wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 11, 2023 – 12:15 AM EDT

PDB ID	:	4J70
Title	:	Yeast 20S proteasome in complex with the belactosin derivative 3e
Authors	:	Kawamura, S.; Unno, Y.; List, A.; Tanaka, M.; Sasaki, T.; Arisawa, M.; Asai, A.; Groll, M.; Shuto, S.
Deposited on	:	2013-02-12
Resolution	:	2.80 Å(reported)

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

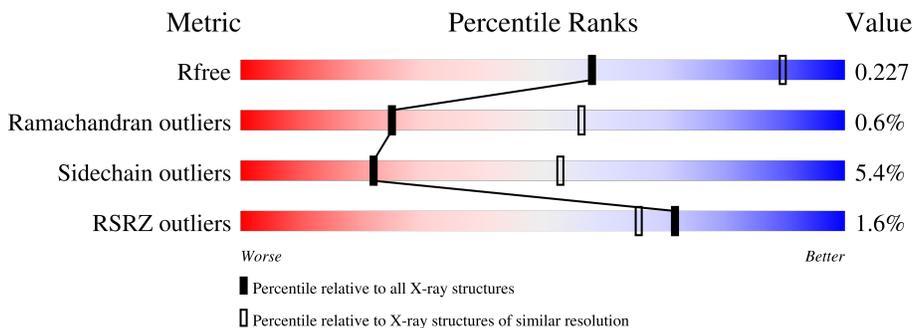|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 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	Percentile Ranks	Value
Rfree		0.227
Ramachandran outliers		0.6%
Sidechain outliers		5.4%
RSRZ outliers		1.6%

Worse   *Better*

▌Percentile relative to all X-ray structures

◨ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (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	.% 97% •
1	O	250	.% 98% •
2	B	258	3% 88% 7% 5%
2	P	258	4% 88% 7% 5%
3	C	254	4% 89% 6% 5%
3	Q	254	4% 88% 7% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	260	3% 87% 6% 7%
4	R	260	3% 87% 6% 7%
5	E	234	2% 89% 10%
5	S	234	2% 88% 11%
6	F	288	2% 80% 5% 15%
6	T	288	2% 80% 5% 15%
7	G	252	2% 92%
7	U	252	% 91% 6%
8	H	232	90% 6%
8	V	232	% 91% 5%
9	I	205	97%
9	W	205	98%
10	J	198	2% 96%
10	X	198	2% 95%
11	K	212	96%
11	Y	212	96%
12	L	222	96%
12	Z	222	95%
13	M	233	96%
13	a	233	96%
14	N	196	97%
14	b	196	95% 5%

2 Entry composition ⓘ

There are 16 unique types of molecules in this entry. The entry contains 50956 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 1795	C 1129	N 312	O 350	S 4	0	0	0
5	S	233	Total 1795	C 1129	N 312	O 350	S 4	0	0	0

- Molecule 6 is a protein called Proteasome component C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	244	Total 1896	C 1205	N 330	O 357	S 4	0	0	0
6	T	244	Total 1896	C 1205	N 330	O 357	S 4	0	0	0

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

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

- Molecule 8 is a protein called Proteasome component PUP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	222	Total 1684	C 1061	N 293	O 323	S 7	0	0	0
8	V	222	Total 1684	C 1061	N 293	O 323	S 7	0	0	0

- Molecule 9 is a protein called Proteasome component PUP3.

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

- Molecule 10 is a protein called Proteasome component C11.

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

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	Total 1644	C 1045	N 280	O 312	S 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	Total 1757	C 1115	N 303	O 335	S 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	a	233	Total 1824	C 1154	N 312	O 351	S 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	b	196	Total 1512	C 955	N 250	O 300	S 7	0	0	0

- Molecule 15 is benzyl [(2S)-1-({(2R)-1-[(1S,2S)-2-{[(2R,3S,4S)-3-formyl-2-hydroxy-4-methylhexanoyl]amino}cyclopropyl]-4-phenylbutan-2-yl}amino)-1-oxopropan-2-yl]carbamate (three-letter code: 1KR) (formula: $C_{32}H_{43}N_3O_6$).
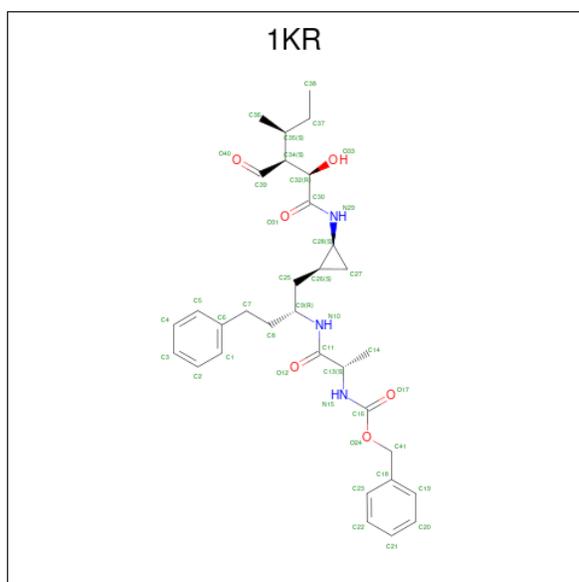
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
15	K	1	Total 41	C 32	N 3	O 6	0	0
15	Y	1	Total 41	C 32	N 3	O 6	0	0

- Molecule 16 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	55	Total 55	O 55	0	0
16	B	37	Total 37	O 37	0	0
16	C	45	Total 45	O 45	0	0
16	D	39	Total 39	O 39	0	0
16	E	22	Total 22	O 22	0	0
16	F	46	Total 46	O 46	0	0
16	G	63	Total 63	O 63	0	0
16	H	52	Total 52	O 52	0	0
16	I	67	Total 67	O 67	0	0
16	J	55	Total 55	O 55	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	K	40	Total 40	O 40	0	0
16	L	57	Total 57	O 57	0	0
16	M	75	Total 75	O 75	0	0
16	N	54	Total 54	O 54	0	0
16	O	31	Total 31	O 31	0	0
16	P	31	Total 31	O 31	0	0
16	Q	26	Total 26	O 26	0	0
16	R	30	Total 30	O 30	0	0
16	S	23	Total 23	O 23	0	0
16	T	43	Total 43	O 43	0	0
16	U	59	Total 59	O 59	0	0
16	V	51	Total 51	O 51	0	0
16	W	62	Total 62	O 62	0	0
16	X	46	Total 46	O 46	0	0
16	Y	49	Total 49	O 49	0	0
16	Z	51	Total 51	O 51	0	0
16	a	72	Total 72	O 72	0	0
16	b	55	Total 55	O 55	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.

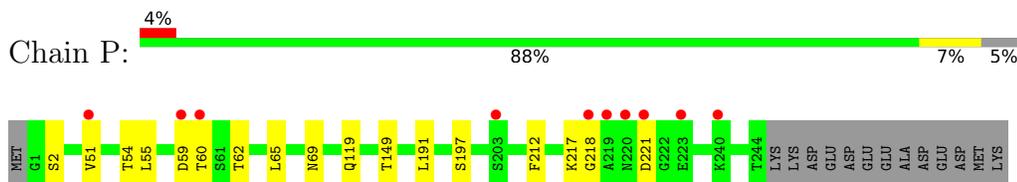• Molecule 1: Proteasome component Y7
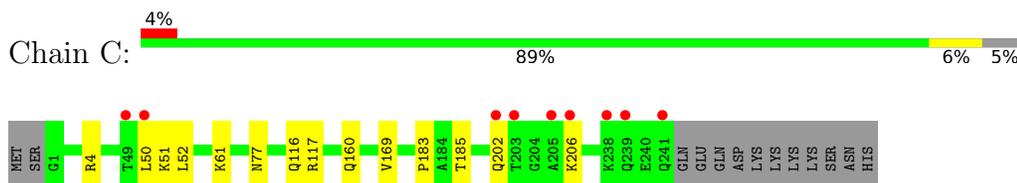
Chain A:

• 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:

MET SER G1 R4 K37 S48 T49 L50 K51 L52 K61 N77 R117 Q147 Q160 V169 P183 A184 T185 Q202 T203 G204 A205 K206 V213 S224 Q229 K238 Q239 E240 Q241 GLN GLN ASP LYS LYS LYS SER ASN HIS

• Molecule 4: Proteasome component PUP2

Chain D:

MET PHE LEU THR ARG SER GLU TYR D1 V4 S5 L20 K44 S48 L51 D54 G118 A119 S120 G121 E122 E123 R124 L125 M126 E169 L176 L193 I214 L235 E242 SER PRO GLU GLU ALA ASP VAL GLU MET SER

• Molecule 4: Proteasome component PUP2

Chain R:

MET PHE LEU THR ARG SER GLU TYR D1 V4 S5 K44 S48 L51 G118 A119 S120 G121 E122 E123 R124 L125 M126 T155 E169 L176 L193 L211 I214 E242 SER PRO GLU GLU ALA ASP VAL GLU MET SER

• Molecule 5: Proteasome component PRE5

Chain E:

MET F1 R2 N3 D8 T9 V10 K29 E54 L55 L71 N92 R106 Q116 T169 R173 T174 N184 L188 Q198 R201 D202 E203 S204 L205 T206 V207 S211 T219 E227 K231 Y232 I233

• Molecule 5: Proteasome component PRE5

Chain S:

MET F1 R2 D8 T9 V10 K29 E54 L55 L71 N92 R106 C112 Q116 T169 R173 T174 N184 P185 L188 Q198 R201 D202 E203 S204 L205 T206 V207 S211 T219 T222 E227 K231 Y232 I233

• Molecule 6: Proteasome component C1

Chain F:

MET THR SER ILE G1 T32 N39 R68 Q117 N123 K139 R165 Q166 S167 L172 H178 E181 R186 E201 D202 N203 K204 W214 K228 N244 GLY ASP ASP ASP GLU GLU ASP GLU ASP SER ASP ASN VAL MET SER SER ASP GLU ASN

ALA PRO VAL ALA THR ASN ALA ASN ALA THR THR ASP GLN GLU GLY ASP ILE HIS LEU GLU

• Molecule 6: Proteasome component C1

Chain T:

MET THR SER ILE **G1 T2** T32 N39 Q117 N123 K139 R165 Q166 S167 L172 P180 E181 R186 E201 D202 N203 W214 K228 G229 D230 K241 N244 GLY ASP ASP ASP GLU ASP GLU ASP ASP SER ASP ASN VAL MET SER SER ASP ASP GLU

ASN ALA PRO VAL ALA THR ASN ALA ASN ALA THR THR ASP GLN GLU GLY ASP ILE HIS LEU GLU

• Molecule 7: Proteasome component C7-alpha

Chain G:

MET SER GLY ALA ALA ALA SER ALA **A1 G2** L34 V44 I45 R68 N83 L115 S116 Q117 K221 R235 L236 **Q242 D243**

• Molecule 7: Proteasome component C7-alpha

Chain U:

MET SER GLY ALA ALA ALA SER ALA **A1 G2** L34 A35 V36 V44 I45 R68 N83 L115 S116 Q117 Y124 M125 R126 T207 K221 R235 L236 D243

• Molecule 8: Proteasome component PUP1

Chain H:

T1 N30 L34 V55 T56 L68 S108 V121 S131 Q144 K153 G181 S216 I220 C221 D222 ILE GLN GLU GLU GLN VAL ASP ILE THR ALA

• Molecule 8: Proteasome component PUP1

Chain V:

T1 N30 C31 L34 V55 T56 L68 V121 Q144 K153 S216 I220 C221 D222 ILE GLN GLU GLU GLN VAL ASP ILE THR ALA

• Molecule 9: Proteasome component PUP3

Chain I:

MET **S1** L28 N37 K133 L171 K191 V195 D204

• Molecule 9: Proteasome component PUP3

Chain W:

MET **S1** N37 K133 L171 V195 D204

• Molecule 10: Proteasome component C11

Chain J:

2%
96%

M1 V9 I25 T35 T53 I69 R70 E71 K110 Q196 A197 Q198

• Molecule 10: Proteasome component C11

Chain X:

2%
95%

M1 I4 V9 T35 T53 I69 R70 E71 Q78 R93 K110 D194 F195 Q196 A197 Q198

• Molecule 11: Proteasome component PRE2

Chain K:

96%

T1 L4 Q9 T21 K32 L65 R73 V87 Y104 K107 G212

• Molecule 11: Proteasome component PRE2

Chain Y:

96%

T1 L4 Q9 T21 K32 R73 V87 Y104 K107 G212

• Molecule 12: Proteasome component C5

Chain L:

96%

Q1 L23 D48 N49 D81 H108 T109 D126 K173 Y174 D222

• Molecule 12: Proteasome component C5

Chain Z:

95%

Q1 G10 L23 S34 D48 N49 N80 H108 T109 D126 K173 Y174 D222

• Molecule 13: Proteasome component PRE4

Chain M:

96%

T1 Q2 Q3 T9 D47 N48 R103 R104 F146 R161 V170 R187 K226 I233

• Molecule 13: Proteasome component PRE4

Chain a:

96%

T1 Q2 Q3 N48 I57 R104 R128 F146 R161 V170 R187 K226 I233

• Molecule 14: Proteasome component PRE3

Chain N:

97%

T1 R36 K83 E88 K105 V119 L178 L196

• Molecule 14: Proteasome component PRE3

Chain b:

95%   5%

T1 T21 R36 S48 K83 E88 K105 L115 V119 R185 L196

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	137.08Å 301.71Å 146.23Å 90.00° 113.69° 90.00°	Depositor
Resolution (Å)	15.00 – 2.80 15.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.9 (15.00-2.80) 98.9 (15.00-2.80)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.79 (at 2.81Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.172 , 0.224 0.175 , 0.227	Depositor DCC
R_{free} test set	13052 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	50.2	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 55.6	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	0.95	EDS
Total number of atoms	50956	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.46% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 1KR

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	0.69	0/1952	0.81	1/2642 (0.0%)
1	O	0.64	0/1952	0.79	0/2642
2	B	0.67	0/1934	0.80	0/2618
2	P	0.66	0/1934	0.82	0/2618
3	C	0.65	0/1919	0.82	2/2598 (0.1%)
3	Q	0.61	0/1919	0.79	3/2598 (0.1%)
4	D	0.65	0/1886	0.84	2/2541 (0.1%)
4	R	0.65	0/1886	0.86	3/2541 (0.1%)
5	E	0.64	0/1823	0.82	0/2463
5	S	0.64	0/1823	0.80	1/2463 (0.0%)
6	F	0.68	0/1936	0.82	3/2614 (0.1%)
6	T	0.65	0/1936	0.81	1/2614 (0.0%)
7	G	0.72	0/1959	0.82	0/2652
7	U	0.69	1/1959 (0.1%)	0.80	1/2652 (0.0%)
8	H	0.71	0/1715	0.83	0/2326
8	V	0.70	0/1715	0.82	0/2326
9	I	0.76	0/1611	0.84	1/2174 (0.0%)
9	W	0.72	0/1611	0.83	0/2174
10	J	0.72	0/1613	0.87	0/2173
10	X	0.72	0/1613	0.87	2/2173 (0.1%)
11	K	0.72	0/1681	0.85	1/2274 (0.0%)
11	Y	0.70	0/1681	0.83	0/2274
12	L	0.73	0/1795	0.84	1/2420 (0.0%)
12	Z	0.72	0/1795	0.83	1/2420 (0.0%)
13	M	0.70	0/1855	0.89	2/2514 (0.1%)
13	a	0.71	0/1855	0.90	2/2514 (0.1%)
14	N	0.75	0/1541	0.84	0/2087
14	b	0.70	0/1541	0.84	1/2087 (0.0%)
All	All	0.69	1/50440 (0.0%)	0.83	28/68192 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if

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

Mol	Chain	#Chirality outliers	#Planarity outliers
8	H	0	1
12	L	0	1
12	Z	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	U	124	TYR	CE1-CZ	5.06	1.45	1.38

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	a	128	ARG	NE-CZ-NH1	7.11	123.85	120.30
11	K	65	LEU	CB-CG-CD2	-6.93	99.22	111.00
13	M	104	ARG	NE-CZ-NH1	6.61	123.61	120.30
3	C	50	LEU	CA-CB-CG	6.26	129.69	115.30
6	T	186	ARG	NE-CZ-NH1	6.21	123.40	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
8	H	181	GLY	Peptide
12	L	174	TYR	Peptide
12	Z	174	TYR	Peptide

5.2   Too-close contacts ⓘ

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3   Torsion angles ⓘ

5.3.1   Protein backbone ⓘ

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%)	240 (97%)	7 (3%)	1 (0%)	34	66
1	O	248/250 (99%)	238 (96%)	9 (4%)	1 (0%)	34	66
2	B	242/258 (94%)	226 (93%)	13 (5%)	3 (1%)	13	39
2	P	242/258 (94%)	225 (93%)	14 (6%)	3 (1%)	13	39
3	C	239/254 (94%)	232 (97%)	4 (2%)	3 (1%)	12	36
3	Q	239/254 (94%)	231 (97%)	5 (2%)	3 (1%)	12	36
4	D	240/260 (92%)	225 (94%)	13 (5%)	2 (1%)	19	49
4	R	240/260 (92%)	229 (95%)	9 (4%)	2 (1%)	19	49
5	E	231/234 (99%)	217 (94%)	11 (5%)	3 (1%)	12	36
5	S	231/234 (99%)	216 (94%)	11 (5%)	4 (2%)	9	29
6	F	242/288 (84%)	230 (95%)	12 (5%)	0	100	100
6	T	242/288 (84%)	228 (94%)	14 (6%)	0	100	100
7	G	241/252 (96%)	233 (97%)	8 (3%)	0	100	100
7	U	241/252 (96%)	232 (96%)	9 (4%)	0	100	100
8	H	220/232 (95%)	211 (96%)	8 (4%)	1 (0%)	29	61
8	V	220/232 (95%)	212 (96%)	8 (4%)	0	100	100
9	I	202/205 (98%)	194 (96%)	8 (4%)	0	100	100
9	W	202/205 (98%)	194 (96%)	8 (4%)	0	100	100
10	J	196/198 (99%)	190 (97%)	5 (3%)	1 (0%)	29	61
10	X	196/198 (99%)	191 (97%)	4 (2%)	1 (0%)	29	61
11	K	210/212 (99%)	204 (97%)	5 (2%)	1 (0%)	29	61
11	Y	210/212 (99%)	205 (98%)	4 (2%)	1 (0%)	29	61
12	L	220/222 (99%)	203 (92%)	15 (7%)	2 (1%)	17	46
12	Z	220/222 (99%)	204 (93%)	14 (6%)	2 (1%)	17	46
13	M	231/233 (99%)	219 (95%)	11 (5%)	1 (0%)	34	66
13	a	231/233 (99%)	222 (96%)	9 (4%)	0	100	100
14	N	194/196 (99%)	187 (96%)	7 (4%)	0	100	100
14	b	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
All	All	6312/6588 (96%)	6027 (96%)	250 (4%)	35 (1%)	25	56

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
3	C	52	LEU
2	P	51	VAL
3	Q	52	LEU
2	B	218	GLY

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%)	204 (98%)	5 (2%)	49	81
1	O	209/209 (100%)	204 (98%)	5 (2%)	49	81
2	B	203/216 (94%)	188 (93%)	15 (7%)	13	37
2	P	203/216 (94%)	189 (93%)	14 (7%)	15	41
3	C	213/226 (94%)	204 (96%)	9 (4%)	30	63
3	Q	213/226 (94%)	201 (94%)	12 (6%)	21	51
4	D	198/215 (92%)	186 (94%)	12 (6%)	18	48
4	R	198/215 (92%)	187 (94%)	11 (6%)	21	51
5	E	192/193 (100%)	169 (88%)	23 (12%)	5	15
5	S	192/193 (100%)	167 (87%)	25 (13%)	4	13
6	F	201/239 (84%)	187 (93%)	14 (7%)	15	40
6	T	201/239 (84%)	187 (93%)	14 (7%)	15	40
7	G	207/210 (99%)	197 (95%)	10 (5%)	25	58
7	U	207/210 (99%)	195 (94%)	12 (6%)	20	50
8	H	181/190 (95%)	168 (93%)	13 (7%)	14	38
8	V	181/190 (95%)	169 (93%)	12 (7%)	16	44
9	I	172/173 (99%)	167 (97%)	5 (3%)	42	76
9	W	172/173 (99%)	168 (98%)	4 (2%)	50	82
10	J	175/175 (100%)	169 (97%)	6 (3%)	37	71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	X	175/175 (100%)	168 (96%)	7 (4%)	31	65
11	K	169/169 (100%)	160 (95%)	9 (5%)	22	54
11	Y	169/169 (100%)	161 (95%)	8 (5%)	26	59
12	L	185/185 (100%)	180 (97%)	5 (3%)	44	78
12	Z	185/185 (100%)	178 (96%)	7 (4%)	33	67
13	M	199/199 (100%)	191 (96%)	8 (4%)	31	65
13	a	199/199 (100%)	191 (96%)	8 (4%)	31	65
14	N	162/162 (100%)	156 (96%)	6 (4%)	34	68
14	b	162/162 (100%)	154 (95%)	8 (5%)	25	57
All	All	5332/5522 (97%)	5045 (95%)	287 (5%)	22	53

5 of 287 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	U	235	ARG
14	b	185	ARG
8	V	68	LEU
11	Y	73	ARG
8	H	153	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 168 such sidechains are listed below:

Mol	Chain	Res	Type
6	T	19	GLN
11	Y	85	ASN
6	T	123	ASN
8	V	66	HIS
12	Z	55	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	1KR	K	301	11	43,43,43	1.66	5 (11%)	51,57,57	1.86	11 (21%)
15	1KR	Y	301	11	43,43,43	1.69	7 (16%)	51,57,57	1.87	10 (19%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
15	1KR	K	301	11	-	18/46/51/51	0/3/3/3
15	1KR	Y	301	11	-	16/46/51/51	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	Y	301	1KR	C5-C6	5.62	1.51	1.38
15	K	301	1KR	C5-C6	5.22	1.50	1.38
15	Y	301	1KR	O40-C39	3.61	1.34	1.19
15	K	301	1KR	C27-C28	3.46	1.54	1.49
15	Y	301	1KR	C27-C28	3.35	1.54	1.49

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	Y	301	1KR	C28-N29-C30	-6.75	111.15	123.07

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	K	301	1KR	C28-N29-C30	-5.86	112.71	123.07
15	K	301	1KR	C14-C13-C11	4.90	119.46	110.14
15	Y	301	1KR	C26-C25-C9	-4.66	107.84	115.56
15	Y	301	1KR	C13-N15-C16	4.08	128.14	120.49

There are no chirality outliers.

5 of 34 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	K	301	1KR	O17-C16-O24-C41
15	K	301	1KR	N15-C16-O24-C41
15	K	301	1KR	O24-C16-N15-C13
15	K	301	1KR	O12-C11-C13-C14
15	K	301	1KR	N10-C11-C13-C14

There are no ring outliers.

No monomer is involved in short contacts.

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 1KR K 301

Bond lengths

Bond angles

Torsions

Rings

Ligand 1KR Y 301

Bond lengths

Bond angles

Torsions

Rings
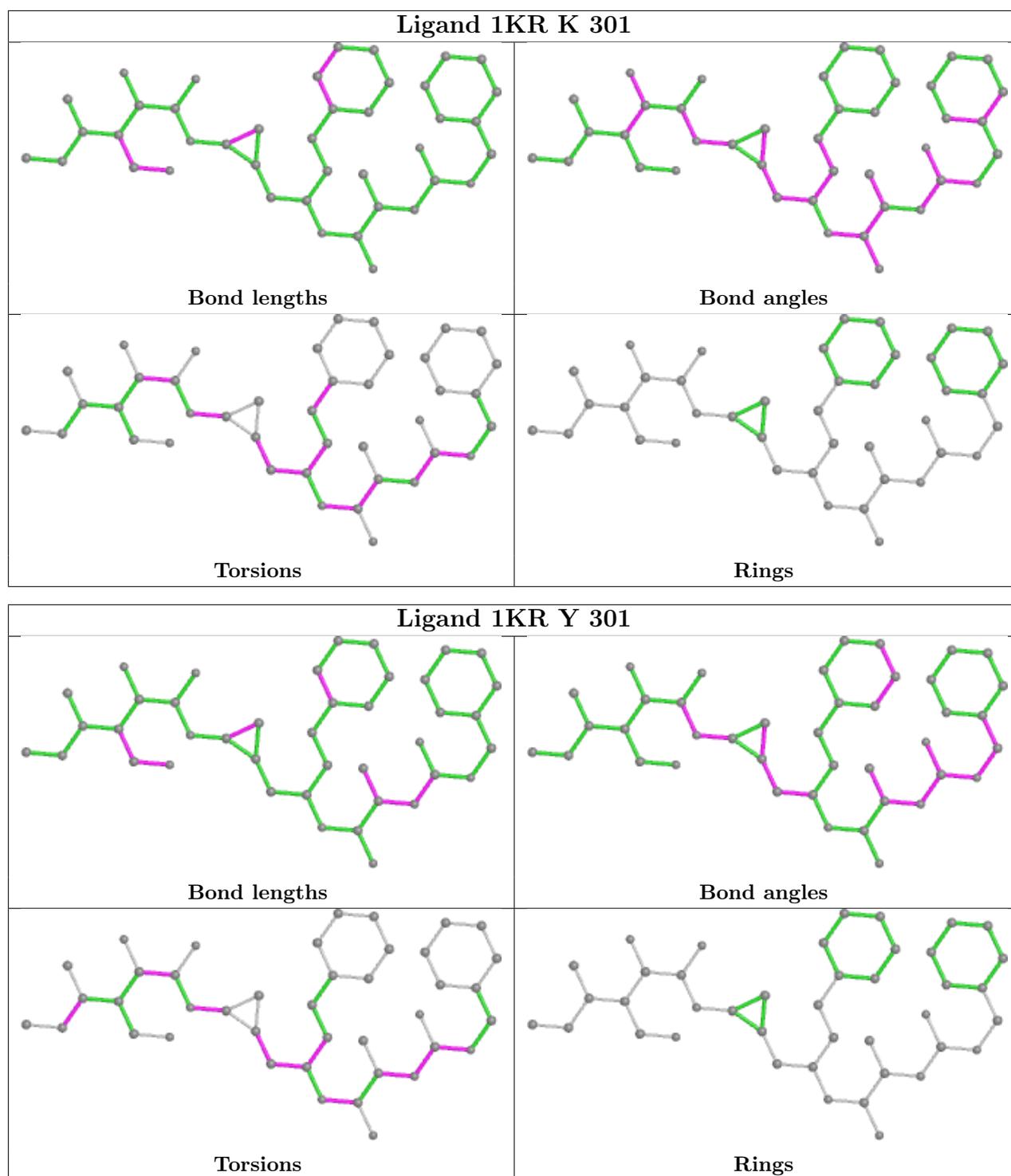
5.7 Other polymers ⓘ

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%)	-0.77	3 (1%)	79	73	28, 43, 74, 146	0
1	O	250/250 (100%)	-0.68	2 (0%)	86	81	30, 51, 87, 142	0
2	B	244/258 (94%)	-0.58	8 (3%)	46	36	27, 48, 104, 154	0
2	P	244/258 (94%)	-0.53	10 (4%)	37	27	33, 52, 105, 156	0
3	C	241/254 (94%)	-0.49	9 (3%)	41	31	28, 52, 111, 147	0
3	Q	241/254 (94%)	-0.33	11 (4%)	32	22	33, 58, 125, 153	0
4	D	242/260 (93%)	-0.55	9 (3%)	41	31	32, 51, 89, 185	0
4	R	242/260 (93%)	-0.51	8 (3%)	46	36	34, 54, 94, 195	0
5	E	233/234 (99%)	-0.58	5 (2%)	63	54	36, 55, 87, 152	0
5	S	233/234 (99%)	-0.52	5 (2%)	63	54	35, 58, 97, 146	0
6	F	244/288 (84%)	-0.71	5 (2%)	65	56	30, 47, 92, 128	0
6	T	244/288 (84%)	-0.59	6 (2%)	57	47	31, 52, 98, 134	0
7	G	243/252 (96%)	-0.74	4 (1%)	72	66	24, 43, 80, 157	0
7	U	243/252 (96%)	-0.70	3 (1%)	79	73	29, 46, 83, 149	0
8	H	222/232 (95%)	-0.87	1 (0%)	91	88	26, 40, 65, 99	0
8	V	222/232 (95%)	-0.84	2 (0%)	84	80	29, 42, 70, 112	0
9	I	204/205 (99%)	-0.98	1 (0%)	91	88	26, 40, 64, 99	0
9	W	204/205 (99%)	-0.95	1 (0%)	91	88	27, 41, 66, 115	0
10	J	198/198 (100%)	-0.79	3 (1%)	73	68	28, 42, 68, 184	0
10	X	198/198 (100%)	-0.76	4 (2%)	65	56	30, 43, 71, 162	0
11	K	212/212 (100%)	-0.91	0	100	100	27, 42, 61, 76	0
11	Y	212/212 (100%)	-0.90	0	100	100	29, 43, 64, 80	0
12	L	222/222 (100%)	-0.87	1 (0%)	91	88	27, 44, 73, 117	0
12	Z	222/222 (100%)	-0.88	1 (0%)	91	88	28, 43, 70, 108	0

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
13	M	233/233 (100%)	-0.89	1 (0%)	92	91	29, 44, 61, 83	0
13	a	233/233 (100%)	-0.86	0	100	100	26, 43, 60, 77	0
14	N	196/196 (100%)	-0.92	0	100	100	27, 39, 63, 90	0
14	b	196/196 (100%)	-0.92	0	100	100	24, 40, 66, 90	0
All	All	6368/6588 (96%)	-0.73	103 (1%)	72	66	24, 46, 87, 195	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
10	J	197	ALA	9.9
2	P	220	ASN	9.2
4	D	119	ALA	9.1
7	U	1	ALA	8.6
4	R	121	GLY	8.6

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.
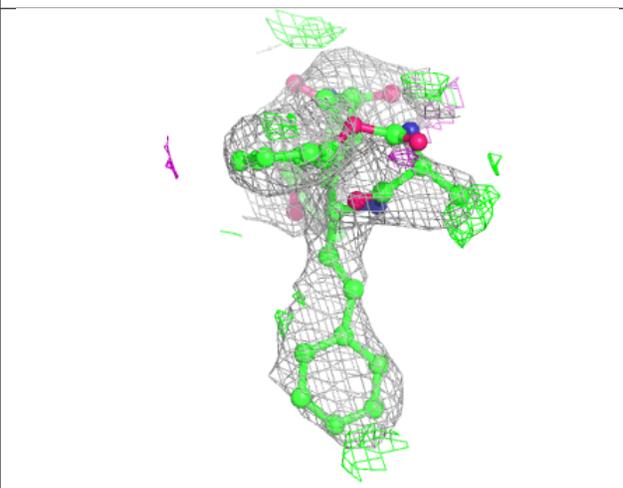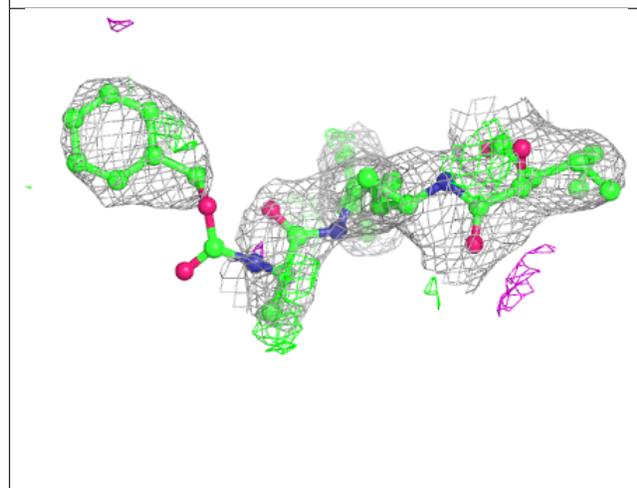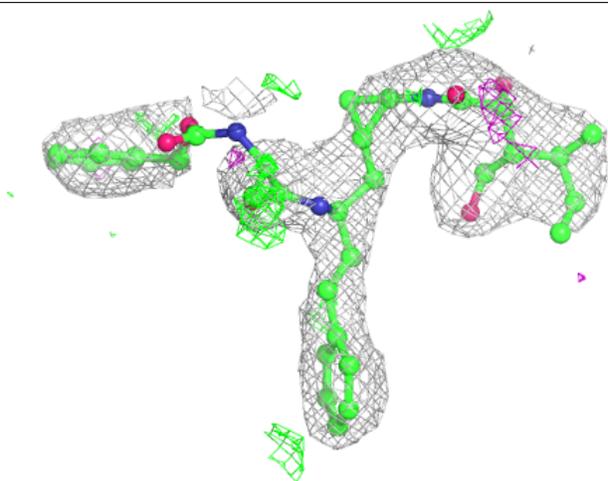
6.4 Ligands ⓘ

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.
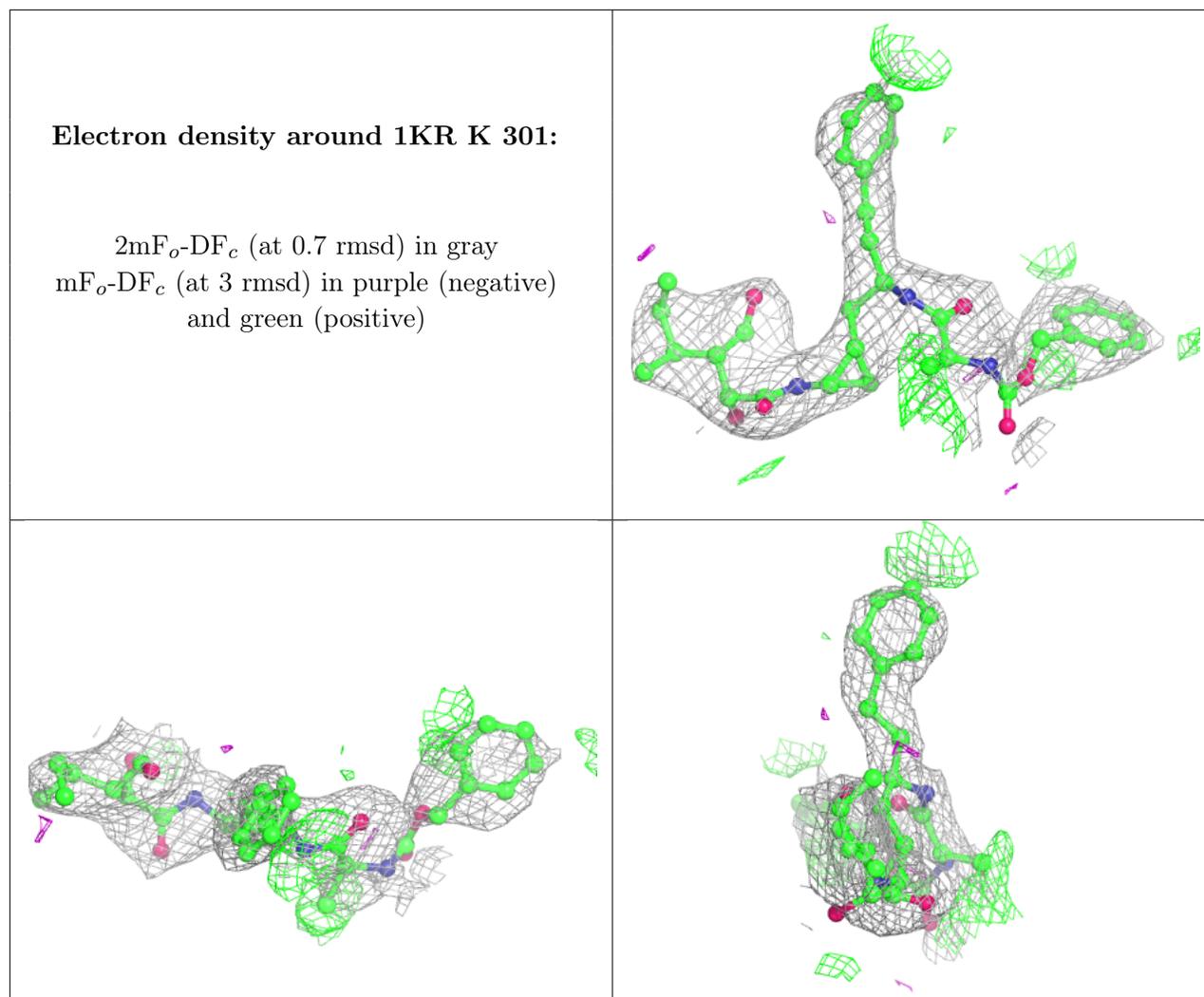
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
15	1KR	Y	301	41/41	0.88	0.25	37,81,137,168	0
15	1KR	K	301	41/41	0.90	0.23	40,81,141,150	0

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 1KR Y 301:

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

Electron density around 1KR 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.