wwPDB X-ray Structure Validation Summary Report (i)

Feb 3, 2024 – 06:46 PM EST

PDB ID	:	1MOW
Title	:	E-DreI
Authors	:	Chevalier, B.S.; Kortemme, T.; Chadsey, M.S.; Baker, D.; Monnat Jr., R.J.; Stoddard, B.L.
Deposited on	:	2002-09-10
Resolution	:	2.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (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:

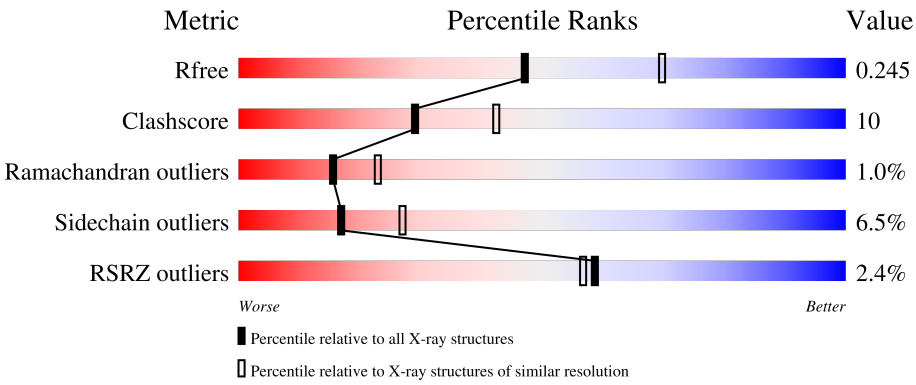
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	B	23	65% 35%
1	E	23	61% 39%
1	H	23	65% 35%
1	K	23	35% 65%
2	C	23	48% 52%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	23	48% 52%
2	I	23	4% 52% 48%
2	L	23	57% 43%
3	A	260	73% 19% 5%
3	D	260	72% 21% 5%
3	G	260	6% 82% 8% 10%
3	J	260	4% 86% 10%

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
4	GOL	A	765	-	X	-	-
4	GOL	A	767	-	X	-	-
4	GOL	A	768	-	X	-	-
4	GOL	A	769	-	X	-	-
4	GOL	A	770	-	X	-	-
4	GOL	A	772	-	X	-	-
4	GOL	A	773	-	X	-	-
4	GOL	B	762	-	X	-	-
4	GOL	B	775	-	X	-	-
4	GOL	D	761	-	X	-	-
4	GOL	D	763	-	X	-	-
4	GOL	D	764	-	X	-	-
4	GOL	D	766	-	X	-	-
4	GOL	D	776	-	X	-	-
4	GOL	D	777	-	X	-	-
4	GOL	D	778	-	X	-	-
4	GOL	E	771	-	X	-	-
4	GOL	F	774	-	X	-	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 10717 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 DNA chain called 5'-D(*CP*CP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*AP*AP*GP*TP*TP*CP*CP*GP*GP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	23	464	222	84	136	22	0	0	0
1	E	23	464	222	84	136	22	0	0	0
1	H	23	463	221	84	136	22	0	0	0
1	K	23	464	222	84	136	22	0	0	0

- Molecule 2 is a DNA chain called 5'-D(*CP*GP*CP*CP*GP*GP*AP*AP*CP*TP*TP*GP*AP*GP*AP*CP*AP*GP*TP*TP*TP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	23	473	225	90	136	22	0	0	0
2	F	23	473	225	90	136	22	0	0	0
2	I	23	473	225	90	136	22	0	0	0
2	L	23	473	225	90	136	22	0	0	0

- Molecule 3 is a protein called chimera of homing endonuclease I-DmoI and DNA endonuclease I-CreI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	248	2022	1304	347	368	3	0	0	0
3	D	248	2019	1303	347	366	3	0	0	0
3	G	233	Total	C	N	O		0	0	0
			1153	687	233	233				

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	J	234	1158	690	234	234	0	0	0

There are 56 discrepancies between the modelled and reference sequences:

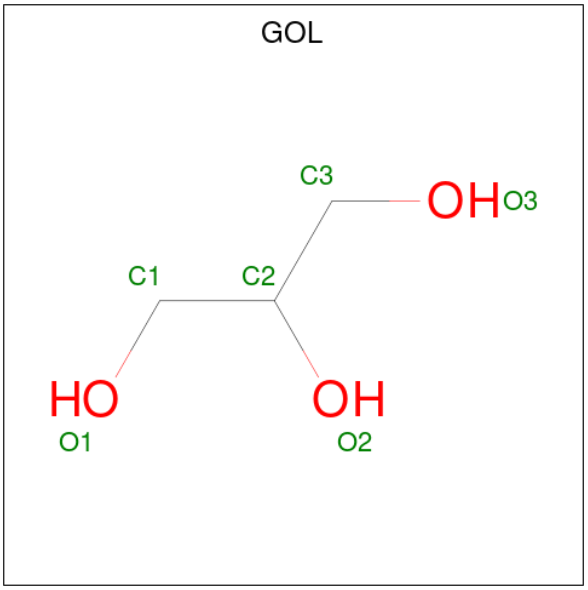
Chain	Residue	Modelled	Actual	Comment	Reference
A	19	TRP	ILE	SEE REMARK 999	UNP p21505
A	51	PHE	HIS	SEE REMARK 999	UNP p21505
A	14	LEU	ARG	SEE REMARK 999	UNP p21505
A	103	GLY	-	linker	UNP p21505
A	104	ASN	-	linker	UNP p21505
A	105	ARG	-	linker	UNP p21505
A	108	ALA	LEU	SEE REMARK 999	UNP p05725
A	113	ILE	PHE	SEE REMARK 999	UNP p05725
A	139	THR	ALA	SEE REMARK 999	UNP p05725
A	147	GLU	GLN	SEE REMARK 999	UNP p05725
A	193	ASN	LYS	SEE REMARK 999	UNP p05725
A	194	PHE	LEU	SEE REMARK 999	UNP p05725
A	207	GLU	TRP	SEE REMARK 999	UNP p05725
A	208	GLN	ARG	SEE REMARK 999	UNP p05725
D	519	TRP	ILE	SEE REMARK 999	UNP p21505
D	551	PHE	HIS	SEE REMARK 999	UNP p21505
D	514	LEU	ARG	SEE REMARK 999	UNP p21505
D	603	GLY	-	linker	UNP p21505
D	604	ASN	-	linker	UNP p21505
D	605	ARG	-	linker	UNP p21505
D	608	ALA	LEU	SEE REMARK 999	UNP p05725
D	613	ILE	PHE	SEE REMARK 999	UNP p05725
D	639	THR	ALA	SEE REMARK 999	UNP p05725
D	647	GLU	GLN	SEE REMARK 999	UNP p05725
D	693	ASN	LYS	SEE REMARK 999	UNP p05725
D	694	PHE	LEU	SEE REMARK 999	UNP p05725
D	707	GLU	TRP	SEE REMARK 999	UNP p05725
D	708	GLN	ARG	SEE REMARK 999	UNP p05725
G	1019	TRP	ILE	SEE REMARK 999	UNP p21505
G	1051	PHE	HIS	SEE REMARK 999	UNP p21505
G	1014	LEU	ARG	SEE REMARK 999	UNP p21505
G	1103	GLY	-	linker	UNP p21505
G	1104	ASN	-	linker	UNP p21505
G	1105	ARG	-	linker	UNP p21505
G	1108	ALA	LEU	SEE REMARK 999	UNP p05725
G	1113	ILE	PHE	SEE REMARK 999	UNP p05725

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	1139	THR	ALA	SEE REMARK 999	UNP p05725
G	1147	GLU	GLN	SEE REMARK 999	UNP p05725
G	1193	ASN	LYS	SEE REMARK 999	UNP p05725
G	1194	PHE	LEU	SEE REMARK 999	UNP p05725
G	1207	GLU	TRP	SEE REMARK 999	UNP p05725
G	1208	GLN	ARG	SEE REMARK 999	UNP p05725
J	1519	TRP	ILE	SEE REMARK 999	UNP p21505
J	1551	PHE	HIS	SEE REMARK 999	UNP p21505
J	1514	LEU	ARG	SEE REMARK 999	UNP p21505
J	1603	GLY	-	linker	UNP p21505
J	1604	ASN	-	linker	UNP p21505
J	1605	ARG	-	linker	UNP p21505
J	1608	ALA	LEU	SEE REMARK 999	UNP p05725
J	1613	ILE	PHE	SEE REMARK 999	UNP p05725
J	1639	THR	ALA	SEE REMARK 999	UNP p05725
J	1647	GLU	GLN	SEE REMARK 999	UNP p05725
J	1693	ASN	LYS	SEE REMARK 999	UNP p05725
J	1694	PHE	LEU	SEE REMARK 999	UNP p05725
J	1707	GLU	TRP	SEE REMARK 999	UNP p05725
J	1708	GLN	ARG	SEE REMARK 999	UNP p05725

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	B	1	6	3	3	0	0
4	E	1	6	3	3	0	0
4	F	1	6	3	3	0	0
4	A	1	6	3	3	0	0
4	A	1	6	3	3	0	0
4	A	1	6	3	3	0	0
4	A	1	6	3	3	0	0
4	A	1	6	3	3	0	0
4	A	1	6	3	3	0	0
4	A	1	6	3	3	0	0
4	D	1	6	3	3	0	0
4	D	1	6	3	3	0	0
4	D	1	6	3	3	0	0
4	D	1	6	3	3	0	0
4	D	1	6	3	3	0	0
4	D	1	6	3	3	0	0
4	D	1	6	3	3	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

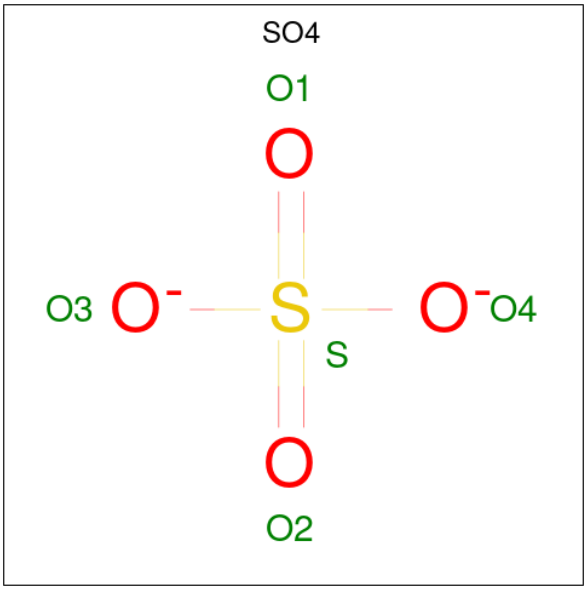
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
5	E	1	1	1	0	0
5	A	3	3	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
5	D	2	2	2	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	A	1	5	4	1	0	0
			Total	O	S		
6	D	1	5	4	1	0	0
			Total	O	S		
6	D	1	5	4	1	0	0
			Total	O	S		
6	D	1	5	4	1	0	0
			Total	O	S		
6	D	1	5	4	1	0	0

- Molecule 7 is water.

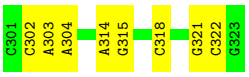
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	41	Total O 41 41	0	0
7	C	39	Total O 39 39	0	0
7	E	61	Total O 61 61	0	0
7	F	37	Total O 37 37	0	0
7	A	137	Total O 137 137	0	0
7	D	139	Total O 139 139	0	0

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.

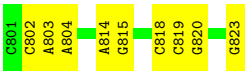
• Molecule 1: 5'-D(*CP*CP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*AP*AP*GP*TP*TP*CP*CP*GP*GP*CP*G)-3'

Chain B: 65% 35%



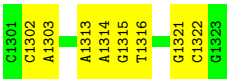
• Molecule 1: 5'-D(*CP*CP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*AP*AP*GP*TP*TP*CP*CP*GP*GP*CP*G)-3'

Chain E: 61% 39%



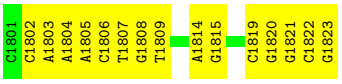
• Molecule 1: 5'-D(*CP*CP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*AP*AP*GP*TP*TP*CP*CP*GP*GP*CP*G)-3'

Chain H: 65% 35%



• Molecule 1: 5'-D(*CP*CP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*AP*AP*GP*TP*TP*CP*CP*GP*GP*CP*G)-3'

Chain K: 35% 65%

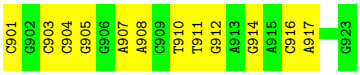


• Molecule 2: 5'-D(*CP*GP*CP*CP*GP*GP*AP*AP*CP*TP*TP*GP*AP*GP*AP*CP*AP*GP*TP*TP*TP*GP*G)-3'

Chain C: 48% 52%

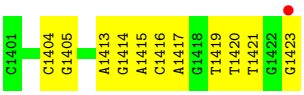
• Molecule 2: 5'-D(*CP*GP*CP*CP*GP*GP*AP*AP*CP*TP*TP*GP*AP*GP*AP*CP*AP*GP*TP*TP*TP*GP*G)-3'

Chain F: 48% 52%



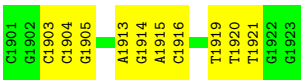
• Molecule 2: 5'-D(*CP*GP*CP*CP*GP*GP*AP*AP*CP*TP*TP*GP*AP*GP*AP*CP*AP*GP*TP*TP*TP*GP*G)-3'

Chain I: 4% 52% 48%



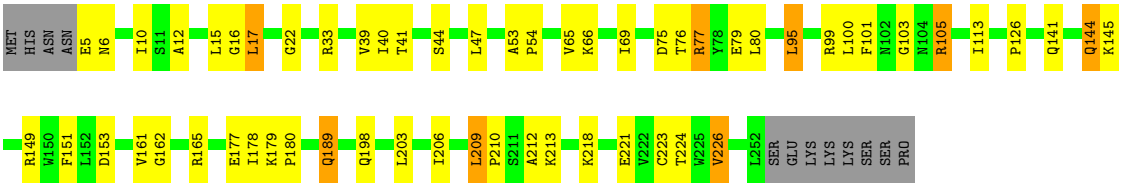
• Molecule 2: 5'-D(*CP*GP*CP*CP*GP*GP*AP*AP*CP*TP*TP*GP*AP*GP*AP*CP*AP*GP*TP*TP*TP*GP*G)-3'

Chain L: 57% 43%



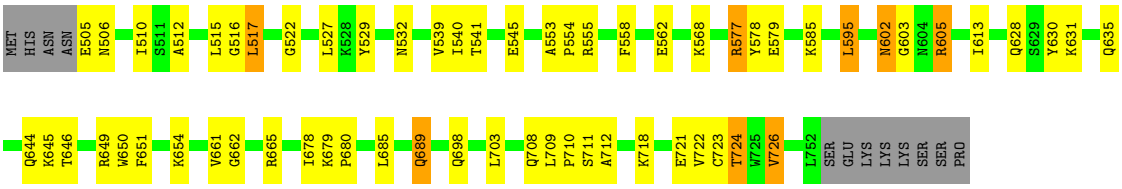
• Molecule 3: chimera of homing endonuclease I-DmoI and DNA endonuclease I-CreI

Chain A: 73% 19% • 5%



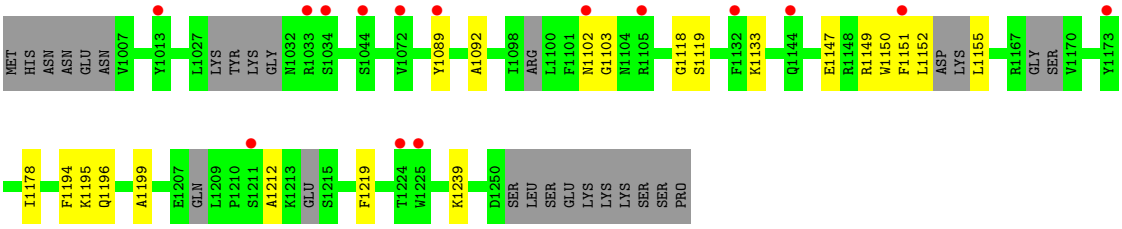
• Molecule 3: chimera of homing endonuclease I-DmoI and DNA endonuclease I-CreI

Chain D: 72% 21% • 5%



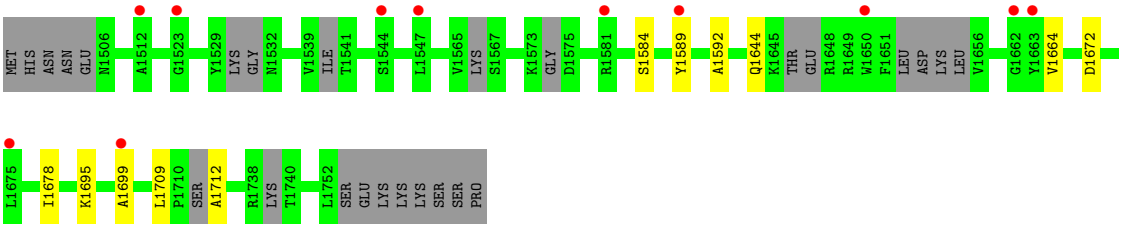
• Molecule 3: chimera of homing endonuclease I-DmoI and DNA endonuclease I-CreI

Chain G:



- Molecule 3: chimera of homing endonuclease I-DmoI and DNA endonuclease I-CreI

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	131.76Å 131.76Å 120.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.86 – 2.40 19.86 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.86-2.40) 99.9 (19.86-2.40)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.76 (at 2.41Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.231 , 0.257 0.215 , 0.245	Depositor DCC
R_{free} test set	4592 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	41.0	Xtriage
Anisotropy	0.005	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 54.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.015 for -h,-k,l 0.418 for h,-h-k,-l 0.017 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10717	wwPDB-VP
Average B, all atoms (Å^2)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.67 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.7017e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: MG, GOL, SO4

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	B	0.39	0/519	0.76	0/798
1	E	0.41	0/519	0.77	0/798
1	H	0.20	0/516	0.64	0/790
1	K	0.18	0/519	0.62	0/798
2	C	0.44	0/531	0.80	0/819
2	F	0.43	0/531	0.81	0/819
2	I	0.20	0/531	0.62	0/819
2	L	0.19	0/531	0.61	0/819
3	A	0.38	0/2059	0.63	0/2774
3	D	0.38	0/2056	0.64	0/2770
3	G	0.23	0/1146	0.50	0/1586
3	J	0.21	0/1149	0.49	0/1587
All	All	0.33	0/10607	0.64	0/15177

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	B	0	1
1	E	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	318	DC	Sidechain
1	E	818	DC	Sidechain

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	B	464	0	260	8	0
1	E	464	0	260	7	0
1	H	463	0	258	7	0
1	K	464	0	260	14	0
2	C	473	0	260	14	0
2	F	473	0	260	13	0
2	I	473	0	260	13	0
2	L	473	0	260	9	0
3	A	2022	0	2083	47	0
3	D	2019	0	2075	41	0
3	G	1153	0	498	10	0
3	J	1158	0	498	6	0
4	A	42	0	28	3	0
4	B	12	0	8	1	0
4	D	42	0	28	1	0
4	E	6	0	4	0	0
4	F	6	0	4	0	0
5	A	3	0	0	0	0
5	D	2	0	0	0	0
5	E	1	0	0	0	0
6	A	30	0	0	2	0
6	D	20	0	0	0	0
7	A	137	0	0	2	0
7	B	41	0	0	0	0
7	C	39	0	0	2	0
7	D	139	0	0	3	0
7	E	61	0	0	1	0
7	F	37	0	0	5	0
All	All	10717	0	7304	169	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 10.

The worst 5 of 169 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:H:1313:DA:C4	1:H:1313:DA:C6	2.38	1.09
2:I:1404:DC:H2"	2:I:1405:DG:H5'	1.43	1.00
2:L:1904:DC:H2"	2:L:1905:DG:H5'	1.45	0.98
2:I:1423:DG:H5'	1:K:1823:DG:H5'	1.50	0.94
2:F:912:DG:H5'	3:D:645:LYS:HD3	1.55	0.87

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	
3	A	246/260 (95%)	239 (97%)	6 (2%)	1 (0%)	34	48
3	D	246/260 (95%)	238 (97%)	7 (3%)	1 (0%)	34	48
3	G	219/260 (84%)	191 (87%)	22 (10%)	6 (3%)	5	5
3	J	216/260 (83%)	198 (92%)	17 (8%)	1 (0%)	29	41
All	All	927/1040 (89%)	866 (93%)	52 (6%)	9 (1%)	15	23

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	1102	ASN
3	D	603	GLY
3	G	1178	ILE
3	J	1678	ILE
3	G	1150	TRP

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	
3	A	223/235 (95%)	210 (94%)	13 (6%)	20	32
3	D	221/235 (94%)	205 (93%)	16 (7%)	14	23
All	All	444/470 (94%)	415 (94%)	29 (6%)	17	27

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	568	LYS
3	D	724	THR
3	D	602	ASN
3	D	703	LEU
3	D	595	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	104	ASN
3	D	602	ASN
3	D	604	ASN
3	D	628	GLN

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 34 ligands modelled in this entry, 6 are monoatomic - leaving 28 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	GOL	A	773	-	5,5,5	4.58	5 (100%)	5,5,5	5.78	3 (60%)
4	GOL	D	776	-	5,5,5	4.52	5 (100%)	5,5,5	5.79	3 (60%)
6	SO4	A	280	-	4,4,4	0.25	0	6,6,6	0.04	0
6	SO4	D	274	-	4,4,4	0.29	0	6,6,6	0.07	0
4	GOL	E	771	-	5,5,5	4.52	5 (100%)	5,5,5	5.76	3 (60%)
6	SO4	A	271	-	4,4,4	0.26	0	6,6,6	0.05	0
6	SO4	D	273	-	4,4,4	0.25	0	6,6,6	0.07	0
4	GOL	B	775	-	5,5,5	4.57	5 (100%)	5,5,5	5.76	3 (60%)
4	GOL	D	764	-	5,5,5	4.57	5 (100%)	5,5,5	5.78	3 (60%)
6	SO4	A	277	-	4,4,4	0.26	0	6,6,6	0.10	0
6	SO4	D	278	-	4,4,4	0.27	0	6,6,6	0.04	0
6	SO4	A	272	-	4,4,4	0.25	0	6,6,6	0.07	0
4	GOL	D	761	-	5,5,5	4.54	5 (100%)	5,5,5	5.77	3 (60%)
4	GOL	A	770	-	5,5,5	4.55	5 (100%)	5,5,5	5.90	3 (60%)
6	SO4	A	276	-	4,4,4	0.25	0	6,6,6	0.06	0
4	GOL	A	765	-	5,5,5	4.56	5 (100%)	5,5,5	5.77	3 (60%)
4	GOL	D	763	-	5,5,5	4.52	5 (100%)	5,5,5	5.75	3 (60%)
4	GOL	D	766	-	5,5,5	4.62	5 (100%)	5,5,5	5.76	3 (60%)
4	GOL	F	774	-	5,5,5	4.52	5 (100%)	5,5,5	5.76	3 (60%)
4	GOL	A	767	-	5,5,5	4.66	5 (100%)	5,5,5	5.77	3 (60%)
4	GOL	D	778	-	5,5,5	4.58	5 (100%)	5,5,5	5.74	3 (60%)
6	SO4	A	275	-	4,4,4	0.22	0	6,6,6	0.08	0
6	SO4	D	279	-	4,4,4	0.24	0	6,6,6	0.08	0
4	GOL	A	772	-	5,5,5	4.52	5 (100%)	5,5,5	5.79	3 (60%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	GOL	D	777	-	5,5,5	4.61	5 (100%)	5,5,5	5.74	3 (60%)
4	GOL	A	768	-	5,5,5	4.56	5 (100%)	5,5,5	5.75	3 (60%)
4	GOL	A	769	-	5,5,5	4.51	5 (100%)	5,5,5	5.78	3 (60%)
4	GOL	B	762	-	5,5,5	4.59	5 (100%)	5,5,5	5.77	3 (60%)

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	GOL	A	770	-	-	2/4/4/4	-
4	GOL	D	778	-	-	3/4/4/4	-
4	GOL	A	772	-	-	2/4/4/4	-
4	GOL	A	773	-	-	2/4/4/4	-
4	GOL	A	765	-	-	3/4/4/4	-
4	GOL	B	775	-	-	2/4/4/4	-
4	GOL	D	763	-	-	2/4/4/4	-
4	GOL	D	776	-	-	3/4/4/4	-
4	GOL	D	766	-	-	3/4/4/4	-
4	GOL	F	774	-	-	2/4/4/4	-
4	GOL	D	777	-	-	3/4/4/4	-
4	GOL	A	768	-	-	3/4/4/4	-
4	GOL	A	767	-	-	3/4/4/4	-
4	GOL	B	762	-	-	2/4/4/4	-
4	GOL	A	769	-	-	3/4/4/4	-
4	GOL	E	771	-	-	2/4/4/4	-
4	GOL	D	764	-	-	2/4/4/4	-
4	GOL	D	761	-	-	3/4/4/4	-

The worst 5 of 90 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	777	GOL	C3-C2	-7.77	1.19	1.51
4	A	770	GOL	C3-C2	-7.66	1.20	1.51
4	A	767	GOL	C3-C2	-7.66	1.20	1.51
4	A	773	GOL	C3-C2	-7.60	1.20	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	766	GOL	C3-C2	-7.57	1.20	1.51

The worst 5 of 54 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	770	GOL	O3-C3-C2	10.68	161.43	110.20
4	D	776	GOL	O3-C3-C2	10.57	160.91	110.20
4	A	773	GOL	O3-C3-C2	10.55	160.78	110.20
4	B	762	GOL	O3-C3-C2	10.52	160.65	110.20
4	D	766	GOL	O3-C3-C2	10.51	160.61	110.20

There are no chirality outliers.

5 of 45 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	762	GOL	C1-C2-C3-O3
4	B	775	GOL	C1-C2-C3-O3
4	E	771	GOL	O1-C1-C2-C3
4	E	771	GOL	C1-C2-C3-O3
4	F	774	GOL	C1-C2-C3-O3

There are no ring outliers.

7 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	775	GOL	1	0
6	A	272	SO4	1	0
4	A	770	GOL	1	0
4	A	767	GOL	1	0
6	A	275	SO4	1	0
4	A	772	GOL	1	0
4	D	777	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 (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	B	23/23 (100%)	-0.59	0	100	100	26, 39, 48, 52	0
1	E	23/23 (100%)	-0.71	0	100	100	25, 38, 49, 54	0
1	H	23/23 (100%)	0.50	0	100	100	120, 133, 144, 146	0
1	K	23/23 (100%)	0.52	0	100	100	130, 138, 151, 154	0
2	C	23/23 (100%)	-0.56	0	100	100	28, 39, 52, 55	0
2	F	23/23 (100%)	-0.64	0	100	100	27, 37, 53, 57	0
2	I	23/23 (100%)	0.44	1 (4%)	35	33	120, 132, 143, 145	0
2	L	23/23 (100%)	0.52	0	100	100	127, 136, 146, 149	0
3	A	248/260 (95%)	-0.52	0	100	100	22, 33, 55, 64	0
3	D	248/260 (95%)	-0.51	0	100	100	22, 33, 54, 64	0
3	G	233/260 (89%)	0.53	15 (6%)	19	18	103, 111, 117, 126	0
3	J	234/260 (90%)	0.41	11 (4%)	31	30	100, 108, 115, 118	0
All	All	1147/1224 (93%)	-0.04	27 (2%)	59	57	22, 56, 133, 154	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	J	1512	ALA	3.9
3	J	1675	LEU	3.3
3	G	1144	GLN	3.1
3	J	1663	TYR	3.1
3	J	1662	GLY	3.0

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, 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
6	SO4	A	275	5/5	0.63	0.28	150,150,151,152	0
4	GOL	B	775	6/6	0.74	0.26	91,97,99,103	0
4	GOL	D	763	6/6	0.76	0.22	84,86,90,92	0
4	GOL	A	770	6/6	0.76	0.28	54,85,86,87	0
4	GOL	A	769	6/6	0.77	0.26	64,75,79,87	0
4	GOL	D	761	6/6	0.77	0.23	79,86,89,90	0
6	SO4	A	277	5/5	0.78	0.26	126,127,130,131	0
4	GOL	D	764	6/6	0.80	0.15	93,98,99,99	0
4	GOL	D	777	6/6	0.81	0.52	112,113,118,122	0
4	GOL	A	768	6/6	0.82	0.23	67,74,75,78	0
4	GOL	A	765	6/6	0.83	0.17	67,73,77,85	0
4	GOL	E	771	6/6	0.84	0.14	62,68,72,74	0
4	GOL	D	776	6/6	0.85	0.28	79,81,83,90	0
4	GOL	F	774	6/6	0.86	0.13	62,67,71,73	0
5	MG	A	373	1/1	0.86	0.11	32,32,32,32	0
4	GOL	A	773	6/6	0.87	0.46	106,108,113,115	0
6	SO4	A	271	5/5	0.87	0.17	115,116,117,119	0
6	SO4	D	279	5/5	0.87	0.27	130,132,133,135	0
5	MG	E	375	1/1	0.88	0.14	33,33,33,33	0
4	GOL	A	772	6/6	0.88	0.17	76,77,79,82	0
6	SO4	A	280	5/5	0.89	0.21	139,140,140,140	0
6	SO4	A	272	5/5	0.89	0.18	113,115,116,116	0
6	SO4	D	273	5/5	0.90	0.24	137,137,139,139	0
6	SO4	A	276	5/5	0.90	0.21	139,139,140,142	0
6	SO4	D	274	5/5	0.91	0.16	90,92,97,103	0
6	SO4	D	278	5/5	0.91	0.28	138,139,139,140	0
4	GOL	B	762	6/6	0.91	0.17	64,68,69,76	0
4	GOL	D	766	6/6	0.92	0.26	72,76,85,86	0
4	GOL	D	778	6/6	0.93	0.19	75,76,78,79	0
5	MG	A	372	1/1	0.95	0.09	37,37,37,37	0
4	GOL	A	767	6/6	0.96	0.24	58,60,71,73	0
5	MG	D	374	1/1	0.96	0.14	20,20,20,20	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	MG	A	371	1/1	0.97	0.09	13,13,13,13	0
5	MG	D	376	1/1	0.99	0.16	16,16,16,16	0

6.5 Other polymers (i)

There are no such residues in this entry.