Full wwPDB X-ray Structure Validation Report (i)

May 10, 2022 – 10:07 pm BST

PDB ID	:	7PQO
Title	:	Catalytic fragment of MASP-1 in complex with P1 site mutant ecotin
Authors	:	Harmat, V.; Fodor, K.; Heja, D.
Deposited on	:	2021-09-17
Resolution	:	3.39 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 following versions of software and data (see references (1)) were used in the production of this report:

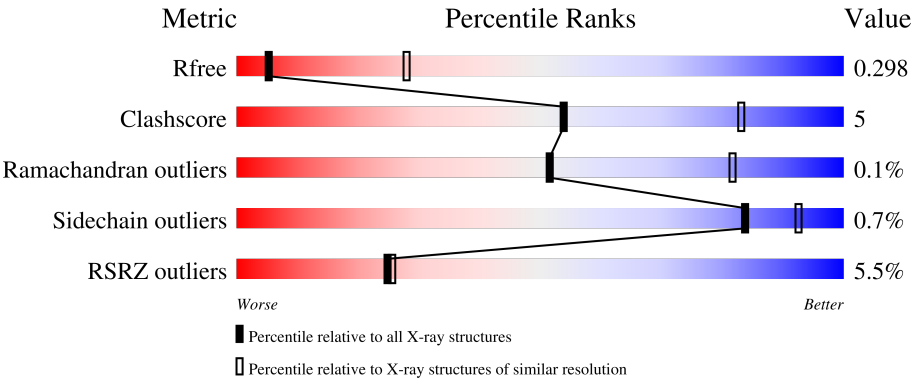
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

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 3.39 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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	406	5% 85% 13% ·
1	B	406	6% 85% 10% 5%
1	C	406	6% 86% 11% ·
2	I	162	2% 69% 17% 14%
2	J	162	7% 74% 12% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	K	162	2% 76% 10% 14%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	703	-	-	-	X
3	GOL	I	202	-	-	-	X
4	NA	C	702	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 11373 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 Mannan-binding lectin serine protease 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	393	Total	C	N	O	S	0	0	0
			2815	1773	477	540	25			
1	A	398	Total	C	N	O	S	0	0	0
			2841	1798	477	541	25			
1	B	385	Total	C	N	O	S	0	0	0
			2585	1621	438	503	23			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	294	ALA	-	expression tag	UNP P48740
C	295	SER	-	expression tag	UNP P48740
C	296	MET	-	expression tag	UNP P48740
C	297	THR	-	expression tag	UNP P48740
C	499	LYS	GLU	variant	UNP P48740
A	294	ALA	-	expression tag	UNP P48740
A	295	SER	-	expression tag	UNP P48740
A	296	MET	-	expression tag	UNP P48740
A	297	THR	-	expression tag	UNP P48740
A	499	LYS	GLU	variant	UNP P48740
B	294	ALA	-	expression tag	UNP P48740
B	295	SER	-	expression tag	UNP P48740
B	296	MET	-	expression tag	UNP P48740
B	297	THR	-	expression tag	UNP P48740
B	499	LYS	GLU	variant	UNP P48740

- Molecule 2 is a protein called Ecotin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	K	139	Total	C	N	O	S	0	0	0
			1043	665	172	201	5			

Continued on next page...

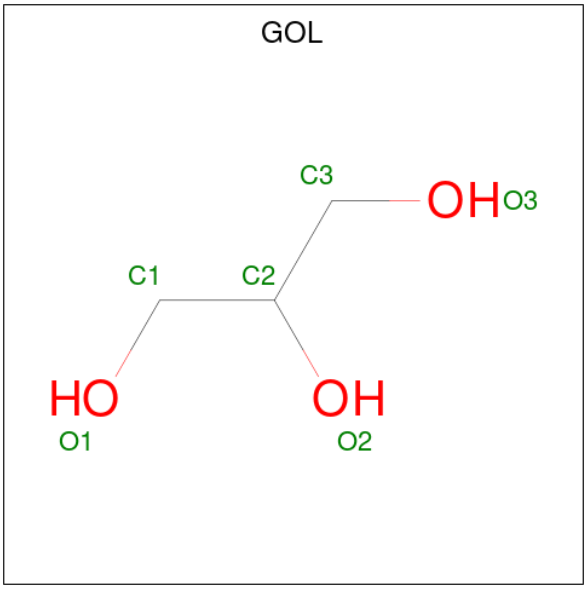
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	139	Total	C	N	O	S	0	0	0
			1026	656	167	198	5			
2	J	139	Total	C	N	O	S	0	0	0
			1020	654	172	189	5			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	84	ARG	MET	engineered mutation	UNP P23827
I	84	ARG	MET	engineered mutation	UNP P23827
J	84	ARG	MET	engineered mutation	UNP P23827

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total	C	O		0	0
			6	3	3			
3	K	1	Total	C	O		0	0
			6	3	3			
3	A	1	Total	C	O		0	0
			6	3	3			
3	A	1	Total	C	O		0	0
			6	3	3			
3	A	1	Total	C	O		0	0
			6	3	3			
3	I	1	Total	C	O		0	0
			6	3	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	I	1	Total C O 6 3 3	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

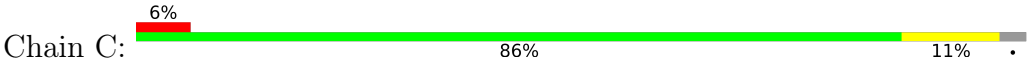
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total Na 1 1	0	0

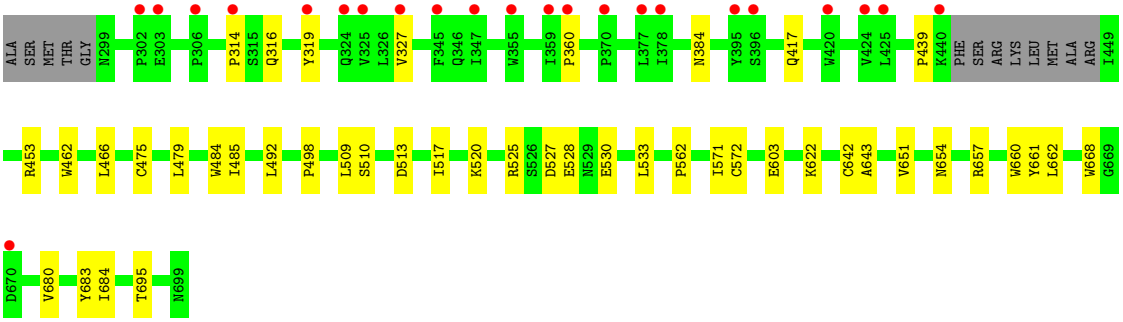
3 Residue-property plots (i)

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: Mannan-binding lectin serine protease 1

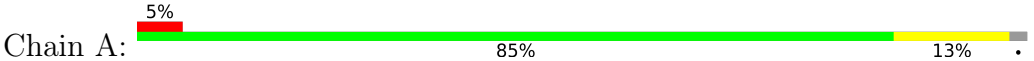
Chain C:

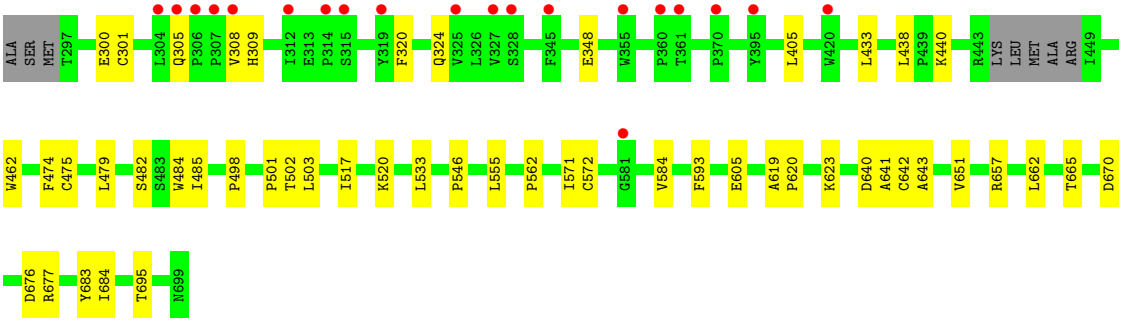




• Molecule 1: Mannan-binding lectin serine protease 1

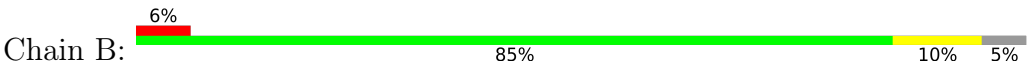
Chain A:

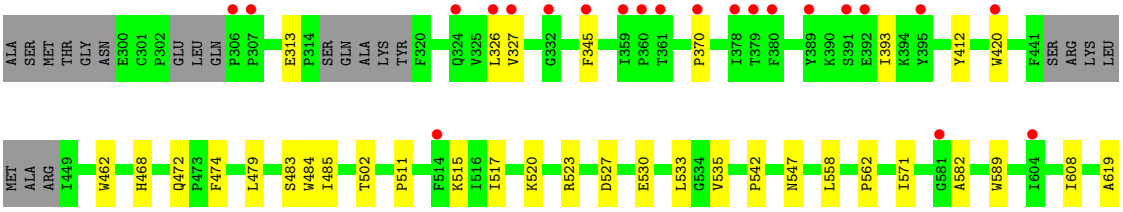


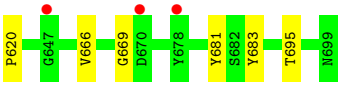


• Molecule 1: Mannan-binding lectin serine protease 1

Chain B:

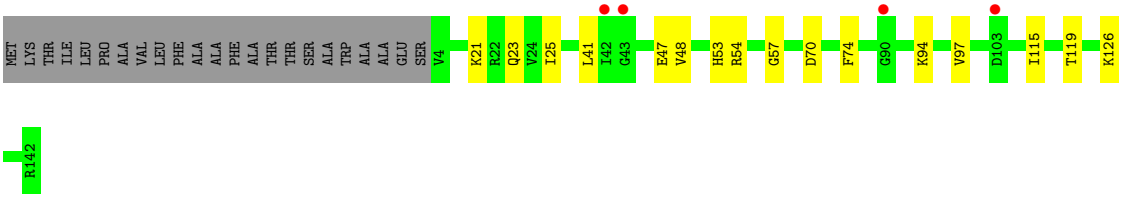






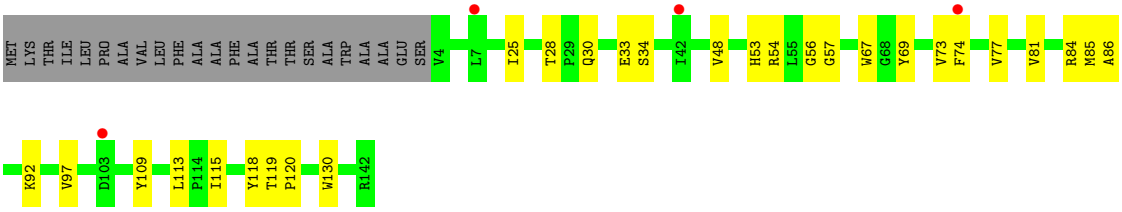
• Molecule 2: Ecotin

Chain K:



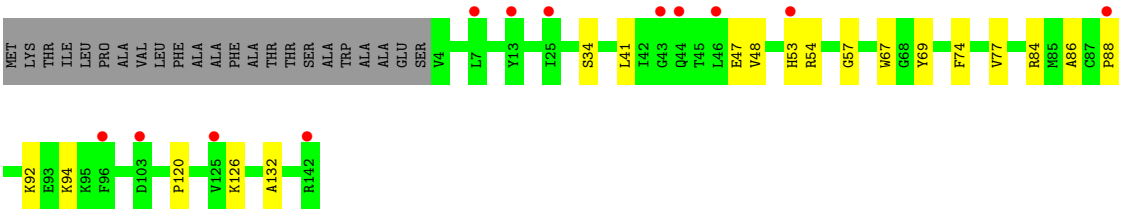
• Molecule 2: Ecotin

Chain I:



• Molecule 2: Ecotin

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α, β, γ	251.07Å 251.07Å 211.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.60 – 3.39 39.60 – 3.39	Depositor EDS
% Data completeness (in resolution range)	96.3 (39.60-3.39) 79.5 (39.60-3.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.74 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660, PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.257 , 0.297 0.258 , 0.298	Depositor DCC
R_{free} test set	2261 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	93.4	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	11373	wwPDB-VP
Average B, all atoms (Å^2)	134.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.15% 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: GOL, NA

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	0.25	0/2919	0.45	0/4002
1	B	0.24	0/2655	0.45	0/3653
1	C	0.26	0/2892	0.45	0/3966
2	I	0.24	0/1048	0.43	0/1436
2	J	0.23	0/1042	0.45	0/1429
2	K	0.24	0/1065	0.44	0/1457
All	All	0.25	0/11621	0.45	0/15943

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2841	0	2440	31	0
1	B	2585	0	2035	24	0
1	C	2815	0	2411	27	0
2	I	1026	0	931	21	0
2	J	1020	0	934	14	0
2	K	1043	0	967	11	0
3	A	18	0	24	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	6	0	8	1	0
3	I	12	0	16	0	0
3	K	6	0	8	1	0
4	C	1	0	0	0	0
All	All	11373	0	9774	112	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 5.

All (112) 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:B:502:THR:OG1	2:I:34:SER:HB2	1.91	0.71
1:B:468:HIS:O	1:B:515:LYS:NZ	2.31	0.64
1:A:502:THR:HG21	2:J:34:SER:HB2	1.80	0.63
1:A:498:PRO:HB2	2:I:97:VAL:HA	1.79	0.62
1:B:533:LEU:HD11	1:B:562:PRO:HB3	1.83	0.61
1:A:503:LEU:HG	1:A:546:PRO:HG3	1.83	0.61
2:K:41:LEU:HB2	2:K:126:LYS:HB2	1.82	0.60
1:A:300:GLU:HA	1:A:320:PHE:HA	1.83	0.60
2:I:74:PHE:HB3	2:I:119:THR:HG22	1.83	0.59
2:I:25:ILE:HB	2:I:115:ILE:HB	1.85	0.59
1:B:547:ASN:O	2:J:54:ARG:HD2	2.03	0.58
1:A:665:THR:HG23	1:A:684:ILE:HD11	1.86	0.58
1:C:417:GLN:OE1	1:C:417:GLN:N	2.37	0.57
1:C:520:LYS:NZ	1:C:525:ARG:O	2.34	0.57
1:B:483:SER:HA	1:B:558:LEU:HD12	1.87	0.56
1:A:474:PHE:HA	2:I:86:ALA:HB3	1.88	0.56
1:A:670:ASP:HA	2:I:81:VAL:HG12	1.88	0.54
1:B:517:ILE:HG21	1:B:520:LYS:HD2	1.89	0.53
1:C:571:ILE:HG12	1:C:662:LEU:HB2	1.90	0.53
1:C:651:VAL:HG12	1:C:662:LEU:HA	1.91	0.52
1:B:582:ALA:HB3	1:B:608:ILE:HD12	1.90	0.52
2:K:74:PHE:HB3	2:K:119:THR:HG22	1.92	0.51
1:C:484:TRP:CD2	1:C:695:THR:HG22	2.46	0.51
1:C:517:ILE:HG21	1:C:520:LYS:HD2	1.92	0.51
1:C:453:ARG:NH2	1:C:603:GLU:OE2	2.45	0.50
1:A:651:VAL:HG12	1:A:662:LEU:HA	1.94	0.50
1:C:527:ASP:OD1	1:C:530:GLU:HG3	2.12	0.50
1:A:484:TRP:CD2	1:A:695:THR:HG22	2.47	0.49
2:K:21:LYS:HE3	2:K:23:GLN:NE2	2.27	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:316:GLN:HB3	1:C:319:TYR:CE2	2.47	0.49
1:B:313:GLU:CB	1:B:326:LEU:H	2.26	0.49
2:I:77:VAL:HG21	2:I:120:PRO:HG2	1.95	0.48
1:C:642:CYS:SG	1:C:643:ALA:N	2.84	0.48
2:K:47:GLU:HA	2:K:94:LYS:HA	1.96	0.48
1:A:517:ILE:HG21	1:A:520:LYS:HD2	1.94	0.48
1:C:622:LYS:HB3	2:K:57:GLY:H	1.79	0.48
1:A:475:CYS:SG	2:I:85:MET:HG3	2.54	0.48
1:C:533:LEU:HD11	1:C:562:PRO:HB3	1.95	0.47
1:C:572:CYS:HB2	1:C:661:TYR:HA	1.96	0.47
1:B:619:ALA:N	1:B:620:PRO:HD2	2.30	0.47
2:J:48:VAL:HB	2:J:53:HIS:NE2	2.29	0.47
1:B:484:TRP:CD2	1:B:695:THR:HG22	2.50	0.47
2:K:70:ASP:OD2	3:K:901:GOL:O3	2.27	0.47
1:A:533:LEU:HD11	1:A:562:PRO:HB3	1.96	0.47
1:A:555:LEU:HD13	1:A:695:THR:HG23	1.96	0.47
1:B:370:PRO:HG3	1:B:420:TRP:CE2	2.49	0.47
2:J:41:LEU:HB2	2:J:126:LYS:HB2	1.96	0.47
1:B:669:GLY:HA3	2:J:84:ARG:NH2	2.30	0.47
1:A:623:LYS:HA	1:A:623:LYS:HD2	1.64	0.46
1:A:482:SER:O	1:A:562:PRO:HD2	2.15	0.46
1:A:571:ILE:HG12	1:A:662:LEU:HB2	1.98	0.46
2:J:47:GLU:HA	2:J:94:LYS:HA	1.98	0.46
2:J:67:TRP:HB3	2:J:69:TYR:CD1	2.50	0.46
1:C:510:SER:OG	1:C:513:ASP:OD2	2.33	0.46
1:C:439:PRO:HG3	1:C:660:TRP:HE1	1.80	0.46
1:C:462:TRP:CE2	1:C:571:ILE:HD13	2.51	0.45
2:K:48:VAL:HB	2:K:53:HIS:NE2	2.32	0.45
1:C:498:PRO:HB2	2:K:97:VAL:HA	1.99	0.45
2:I:67:TRP:HB3	2:I:69:TYR:CD1	2.52	0.45
1:B:527:ASP:OD1	1:B:530:GLU:HG3	2.16	0.44
2:I:48:VAL:HB	2:I:53:HIS:NE2	2.32	0.44
1:B:479:LEU:HA	1:B:485:ILE:HD13	1.99	0.44
1:C:662:LEU:HG	1:C:684:ILE:HD12	1.99	0.44
1:A:642:CYS:SG	1:A:643:ALA:N	2.86	0.44
1:A:619:ALA:N	1:A:620:PRO:HD2	2.32	0.44
2:K:57:GLY:HA3	2:K:74:PHE:CZ	2.53	0.44
1:A:501:PRO:HG3	2:I:54:ARG:CZ	2.48	0.44
2:I:28:THR:O	2:I:109:TYR:OH	2.27	0.44
1:C:492:LEU:HA	1:C:509:LEU:HD12	2.00	0.44
1:A:642:CYS:O	2:I:84:ARG:HB3	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:474:PHE:HA	2:J:86:ALA:HB3	1.98	0.43
1:B:327:VAL:HG12	1:B:345:PHE:HB3	2.00	0.43
2:I:53:HIS:O	2:I:54:ARG:NE	2.49	0.43
2:K:53:HIS:O	2:K:54:ARG:NH1	2.52	0.43
1:A:405:LEU:HD13	1:A:433:LEU:HD11	2.01	0.43
1:A:462:TRP:CE2	1:A:571:ILE:HD13	2.54	0.43
1:A:324:GLN:HB3	1:A:348:GLU:HG2	2.01	0.43
1:B:511:PRO:HB3	1:B:535:VAL:HG12	2.01	0.43
2:I:57:GLY:HA3	2:I:74:PHE:CZ	2.54	0.43
1:B:462:TRP:CE2	1:B:571:ILE:HD13	2.54	0.42
1:B:393:ILE:N	1:B:412:TYR:O	2.46	0.42
2:I:73:VAL:HG22	2:I:118:TYR:HB2	2.00	0.42
2:I:56:GLY:O	2:I:77:VAL:HA	2.20	0.42
1:C:479:LEU:HA	1:C:485:ILE:HD13	2.00	0.42
1:B:666:VAL:HA	1:B:681:TYR:CD2	2.55	0.42
1:A:584:VAL:O	1:A:605:GLU:HA	2.20	0.42
1:A:438:LEU:O	1:A:440:LYS:NZ	2.53	0.42
1:A:479:LEU:HA	1:A:485:ILE:HD13	2.01	0.42
1:A:571:ILE:HG13	1:A:572:CYS:N	2.35	0.42
2:I:130:TRP:CZ2	2:J:132:ALA:HB2	2.55	0.42
2:J:53:HIS:O	2:J:54:ARG:NE	2.53	0.42
1:C:668:TRP:CH2	1:C:680:VAL:HG21	2.55	0.41
2:K:25:ILE:HB	2:K:115:ILE:HB	2.02	0.41
1:A:308:VAL:HG12	1:A:309:HIS:ND1	2.35	0.41
1:C:439:PRO:HD3	1:C:660:TRP:CD1	2.55	0.41
2:I:30:GLN:HB2	2:I:33:GLU:HG2	2.01	0.41
2:I:48:VAL:O	2:I:92:LYS:HA	2.20	0.41
1:A:640:ASP:OD1	1:A:641:ALA:N	2.53	0.41
2:J:48:VAL:O	2:J:92:LYS:HA	2.20	0.41
1:C:466:LEU:HD12	1:C:475:CYS:HB3	2.03	0.41
1:C:654:ASN:HB3	1:C:657:ARG:HB3	2.03	0.41
1:B:472:GLN:HB3	2:J:88:PRO:HB3	2.02	0.41
1:A:305:GLN:OE1	1:A:305:GLN:N	2.51	0.41
2:J:57:GLY:HA3	2:J:74:PHE:CZ	2.55	0.41
1:B:523:ARG:HG3	1:B:589:TRP:HB3	2.03	0.41
1:C:327:VAL:HG11	1:C:360:PRO:HG2	2.03	0.41
1:C:528:GLU:OE2	3:C:701:GOL:O3	2.38	0.41
1:B:542:PRO:HB2	2:I:113:LEU:HD11	2.02	0.41
2:J:77:VAL:HG21	2:J:120:PRO:HG2	2.02	0.41
1:C:571:ILE:HG13	1:C:572:CYS:N	2.36	0.40
1:A:676:ASP:O	1:A:677:ARG:HG3	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:547:ASN:OD1	1:B:547:ASN:N	2.54	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	394/406 (97%)	376 (95%)	18 (5%)	0	100	100
1	B	377/406 (93%)	358 (95%)	19 (5%)	0	100	100
1	C	389/406 (96%)	369 (95%)	19 (5%)	1 (0%)	41	72
2	I	137/162 (85%)	131 (96%)	6 (4%)	0	100	100
2	J	137/162 (85%)	131 (96%)	6 (4%)	0	100	100
2	K	137/162 (85%)	131 (96%)	6 (4%)	0	100	100
All	All	1571/1704 (92%)	1496 (95%)	74 (5%)	1 (0%)	51	82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	314	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	263/355 (74%)	259 (98%)	4 (2%)	65 82
1	B	214/355 (60%)	213 (100%)	1 (0%)	88 94
1	C	265/355 (75%)	263 (99%)	2 (1%)	81 91
2	I	97/139 (70%)	97 (100%)	0	100 100
2	J	95/139 (68%)	95 (100%)	0	100 100
2	K	103/139 (74%)	103 (100%)	0	100 100
All	All	1037/1482 (70%)	1030 (99%)	7 (1%)	84 92

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	384	ASN
1	C	683	TYR
1	A	301	CYS
1	A	593	PHE
1	A	657	ARG
1	A	683	TYR
1	B	683	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 1 is monoatomic - leaving 7 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	$\#\|Z\| > 2$
3	GOL	K	901	-	5,5,5	0.91	0	5,5,5	0.99	0
3	GOL	I	202	-	5,5,5	0.90	0	5,5,5	1.01	0
3	GOL	A	701	-	5,5,5	0.92	0	5,5,5	0.98	0
3	GOL	A	703	-	5,5,5	0.89	0	5,5,5	1.04	0
3	GOL	I	201	-	5,5,5	0.90	0	5,5,5	0.98	0
3	GOL	C	701	-	5,5,5	0.91	0	5,5,5	0.98	0
3	GOL	A	702	-	5,5,5	0.93	0	5,5,5	0.98	0

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
3	GOL	K	901	-	-	0/4/4/4	-
3	GOL	I	202	-	-	0/4/4/4	-
3	GOL	A	701	-	-	0/4/4/4	-
3	GOL	A	703	-	-	0/4/4/4	-
3	GOL	I	201	-	-	0/4/4/4	-
3	GOL	C	701	-	-	0/4/4/4	-
3	GOL	A	702	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	K	901	GOL	1	0
3	C	701	GOL	1	0

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	398/406 (98%)	0.30	20 (5%)	28	29	71, 122, 188, 258	0
1	B	385/406 (94%)	0.37	25 (6%)	18	20	103, 157, 210, 262	0
1	C	393/406 (96%)	0.28	23 (5%)	22	23	68, 123, 194, 228	0
2	I	139/162 (85%)	0.28	4 (2%)	51	50	83, 120, 163, 204	0
2	J	139/162 (85%)	0.37	12 (8%)	10	12	97, 136, 181, 199	0
2	K	139/162 (85%)	0.18	4 (2%)	51	50	82, 115, 163, 193	0
All	All	1593/1704 (93%)	0.31	88 (5%)	25	25	68, 133, 195, 262	0

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	360	PRO	6.5
2	K	103	ASP	6.2
1	C	395	TYR	5.5
1	B	395	TYR	4.8
1	B	360	PRO	4.5
2	I	103	ASP	4.4
1	A	314	PRO	4.2
1	C	396	SER	4.1
1	B	307	PRO	4.0
1	A	361	THR	3.9
2	J	43	GLY	3.9
1	A	315	SER	3.7
1	A	305	GLN	3.7
1	A	360	PRO	3.6
2	J	103	ASP	3.5
1	C	327	VAL	3.5
1	B	380	PHE	3.5
1	A	306	PRO	3.5
1	C	440	LYS	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	327	VAL	3.2
1	C	378	ILE	3.1
1	A	304	LEU	3.1
1	B	391	SER	3.1
1	B	678	TYR	3.0
1	B	370	PRO	3.0
1	B	306	PRO	3.0
1	B	379	THR	2.9
1	C	306	PRO	2.8
1	C	370	PRO	2.8
1	A	370	PRO	2.8
1	C	302	PRO	2.8
1	C	424	VAL	2.8
1	A	328	SER	2.8
1	B	327	VAL	2.7
1	A	420	TRP	2.7
1	B	326	LEU	2.7
1	A	355	TRP	2.7
1	C	325	VAL	2.7
1	C	355	TRP	2.7
1	B	359	ILE	2.7
2	J	142	ARG	2.7
1	B	420	TRP	2.6
1	B	332	GLY	2.6
1	C	345	PHE	2.6
1	C	324	GLN	2.6
1	A	308	VAL	2.6
1	B	361	THR	2.5
2	J	125	VAL	2.5
2	J	53	HIS	2.5
1	A	325	VAL	2.5
1	A	345	PHE	2.5
2	J	44	GLN	2.5
1	A	307	PRO	2.5
2	I	7	LEU	2.4
1	C	314	PRO	2.4
2	J	25	ILE	2.4
2	J	96	PHE	2.4
1	B	378	ILE	2.3
2	J	13	TYR	2.3
1	B	345	PHE	2.3
1	C	359	ILE	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	514	PHE	2.3
1	A	395	TYR	2.3
1	B	389	TYR	2.3
1	B	670	ASP	2.3
1	A	581	GLY	2.3
2	I	42	ILE	2.2
1	C	319	TYR	2.2
2	K	43	GLY	2.2
2	I	74	PHE	2.2
1	C	670	ASP	2.2
1	B	324	GLN	2.2
1	B	581	GLY	2.2
1	A	312	ILE	2.2
1	C	303	GLU	2.1
2	K	42	ILE	2.1
1	C	420	TRP	2.1
1	B	647	GLY	2.1
1	B	604	ILE	2.1
2	J	46	LEU	2.1
1	A	319	TYR	2.1
2	K	90	GLY	2.1
1	B	392	GLU	2.1
2	J	88	PRO	2.1
1	C	425	LEU	2.1
1	C	347	ILE	2.0
2	J	7	LEU	2.0
1	C	377	LEU	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
3	GOL	A	703	6/6	0.39	0.46	142,147,155,157	0
4	NA	C	702	1/1	0.64	0.48	159,159,159,159	0
3	GOL	C	701	6/6	0.66	0.35	150,157,166,173	0
3	GOL	A	701	6/6	0.66	0.39	148,165,179,184	0
3	GOL	I	202	6/6	0.76	0.43	128,143,148,150	0
3	GOL	K	901	6/6	0.83	1.09	124,134,144,150	0
3	GOL	A	702	6/6	0.88	0.23	131,141,144,148	0
3	GOL	I	201	6/6	0.90	0.42	161,166,173,176	0

6.5 Other polymers (i)

There are no such residues in this entry.