Full wwPDB X-ray Structure Validation Report (i)

Feb 29, 2024 – 12:17 PM JST

PDB ID	:	5X30
Title	:	Crystal structure of Pseudomonas putida methionine gamma-lyase C116H mutant with L-homocysteine intermediates.
Authors	:	Shiba, T.; Sato, D.; Harada, S.
Deposited on	:	2017-02-02
Resolution	:	1.70 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

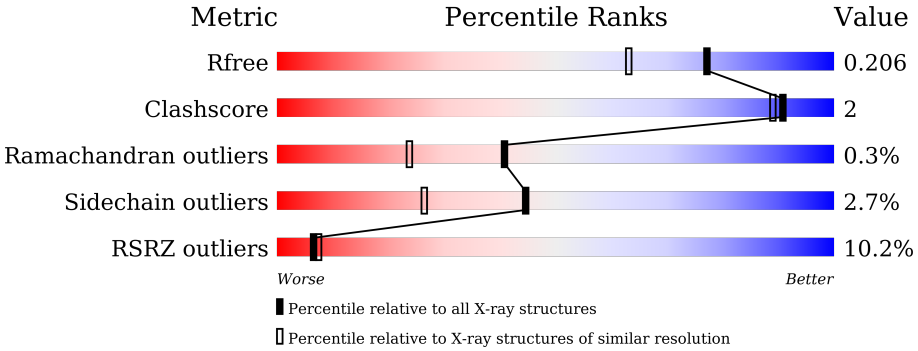
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 <=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	398	11% 94% • ••
1	C	398	9% 92% 6% ••
2	B	398	8% 91% 7% ••
2	D	398	11% 95% 5%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 13001 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 L-methionine gamma-lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	392	2970	1874	525	554	1	16	0	0	0
1	C	393	2975	1877	526	555	1	16	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	116	HIS	CYS	engineered mutation	UNP P13254
C	116	HIS	CYS	engineered mutation	UNP P13254

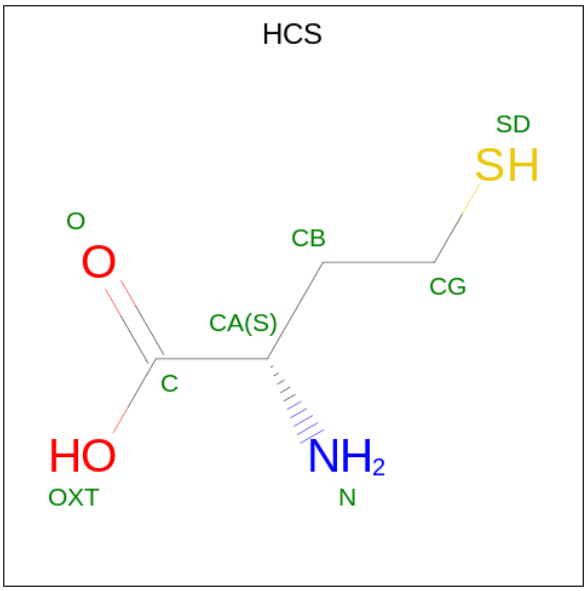
- Molecule 2 is a protein called L-methionine gamma-lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	392	2963	1870	525	552	16	0	1	0
2	D	398	3010	1898	537	558	17	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

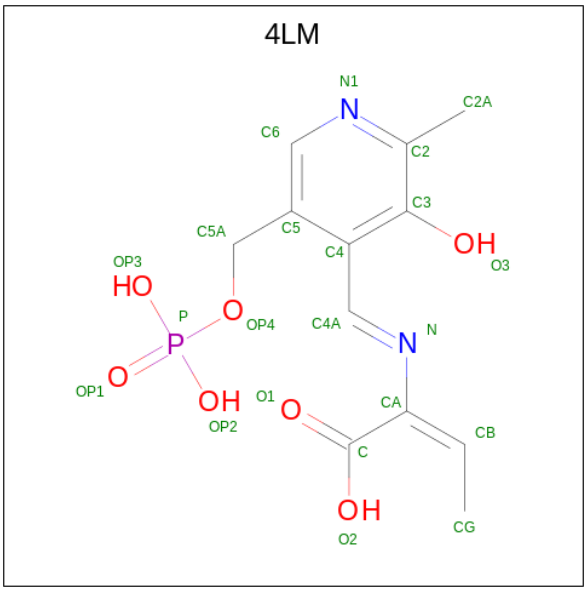
Chain	Residue	Modelled	Actual	Comment	Reference
B	116	HIS	CYS	engineered mutation	UNP P13254
D	116	HIS	CYS	engineered mutation	UNP P13254

- Molecule 3 is 2-AMINO-4-MERCAPTO-BUTYRIC ACID (three-letter code: HCS) (formula: $C_4H_9NO_2S$).



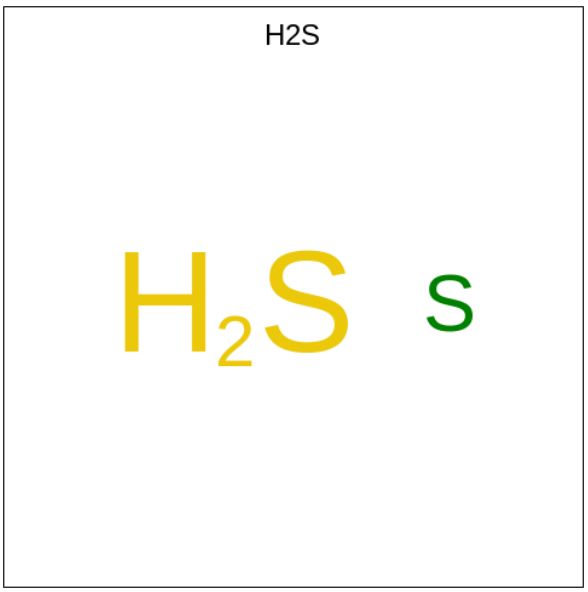
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	8	4	1	2	1	0	0
			Total	C	N	O	S		
3	C	1	8	4	1	2	1	0	0

- Molecule 4 is (2E)-2-{[(1E)-{3-hydroxy-2-methyl-5-[(phosphonooxy)methyl]pyridin-4-yl}methylidene]amino}but-2-enoic acid (three-letter code: 4LM) (formula: $C_{12}H_{15}N_2O_7P$).



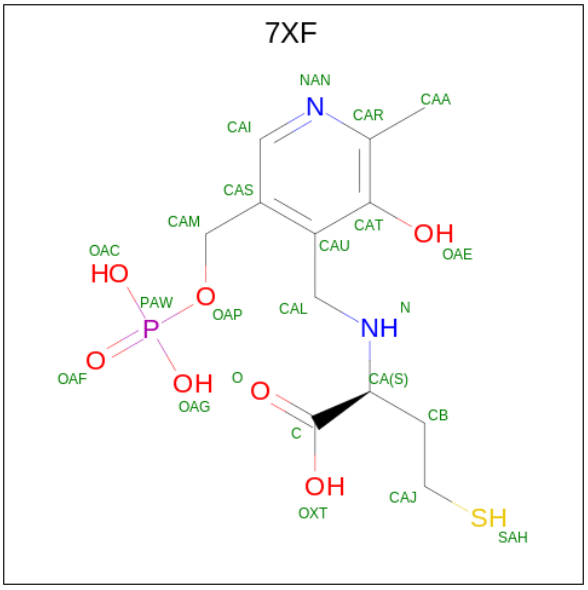
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	B	1	22	12	2	7	1	0	0

- Molecule 5 is HYDROSULFURIC ACID (three-letter code: H2S) (formula: H_2S).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	S		
5	B	1	1	1	0	0

- Molecule 6 is (2 {S})-2-[[2-methyl-3-oxidanyl-5-(phosphonooxymethyl)pyridin-4-yl]methylamino]-4-sulfanyl-butanoic acid (three-letter code: 7XF) (formula: $C_{12}H_{19}N_2O_7PS$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	N	O	P	S		
6	D	1	23	12	2	7	1	1	0	0

- Molecule 7 is water.

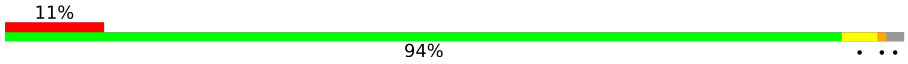
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	194	194	194	0	0
			Total	O		
7	B	292	292	292	0	0
			Total	O		
7	C	268	268	268	0	0
			Total	O		
7	D	267	267	267	0	0

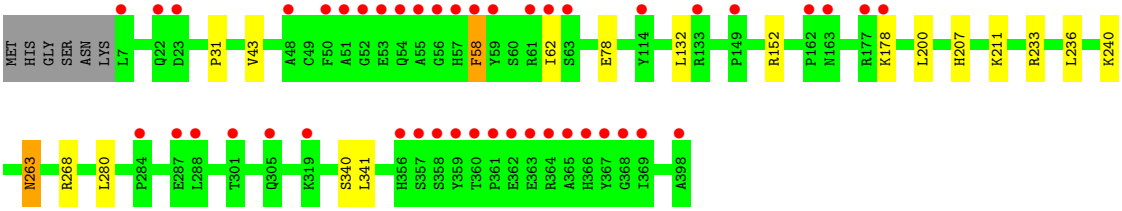
3 Residue-property plots (i)

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: L-methionine gamma-lyase

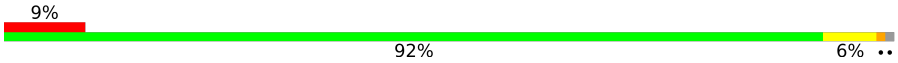
Chain A:

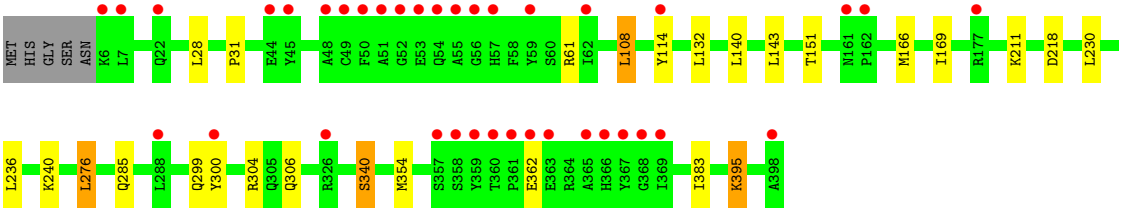




• Molecule 1: L-methionine gamma-lyase

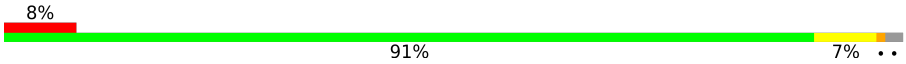
Chain C:

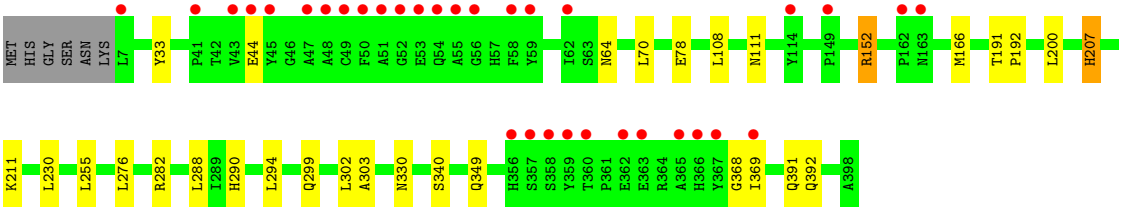




• Molecule 2: L-methionine gamma-lyase

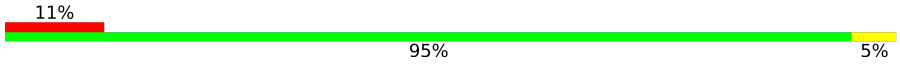
Chain B:

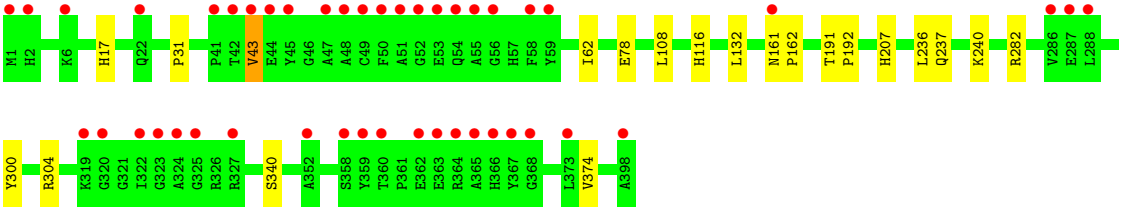




• Molecule 2: L-methionine gamma-lyase

Chain D:




M1
H2
K6
H17
Q22
P31
P41
T42
V43
E44
Y45
G46
A47
A48
C49
F50
A51
G52
E53
Q54
A55
G56
H57
F58
Y59
I62
E78
L108
H116
L132
N161
P162
T191
P192
H207
L236
Q237
K240
R282
V286
E287
L288
Y300
R304
K319
G320
G321
I322
G323
A324
G325
R326
R327
S340
A352
S358
Y359
T360
P361
E362
E363
R364
A365
H366
Y367
G368
L373
V374
A398

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	153.75Å 152.97Å 80.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.70 19.98 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.4 (20.00-1.70) 99.5 (19.98-1.70)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.28 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.170 , 0.195 0.180 , 0.206	Depositor DCC
R_{free} test set	10413 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	20.6	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 45.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.009 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	13001	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: LLP, HCS, 4LM, 7XF, H2S

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	0.45	0/3010	0.71	2/4086 (0.0%)
1	C	0.47	0/3015	0.71	1/4093 (0.0%)
2	B	0.49	0/3028	0.72	2/4111 (0.0%)
2	D	0.49	0/3077	0.70	1/4175 (0.0%)
All	All	0.48	0/12130	0.71	6/16465 (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
2	B	0	1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	282	ARG	NE-CZ-NH1	6.09	123.35	120.30
1	A	268	ARG	NE-CZ-NH1	5.56	123.08	120.30
2	B	152	ARG	NE-CZ-NH2	5.48	123.04	120.30
1	A	233	ARG	NE-CZ-NH1	5.39	123.00	120.30
2	D	282	ARG	NE-CZ-NH2	5.38	122.99	120.30
1	C	218	ASP	CB-CG-OD2	5.16	122.94	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	207	HIS	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2970	0	2926	6	0
1	C	2975	0	2928	13	0
2	B	2963	0	2926	15	0
2	D	3010	0	2975	12	0
3	A	8	0	8	0	0
3	C	8	0	8	2	0
4	B	22	0	11	2	0
5	B	1	0	0	0	0
6	D	23	0	0	0	0
7	A	194	0	0	0	0
7	B	292	0	0	1	0
7	C	268	0	0	1	0
7	D	267	0	0	3	0
All	All	13001	0	11782	41	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 2.

All (41) 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)
1:C:169:ILE:H	1:C:306:GLN:HE22	1.37	0.73
1:C:61:ARG:HH22	2:D:116:HIS:CD2	2.10	0.69
2:B:288:LEU:HD21	2:B:290:HIS:NE2	2.08	0.68
2:B:166:MET:H	2:B:299:GLN:HE22	1.43	0.65
2:B:211:LYS:HZ3	4:B:401:4LM:C4A	2.10	0.65
1:A:43:VAL:H	2:B:330:ASN:HD21	1.46	0.64
2:B:211:LYS:NZ	4:B:401:4LM:C4A	2.65	0.60
1:C:166:MET:H	1:C:299:GLN:HE22	1.51	0.59
2:B:166:MET:HE1	2:B:303:ALA:HA	1.86	0.57
1:C:236:LEU:O	1:C:240:LYS:HE3	2.08	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:33:TYR:CE2	1:C:31:PRO:HG3	2.45	0.52
2:D:300:TYR:CZ	2:D:304:ARG:HD2	2.45	0.51
1:C:276:LEU:HD13	1:C:383:ILE:HD11	1.93	0.50
1:C:354:MET:HG3	2:D:43:VAL:HG22	1.93	0.49
2:B:288:LEU:CD2	2:B:290:HIS:CE1	2.95	0.48
2:D:78:GLU:OE1	2:D:207:HIS:HE1	1.97	0.48
2:B:78:GLU:OE2	2:B:207:HIS:HE1	1.98	0.47
1:A:78:GLU:OE2	1:A:207:HIS:CE1	2.69	0.46
2:B:78:GLU:OE2	2:B:207:HIS:CE1	2.69	0.45
2:D:236:LEU:O	2:D:240:LYS:HE3	2.15	0.45
1:C:114:TYR:CE1	3:C:501:HCS:HG2	2.51	0.45
2:D:78:GLU:OE1	2:D:207:HIS:CE1	2.69	0.45
2:D:237:GLN:NE2	7:D:602:HOH:O	2.47	0.45
2:D:17:HIS:HD2	7:D:845:HOH:O	1.99	0.45
1:A:31:PRO:HB2	2:D:31:PRO:HB2	1.99	0.45
2:D:191:THR:HB	2:D:192:PRO:HD2	1.98	0.44
1:A:58:PHE:CZ	1:A:62:ILE:HG21	2.53	0.44
2:B:288:LEU:HD21	2:B:290:HIS:CE1	2.53	0.44
1:A:263:ASN:HD22	1:A:263:ASN:H	1.66	0.43
2:B:191:THR:HB	2:B:192:PRO:HD2	2.01	0.43
1:A:236:LEU:O	1:A:240:LYS:HE3	2.18	0.43
1:C:108:LEU:HD22	1:C:151:THR:HG21	2.01	0.42
1:C:340:SER:HA	3:C:501:HCS:CB	2.49	0.42
2:B:70:LEU:HD21	2:B:255:LEU:HD23	2.01	0.42
1:C:300:TYR:CZ	1:C:304:ARG:HD2	2.54	0.42
2:B:349:GLN:NE2	7:B:505:HOH:O	2.51	0.42
2:B:368:GLY:C	2:B:369:ILE:HD12	2.39	0.42
2:D:62:ILE:HA	7:D:651:HOH:O	2.19	0.42
1:C:230:LEU:HD12	7:C:617:HOH:O	2.20	0.41
1:C:395:LYS:N	1:C:395:LYS:HD3	2.34	0.41
2:D:161:ASN:HA	2:D:162:PRO:HA	1.92	0.40

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	389/398 (98%)	376 (97%)	12 (3%)	1 (0%)	41	24
1	C	390/398 (98%)	380 (97%)	9 (2%)	1 (0%)	41	24
2	B	391/398 (98%)	380 (97%)	10 (3%)	1 (0%)	41	24
2	D	397/398 (100%)	386 (97%)	10 (2%)	1 (0%)	41	24
All	All	1567/1592 (98%)	1522 (97%)	41 (3%)	4 (0%)	41	24

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	340	SER
1	C	340	SER
2	D	340	SER
1	A	340	SER

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	301/306 (98%)	293 (97%)	8 (3%)	44	26
1	C	301/306 (98%)	292 (97%)	9 (3%)	41	22
2	B	303/307 (99%)	291 (96%)	12 (4%)	31	13
2	D	308/307 (100%)	304 (99%)	4 (1%)	69	56
All	All	1213/1226 (99%)	1180 (97%)	33 (3%)	44	26

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	PHE
1	A	132	LEU
1	A	152	ARG
1	A	178	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	200	LEU
1	A	263	ASN
1	A	280	LEU
1	A	341	LEU
2	B	44	GLU
2	B	64	ASN
2	B	108	LEU
2	B	111	ASN
2	B	152	ARG
2	B	200	LEU
2	B	230	LEU
2	B	276	LEU
2	B	294	LEU
2	B	302	LEU
2	B	391	GLN
2	B	392	GLN
1	C	28	LEU
1	C	108	LEU
1	C	132	LEU
1	C	140	LEU
1	C	143	LEU
1	C	276	LEU
1	C	285	GLN
1	C	362	GLU
1	C	395	LYS
2	D	43	VAL
2	D	108	LEU
2	D	132	LEU
2	D	374	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (38) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	HIS
1	A	22	GLN
1	A	134	HIS
1	A	163	ASN
1	A	207	HIS
1	A	250	HIS
1	A	263	ASN
1	A	274	GLN
1	A	305	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	57	HIS
2	B	64	ASN
2	B	68	ASN
2	B	111	ASN
2	B	207	HIS
2	B	228	GLN
2	B	299	GLN
2	B	330	ASN
2	B	349	GLN
2	B	392	GLN
1	C	34	GLN
1	C	57	HIS
1	C	68	ASN
1	C	165	HIS
1	C	187	ASN
1	C	250	HIS
1	C	285	GLN
1	C	299	GLN
1	C	306	GLN
2	D	17	HIS
2	D	34	GLN
2	D	68	ASN
2	D	116	HIS
2	D	187	ASN
2	D	195	GLN
2	D	207	HIS
2	D	237	GLN
2	D	274	GLN
2	D	330	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)

2 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	LLP	C	211	1	23,24,25	2.85	5 (21%)	25,32,34	1.38	6 (24%)
1	LLP	A	211	1	23,24,25	2.86	5 (21%)	25,32,34	1.39	4 (16%)

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
1	LLP	C	211	1	-	5/16/17/19	0/1/1/1
1	LLP	A	211	1	-	4/16/17/19	0/1/1/1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	211	LLP	C3-C2	7.64	1.48	1.40
1	A	211	LLP	C3-C2	7.02	1.47	1.40
1	A	211	LLP	C4-C5	6.75	1.50	1.42
1	C	211	LLP	C4-C3	6.18	1.50	1.40
1	C	211	LLP	C4'-NZ	5.99	1.47	1.27
1	A	211	LLP	C4'-NZ	5.90	1.47	1.27
1	A	211	LLP	C4-C3	5.84	1.49	1.40
1	C	211	LLP	C4-C5	5.57	1.49	1.42
1	C	211	LLP	C4-C4'	3.51	1.53	1.46
1	A	211	LLP	C4-C4'	3.11	1.52	1.46

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	211	LLP	C6-N1-C2	3.10	124.91	119.17
1	A	211	LLP	C2'-C2-N1	2.68	122.90	117.67
1	C	211	LLP	C6-N1-C2	2.64	124.05	119.17
1	C	211	LLP	C3-C4-C5	-2.43	116.39	118.26
1	A	211	LLP	OP3-P-OP2	2.38	116.73	107.64
1	C	211	LLP	C4-C4'-NZ	-2.28	113.84	124.31
1	A	211	LLP	C4-C3-C2	-2.24	118.80	120.19
1	C	211	LLP	OP4-P-OP1	-2.15	100.43	106.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	211	LLP	OP2-P-OP1	2.13	119.02	110.68
1	C	211	LLP	C3-C2-N1	-2.02	118.16	120.77

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	211	LLP	O-C-CA-CB
1	C	211	LLP	O-C-CA-CB
1	C	211	LLP	C4-C4'-NZ-CE
1	A	211	LLP	C4-C4'-NZ-CE
1	C	211	LLP	CA-CB-CG-CD
1	A	211	LLP	CG-CD-CE-NZ
1	C	211	LLP	CD-CE-NZ-C4'
1	C	211	LLP	C3-C4-C4'-NZ
1	A	211	LLP	CD-CE-NZ-C4'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 1 is modelled with single atom - leaving 4 for Mogul analysis.

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	$\#\|Z\| > 2$
4	4LM	B	401	-	21,22,22	1.73	4 (19%)	25,31,31	1.88	7 (28%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	HCS	C	501	-	6,7,7	1.14	1 (16%)	7,8,8	2.61	3 (42%)
6	7XF	D	501	-	23,23,23	2.19	5 (21%)	27,32,32	2.53	5 (18%)
3	HCS	A	501	-	6,7,7	0.98	1 (16%)	7,8,8	3.47	3 (42%)

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
4	4LM	B	401	-	-	0/15/17/17	0/1/1/1
3	HCS	C	501	-	-	1/7/7/7	-
6	7XF	D	501	-	-	5/18/18/18	0/1/1/1
3	HCS	A	501	-	-	2/7/7/7	-

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	501	7XF	CB-CA	-7.30	1.36	1.53
4	B	401	4LM	C4-C4A	3.75	1.53	1.46
6	D	501	7XF	CAR-NAN	3.62	1.40	1.33
4	B	401	4LM	C2-N1	3.32	1.40	1.33
4	B	401	4LM	P-OP2	-2.61	1.44	1.54
6	D	501	7XF	CAT-CAR	-2.50	1.38	1.40
6	D	501	7XF	CA-C	-2.43	1.46	1.52
3	C	501	HCS	OXT-C	-2.39	1.22	1.30
4	B	401	4LM	C3-C2	-2.27	1.38	1.40
3	A	501	HCS	OXT-C	-2.24	1.23	1.30
6	D	501	7XF	CAL-CAU	2.21	1.54	1.51

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	D	501	7XF	CB-CAJ-SAH	-9.96	103.36	113.74
3	A	501	HCS	CB-CG-SD	-8.23	105.16	113.74
4	B	401	4LM	O2-C-CA	5.69	122.80	113.42
3	C	501	HCS	CG-CB-CA	5.54	122.19	113.14
6	D	501	7XF	CAL-N-CA	5.03	123.45	113.92
4	B	401	4LM	O2-C-O1	-3.47	115.67	123.61
6	D	501	7XF	OXT-C-O	-3.24	116.74	124.09
3	C	501	HCS	CB-CA-N	3.17	118.48	110.17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	401	4LM	C3-C2-N1	-3.02	116.87	120.77
4	B	401	4LM	C4-C3-C2	2.87	121.97	120.19
6	D	501	7XF	CAT-CAR-NAN	-2.64	117.35	120.77
3	A	501	HCS	OXT-C-O	-2.63	118.12	124.09
6	D	501	7XF	OXT-C-CA	2.45	121.55	113.40
4	B	401	4LM	OP2-P-OP4	2.36	113.02	106.73
4	B	401	4LM	C2A-C2-N1	2.32	122.20	117.67
4	B	401	4LM	OP4-C5A-C5	-2.12	105.31	109.35
3	C	501	HCS	OXT-C-O	-2.12	119.28	124.09
3	A	501	HCS	OXT-C-CA	2.00	120.20	113.38

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	HCS	N-CA-CB-CG
3	A	501	HCS	C-CA-CB-CG
3	C	501	HCS	CA-CB-CG-SD
6	D	501	7XF	N-CAL-CAU-CAS
6	D	501	7XF	CB-CA-N-CAL
6	D	501	7XF	C-CA-N-CAL
6	D	501	7XF	N-CAL-CAU-CAT
6	D	501	7XF	N-CA-CB-CAJ

There are no ring outliers.

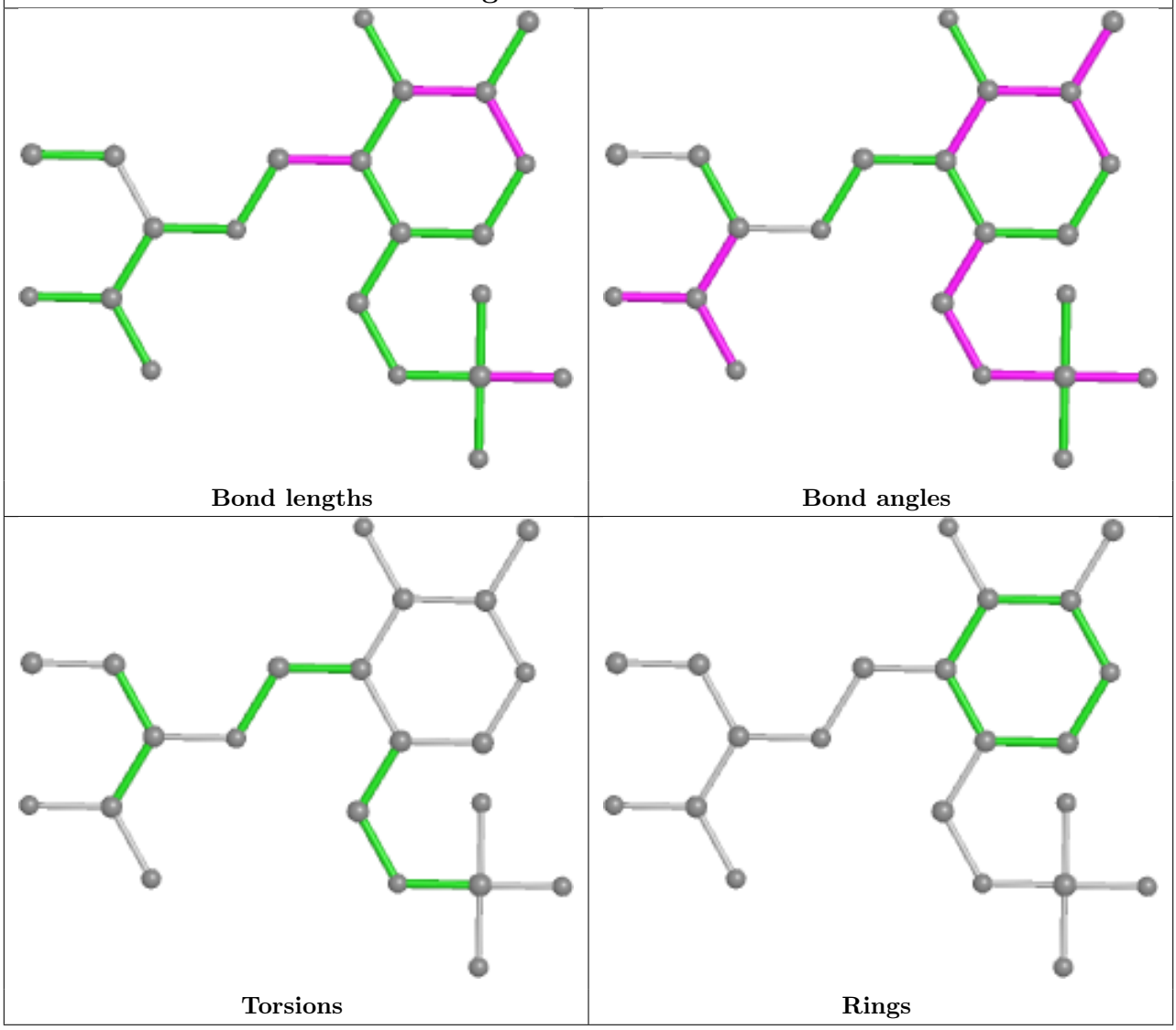
2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	401	4LM	2	0
3	C	501	HCS	2	0

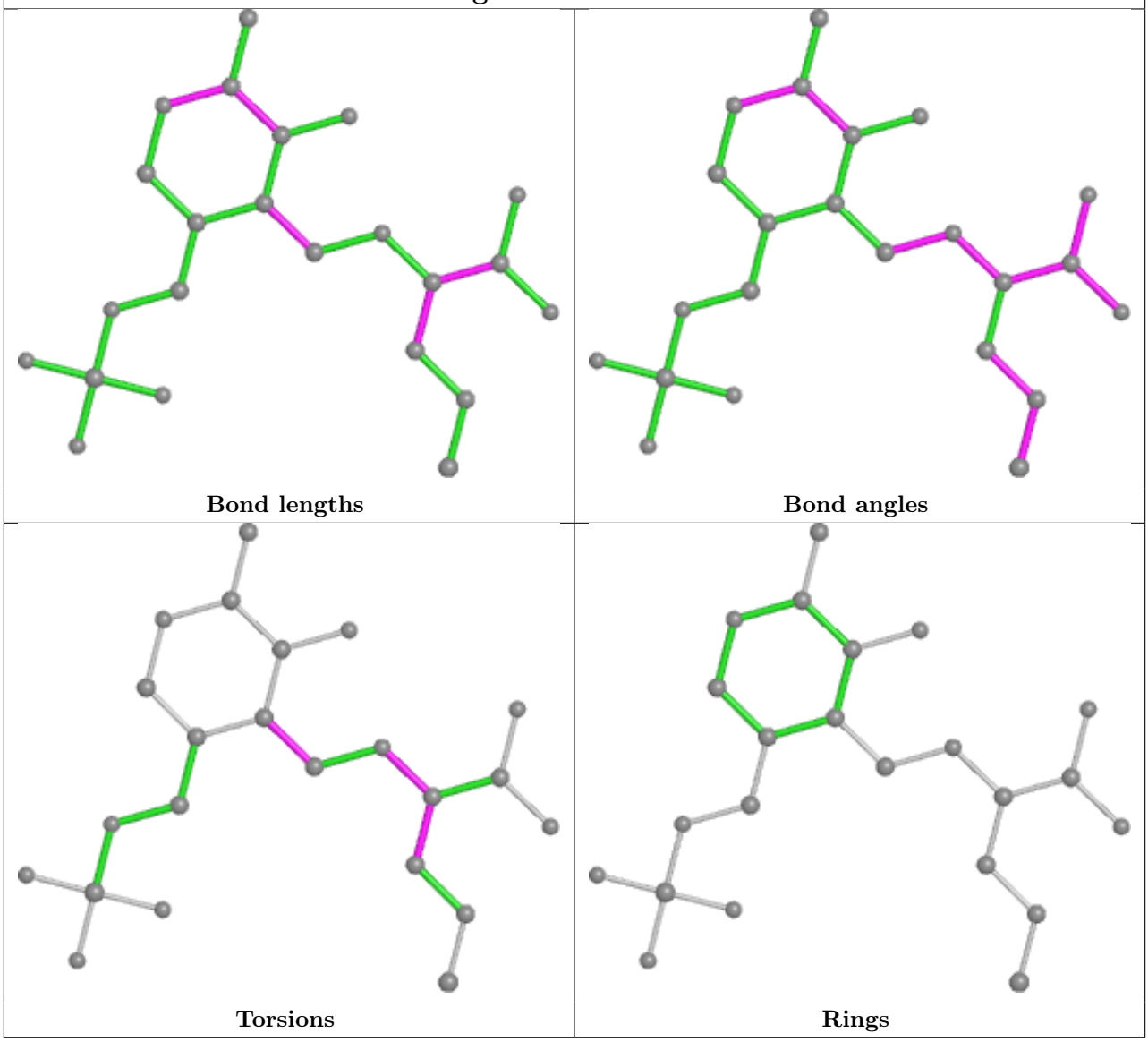
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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 4LM B 401



Ligand 7XF D 501



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	391/398 (98%)	0.59	45 (11%) 4 5	15, 29, 71, 119	0
1	C	392/398 (98%)	0.29	37 (9%) 8 9	13, 23, 68, 98	0
2	B	392/398 (98%)	0.22	33 (8%) 11 12	13, 21, 57, 112	0
2	D	398/398 (100%)	0.44	45 (11%) 5 6	12, 21, 74, 111	0
All	All	1573/1592 (98%)	0.39	160 (10%) 6 7	12, 23, 68, 119	0

All (160) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	55	ALA	14.7
2	D	51	ALA	14.3
2	D	365	ALA	12.6
2	B	52	GLY	12.3
2	B	55	ALA	12.2
2	D	55	ALA	11.4
2	B	51	ALA	11.1
1	C	367	TYR	10.6
1	A	366	HIS	10.1
1	A	365	ALA	9.3
1	C	55	ALA	9.3
2	D	52	GLY	8.4
1	A	54	GLN	8.3
1	A	52	GLY	8.2
2	D	368	GLY	8.0
1	C	51	ALA	7.9
2	D	367	TYR	7.9
1	A	53	GLU	7.7
2	D	2	HIS	7.6
1	A	360	THR	7.5
1	A	358	SER	7.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	366	HIS	7.2
1	A	59	TYR	7.1
2	D	1	MET	7.0
2	D	360	THR	7.0
1	C	366	HIS	6.9
2	B	359	TYR	6.9
1	A	363	GLU	6.9
2	B	53	GLU	6.8
1	C	53	GLU	6.7
1	A	367	TYR	6.6
1	C	52	GLY	6.1
1	C	362	GLU	6.1
2	B	45	TYR	6.1
1	C	360	THR	6.0
2	B	366	HIS	6.0
2	B	48	ALA	5.8
2	D	359	TYR	5.8
2	D	53	GLU	5.7
2	D	362	GLU	5.5
1	C	359	TYR	5.3
2	B	367	TYR	5.2
1	C	358	SER	5.1
2	B	7	LEU	5.1
1	A	359	TYR	5.1
1	C	54	GLN	5.0
2	D	398	ALA	5.0
1	A	7	LEU	5.0
1	C	365	ALA	4.9
2	D	49	CYS	4.9
1	A	56	GLY	4.9
1	C	48	ALA	4.9
1	C	162	PRO	4.8
1	A	362	GLU	4.8
1	C	56	GLY	4.7
1	C	398	ALA	4.7
2	B	358	SER	4.7
2	D	50	PHE	4.6
2	D	358	SER	4.6
1	A	288	LEU	4.4
2	B	50	PHE	4.4
1	A	58	PHE	4.3
1	C	6	LYS	4.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	320	GLY	4.3
2	B	49	CYS	4.2
1	C	45	TYR	4.2
1	A	398	ALA	4.1
2	D	48	ALA	4.1
2	D	45	TYR	4.0
2	D	54	GLN	4.0
2	D	324	ALA	4.0
1	A	287	GLU	3.9
2	B	365	ALA	3.9
1	C	50	PHE	3.8
1	A	178	LYS	3.8
1	A	364	ARG	3.6
2	D	43	VAL	3.6
2	D	59	TYR	3.5
1	A	51	ALA	3.5
1	C	49	CYS	3.5
1	A	62	ILE	3.5
1	A	361	PRO	3.5
2	B	56	GLY	3.4
1	A	163	ASN	3.3
1	A	369	ILE	3.3
1	C	62	ILE	3.3
2	B	362	GLU	3.3
1	A	357	SER	3.2
1	A	50	PHE	3.2
2	B	114	TYR	3.2
2	D	44	GLU	3.1
1	C	363	GLU	3.1
2	B	149	PRO	3.1
1	A	356	HIS	3.1
1	A	368	GLY	3.1
1	C	7	LEU	3.1
2	D	323	GLY	3.0
2	B	162	PRO	3.0
1	A	61	ARG	3.0
2	D	56	GLY	3.0
2	D	58	PHE	3.0
1	A	284	PRO	2.9
2	D	47	ALA	2.9
2	B	54	GLN	2.9
2	D	322	ILE	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	368	GLY	2.8
2	B	369	ILE	2.8
2	D	288	LEU	2.8
1	A	63	SER	2.8
1	C	288	LEU	2.8
2	D	42	THR	2.8
1	A	319	LYS	2.7
2	B	360	THR	2.7
1	C	161	ASN	2.7
1	A	48	ALA	2.7
2	B	58	PHE	2.6
1	C	361	PRO	2.6
2	B	47	ALA	2.6
2	B	363	GLU	2.6
1	A	177	ARG	2.6
1	C	177	ARG	2.6
2	B	41	PRO	2.6
2	B	163	ASN	2.5
2	D	161	ASN	2.5
1	A	149	PRO	2.5
1	C	357	SER	2.5
2	D	286	VAL	2.5
2	D	327	ARG	2.4
1	C	114	TYR	2.4
2	D	287	GLU	2.4
1	A	301	THR	2.4
1	C	44	GLU	2.4
2	B	357	SER	2.4
2	B	43	VAL	2.3
2	D	41	PRO	2.3
1	A	23	ASP	2.3
1	C	22	GLN	2.3
1	A	22	GLN	2.3
2	D	363	GLU	2.2
1	A	57	HIS	2.2
1	A	114	TYR	2.2
1	C	59	TYR	2.2
2	B	59	TYR	2.2
2	D	364	ARG	2.2
2	D	6	LYS	2.2
2	D	325	GLY	2.2
1	A	162	PRO	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	133	ARG	2.1
2	B	62	ILE	2.1
2	B	356	HIS	2.1
2	D	22	GLN	2.1
1	C	57	HIS	2.1
2	D	352	ALA	2.1
1	C	300	TYR	2.1
2	D	319	LYS	2.1
2	D	373	LEU	2.1
1	C	369	ILE	2.1
1	C	326	ARG	2.1
1	A	305	GLN	2.0
2	B	44	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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
1	LLP	A	211	24/25	0.95	0.10	17,26,31,34	0
1	LLP	C	211	24/25	0.96	0.09	14,21,26,30	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	HCS	C	501	8/8	0.68	0.27	60,63,64,76	0
3	HCS	A	501	8/8	0.79	0.25	58,63,69,79	0
5	H2S	B	402	1/1	0.84	0.08	74,74,74,74	0

Continued on next page...

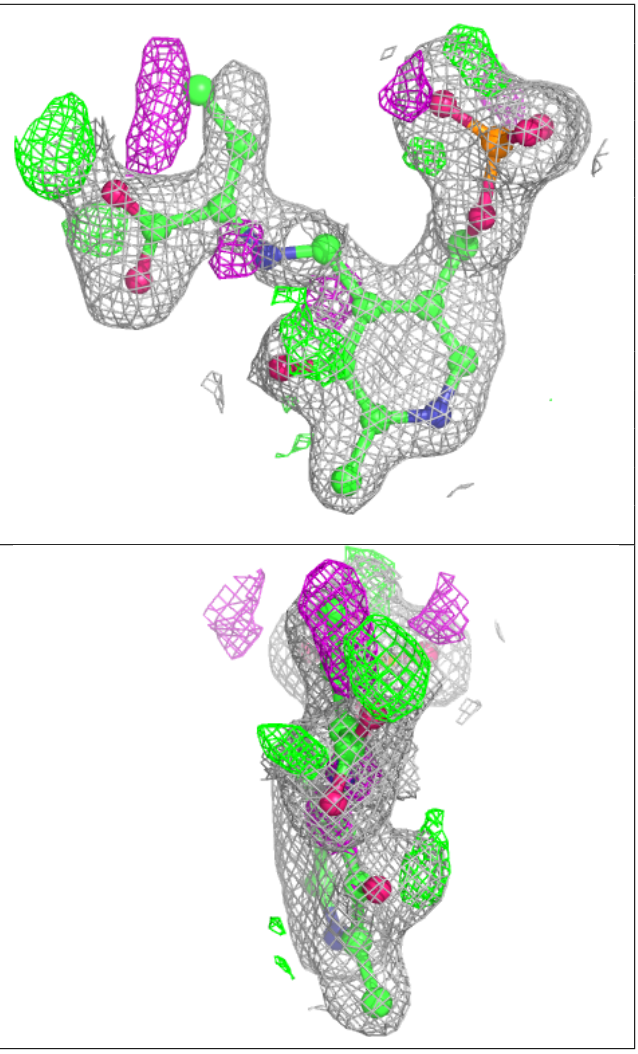
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	4LM	B	401	22/22	0.95	0.13	19,32,47,50	0
6	7XF	D	501	23/23	0.95	0.12	16,26,54,71	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.

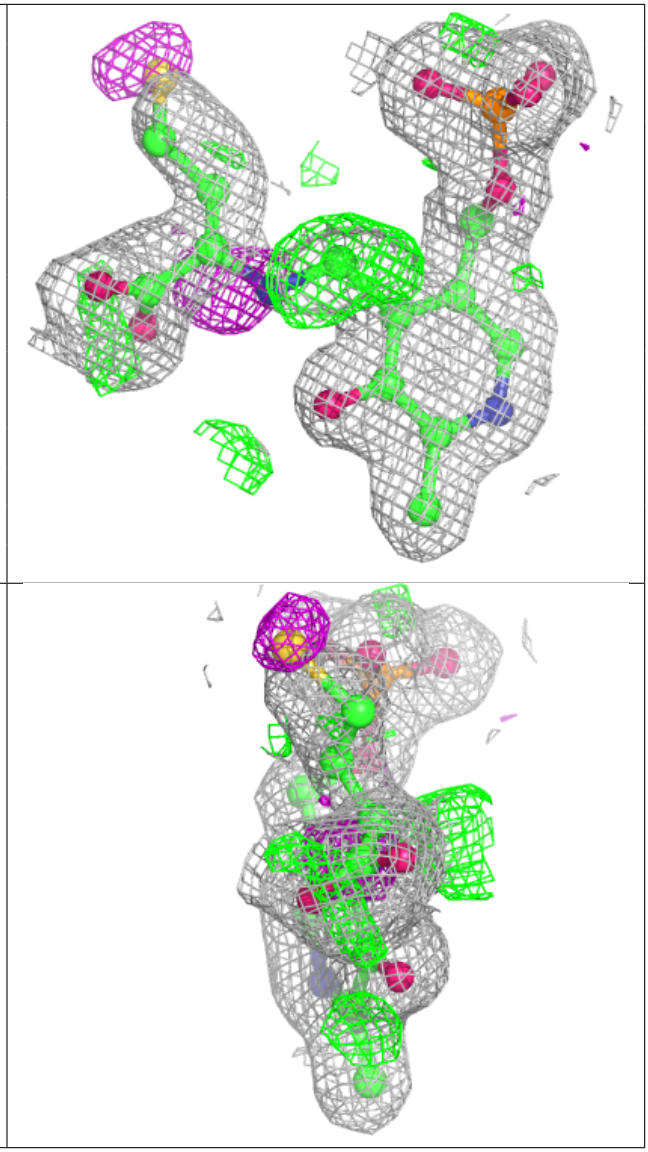
Electron density around 4LM B 401:

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 7XF D 501:

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



6.5 Other polymers (i)

There are no such residues in this entry.