Full wwPDB X-ray Structure Validation Report (i)

Dec 14, 2021 – 10:10 AM JST

PDB ID	:	7E0M
Title	:	Crystal structure of phospholipase D
Authors	:	Wang, F.H.
Deposited on	:	2021-01-28
Resolution	:	1.79 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:

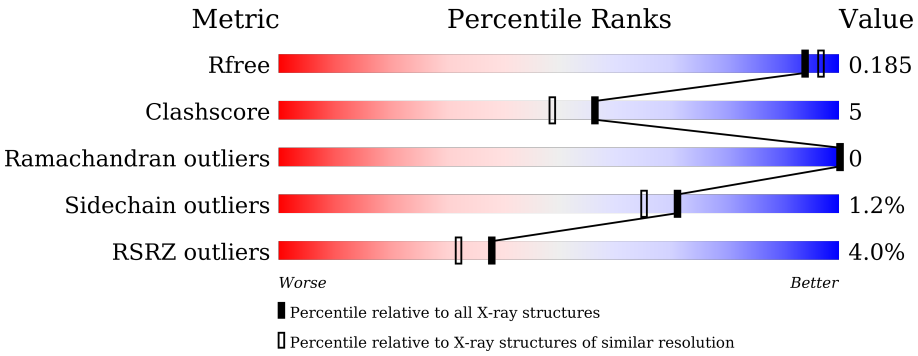
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.24
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

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 1.79 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	391	3% 87% 12%
1	B	391	5% 84% 13%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6902 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 Phospholipase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	0	0	0
			3068	1959	541	562	6			
1	B	384	Total	C	N	O	S	0	1	0
			3036	1940	530	559	7			

There are 26 discrepancies between the modelled and reference sequences:

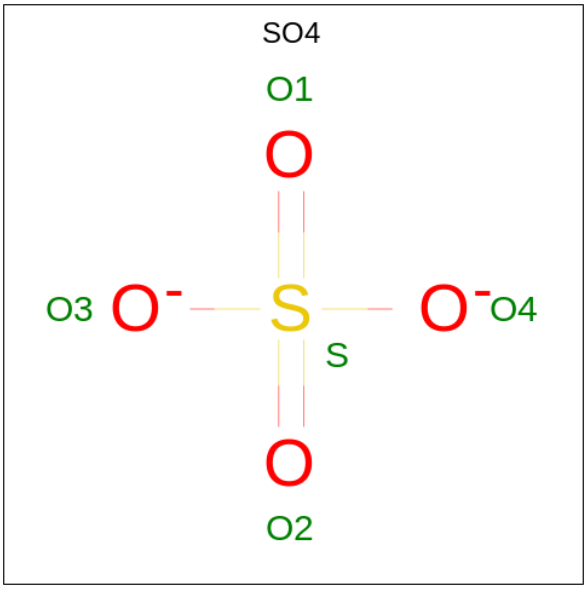
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	expression tag	UNP S4YJ13
A	-5	HIS	-	expression tag	UNP S4YJ13
A	-4	HIS	-	expression tag	UNP S4YJ13
A	-3	HIS	-	expression tag	UNP S4YJ13
A	-2	HIS	-	expression tag	UNP S4YJ13
A	-1	HIS	-	expression tag	UNP S4YJ13
A	0	HIS	-	expression tag	UNP S4YJ13
A	3	ASP	ALA	conflict	UNP S4YJ13
A	85	LEU	ILE	conflict	UNP S4YJ13
A	108	ARG	GLN	conflict	UNP S4YJ13
A	182	THR	ALA	conflict	UNP S4YJ13
A	301	HIS	GLN	conflict	UNP S4YJ13
A	369	ILE	LEU	conflict	UNP S4YJ13
B	-6	MET	-	expression tag	UNP S4YJ13
B	-5	HIS	-	expression tag	UNP S4YJ13
B	-4	HIS	-	expression tag	UNP S4YJ13
B	-3	HIS	-	expression tag	UNP S4YJ13
B	-2	HIS	-	expression tag	UNP S4YJ13
B	-1	HIS	-	expression tag	UNP S4YJ13
B	0	HIS	-	expression tag	UNP S4YJ13
B	3	ASP	ALA	conflict	UNP S4YJ13
B	85	LEU	ILE	conflict	UNP S4YJ13
B	108	ARG	GLN	conflict	UNP S4YJ13
B	182	THR	ALA	conflict	UNP S4YJ13
B	301	HIS	GLN	conflict	UNP S4YJ13

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	369	ILE	LEU	conflict	UNP S4YJ13

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is water.

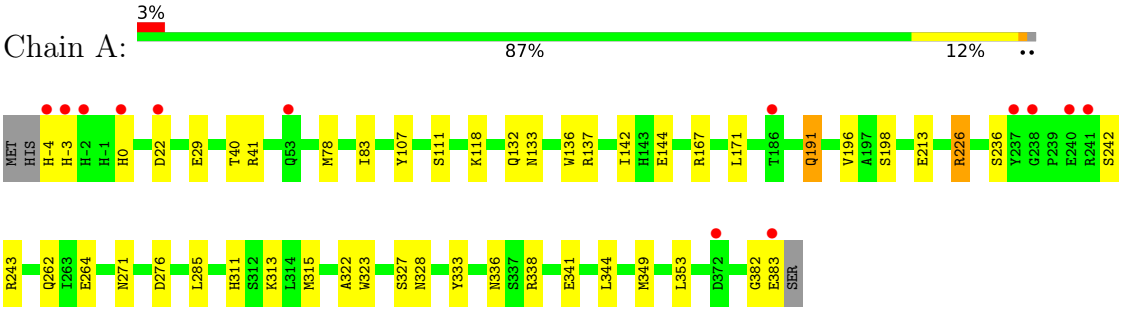
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	425	Total O 425 425	0	0
3	B	368	Total O 368 368	0	0

3 Residue-property plots

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.

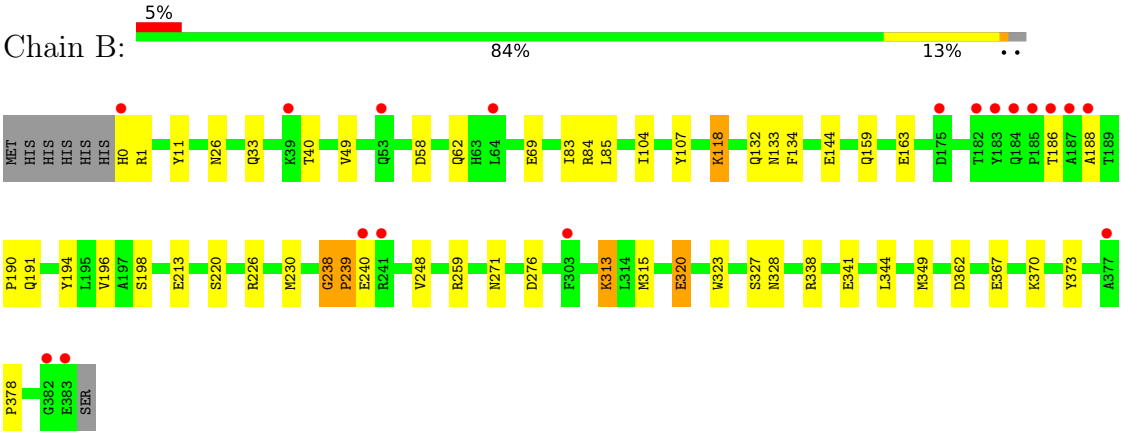
• Molecule 1: Phospholipase

Chain A:



• Molecule 1: Phospholipase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	78.78Å 97.57Å 117.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.88 – 1.79 23.88 – 1.79	Depositor EDS
% Data completeness (in resolution range)	97.0 (23.88-1.79) 97.0 (23.88-1.79)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.91 (at 1.79Å)	Xtriage
Refinement program	PHENIX 1.18.1_3865	Depositor
R, R_{free}	0.170 , 0.185 0.170 , 0.185	Depositor DCC
R_{free} test set	4150 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	25.4	Xtriage
Anisotropy	0.381	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 38.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	6902	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 9.37% 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: 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	A	0.51	0/3146	0.81	9/4289 (0.2%)
1	B	0.60	2/3110 (0.1%)	0.91	14/4239 (0.3%)
All	All	0.55	2/6256 (0.0%)	0.86	23/8528 (0.3%)

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	A	0	1
1	B	0	4
All	All	0	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	239	PRO	N-CA	12.94	1.69	1.47
1	B	238	GLY	C-N	5.47	1.44	1.34

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	107	TYR	CB-CG-CD1	-15.96	111.43	121.00
1	B	107	TYR	CB-CG-CD2	13.84	129.31	121.00
1	A	107	TYR	CB-CG-CD1	-12.98	113.21	121.00
1	B	320	GLU	CG-CD-OE2	-12.01	94.29	118.30
1	A	107	TYR	CB-CG-CD2	11.12	127.67	121.00
1	B	320	GLU	CG-CD-OE1	9.18	136.65	118.30
1	B	259	ARG	NE-CZ-NH2	-8.69	115.96	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	276	ASP	CB-CG-OD2	-8.68	110.49	118.30
1	B	320	GLU	OE1-CD-OE2	-8.57	113.02	123.30
1	B	118	LYS	CD-CE-NZ	-7.28	94.95	111.70
1	A	276	ASP	CB-CG-OD1	6.80	124.42	118.30
1	B	239	PRO	CA-N-CD	-6.68	102.15	111.50
1	B	313	LYS	CB-CG-CD	6.58	128.71	111.60
1	A	226	ARG	NE-CZ-NH2	-6.34	117.13	120.30
1	B	259	ARG	NE-CZ-NH1	6.34	123.47	120.30
1	B	313	LYS	CD-CE-NZ	-6.26	97.30	111.70
1	A	383	GLU	CA-CB-CG	-5.89	100.44	113.40
1	A	226	ARG	NE-CZ-NH1	5.72	123.16	120.30
1	B	1	ARG	NE-CZ-NH1	-5.60	117.50	120.30
1	B	1	ARG	CG-CD-NE	5.56	123.47	111.80
1	B	259	ARG	CG-CD-NE	-5.36	100.54	111.80
1	A	333	TYR	CB-CG-CD2	-5.32	117.81	121.00
1	A	285	LEU	CB-CG-CD1	-5.12	102.30	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	382	GLY	Peptide
1	B	191	GLN	Sidechain
1	B	230[B]	MET	Mainchain
1	B	238	GLY	Mainchain
1	B	320	GLU	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	A	3068	0	3054	30	0
1	B	3036	0	3034	33	0
2	A	5	0	0	0	0
3	A	425	0	0	8	1
3	B	368	0	0	10	1
All	All	6902	0	6088	61	2

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 (61) 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:239:PRO:N	1:B:239:PRO:CA	1.69	1.41
1:B:84:ARG:NH2	3:B:406:HOH:O	1.98	0.97
1:B:378:PRO:O	3:B:401:HOH:O	1.81	0.97
1:B:220:SER:OG	3:B:402:HOH:O	1.85	0.94
1:B:69:GLU:OE1	3:B:404:HOH:O	1.89	0.90
1:B:362:ASP:OD2	3:B:403:HOH:O	1.88	0.88
1:A:118:LYS:HD2	1:A:133:ASN:OD1	1.81	0.80
1:A:242:SER:OG	3:A:501:HOH:O	1.99	0.80
1:B:118:LYS:HD2	1:B:133:ASN:OD1	1.82	0.78
1:B:313:LYS:HG2	1:B:327:SER:O	1.86	0.76
1:A:29:GLU:OE2	3:A:502:HOH:O	2.04	0.75
1:A:-4:HIS:ND1	1:B:85:LEU:HD22	2.03	0.73
1:A:-4:HIS:CE1	1:B:85:LEU:HD22	2.27	0.69
1:B:186:THR:HG22	1:B:188:ALA:H	1.58	0.67
1:A:191:GLN:OE1	3:A:503:HOH:O	2.13	0.67
1:B:315:MET:HB3	1:B:323:TRP:HB3	1.76	0.66
1:B:83:ILE:HD11	1:B:104:ILE:HD13	1.77	0.65
1:A:167:ARG:NH1	3:A:507:HOH:O	2.26	0.65
1:A:-3:HIS:HB3	3:A:809:HOH:O	1.96	0.64
1:A:313:LYS:HG2	1:A:327:SER:O	1.98	0.63
1:B:239:PRO:N	1:B:239:PRO:C	2.50	0.63
1:A:132:GLN:HG3	1:A:144:GLU:HB3	1.83	0.61
1:B:58:ASP:O	1:B:62:GLN:HG2	2.01	0.60
1:B:248:VAL:HG23	3:B:673:HOH:O	2.02	0.60
1:B:40:THR:OG1	3:B:405:HOH:O	1.90	0.58
1:A:111:SER:O	1:A:336:ASN:HB2	2.05	0.57
1:B:313:LYS:HD2	1:B:328:ASN:OD1	2.05	0.56
1:A:315:MET:HB3	1:A:323:TRP:HB3	1.88	0.55
1:B:33:GLN:NE2	3:B:410:HOH:O	2.34	0.54
1:B:132:GLN:HG3	1:B:144:GLU:HB3	1.89	0.54
1:B:367:GLU:HG3	3:B:509:HOH:O	2.08	0.53
1:A:167:ARG:NH1	1:A:171:LEU:HD11	2.24	0.52
1:A:40:THR:HG22	1:A:41:ARG:HG3	1.92	0.52
1:B:196:VAL:HG11	1:B:213:GLU:HG2	1.92	0.52
1:A:78:MET:HE1	1:A:83:ILE:HD13	1.92	0.51
1:A:78:MET:CE	1:A:83:ILE:HD13	2.41	0.50
1:B:83:ILE:HD11	1:B:104:ILE:CD1	2.41	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:262:GLN:OE1	3:A:505:HOH:O	2.20	0.48
1:A:313:LYS:HE3	1:A:328:ASN:OD1	2.14	0.48
1:B:190:PRO:HB2	1:B:194:TYR:CZ	2.49	0.47
1:B:0:HIS:N	3:B:413:HOH:O	2.45	0.47
1:B:226:ARG:O	1:B:315:MET:HA	2.15	0.46
1:A:144:GLU:OE2	3:A:506:HOH:O	2.21	0.45
1:B:315:MET:HB3	1:B:323:TRP:CB	2.45	0.45
1:A:0:HIS:HB2	3:A:538:HOH:O	2.16	0.45
1:A:322:ALA:HB1	1:A:353:LEU:HD12	1.99	0.45
1:A:344:LEU:HD12	1:A:349:MET:HG2	1.98	0.45
1:B:198:SER:O	1:B:338:ARG:HD3	2.16	0.45
1:A:142:ILE:CG2	1:A:311:HIS:HB2	2.47	0.44
1:B:49:VAL:HG13	1:B:134:PHE:HD2	1.83	0.44
1:B:159:GLN:O	1:B:163:GLU:HG2	2.18	0.43
1:B:11:TYR:O	1:B:26:ASN:HA	2.19	0.42
1:B:344:LEU:HD12	1:B:349:MET:HG2	2.01	0.42
1:A:198:SER:O	1:A:338:ARG:HA	2.19	0.42
1:A:311:HIS:O	1:A:313:LYS:HD2	2.21	0.41
1:A:136:TRP:CE2	1:A:137:ARG:HG3	2.55	0.41
1:A:226:ARG:O	1:A:315:MET:HA	2.21	0.41
1:A:196:VAL:HG11	1:A:213:GLU:HG2	2.02	0.41
1:A:226:ARG:HD2	1:A:264:GLU:OE1	2.21	0.41
1:A:236:SER:O	1:A:243:ARG:HA	2.21	0.41
1:B:370:LYS:HE3	1:B:373:TYR:CD2	2.56	0.41

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:530:HOH:O	3:B:668:HOH:O[3_655]	2.12	0.08
3:A:829:HOH:O	3:A:835:HOH:O[3_745]	2.15	0.05

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	386/391 (99%)	375 (97%)	11 (3%)	0	100	100
1	B	383/391 (98%)	372 (97%)	11 (3%)	0	100	100
All	All	769/782 (98%)	747 (97%)	22 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	323/326 (99%)	319 (99%)	4 (1%)	71	65
1	B	320/326 (98%)	316 (99%)	4 (1%)	69	62
All	All	643/652 (99%)	635 (99%)	8 (1%)	71	65

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	22	ASP
1	A	191	GLN
1	A	271	ASN
1	A	341	GLU
1	B	240	GLU
1	B	271	ASN
1	B	276	ASP
1	B	341	GLU

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)

1 ligand is 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	SO4	A	401	-	4,4,4	0.23	0	6,6,6	0.44	0

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.

No monomer is involved in short contacts.

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	388/391 (99%)	0.05	13 (3%)	45	39	18, 26, 45, 80	0
1	B	384/391 (98%)	0.12	18 (4%)	31	25	19, 26, 45, 83	0
All	All	772/782 (98%)	0.08	31 (4%)	38	32	18, 26, 45, 83	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	187	ALA	6.1
1	B	186	THR	4.4
1	B	0	HIS	4.3
1	A	383	GLU	4.2
1	B	382	GLY	4.1
1	A	237	TYR	4.0
1	B	183	TYR	4.0
1	A	240	GLU	4.0
1	A	241	ARG	3.9
1	A	186	THR	3.7
1	B	185	PRO	3.6
1	A	-4	HIS	3.1
1	B	53	GLN	2.9
1	B	241	ARG	2.9
1	A	372	ASP	2.8
1	B	182	THR	2.7
1	A	-2	HIS	2.7
1	A	53	GLN	2.7
1	B	64	LEU	2.6
1	A	22	ASP	2.6
1	B	383	GLU	2.5
1	A	0	HIS	2.4
1	B	303	PHE	2.4
1	B	188	ALA	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	377	ALA	2.4
1	B	175	ASP	2.2
1	A	238	GLY	2.2
1	B	240	GLU	2.1
1	A	-3	HIS	2.0
1	B	39	LYS	2.0
1	B	184	GLN	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	SO4	A	401	5/5	0.96	0.13	31,33,42,43	0

6.5 Other polymers (i)

There are no such residues in this entry.