Full wwPDB X-ray Structure Validation Report ⓘ

Sep 9, 2024 – 12:23 PM EDT

PDB ID	:	8SZM
Title	:	Crystal structure of E. coli ClpP protease in complex with phosphine oxide compound ACP6-12
Authors	:	Mabanglo, M.F.; Houry, W.A.
Deposited on	:	2023-05-30
Resolution	:	2.35 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ) were used in the production of this report:

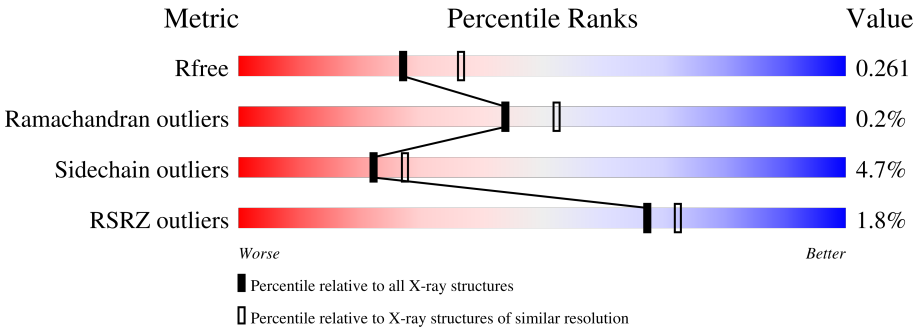
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.3

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 2.35 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}	164625	1460 (2.36-2.36)
Ramachandran outliers	177936	1559 (2.36-2.36)
Sidechain outliers	177891	1559 (2.36-2.36)
RSRZ outliers	164620	1460 (2.36-2.36)

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	207	2% 84% • 13%
1	B	207	81% • 15%
1	C	207	81% • 15%
1	D	207	% 85% • 13%
1	E	207	2% 85% • 12%
1	F	207	% 81% 6% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	207	2% 84% 13%
1	H	207	3% 87% 12%
1	I	207	83% 15%
1	J	207	2% 83% 13%
1	K	207	% 84% 13%
1	L	207	2% 83% 13%
1	M	207	% 84% 13%
1	N	207	2% 84% 13%
1	O	207	% 83% 5% 13%
1	P	207	82% 15%
1	Q	207	% 83% 13%
1	R	207	% 85% 13%
1	S	207	2% 84% 13%
1	T	207	% 83% 13%
1	U	207	2% 81% 5% 14%
1	V	207	2% 82% 6% 13%
1	W	207	84% 15%
1	X	207	% 81% 6% 13%
1	Y	207	2% 83% 13%
1	Z	207	3% 84% 12%
1	a	207	2% 84% 13%
1	b	207	% 84% 13%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 39622 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 ATP-dependent Clp protease proteolytic subunit.

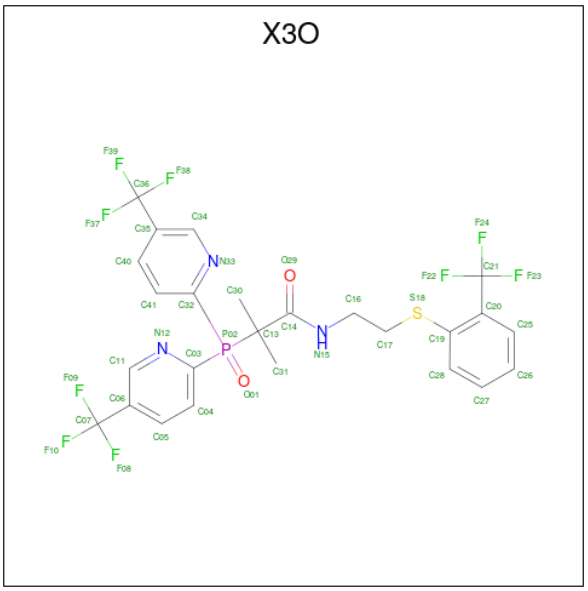
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	182	Total	C	N	O	S	0	0	0
			1424	902	244	266	12			
1	I	176	Total	C	N	O	S	0	0	0
			1376	867	238	260	11			
1	J	180	Total	C	N	O	S	0	0	0
			1405	887	242	264	12			
1	K	181	Total	C	N	O	S	0	0	0
			1413	893	243	265	12			
1	L	181	Total	C	N	O	S	0	0	0
			1416	896	243	265	12			
1	M	180	Total	C	N	O	S	0	0	0
			1409	891	242	264	12			
1	N	180	Total	C	N	O	S	0	0	0
			1409	891	242	264	12			
1	A	181	Total	C	N	O	S	0	0	0
			1416	896	243	265	12			
1	B	176	Total	C	N	O	S	0	0	0
			1376	867	238	260	11			
1	C	176	Total	C	N	O	S	0	0	0
			1376	867	238	260	11			
1	D	181	Total	C	N	O	S	0	0	0
			1413	893	243	265	12			
1	E	182	Total	C	N	O	S	0	0	0
			1424	902	244	266	12			
1	F	180	Total	C	N	O	S	0	0	0
			1409	891	242	264	12			
1	G	180	Total	C	N	O	S	0	0	0
			1409	891	242	264	12			
1	O	181	Total	C	N	O	S	0	0	0
			1416	896	243	265	12			
1	P	176	Total	C	N	O	S	0	0	0
			1376	867	238	260	11			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	180	Total	C	N	O	S	0	0	0
			1405	887	242	264	12			
1	R	181	Total	C	N	O	S	0	0	0
			1413	893	243	265	12			
1	S	181	Total	C	N	O	S	0	0	0
			1416	896	243	265	12			
1	T	180	Total	C	N	O	S	0	0	0
			1409	891	242	264	12			
1	U	179	Total	C	N	O	S	0	0	0
			1401	886	241	263	11			
1	V	181	Total	C	N	O	S	0	0	0
			1416	896	243	265	12			
1	W	176	Total	C	N	O	S	0	0	0
			1376	867	238	260	11			
1	X	180	Total	C	N	O	S	0	0	0
			1405	887	242	264	12			
1	Y	181	Total	C	N	O	S	0	0	0
			1413	893	243	265	12			
1	Z	182	Total	C	N	O	S	0	0	0
			1424	902	244	266	12			
1	a	180	Total	C	N	O	S	0	0	0
			1409	891	242	264	12			
1	b	180	Total	C	N	O	S	0	0	0
			1409	891	242	264	12			

- Molecule 2 is 2-{bis[5-(trifluoromethyl)pyridin-2-yl]phosphoryl}-2-methyl-N-(2-{[2-(trifluoromethyl)phenyl]sulfanyl}ethyl)propanamide (three-letter code: X3O) (formula: $C_{25}H_{21}F_9N_3O_2PS$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
2	G	1	Total	C	F	N	O	P	S	0	0
			41	25	9	3	2	1	1		
2	X	1	Total	C	F	N	O	P	S	0	0
			41	25	9	3	2	1	1		
2	Z	1	Total	C	F	N	O	P	S	0	0
			41	25	9	3	2	1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	H	4	Total	O	0	0
			4	4		
3	I	5	Total	O	0	0
			5	5		
3	J	5	Total	O	0	0
			5	5		
3	K	7	Total	O	0	0
			7	7		
3	L	7	Total	O	0	0
			7	7		
3	M	6	Total	O	0	0
			6	6		
3	N	5	Total	O	0	0
			5	5		
3	A	5	Total	O	0	0
			5	5		
3	B	2	Total	O	0	0
			2	2		

Continued on next page...

Continued from previous page...

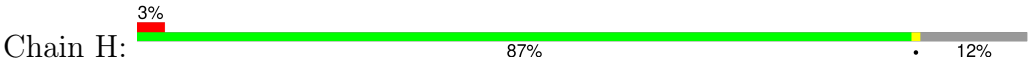
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	9	Total O 9 9	0	0
3	D	10	Total O 10 10	0	0
3	E	4	Total O 4 4	0	0
3	F	6	Total O 6 6	0	0
3	G	10	Total O 10 10	0	0
3	O	5	Total O 5 5	0	0
3	P	5	Total O 5 5	0	0
3	Q	6	Total O 6 6	0	0
3	R	2	Total O 2 2	0	0
3	S	3	Total O 3 3	0	0
3	T	2	Total O 2 2	0	0
3	U	3	Total O 3 3	0	0
3	V	3	Total O 3 3	0	0
3	W	4	Total O 4 4	0	0
3	X	5	Total O 5 5	0	0
3	Y	4	Total O 4 4	0	0
3	Z	3	Total O 3 3	0	0
3	a	2	Total O 2 2	0	0
3	b	4	Total O 4 4	0	0

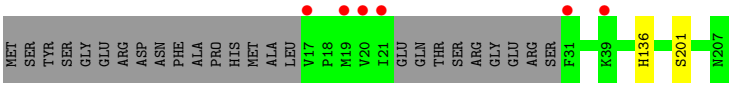
3 Residue-property plots (i)

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: ATP-dependent Clp protease proteolytic subunit

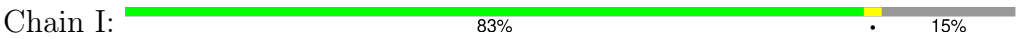
Chain H:





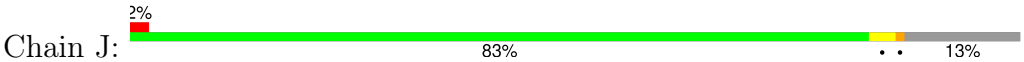
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

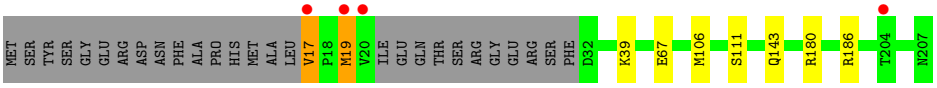
Chain I:



• Molecule 1: ATP-dependent Clp protease proteolytic subunit

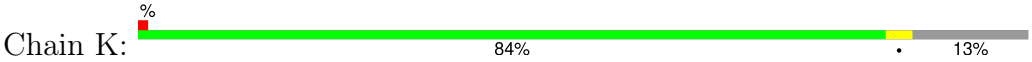
Chain J:

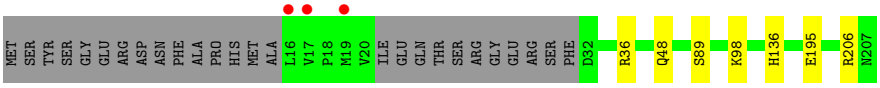




• Molecule 1: ATP-dependent Clp protease proteolytic subunit

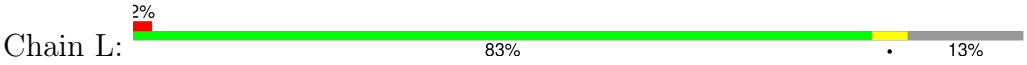
Chain K:

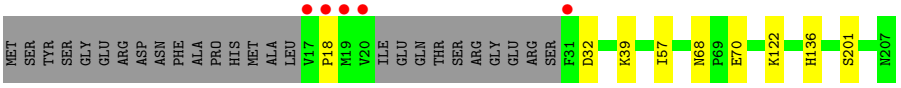




• Molecule 1: ATP-dependent Clp protease proteolytic subunit

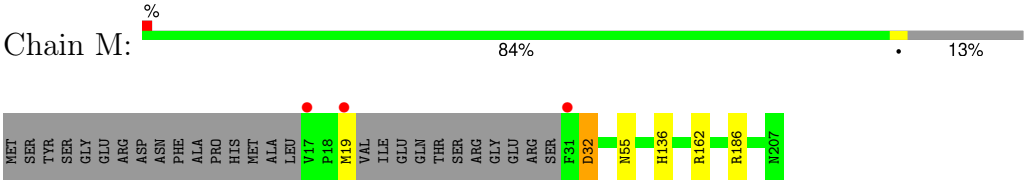
Chain L:





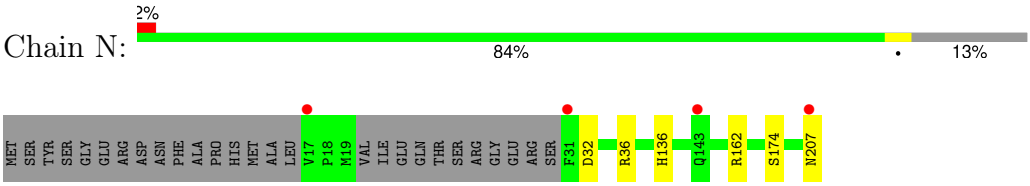
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain M:



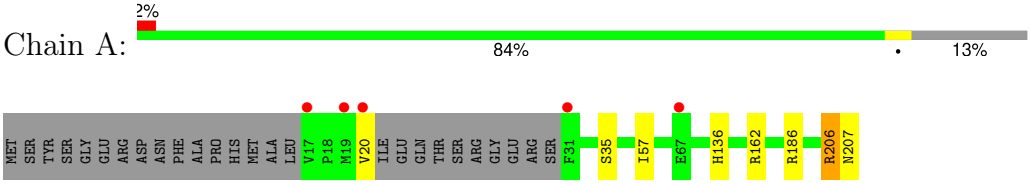
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain N:



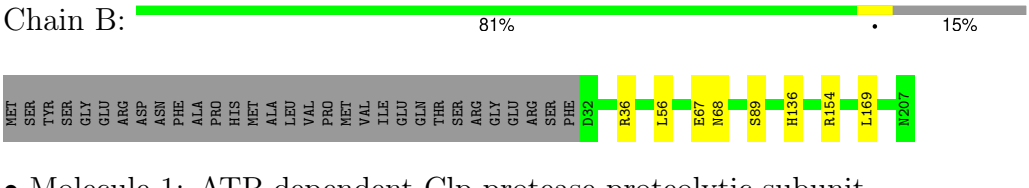
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain A:



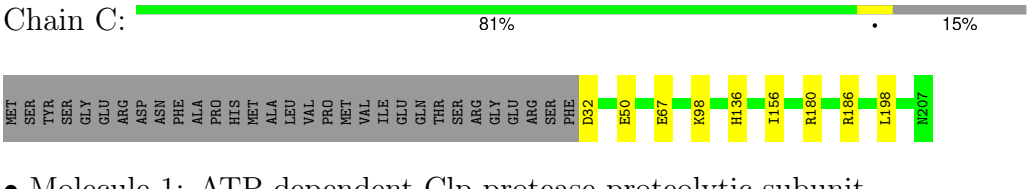
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain B:



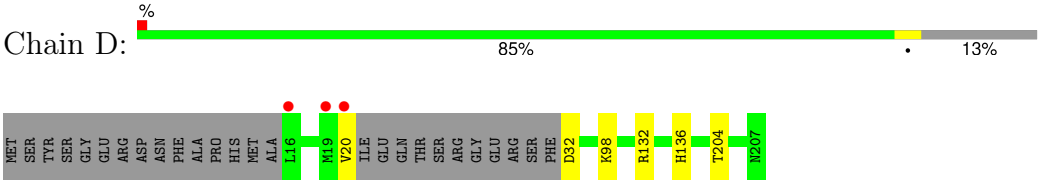
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain C:



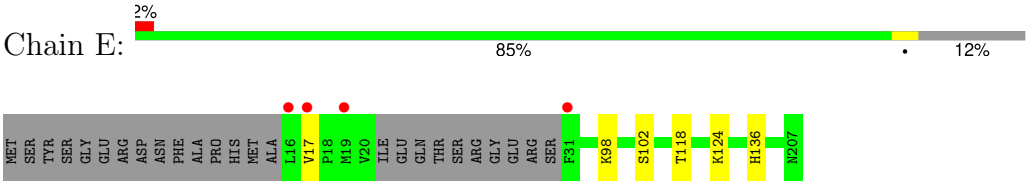
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain D:



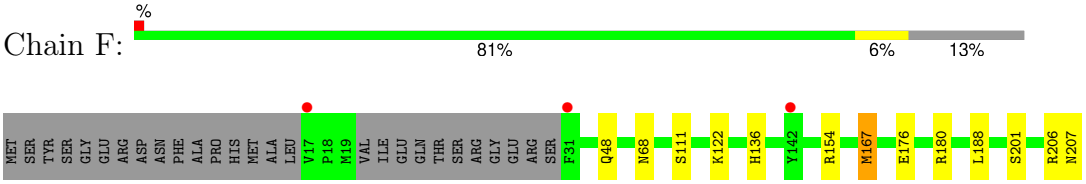
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain E:



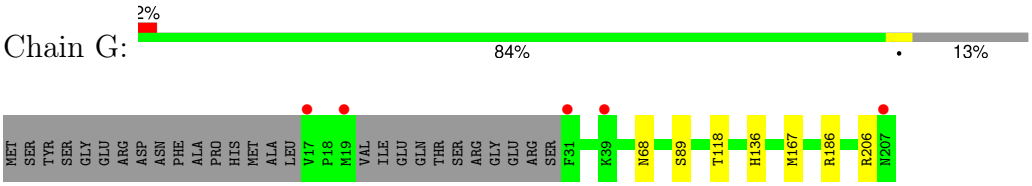
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain F:



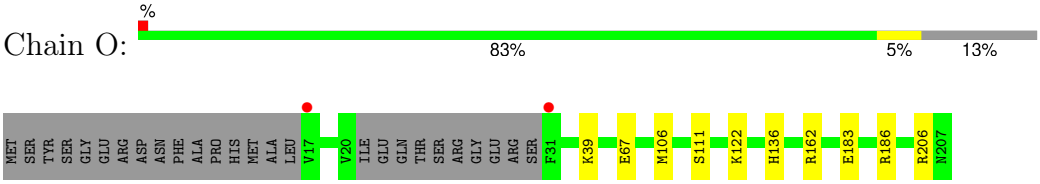
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain G:



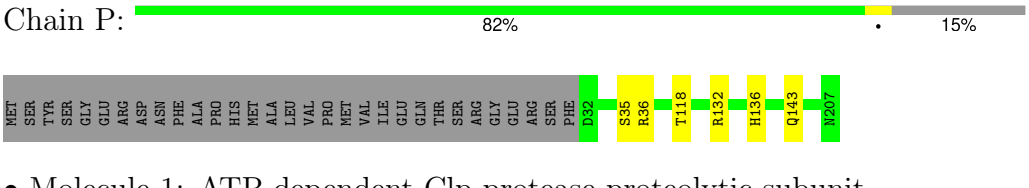
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain O:



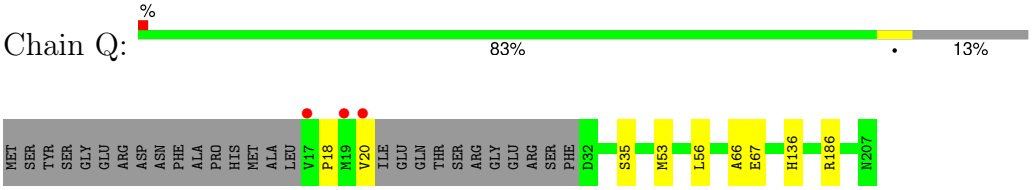
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain P:



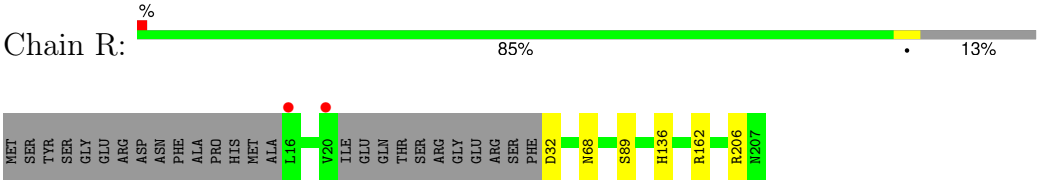
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain Q:



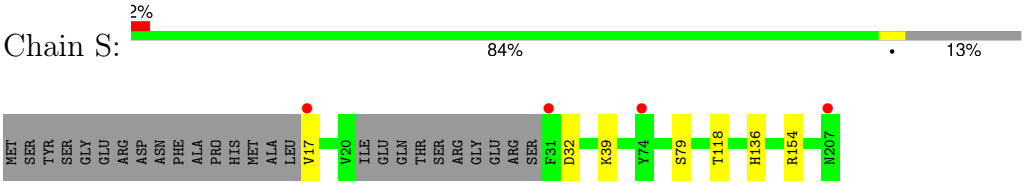
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain R:



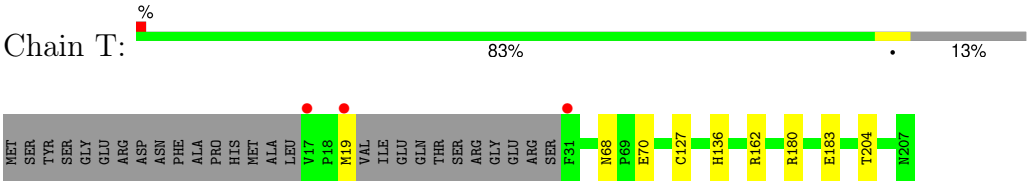
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain S:



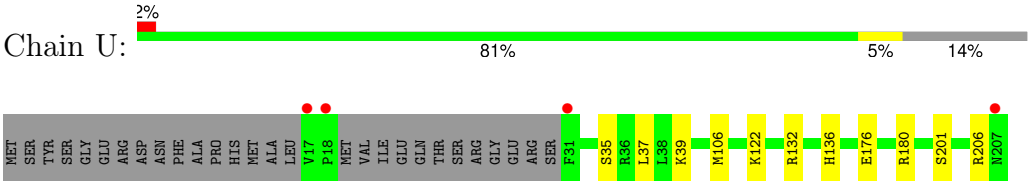
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain T:



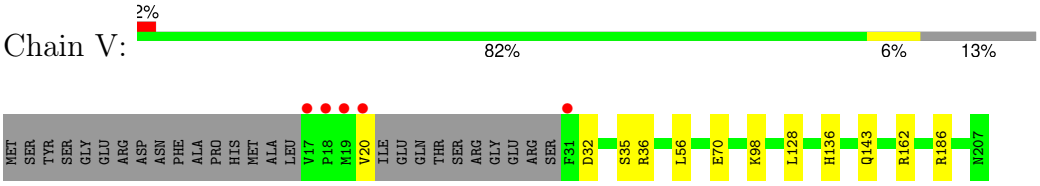
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain U:



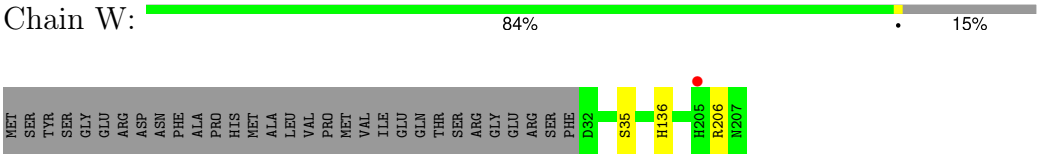
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain V:



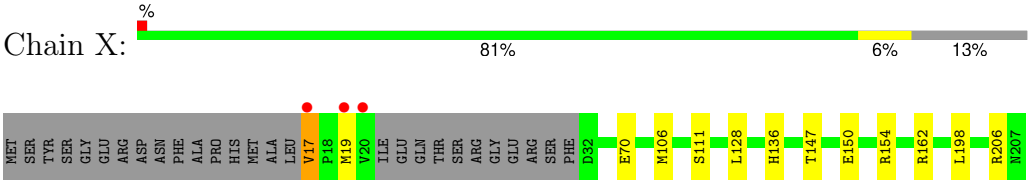
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain W:



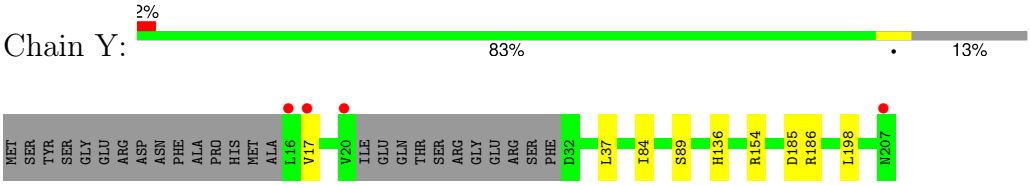
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain X:



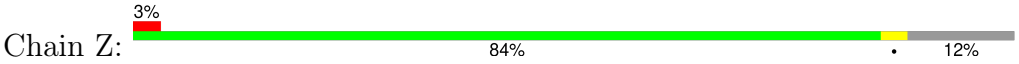
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

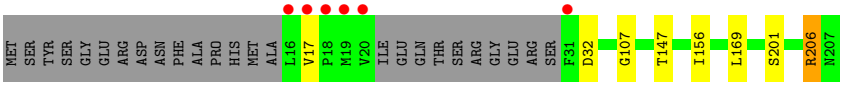
Chain Y:



• Molecule 1: ATP-dependent Clp protease proteolytic subunit

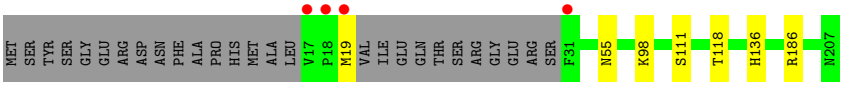
Chain Z:





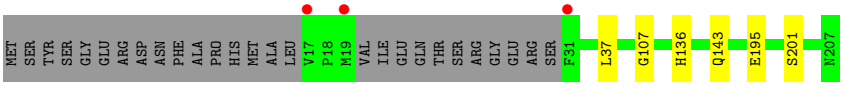
• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain a:



• Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain b:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.97Å 187.59Å 169.41Å 90.00° 93.72° 90.00°	Depositor
Resolution (Å)	79.83 – 2.35 79.83 – 2.35	Depositor EDS
% Data completeness (in resolution range)	97.2 (79.83-2.35) 97.3 (79.83-2.35)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.39 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, R_{free}	0.203 , 0.260 0.204 , 0.261	Depositor DCC
R_{free} test set	12141 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	43.4	Xtriage
Anisotropy	0.314	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 38.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	39622	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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: X3O

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	0/1439	0.62	0/1938
1	B	0.43	0/1398	0.63	0/1883
1	C	0.46	0/1398	0.66	0/1883
1	D	0.43	0/1435	0.65	0/1933
1	E	0.42	0/1447	0.63	0/1949
1	F	0.44	0/1432	0.65	1/1928 (0.1%)
1	G	0.45	0/1432	0.65	0/1928
1	H	0.45	0/1447	0.62	0/1949
1	I	0.46	0/1398	0.64	0/1883
1	J	0.44	0/1427	0.65	0/1922
1	K	0.42	0/1435	0.65	0/1933
1	L	0.43	0/1439	0.67	0/1938
1	M	0.44	0/1432	0.65	1/1928 (0.1%)
1	N	0.41	0/1432	0.66	0/1928
1	O	0.41	0/1439	0.64	0/1938
1	P	0.43	0/1398	0.62	0/1883
1	Q	0.42	0/1427	0.66	0/1922
1	R	0.42	0/1435	0.63	0/1933
1	S	0.38	0/1439	0.63	0/1938
1	T	0.42	0/1432	0.62	0/1928
1	U	0.44	0/1424	0.64	0/1918
1	V	0.45	0/1439	0.65	1/1938 (0.1%)
1	W	0.42	0/1398	0.67	0/1883
1	X	0.44	0/1427	0.68	0/1922
1	Y	0.41	0/1435	0.66	0/1933
1	Z	0.42	0/1447	0.62	0/1949
1	a	0.39	0/1432	0.61	0/1928
1	b	0.40	0/1432	0.62	0/1928
All	All	0.43	0/39995	0.64	3/53864 (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
1	G	0	1
1	J	0	1
1	X	0	1
1	Z	0	1
All	All	0	4

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	M	32	ASP	CB-CG-OD1	5.51	123.26	118.30
1	V	128	LEU	CA-CB-CG	-5.50	102.64	115.30
1	F	167	MET	CA-CB-CG	5.26	122.25	113.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	206	ARG	Sidechain
1	J	17	VAL	Peptide
1	X	17	VAL	Peptide
1	Z	206	ARG	Sidechain

5.2 Too-close contacts (i)

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

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	177/207 (86%)	171 (97%)	6 (3%)	0	100	100
1	B	174/207 (84%)	171 (98%)	2 (1%)	1 (1%)	22	24
1	C	174/207 (84%)	170 (98%)	4 (2%)	0	100	100
1	D	177/207 (86%)	170 (96%)	7 (4%)	0	100	100
1	E	178/207 (86%)	173 (97%)	5 (3%)	0	100	100
1	F	176/207 (85%)	171 (97%)	4 (2%)	1 (1%)	22	24
1	G	176/207 (85%)	168 (96%)	8 (4%)	0	100	100
1	H	178/207 (86%)	174 (98%)	4 (2%)	0	100	100
1	I	174/207 (84%)	170 (98%)	4 (2%)	0	100	100
1	J	176/207 (85%)	169 (96%)	6 (3%)	1 (1%)	22	24
1	K	177/207 (86%)	172 (97%)	5 (3%)	0	100	100
1	L	177/207 (86%)	172 (97%)	3 (2%)	2 (1%)	12	11
1	M	176/207 (85%)	172 (98%)	4 (2%)	0	100	100
1	N	176/207 (85%)	171 (97%)	5 (3%)	0	100	100
1	O	177/207 (86%)	170 (96%)	7 (4%)	0	100	100
1	P	174/207 (84%)	170 (98%)	4 (2%)	0	100	100
1	Q	176/207 (85%)	166 (94%)	7 (4%)	3 (2%)	7	6
1	R	177/207 (86%)	170 (96%)	6 (3%)	1 (1%)	22	24
1	S	177/207 (86%)	173 (98%)	4 (2%)	0	100	100
1	T	176/207 (85%)	171 (97%)	4 (2%)	1 (1%)	22	24
1	U	175/207 (84%)	168 (96%)	7 (4%)	0	100	100
1	V	177/207 (86%)	170 (96%)	7 (4%)	0	100	100
1	W	174/207 (84%)	169 (97%)	5 (3%)	0	100	100
1	X	176/207 (85%)	169 (96%)	7 (4%)	0	100	100
1	Y	177/207 (86%)	173 (98%)	4 (2%)	0	100	100
1	Z	178/207 (86%)	171 (96%)	6 (3%)	1 (1%)	22	24
1	a	176/207 (85%)	168 (96%)	8 (4%)	0	100	100
1	b	176/207 (85%)	168 (96%)	7 (4%)	1 (1%)	22	24
All	All	4932/5796 (85%)	4770 (97%)	150 (3%)	12 (0%)	44	52

All (12) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	Q	67	GLU
1	Z	107	GLY
1	L	18	PRO
1	L	68	ASN
1	F	68	ASN
1	Q	66	ALA
1	J	19	MET
1	R	68	ASN
1	T	68	ASN
1	b	107	GLY
1	B	68	ASN
1	Q	18	PRO

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	153/175 (87%)	145 (95%)	8 (5%)	19	23
1	B	148/175 (85%)	141 (95%)	7 (5%)	22	27
1	C	148/175 (85%)	139 (94%)	9 (6%)	15	17
1	D	153/175 (87%)	147 (96%)	6 (4%)	27	36
1	E	154/175 (88%)	148 (96%)	6 (4%)	27	36
1	F	152/175 (87%)	140 (92%)	12 (8%)	10	10
1	G	152/175 (87%)	146 (96%)	6 (4%)	27	36
1	H	154/175 (88%)	152 (99%)	2 (1%)	65	77
1	I	148/175 (85%)	144 (97%)	4 (3%)	40	50
1	J	152/175 (87%)	143 (94%)	9 (6%)	16	18
1	K	153/175 (87%)	146 (95%)	7 (5%)	23	28
1	L	153/175 (87%)	147 (96%)	6 (4%)	27	36
1	M	152/175 (87%)	146 (96%)	6 (4%)	27	36
1	N	152/175 (87%)	146 (96%)	6 (4%)	27	36
1	O	153/175 (87%)	144 (94%)	9 (6%)	16	18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	P	148/175 (85%)	142 (96%)	6 (4%)	26	33
1	Q	152/175 (87%)	146 (96%)	6 (4%)	27	36
1	R	153/175 (87%)	148 (97%)	5 (3%)	33	42
1	S	153/175 (87%)	146 (95%)	7 (5%)	23	28
1	T	152/175 (87%)	145 (95%)	7 (5%)	23	28
1	U	151/175 (86%)	140 (93%)	11 (7%)	11	12
1	V	153/175 (87%)	142 (93%)	11 (7%)	12	12
1	W	148/175 (85%)	145 (98%)	3 (2%)	50	63
1	X	152/175 (87%)	140 (92%)	12 (8%)	10	10
1	Y	153/175 (87%)	144 (94%)	9 (6%)	16	18
1	Z	154/175 (88%)	148 (96%)	6 (4%)	27	36
1	a	152/175 (87%)	145 (95%)	7 (5%)	23	28
1	b	152/175 (87%)	147 (97%)	5 (3%)	33	42
All	All	4250/4900 (87%)	4052 (95%)	198 (5%)	22	27

All (198) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	136	HIS
1	H	201	SER
1	I	84	ILE
1	I	111	SER
1	I	136	HIS
1	I	207	ASN
1	J	17	VAL
1	J	19	MET
1	J	39	LYS
1	J	67	GLU
1	J	106	MET
1	J	111	SER
1	J	143	GLN
1	J	180	ARG
1	J	186	ARG
1	K	36	ARG
1	K	48	GLN
1	K	89	SER
1	K	98	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	K	136	HIS
1	K	195	GLU
1	K	206	ARG
1	L	32	ASP
1	L	39	LYS
1	L	57	ILE
1	L	122	LYS
1	L	136	HIS
1	L	201	SER
1	M	19	MET
1	M	32	ASP
1	M	55	ASN
1	M	136	HIS
1	M	162	ARG
1	M	186	ARG
1	N	32	ASP
1	N	36	ARG
1	N	136	HIS
1	N	162	ARG
1	N	174	SER
1	N	207	ASN
1	A	20	VAL
1	A	35	SER
1	A	57	ILE
1	A	136	HIS
1	A	162	ARG
1	A	186	ARG
1	A	206	ARG
1	A	207	ASN
1	B	36	ARG
1	B	56	LEU
1	B	67	GLU
1	B	89	SER
1	B	136	HIS
1	B	154	ARG
1	B	169	LEU
1	C	32	ASP
1	C	50	GLU
1	C	67	GLU
1	C	98	LYS
1	C	136	HIS
1	C	156	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	180	ARG
1	C	186	ARG
1	C	198	LEU
1	D	20	VAL
1	D	32	ASP
1	D	98	LYS
1	D	132	ARG
1	D	136	HIS
1	D	204	THR
1	E	17	VAL
1	E	98	LYS
1	E	102	SER
1	E	118	THR
1	E	124	LYS
1	E	136	HIS
1	F	48	GLN
1	F	111	SER
1	F	122	LYS
1	F	136	HIS
1	F	154	ARG
1	F	167	MET
1	F	176	GLU
1	F	180	ARG
1	F	188	LEU
1	F	201	SER
1	F	206	ARG
1	F	207	ASN
1	G	68	ASN
1	G	89	SER
1	G	118	THR
1	G	136	HIS
1	G	167	MET
1	G	186	ARG
1	O	39	LYS
1	O	67	GLU
1	O	106	MET
1	O	111	SER
1	O	122	LYS
1	O	136	HIS
1	O	162	ARG
1	O	186	ARG
1	O	206	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	P	35	SER
1	P	36	ARG
1	P	118	THR
1	P	132	ARG
1	P	136	HIS
1	P	143	GLN
1	Q	20	VAL
1	Q	35	SER
1	Q	53	MET
1	Q	56	LEU
1	Q	136	HIS
1	Q	186	ARG
1	R	32	ASP
1	R	89	SER
1	R	136	HIS
1	R	162	ARG
1	R	206	ARG
1	S	17	VAL
1	S	32	ASP
1	S	39	LYS
1	S	79	SER
1	S	118	THR
1	S	136	HIS
1	S	154	ARG
1	T	19	MET
1	T	70	GLU
1	T	127	CYS
1	T	136	HIS
1	T	162	ARG
1	T	180	ARG
1	T	204	THR
1	U	35	SER
1	U	37	LEU
1	U	39	LYS
1	U	106	MET
1	U	122	LYS
1	U	132	ARG
1	U	136	HIS
1	U	176	GLU
1	U	180	ARG
1	U	201	SER
1	U	206	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	V	20	VAL
1	V	32	ASP
1	V	35	SER
1	V	36	ARG
1	V	56	LEU
1	V	70	GLU
1	V	98	LYS
1	V	136	HIS
1	V	143	GLN
1	V	162	ARG
1	V	186	ARG
1	W	35	SER
1	W	136	HIS
1	W	206	ARG
1	X	17	VAL
1	X	19	MET
1	X	70	GLU
1	X	106	MET
1	X	111	SER
1	X	128	LEU
1	X	136	HIS
1	X	150	GLU
1	X	154	ARG
1	X	162	ARG
1	X	198	LEU
1	X	206	ARG
1	Y	17	VAL
1	Y	37	LEU
1	Y	84	ILE
1	Y	89	SER
1	Y	136	HIS
1	Y	154	ARG
1	Y	185	ASP
1	Y	186	ARG
1	Y	198	LEU
1	Z	17	VAL
1	Z	32	ASP
1	Z	156	ILE
1	Z	169	LEU
1	Z	201	SER
1	Z	206	ARG
1	a	19	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	a	55	ASN
1	a	98	LYS
1	a	111	SER
1	a	118	THR
1	a	136	HIS
1	a	186	ARG
1	b	37	LEU
1	b	136	HIS
1	b	143	GLN
1	b	195	GLU
1	b	201	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

Mol	Chain	Res	Type
1	I	55	ASN
1	J	143	GLN
1	K	143	GLN
1	M	207	ASN
1	A	205	HIS
1	C	48	GLN
1	D	55	ASN
1	E	48	GLN
1	E	136	HIS
1	F	55	ASN
1	G	48	GLN
1	G	68	ASN
1	P	177	GLN
1	Q	55	ASN
1	Q	205	HIS
1	R	55	ASN
1	U	55	ASN
1	U	207	ASN
1	V	55	ASN
1	Z	60	GLN
1	Z	143	GLN
1	b	48	GLN
1	b	55	ASN
1	b	68	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)

3 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	$\#\|Z\| > 2$
2	X3O	X	301	-	39,43,43	1.87	8 (20%)	53,67,67	1.79	11 (20%)
2	X3O	Z	301	-	39,43,43	1.56	5 (12%)	53,67,67	1.74	12 (22%)
2	X3O	G	301	-	39,43,43	1.64	7 (17%)	53,67,67	1.49	8 (15%)

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
2	X3O	X	301	-	-	2/47/52/52	0/3/3/3
2	X3O	Z	301	-	-	5/47/52/52	0/3/3/3
2	X3O	G	301	-	-	7/47/52/52	0/3/3/3

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	301	X3O	C19-S18	5.85	1.86	1.77
2	X	301	X3O	C14-N15	5.34	1.46	1.33
2	Z	301	X3O	C19-S18	4.94	1.84	1.77
2	Z	301	X3O	C14-N15	4.92	1.45	1.33
2	G	301	X3O	C14-N15	4.88	1.45	1.33
2	G	301	X3O	C19-S18	4.81	1.84	1.77
2	X	301	X3O	C07-C06	3.14	1.56	1.49
2	G	301	X3O	C21-C20	2.97	1.57	1.50
2	X	301	X3O	C21-C20	2.73	1.56	1.50
2	G	301	X3O	C07-C06	2.64	1.55	1.49
2	X	301	X3O	C36-C35	2.59	1.55	1.49
2	Z	301	X3O	O29-C14	-2.39	1.18	1.22
2	Z	301	X3O	C21-C20	2.38	1.55	1.50
2	X	301	X3O	O29-C14	-2.26	1.19	1.22
2	X	301	X3O	C19-C20	2.18	1.42	1.40
2	G	301	X3O	C19-C20	2.17	1.42	1.40
2	G	301	X3O	P02-O01	-2.16	1.47	1.49
2	Z	301	X3O	C07-C06	2.15	1.54	1.49
2	G	301	X3O	O29-C14	-2.09	1.19	1.22
2	X	301	X3O	C05-C04	2.00	1.42	1.38

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	X	301	X3O	C17-S18-C19	6.50	116.58	103.05
2	Z	301	X3O	C35-C34-N33	-4.19	119.45	123.38
2	Z	301	X3O	C17-S18-C19	3.79	110.94	103.05
2	G	301	X3O	C35-C34-N33	-3.72	119.89	123.38
2	Z	301	X3O	C13-C14-N15	3.71	121.30	116.95
2	G	301	X3O	C17-S18-C19	3.69	110.74	103.05
2	X	301	X3O	F09-C07-C06	-3.50	105.40	112.90
2	X	301	X3O	C04-C03-N12	-3.49	118.33	123.22
2	G	301	X3O	C17-C16-N15	-3.14	105.87	112.41
2	Z	301	X3O	C04-C03-N12	-3.01	119.01	123.22
2	X	301	X3O	C05-C06-C07	2.97	124.73	119.96
2	G	301	X3O	F39-C36-C35	-2.77	106.96	112.90
2	Z	301	X3O	C40-C35-C34	2.75	120.72	117.73
2	X	301	X3O	C28-C19-C20	-2.67	117.42	120.40
2	Z	301	X3O	C41-C40-C35	-2.59	117.59	121.17
2	X	301	X3O	C35-C34-N33	-2.51	121.02	123.38
2	G	301	X3O	C41-C40-C35	-2.51	117.70	121.17
2	Z	301	X3O	F39-C36-C35	-2.48	107.58	112.90
2	X	301	X3O	C13-C14-N15	2.44	119.81	116.95

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	301	X3O	C40-C35-C34	2.44	120.38	117.73
2	G	301	X3O	C04-C03-N12	-2.40	119.86	123.22
2	Z	301	X3O	F10-C07-C06	-2.35	107.87	112.90
2	Z	301	X3O	C05-C06-C07	2.29	123.63	119.96
2	Z	301	X3O	C25-C20-C19	2.27	119.43	117.44
2	X	301	X3O	C05-C06-C11	-2.24	115.29	117.73
2	X	301	X3O	P02-C03-C04	2.21	125.83	120.71
2	Z	301	X3O	F37-C36-C35	-2.08	108.43	112.90
2	X	301	X3O	C41-C32-N33	-2.05	120.35	123.22
2	Z	301	X3O	O29-C14-N15	-2.04	118.29	122.64
2	X	301	X3O	O29-C14-N15	-2.03	118.31	122.64
2	G	301	X3O	F10-C07-C06	-2.02	108.57	112.90

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	301	X3O	N33-C32-P02-C13
2	Z	301	X3O	C20-C19-S18-C17
2	G	301	X3O	C19-C20-C21-F23
2	G	301	X3O	C19-C20-C21-F24
2	G	301	X3O	C19-C20-C21-F22
2	G	301	X3O	N12-C03-P02-O01
2	Z	301	X3O	N12-C03-P02-O01
2	G	301	X3O	C41-C32-P02-C13
2	Z	301	X3O	N15-C16-C17-S18
2	X	301	X3O	C20-C19-S18-C17
2	Z	301	X3O	C28-C19-S18-C17
2	G	301	X3O	N33-C32-P02-O01
2	X	301	X3O	N12-C03-P02-C32
2	Z	301	X3O	N12-C03-P02-C32

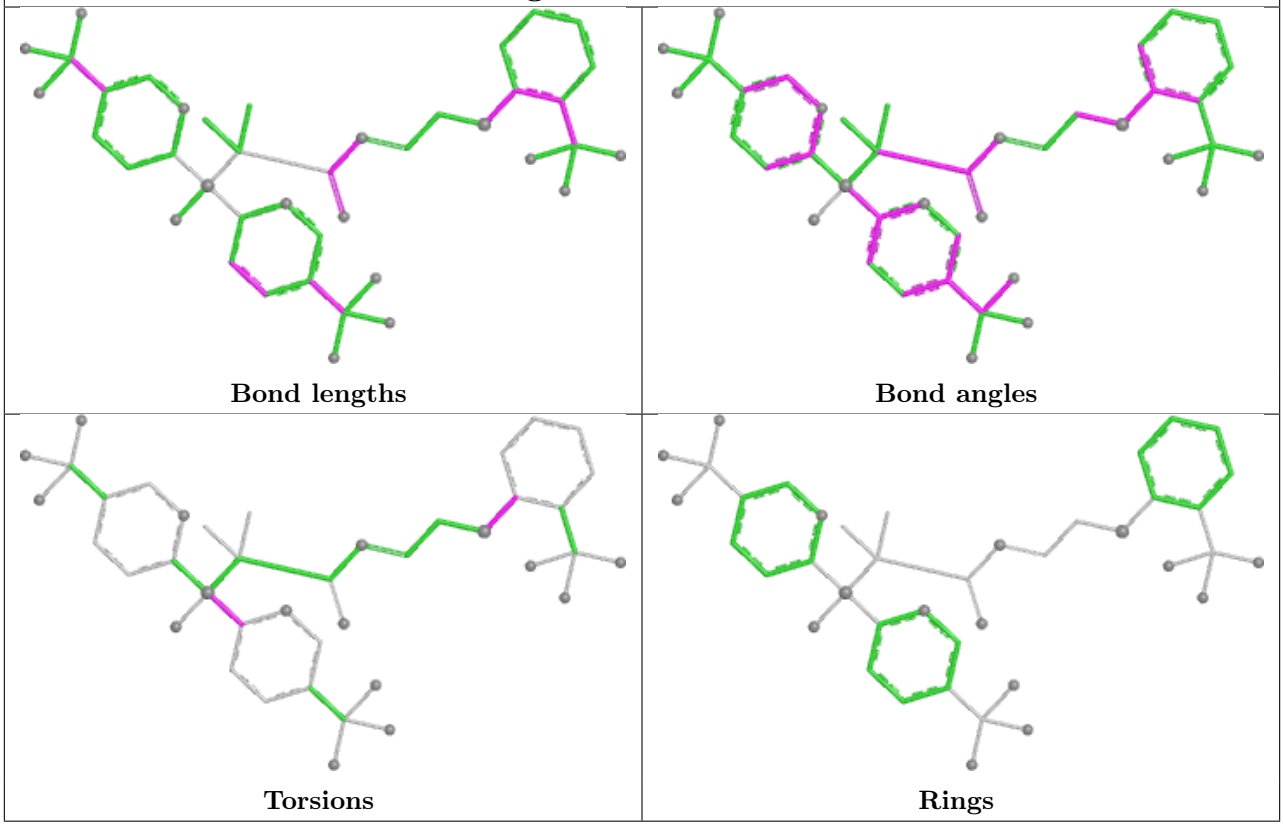
There are no ring outliers.

No monomer is involved in short contacts.

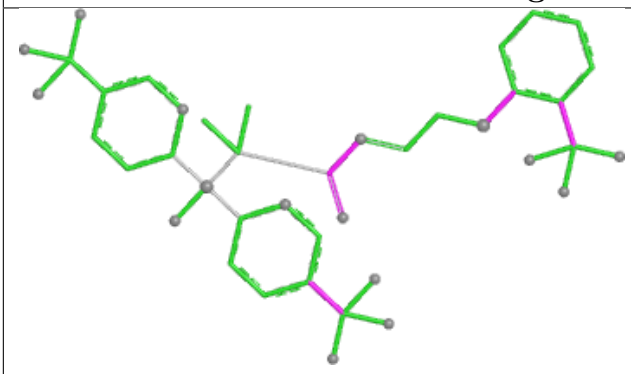
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

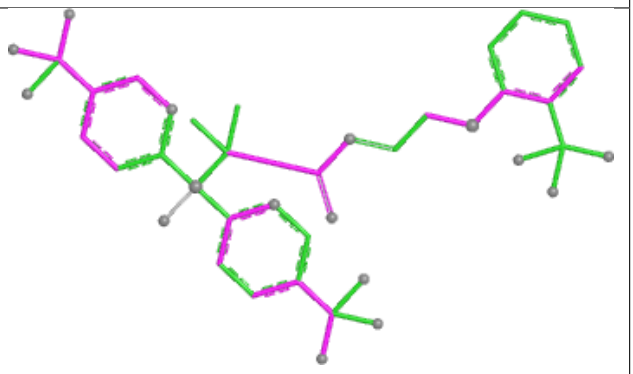
Ligand X3O X 301



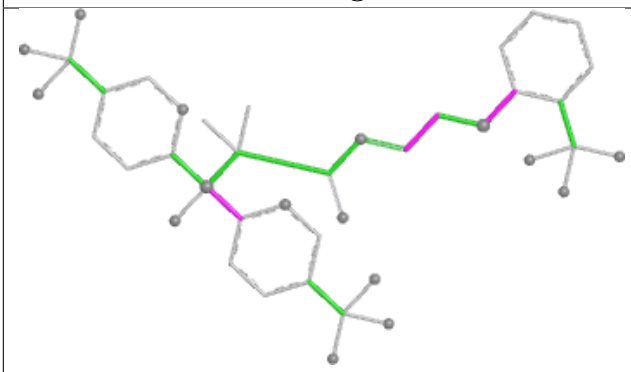
Ligand X3O Z 301



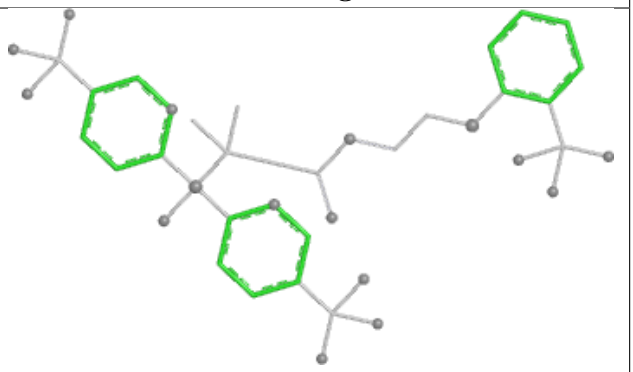
Bond lengths



Bond angles

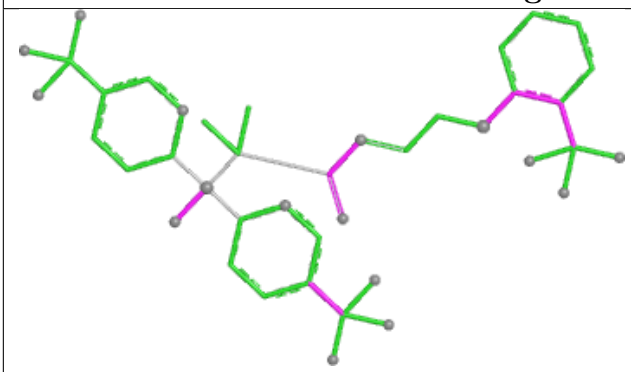


Torsions

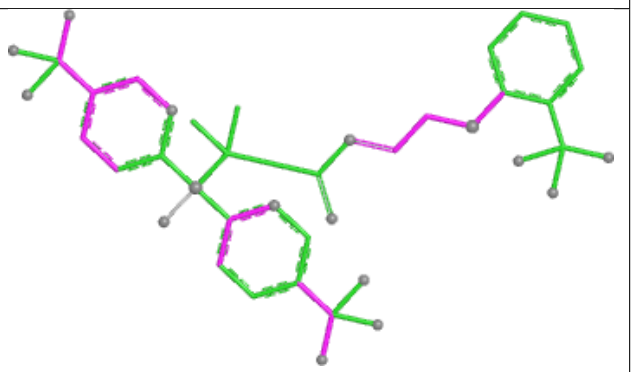


Rings

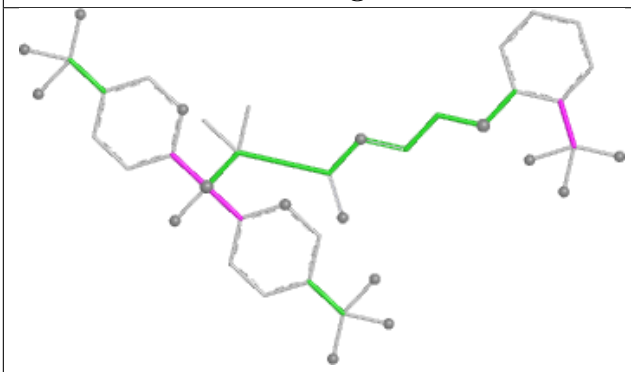
Ligand X3O G 301



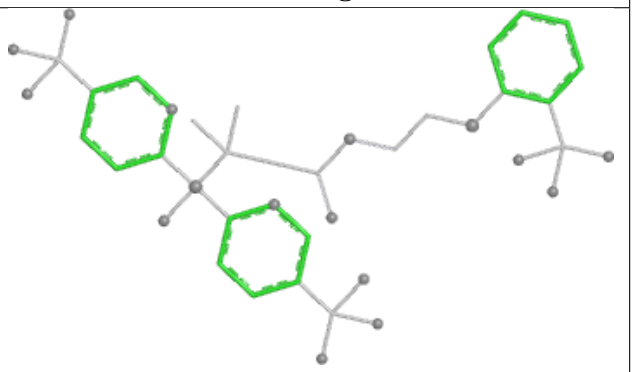
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	181/207 (87%)	-0.06	5 (2%)	55	61	31, 43, 67, 85	0
1	B	176/207 (85%)	-0.33	0	100	100	29, 38, 57, 77	1 (0%)
1	C	176/207 (85%)	-0.40	0	100	100	27, 37, 53, 89	0
1	D	181/207 (87%)	-0.17	3 (1%)	69	73	30, 40, 61, 94	0
1	E	182/207 (87%)	-0.08	4 (2%)	62	67	32, 41, 63, 86	2 (1%)
1	F	180/207 (86%)	-0.35	3 (1%)	69	73	28, 38, 60, 81	1 (0%)
1	G	180/207 (86%)	-0.20	5 (2%)	55	61	28, 40, 62, 85	0
1	H	182/207 (87%)	0.00	6 (3%)	49	55	32, 43, 63, 89	0
1	I	176/207 (85%)	-0.26	0	100	100	30, 40, 57, 80	1 (0%)
1	J	180/207 (86%)	-0.20	4 (2%)	62	67	31, 41, 60, 89	1 (0%)
1	K	181/207 (87%)	-0.15	3 (1%)	69	73	31, 41, 58, 80	0
1	L	181/207 (87%)	-0.25	5 (2%)	55	61	28, 39, 60, 83	0
1	M	180/207 (86%)	-0.32	3 (1%)	69	73	28, 37, 58, 87	0
1	N	180/207 (86%)	-0.11	4 (2%)	62	67	32, 42, 61, 91	0
1	O	181/207 (87%)	-0.28	2 (1%)	77	82	34, 43, 59, 74	0
1	P	176/207 (85%)	-0.30	0	100	100	34, 43, 60, 85	0
1	Q	180/207 (86%)	-0.20	3 (1%)	69	73	34, 43, 61, 89	0
1	R	181/207 (87%)	0.08	2 (1%)	77	82	36, 46, 62, 82	0
1	S	181/207 (87%)	0.13	4 (2%)	62	67	36, 47, 67, 92	0
1	T	180/207 (86%)	-0.04	3 (1%)	69	73	33, 45, 71, 90	0
1	U	179/207 (86%)	-0.20	4 (2%)	62	67	32, 42, 61, 76	0
1	V	181/207 (87%)	-0.19	5 (2%)	55	61	29, 40, 62, 84	0
1	W	176/207 (85%)	-0.24	1 (0%)	85	88	31, 41, 56, 78	1 (0%)
1	X	180/207 (86%)	-0.29	3 (1%)	69	73	30, 38, 62, 87	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	Y	181/207 (87%)	-0.03	4 (2%)	62	67	34, 43, 61, 90	0
1	Z	182/207 (87%)	0.14	6 (3%)	49	55	33, 46, 73, 92	2 (1%)
1	a	180/207 (86%)	0.08	4 (2%)	62	67	37, 46, 69, 89	0
1	b	180/207 (86%)	-0.03	3 (1%)	69	73	35, 46, 65, 85	0
All	All	5034/5796 (86%)	-0.15	89 (1%)	67	72	27, 42, 62, 94	9 (0%)

All (89) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	21	ILE	6.5
1	D	16	LEU	5.7
1	N	17	VAL	5.7
1	Y	16	LEU	5.3
1	E	16	LEU	5.3
1	T	31	PHE	5.3
1	X	20	VAL	5.2
1	Q	20	VAL	5.2
1	Z	16	LEU	5.2
1	L	17	VAL	5.0
1	U	17	VAL	4.9
1	b	17	VAL	4.8
1	a	31	PHE	4.7
1	K	16	LEU	4.6
1	R	16	LEU	4.4
1	E	17	VAL	4.2
1	b	31	PHE	4.1
1	T	17	VAL	4.1
1	G	31	PHE	4.0
1	X	17	VAL	4.0
1	Q	17	VAL	3.9
1	a	17	VAL	3.9
1	E	31	PHE	3.9
1	Z	31	PHE	3.8
1	H	20	VAL	3.8
1	L	20	VAL	3.7
1	A	67	GLU	3.7
1	a	19	MET	3.7
1	U	18	PRO	3.6
1	V	31	PHE	3.6
1	L	31	PHE	3.5
1	M	31	PHE	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	Z	20	VAL	3.5
1	Z	17	VAL	3.4
1	J	19	MET	3.4
1	Y	17	VAL	3.4
1	G	17	VAL	3.3
1	Y	20	VAL	3.3
1	A	19	MET	3.3
1	J	17	VAL	3.2
1	F	31	PHE	3.2
1	H	31	PHE	3.1
1	X	19	MET	3.1
1	H	17	VAL	3.1
1	K	17	VAL	3.1
1	Y	207	ASN	3.0
1	J	20	VAL	3.0
1	S	207	ASN	2.9
1	J	204	THR	2.9
1	A	17	VAL	2.8
1	V	20	VAL	2.8
1	b	19	MET	2.8
1	U	31	PHE	2.8
1	H	19	MET	2.8
1	D	19	MET	2.8
1	M	17	VAL	2.8
1	D	20	VAL	2.7
1	F	17	VAL	2.7
1	F	142	TYR	2.7
1	S	17	VAL	2.7
1	V	17	VAL	2.7
1	T	19	MET	2.6
1	N	143	GLN	2.6
1	O	17	VAL	2.6
1	Z	19	MET	2.6
1	L	19	MET	2.5
1	N	31	PHE	2.5
1	Q	19	MET	2.4
1	V	18	PRO	2.4
1	H	39	LYS	2.4
1	U	207	ASN	2.4
1	G	19	MET	2.4
1	O	31	PHE	2.3
1	G	39	LYS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	31	PHE	2.3
1	S	74	TYR	2.3
1	V	19	MET	2.3
1	M	19	MET	2.2
1	a	18	PRO	2.2
1	N	207	ASN	2.2
1	L	18	PRO	2.2
1	K	19	MET	2.2
1	Z	18	PRO	2.1
1	A	20	VAL	2.1
1	R	20	VAL	2.0
1	S	31	PHE	2.0
1	E	19	MET	2.0
1	W	205	HIS	2.0
1	G	207	ASN	2.0

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)

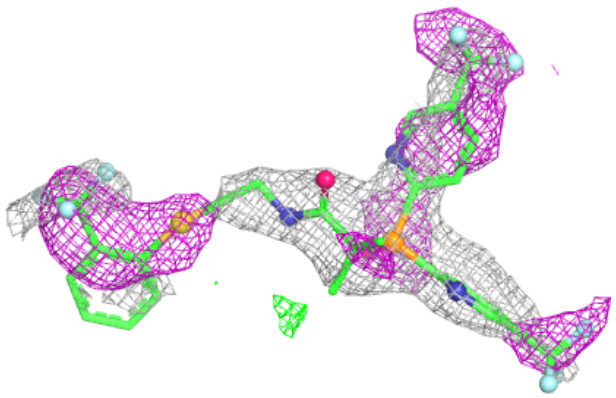
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

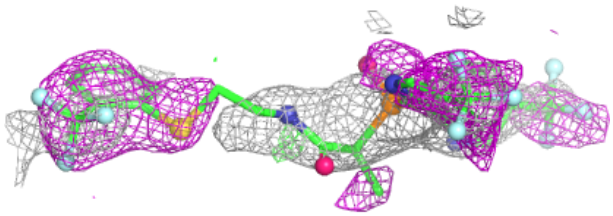
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	X3O	G	301	41/41	0.66	0.18	71,82,94,102	0
2	X3O	X	301	41/41	0.66	0.20	57,70,86,97	0
2	X3O	Z	301	41/41	0.67	0.20	77,88,100,105	0

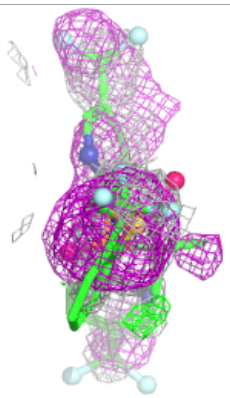
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around X3O G 301:

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

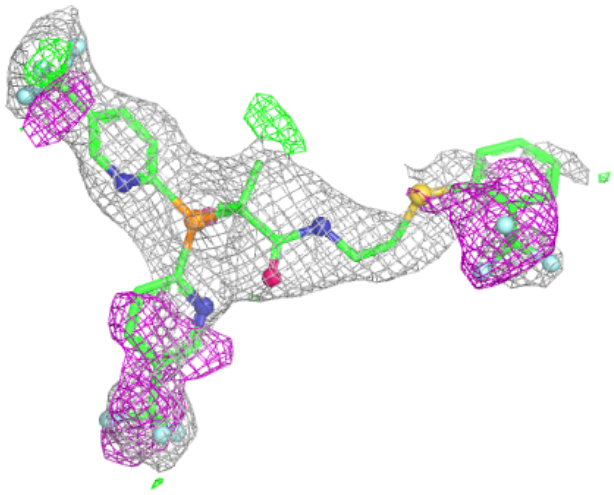


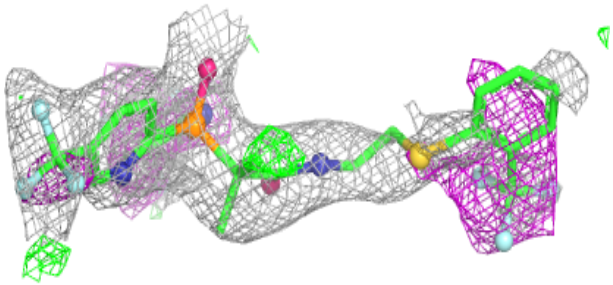


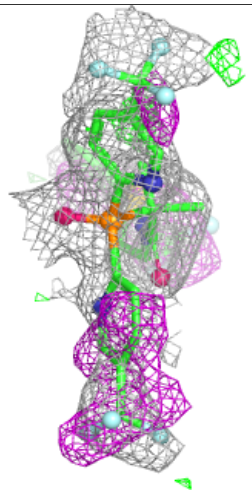


Electron density around X3O X 301:

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

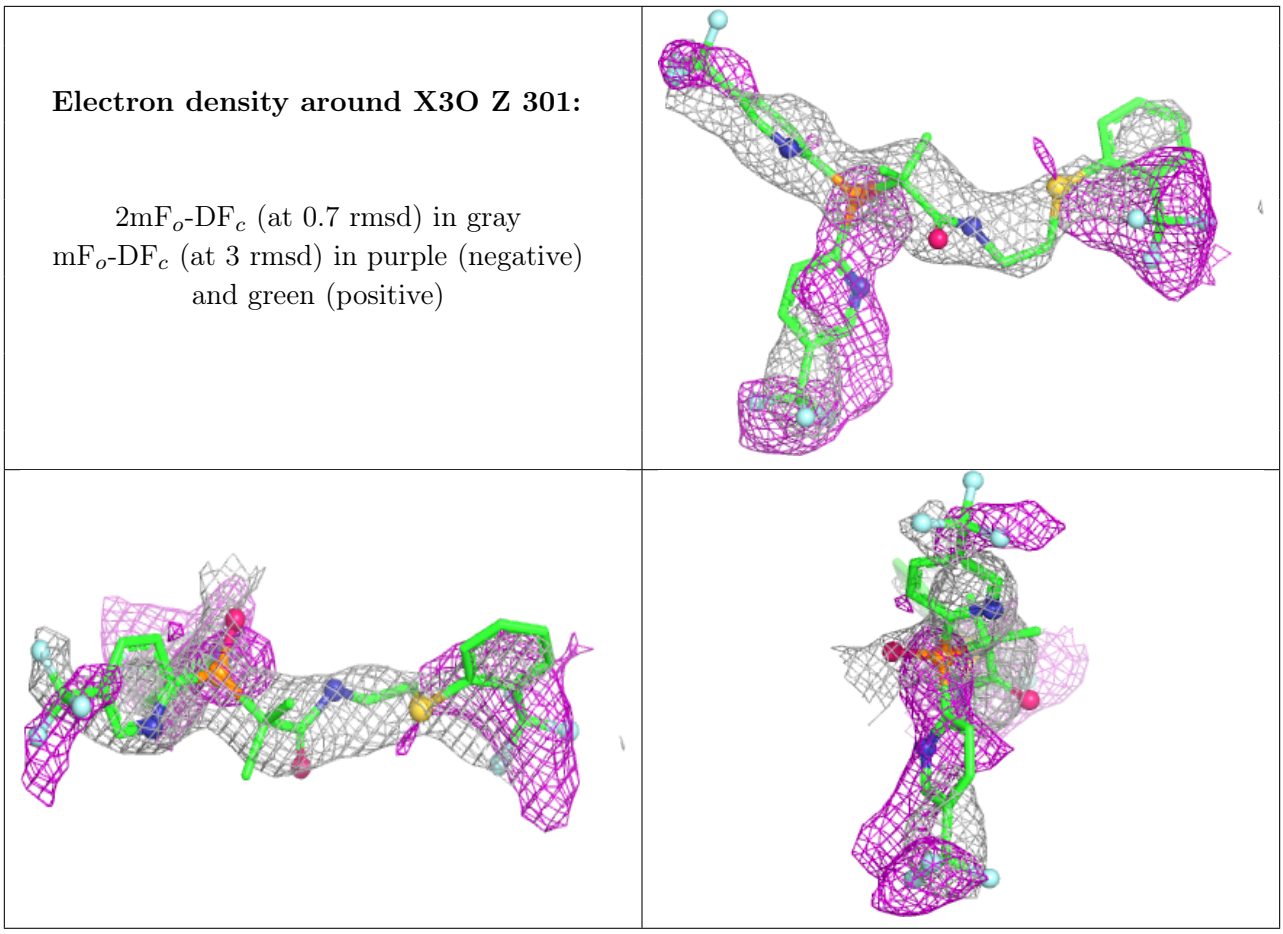






Electron density around X3O Z 301:

$2mF_o$-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.