Full wwPDB X-ray Structure Validation Report (i)

Apr 6, 2016 – 08:55 AM EDT

PDB ID	:	5EN5
Title	:	Apo structure of bacterial efflux pump.
Authors	:	Sjuts, H.; Ornik, A.R.; Pos, K.M.
Deposited on	:	2015-11-09
Resolution	:	2.30 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
http://wwpdb.org/validation/2016/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:

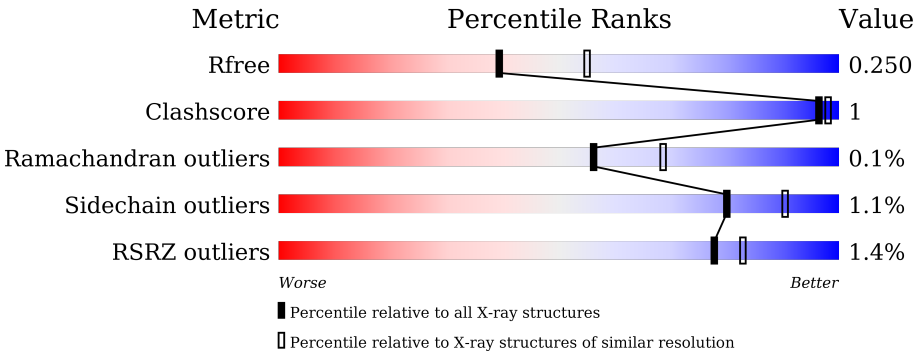
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027107
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027107

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 2.30 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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	609	2% 91% • 5%
1	B	609	89% 6% 5%
1	C	609	% 91% 5% •
2	D	169	% 91% • 5%
2	E	169	2% 85% 7% 8%
2	F	169	2% 92% • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 17497 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 Multidrug efflux pump subunit AcrB,Multidrug efflux pump subunit AcrB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	580	Total	C	N	O	S	0	0	0
			4417	2773	748	874	22			
1	B	578	Total	C	N	O	S	0	0	0
			4405	2766	746	871	22			
1	C	584	Total	C	N	O	S	0	0	0
			4441	2789	752	878	22			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	552	GLY	-	linker	UNP P31224
A	553	GLY	-	linker	UNP P31224
A	554	SER	-	linker	UNP P31224
A	555	GLY	-	linker	UNP P31224
A	556	GLY	-	linker	UNP P31224
A	557	SER	-	linker	UNP P31224
A	558	GLY	-	linker	UNP P31224
A	559	GLY	-	linker	UNP P31224
A	560	SER	-	linker	UNP P31224
B	552	GLY	-	linker	UNP P31224
B	553	GLY	-	linker	UNP P31224
B	554	SER	-	linker	UNP P31224
B	555	GLY	-	linker	UNP P31224
B	556	GLY	-	linker	UNP P31224
B	557	SER	-	linker	UNP P31224
B	558	GLY	-	linker	UNP P31224
B	559	GLY	-	linker	UNP P31224
B	560	SER	-	linker	UNP P31224
C	552	GLY	-	linker	UNP P31224
C	553	GLY	-	linker	UNP P31224
C	554	SER	-	linker	UNP P31224
C	555	GLY	-	linker	UNP P31224

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	556	GLY	-	linker	UNP P31224
C	557	SER	-	linker	UNP P31224
C	558	GLY	-	linker	UNP P31224
C	559	GLY	-	linker	UNP P31224
C	560	SER	-	linker	UNP P31224

- Molecule 2 is a protein called DARPin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	161	Total	C	N	O	S	0	0	0
			1227	771	221	234	1			
2	E	156	Total	C	N	O	S	0	0	0
			1177	741	206	229	1			
2	F	162	Total	C	N	O	S	0	0	0
			1237	777	224	235	1			

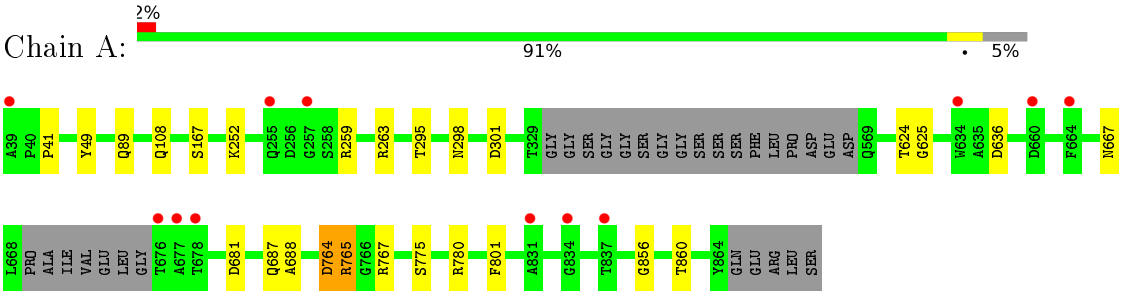
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	169	Total	O	0	0
			169	169		
3	B	193	Total	O	0	0
			193	193		
3	C	173	Total	O	0	0
			173	173		
3	D	23	Total	O	0	0
			23	23		
3	E	17	Total	O	0	0
			17	17		
3	F	18	Total	O	0	0
			18	18		

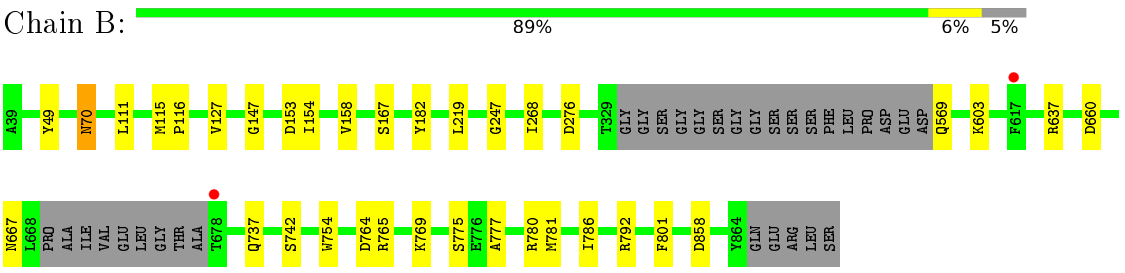
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

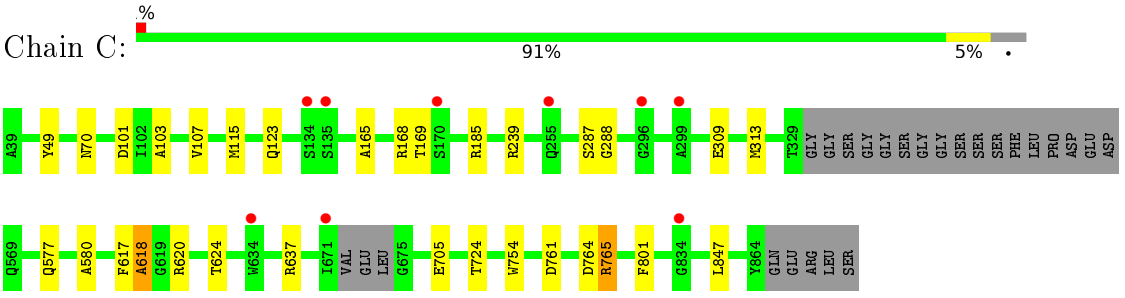
- Molecule 1: Multidrug efflux pump subunit AcrB,Multidrug efflux pump subunit AcrB



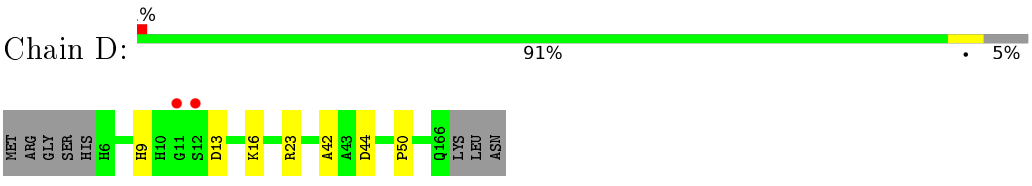
- Molecule 1: Multidrug efflux pump subunit AcrB,Multidrug efflux pump subunit AcrB



- Molecule 1: Multidrug efflux pump subunit AcrB,Multidrug efflux pump subunit AcrB

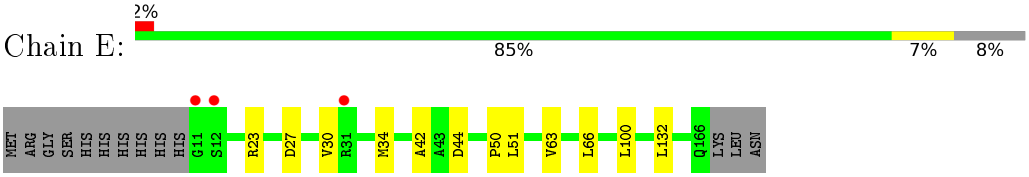


- Molecule 2: DARPin



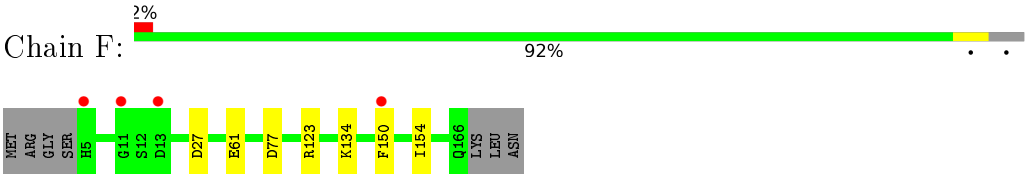
• Molecule 2: DARPin

Chain E:



• Molecule 2: DARPin

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	108.34Å 145.29Å 174.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.30 49.63 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-2.30) 99.7 (49.63-2.30)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.29 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.197 , 0.249 0.202 , 0.250	Depositor DCC
R_{free} test set	5991 reflections (5.15%)	DCC
Wilson B-factor (Å^2)	29.7	Xtriage
Anisotropy	0.297	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 28.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 122370 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	17497	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.79	0/4492	0.87	7/6087 (0.1%)
1	B	0.80	0/4480	0.88	6/6070 (0.1%)
1	C	0.79	0/4517	0.86	9/6122 (0.1%)
2	D	0.71	0/1251	0.80	1/1701 (0.1%)
2	E	0.69	0/1196	0.82	1/1626 (0.1%)
2	F	0.69	0/1262	0.77	2/1716 (0.1%)
All	All	0.77	0/17198	0.85	26/23322 (0.1%)

There are no bond length outliers.

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	765	ARG	NE-CZ-NH1	8.48	124.54	120.30
1	B	765	ARG	NE-CZ-NH1	8.26	124.43	120.30
1	A	263	ARG	NE-CZ-NH1	7.56	124.08	120.30
1	C	637	ARG	NE-CZ-NH1	7.17	123.89	120.30
1	C	765	ARG	NE-CZ-NH1	6.93	123.76	120.30
2	F	77	ASP	CB-CG-OD1	6.44	124.10	118.30
1	C	765	ARG	NE-CZ-NH2	-6.17	117.22	120.30
1	B	765	ARG	NE-CZ-NH2	-6.13	117.23	120.30
2	D	44	ASP	CB-CG-OD1	5.90	123.61	118.30
1	C	168	ARG	NE-CZ-NH1	5.73	123.17	120.30
2	E	44	ASP	CB-CG-OD1	5.69	123.42	118.30
1	A	765	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	C	761	ASP	CB-CG-OD2	-5.65	113.21	118.30
2	F	123	ARG	NE-CZ-NH1	5.63	123.11	120.30
1	B	792	ARG	NE-CZ-NH1	5.57	123.08	120.30
1	B	637	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	C	185	ARG	NE-CZ-NH2	-5.53	117.54	120.30
1	B	603	LYS	CB-CA-C	-5.50	99.40	110.40
1	B	276	ASP	CB-CG-OD2	-5.49	113.36	118.30
1	A	263	ARG	NE-CZ-NH2	-5.41	117.59	120.30
1	C	620	ARG	NE-CZ-NH1	5.21	122.91	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	259	ARG	NE-CZ-NH2	-5.21	117.70	120.30
1	A	767	ARG	NE-CZ-NH2	5.13	122.87	120.30
1	A	764	ASP	CB-CG-OD1	5.10	122.89	118.30
1	C	239	ARG	NE-CZ-NH2	-5.07	117.77	120.30
1	C	101	ASP	CB-CG-OD1	5.04	122.84	118.30

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	4417	0	4355	13	0
1	B	4405	0	4343	14	0
1	C	4441	0	4381	16	0
2	D	1227	0	1194	1	0
2	E	1177	0	1159	5	0
2	F	1237	0	1201	2	0
3	A	169	0	0	2	0
3	B	193	0	0	0	0
3	C	173	0	0	3	0
3	D	23	0	0	0	0
3	E	17	0	0	0	0
3	F	18	0	0	0	0
All	All	17497	0	16633	47	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 1.

All (47) 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:A:108:GLN:OE1	3:A:901:HOH:O	2.08	0.72
1:B:569:GLN:HG2	1:B:667:ASN:HD21	1.64	0.63
1:C:580:ALA:HB1	1:C:724:THR:HG22	1.81	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:167:SER:HB3	1:B:70:ASN:HB3	1.86	0.57
1:C:70:ASN:HB3	3:C:972:HOH:O	2.04	0.57
1:C:617:PHE:O	1:C:618:ALA:HB3	2.07	0.55
1:A:681:ASP:HB3	1:A:860:THR:O	2.07	0.54
1:B:153:ASP:OD1	1:B:182:TYR:OH	2.27	0.53
1:C:705:GLU:HB3	1:C:847:LEU:HD22	1.93	0.50
1:C:618:ALA:HA	3:C:1018:HOH:O	2.11	0.49
1:C:764:ASP:OD1	1:C:765:ARG:HD3	2.12	0.48
1:A:687:GLN:HE21	1:A:856:GLY:HA3	1.77	0.48
2:E:30:VAL:O	2:E:34:MET:HG2	2.15	0.47
1:C:165:ALA:HB3	1:C:313:MET:CE	2.45	0.46
1:C:115:MET:HE3	1:C:123:GLN:HG2	1.97	0.46
1:C:577:GLN:OE1	1:C:624:THR:HG22	2.15	0.46
2:E:51:LEU:HD11	2:E:63:VAL:HG13	1.97	0.45
1:C:617:PHE:O	1:C:618:ALA:CB	2.64	0.45
1:B:247:GLY:HA2	1:B:268:ILE:CD1	2.47	0.45
1:B:167:SER:HB3	1:C:70:ASN:HB3	1.97	0.45
1:B:777:ALA:O	1:B:781:MET:HG2	2.17	0.44
1:A:667:ASN:CG	1:A:667:ASN:O	2.55	0.44
2:F:150:PHE:CE2	2:F:154:ILE:HD11	2.53	0.44
1:B:775:SER:HB3	1:B:780:ARG:HD3	2.01	0.43
1:A:298:ASN:HD22	1:A:301:ASP:H	1.66	0.43
1:A:688:ALA:HA	3:C:1036:HOH:O	2.18	0.43
2:E:42:ALA:O	2:E:50:PRO:HD3	2.19	0.43
1:B:111:LEU:HD21	1:B:127:VAL:HG11	2.02	0.42
1:B:115:MET:N	1:B:116:PRO:CD	2.82	0.42
2:F:150:PHE:CZ	2:F:154:ILE:HD11	2.54	0.42
1:A:764:ASP:OD1	1:A:765:ARG:HD3	2.20	0.42
1:C:103:ALA:O	1:C:107:VAL:HG23	2.19	0.42
1:B:754:TRP:CZ2	1:B:786:ILE:HD13	2.54	0.42
1:B:764:ASP:HB3	1:B:769:LYS:HD2	2.01	0.42
1:B:219:LEU:HD23	1:C:754:TRP:CZ3	2.55	0.42
1:A:775:SER:HB3	1:A:780:ARG:HD3	2.03	0.41
1:A:252:LYS:HA	1:B:737:GLN:OE1	2.21	0.41
2:E:100:LEU:HD11	2:E:132:LEU:HD23	2.02	0.41
1:A:41:PRO:HB3	1:A:295:THR:HG21	2.02	0.41
2:E:34:MET:HE2	2:E:66:LEU:HD23	2.02	0.41
1:B:154:ILE:O	1:B:158:VAL:HG23	2.20	0.41
1:C:287:SER:OG	1:C:288:GLY:N	2.51	0.41
2:D:42:ALA:O	2:D:50:PRO:HD3	2.20	0.41
1:C:580:ALA:CB	1:C:724:THR:HG22	2.50	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:169:THR:OG1	1:C:309:GLU:HG3	2.21	0.40
1:A:624:THR:HG22	1:A:625:GLY:N	2.36	0.40
1:A:89:GLN:NE2	3:A:902:HOH:O	2.25	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	574/609 (94%)	556 (97%)	18 (3%)	0	100	100
1	B	572/609 (94%)	556 (97%)	15 (3%)	1 (0%)	52	64
1	C	578/609 (95%)	558 (96%)	19 (3%)	1 (0%)	52	64
2	D	159/169 (94%)	156 (98%)	2 (1%)	1 (1%)	30	36
2	E	154/169 (91%)	150 (97%)	4 (3%)	0	100	100
2	F	160/169 (95%)	157 (98%)	3 (2%)	0	100	100
All	All	2197/2334 (94%)	2133 (97%)	61 (3%)	3 (0%)	56	68

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	13	ASP
1	C	618	ALA
1	B	147	GLY

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	471/492 (96%)	468 (99%)	3 (1%)	90 96
1	B	470/492 (96%)	464 (99%)	6 (1%)	76 87
1	C	473/492 (96%)	471 (100%)	2 (0%)	93 97
2	D	125/132 (95%)	122 (98%)	3 (2%)	57 74
2	E	120/132 (91%)	118 (98%)	2 (2%)	68 83
2	F	126/132 (96%)	123 (98%)	3 (2%)	57 74
All	All	1785/1872 (95%)	1766 (99%)	19 (1%)	80 90

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	TYR
1	A	636	ASP
1	A	801	PHE
1	B	49	TYR
1	B	70	ASN
1	B	660	ASP
1	B	742	SER
1	B	801	PHE
1	B	858	ASP
1	C	49	TYR
1	C	801	PHE
2	D	9	HIS
2	D	16	LYS
2	D	23	ARG
2	E	23	ARG
2	E	27	ASP
2	F	27	ASP
2	F	61	GLU
2	F	134	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	298	ASN
1	A	687	GLN
1	B	231	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	622	GLN
1	B	667	ASN
1	C	112	GLN
1	C	298	ASN
2	D	10	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	580/609 (95%)	-0.32	12 (2%) 67 74	16, 30, 66, 102	0
1	B	578/609 (94%)	-0.47	2 (0%) 94 96	15, 28, 57, 96	0
1	C	584/609 (95%)	-0.30	9 (1%) 76 81	17, 32, 71, 107	0
2	D	161/169 (95%)	-0.29	2 (1%) 81 85	22, 37, 67, 106	0
2	E	156/169 (92%)	-0.22	3 (1%) 70 76	24, 37, 75, 110	0
2	F	162/169 (95%)	-0.18	4 (2%) 61 70	27, 43, 80, 112	0
All	All	2221/2334 (95%)	-0.34	32 (1%) 78 83	15, 32, 67, 112	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	11	GLY	7.6
1	A	676	THR	5.5
1	A	677	ALA	4.2
2	E	12	SER	4.1
2	D	11	GLY	3.8
2	E	31	ARG	3.3
1	A	634	TRP	3.2
1	C	170	SER	3.2
2	F	11	GLY	3.2
1	C	134	SER	3.0
1	A	257	GLY	3.0
1	C	255	GLN	2.8
2	F	150	PHE	2.8
1	C	296	GLY	2.7
1	A	255	GLN	2.7
1	A	831	ALA	2.6
1	C	634	TRP	2.5
1	A	678	THR	2.4
1	C	834	GLY	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	837	THR	2.3
1	B	678	THR	2.3
1	A	660	ASP	2.3
2	D	12	SER	2.2
1	C	135	SER	2.2
2	F	5	HIS	2.2
1	A	834	GLY	2.2
1	C	299	ALA	2.1
2	F	13	ASP	2.1
1	B	617	PHE	2.1
1	C	671	ILE	2.1
1	A	39	ALA	2.0
1	A	664	PHE	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 carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.