wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 08:19 AM EDT

PDB ID	:	3RFZ
Title	:	Crystal structure of the FimD usher bound to its cognate FimC:FimH substrate
Authors	:	Phan, G.; Remaut, H.; Lebedev, A.; Geibel, S.; Waksman, G.
Deposited on	:	2011-04-07
Resolution	:	2.80 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:

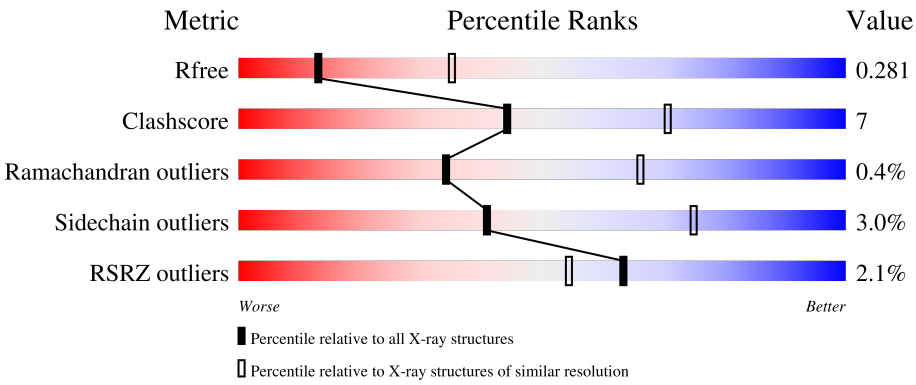
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	279	% 81% 19% •
1	D	279	% 82% 18%
2	B	843	2% 79% 12% • 8%
2	E	843	3% 76% 15% • 8%
3	C	211	% 77% 17% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	211	3% 77% 17% 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 19119 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 Type 1 fimbrial adhesin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	279	2034	1290	335	405	4	0	0	0
1	D	279	2036	1288	336	408	4	0	0	0

- Molecule 2 is a protein called Outer membrane usher protein, type 1 fimbrial synthesis.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	778	5948	3712	1054	1163	19	0	0	0
2	E	778	5969	3725	1057	1168	19	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	834	SER	-	EXPRESSION TAG	UNP C6UL88
B	835	ALA	-	EXPRESSION TAG	UNP C6UL88
B	836	TRP	-	EXPRESSION TAG	UNP C6UL88
B	837	SER	-	EXPRESSION TAG	UNP C6UL88
B	838	HIS	-	EXPRESSION TAG	UNP C6UL88
B	839	PRO	-	EXPRESSION TAG	UNP C6UL88
B	840	GLN	-	EXPRESSION TAG	UNP C6UL88
B	841	PHE	-	EXPRESSION TAG	UNP C6UL88
B	842	GLU	-	EXPRESSION TAG	UNP C6UL88
B	843	LYS	-	EXPRESSION TAG	UNP C6UL88
E	834	SER	-	EXPRESSION TAG	UNP C6UL88
E	835	ALA	-	EXPRESSION TAG	UNP C6UL88
E	836	TRP	-	EXPRESSION TAG	UNP C6UL88
E	837	SER	-	EXPRESSION TAG	UNP C6UL88
E	838	HIS	-	EXPRESSION TAG	UNP C6UL88
E	839	PRO	-	EXPRESSION TAG	UNP C6UL88

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	840	GLN	-	EXPRESSION TAG	UNP C6UL88
E	841	PHE	-	EXPRESSION TAG	UNP C6UL88
E	842	GLU	-	EXPRESSION TAG	UNP C6UL88
E	843	LYS	-	EXPRESSION TAG	UNP C6UL88

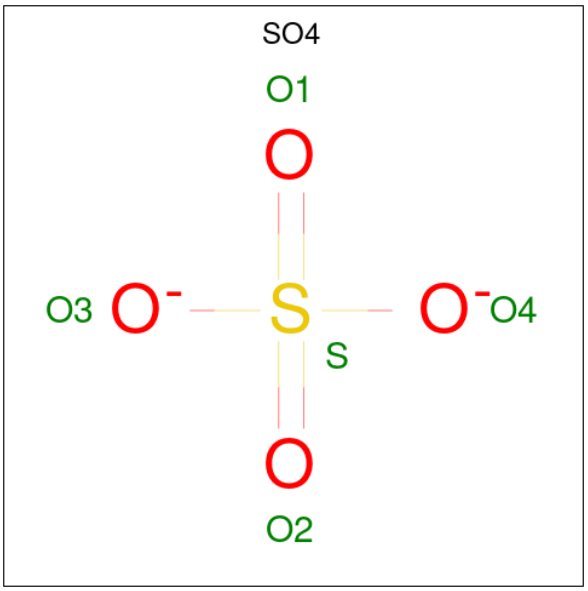
- Molecule 3 is a protein called Chaperone protein fimC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	201	1539	978	262	293	6	0	0	0
			Total	C	N	O	S			
3	F	201	1542	982	262	293	5	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	206	HIS	-	EXPRESSION TAG	UNP P59590
C	207	HIS	-	EXPRESSION TAG	UNP P59590
C	208	HIS	-	EXPRESSION TAG	UNP P59590
C	209	HIS	-	EXPRESSION TAG	UNP P59590
C	210	HIS	-	EXPRESSION TAG	UNP P59590
C	211	HIS	-	EXPRESSION TAG	UNP P59590
F	206	HIS	-	EXPRESSION TAG	UNP P59590
F	207	HIS	-	EXPRESSION TAG	UNP P59590
F	208	HIS	-	EXPRESSION TAG	UNP P59590
F	209	HIS	-	EXPRESSION TAG	UNP P59590
F	210	HIS	-	EXPRESSION TAG	UNP P59590
F	211	HIS	-	EXPRESSION TAG	UNP P59590

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 5 is water.

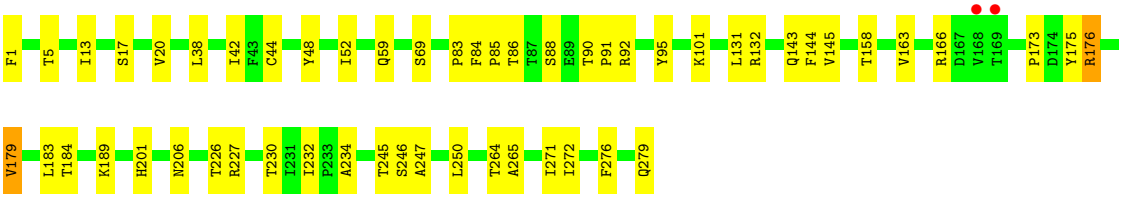
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	2	2	2	0	0
			Total	O		
5	B	22	22	22	0	0
			Total	O		
5	C	4	4	4	0	0
			Total	O		
5	D	1	1	1	0	0
			Total	O		
5	E	8	8	8	0	0
			Total	O		
5	F	4	4	4	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

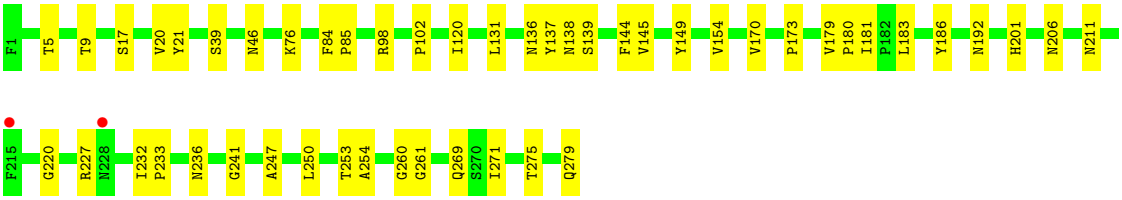
• Molecule 1: Type 1 fimbrial adhesin

Chain A:



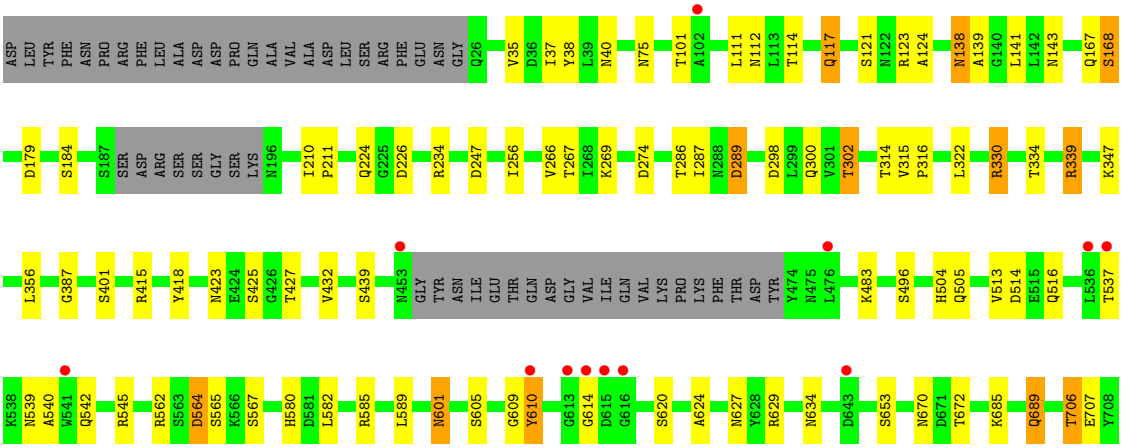
• Molecule 1: Type 1 fimbrial adhesin

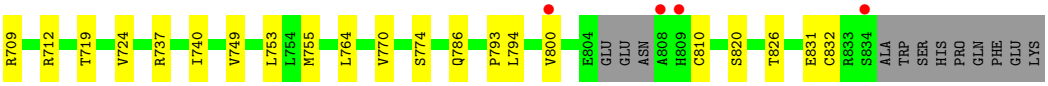
Chain D:



• Molecule 2: Outer membrane usher protein, type 1 fimbrial synthesis

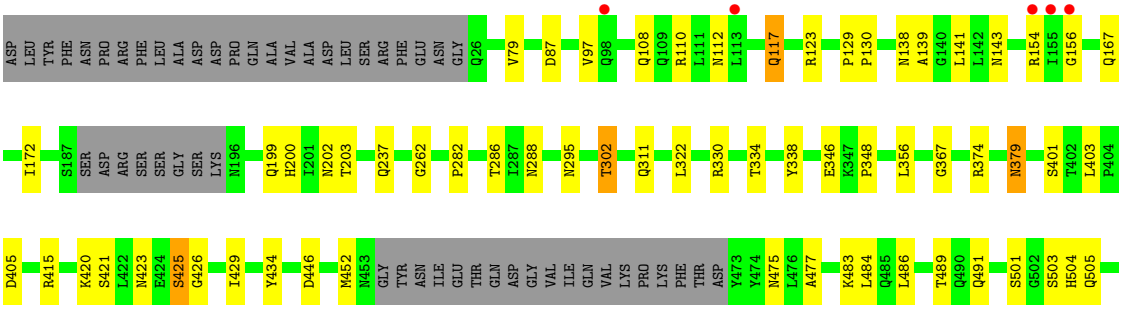
Chain B:

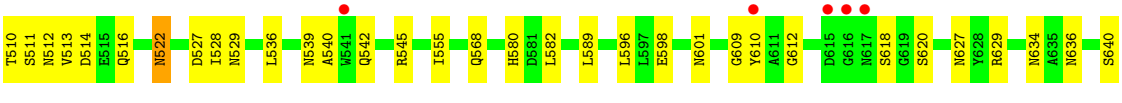


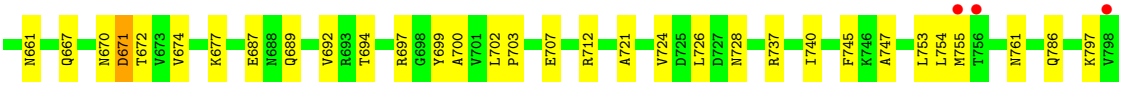


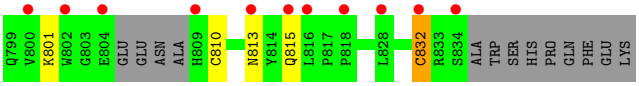
• Molecule 2: Outer membrane usher protein, type 1 fimbrial synthesis

Chain E:



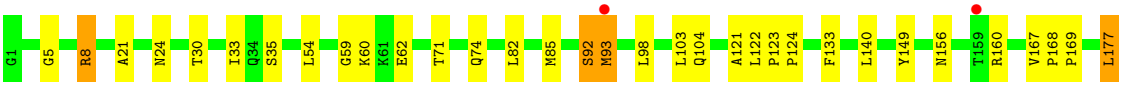


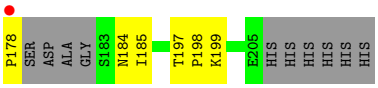




• Molecule 3: Chaperone protein fimC

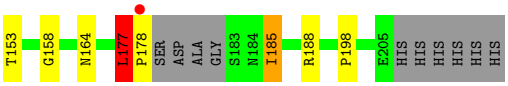
Chain C:





• Molecule 3: Chaperone protein fimC

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	126.19Å 95.90Å 144.57Å 90.00° 112.13° 90.00°	Depositor
Resolution (Å)	59.13 – 2.80 59.15 – 2.80	Depositor EDS
% Data completeness (in resolution range)	97.3 (59.13-2.80) 97.3 (59.15-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.6.0077	Depositor
R, R_{free}	0.223 , 0.282 0.222 , 0.281	Depositor DCC
R_{free} test set	3845 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	53.7	Xtriage
Anisotropy	0.141	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 43.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	19119	wwPDB-VP
Average B, all atoms (Å^2)	61.0	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 35.02 % 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 6.2235e-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: 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	0.39	0/2079	0.54	0/2859
1	D	0.39	0/2081	0.54	0/2861
2	B	0.39	0/6076	0.58	0/8273
2	E	0.38	0/6097	0.56	0/8299
3	C	0.35	0/1567	0.54	0/2138
3	F	0.35	0/1570	0.56	0/2140
All	All	0.38	0/19470	0.56	0/26570

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 (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	2034	0	1969	31	0
1	D	2036	0	1961	29	0
2	B	5948	0	5647	72	0
2	E	5969	0	5692	77	0
3	C	1539	0	1542	28	0
3	F	1542	0	1560	26	0
4	B	10	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	2	0	0	0	0
5	B	22	0	0	0	0
5	C	4	0	0	0	0
5	D	1	0	0	0	0
5	E	8	0	0	0	0
5	F	4	0	0	0	0
All	All	19119	0	18371	246	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 7.

The worst 5 of 246 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)
2:B:516:GLN:HG3	2:B:537:THR:HG22	1.53	0.91
2:B:514:ASP:OD1	2:B:540:ALA:HB2	1.74	0.87
1:A:173:PRO:HD3	1:A:179:VAL:HG22	1.56	0.86
2:E:797:LYS:HG2	2:E:813:ASN:HB3	1.62	0.80
2:E:627:ASN:HD22	2:E:636:ASN:HB3	1.47	0.80

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	277/279 (99%)	263 (95%)	12 (4%)	2 (1%)	22 53
1	D	277/279 (99%)	256 (92%)	20 (7%)	1 (0%)	34 66
2	B	770/843 (91%)	729 (95%)	41 (5%)	0	100 100
2	E	770/843 (91%)	722 (94%)	45 (6%)	3 (0%)	34 66
3	C	197/211 (93%)	185 (94%)	11 (6%)	1 (0%)	29 61

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	F	197/211 (93%)	187 (95%)	7 (4%)	3 (2%)	10	33
All	All	2488/2666 (93%)	2342 (94%)	136 (6%)	10 (0%)	34	66

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	176	ARG
2	E	156	GLY
2	E	671	ASP
3	C	177	LEU
3	F	119	LYS

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	220/226 (97%)	215 (98%)	5 (2%)	50	82
1	D	220/226 (97%)	219 (100%)	1 (0%)	88	96
2	B	620/694 (89%)	596 (96%)	24 (4%)	32	66
2	E	627/694 (90%)	605 (96%)	22 (4%)	36	70
3	C	163/181 (90%)	158 (97%)	5 (3%)	40	74
3	F	165/181 (91%)	161 (98%)	4 (2%)	49	81
All	All	2015/2202 (92%)	1954 (97%)	61 (3%)	41	75

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	810	CYS
2	E	801	LYS
2	E	87	ASP
2	E	761	ASN
3	F	177	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	191	GLN
2	E	112	ASN
2	E	689	GLN
1	D	211	ASN
1	D	255	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	SO4	B	844	-	4,4,4	0.24	0	6,6,6	0.10	0
4	SO4	B	845	-	4,4,4	0.21	0	6,6,6	0.23	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 (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	279/279 (100%)	-0.14	2 (0%)	87	84	32, 52, 82, 125	0
1	D	279/279 (100%)	-0.06	2 (0%)	87	84	32, 55, 94, 114	0
2	B	778/843 (92%)	-0.04	16 (2%)	63	54	33, 54, 96, 127	2 (0%)
2	E	778/843 (92%)	0.02	24 (3%)	49	39	35, 59, 103, 136	2 (0%)
3	C	201/211 (95%)	0.01	3 (1%)	73	68	39, 57, 103, 121	0
3	F	201/211 (95%)	0.04	7 (3%)	44	34	42, 63, 107, 132	0
All	All	2516/2666 (94%)	-0.02	54 (2%)	63	54	32, 57, 99, 136	4 (0%)

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	616	GLY	5.7
3	F	178	PRO	5.5
2	B	808	ALA	4.6
2	B	616	GLY	4.6
2	B	610	TYR	3.8

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
4	SO4	B	845	5/5	0.97	0.12	60,61,63,68	0
4	SO4	B	844	5/5	0.99	0.13	65,66,67,71	0

6.5 Other polymers (i)

There are no such residues in this entry.