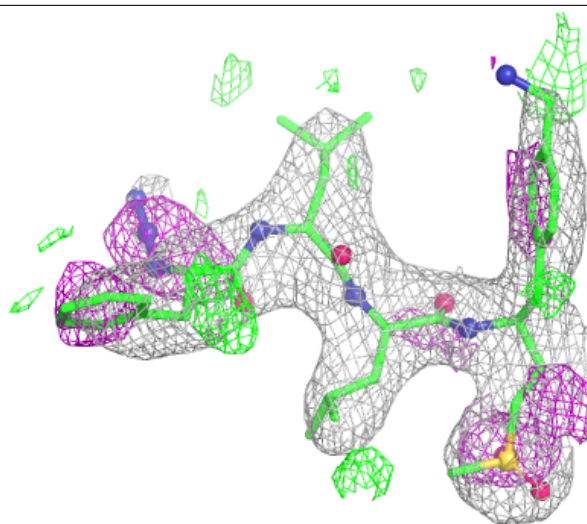
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
15	1G1	H	301	46/46	0.84	0.15	65,70,83,83	0
15	1G1	V	301	46/46	0.85	0.15	65,72,84,84	0
15	1G1	Y	301	46/46	0.87	0.13	55,58,65,65	0
15	1G1	K	301	46/46	0.90	0.12	52,56,69,70	0

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.

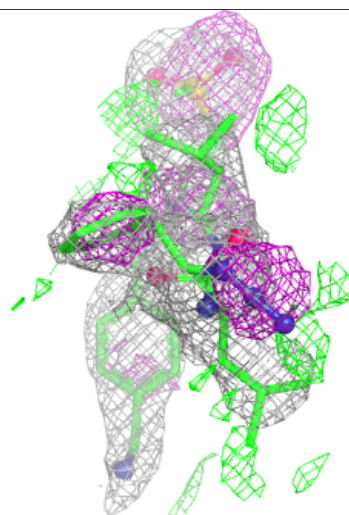
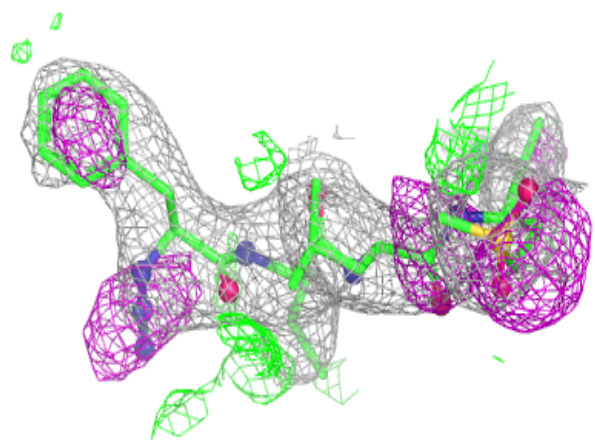
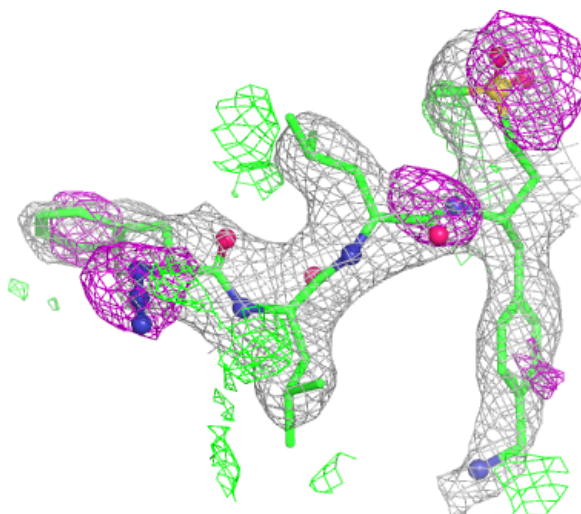
Electron density around 1G1 H 301:

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



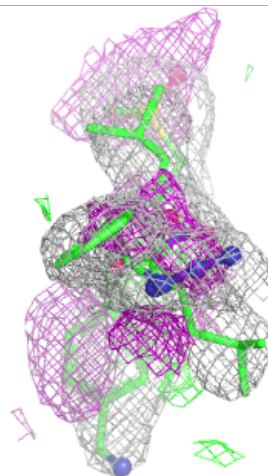
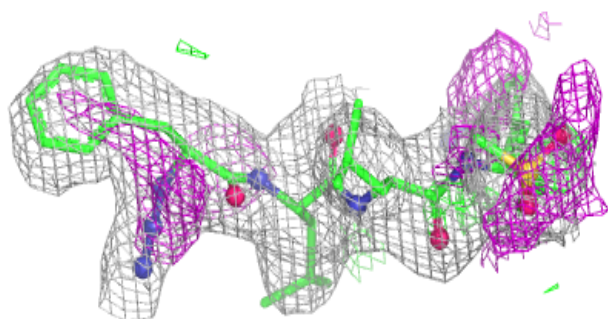
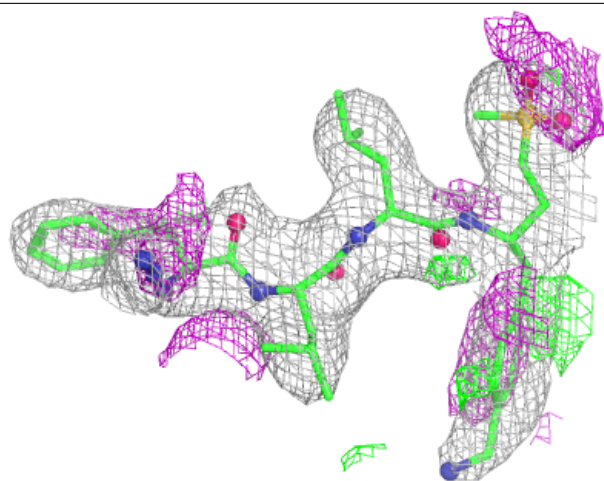
Electron density around 1G1 V 301:

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



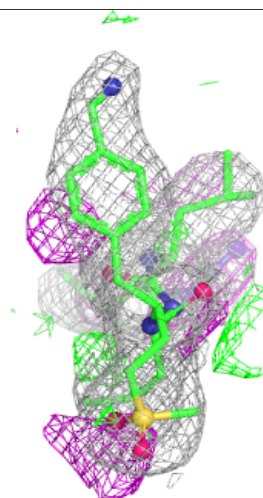
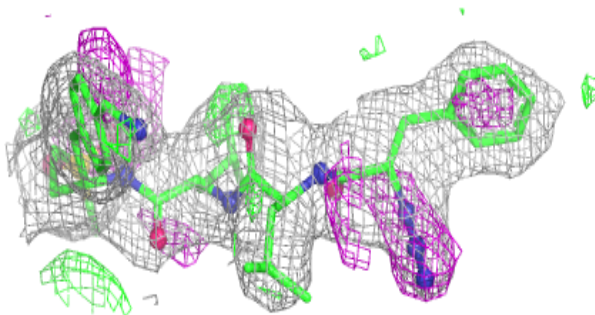
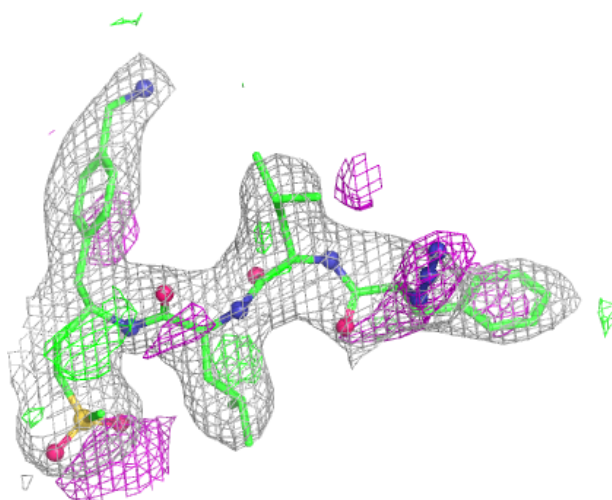
Electron density around 1G1 Y 301:

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



Electron density around 1G1 K 301:

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



6.5 Other polymers ⓘ

There are no such residues in this entry.