wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 01:20 pm BST

PDB ID	:	3PU2
Title	:	Crystal Structure of the Q3J4M4_RHOS4 protein from Rhodobacter sphaeroides. Northeast Structural Genomics Consortium Target RhR263.
Authors	:	Vorobiev, S.; Chen, Y.; Seetharaman, J.; Sahdev, S.; Xiao, R.; Ciccosanti, C.; Wang, D.; Everett, J.K.; Nair, R.; Acton, T.B.; Rost, B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2010-12-03
Resolution	:	2.61 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

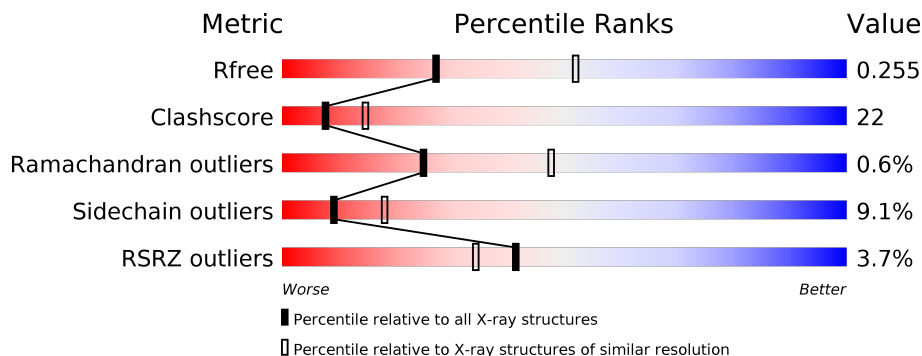|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.61 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.
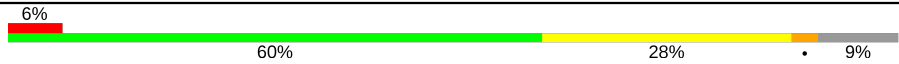
Metric	Percentile Ranks	Value
Rfree		0.255
Clashscore		22
Ramachandran outliers		0.6%
Sidechain outliers		9.1%
RSRZ outliers		3.7%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 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	164	% 60% 27% 5% 8%
1	B	164	2% 56% 34% • 7%
1	C	164	2% 59% 28% 6% 7%
1	D	164	4% 55% 31% • 9%
1	E	164	5% 55% 30% 6% 8%
1	F	164	2% 53% 33% • 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	164	6% · · · 60% · · · 28% · · · 9%
1	H	164	4% · · · 57% · · · 29% · · · 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9371 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 uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	151	Total	C	N	O	S	Se	0	1	0
			1167	744	203	213	1	6			
1	B	153	Total	C	N	O	S	Se	0	1	0
			1181	751	205	218	1	6			
1	C	152	Total	C	N	O	S	Se	0	1	0
			1182	752	204	219	1	6			
1	D	149	Total	C	N	O	S	Se	0	1	0
			1157	739	201	210	1	6			
1	E	151	Total	C	N	O	S	Se	0	0	0
			1169	744	203	216	1	5			
1	F	148	Total	C	N	O	S	Se	0	0	0
			1142	730	197	209	1	5			
1	G	149	Total	C	N	O	S	Se	0	0	0
			1153	734	203	210	1	5			
1	H	148	Total	C	N	O	S	Se	0	0	0
			1148	733	200	209	1	5			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
A	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
A	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
A	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
A	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4
A	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
A	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4
A	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4
B	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
B	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
B	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
B	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
B	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
B	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4
B	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4
C	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
C	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
C	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
C	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
C	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4
C	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
C	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4
C	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4
D	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
D	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
D	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
D	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
D	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4
D	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
D	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4
D	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4
E	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
E	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
E	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
E	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
E	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4
E	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
E	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4
E	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4
F	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
F	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
F	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
F	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
F	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4
F	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
F	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4
F	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4
G	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
G	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
G	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
G	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
G	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4
G	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
G	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4
H	169	LEU	-	EXPRESSION TAG	UNP Q3J4M4
H	170	GLU	-	EXPRESSION TAG	UNP Q3J4M4
H	171	HIS	-	EXPRESSION TAG	UNP Q3J4M4
H	172	HIS	-	EXPRESSION TAG	UNP Q3J4M4
H	173	HIS	-	EXPRESSION TAG	UNP Q3J4M4
H	174	HIS	-	EXPRESSION TAG	UNP Q3J4M4
H	175	HIS	-	EXPRESSION TAG	UNP Q3J4M4
H	176	HIS	-	EXPRESSION TAG	UNP Q3J4M4

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	14	Total 14	O 14	0	0
2	B	14	Total 14	O 14	0	0
2	C	10	Total 10	O 10	0	0
2	D	8	Total 8	O 8	0	0
2	E	7	Total 7	O 7	0	0
2	F	10	Total 10	O 10	0	0
2	G	4	Total 4	O 4	0	0
2	H	5	Total 5	O 5	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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
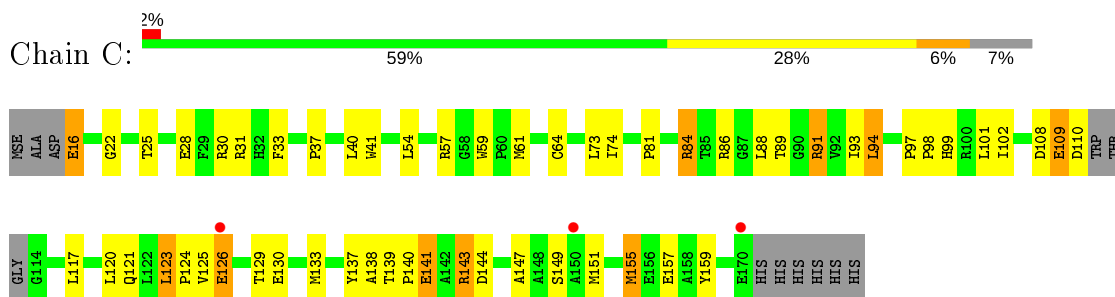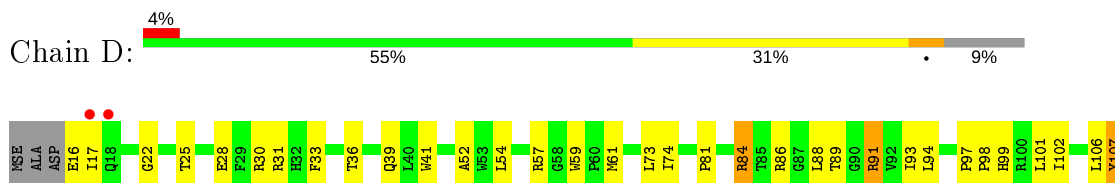sample, but not in the model, are shown in grey.

- Molecule 1: uncharacterized protein
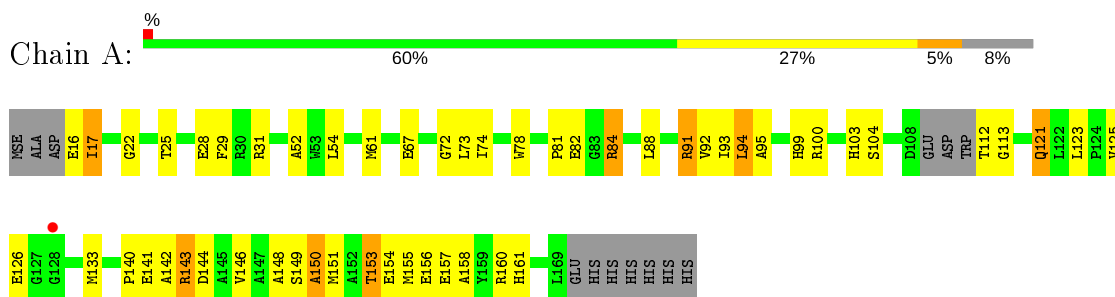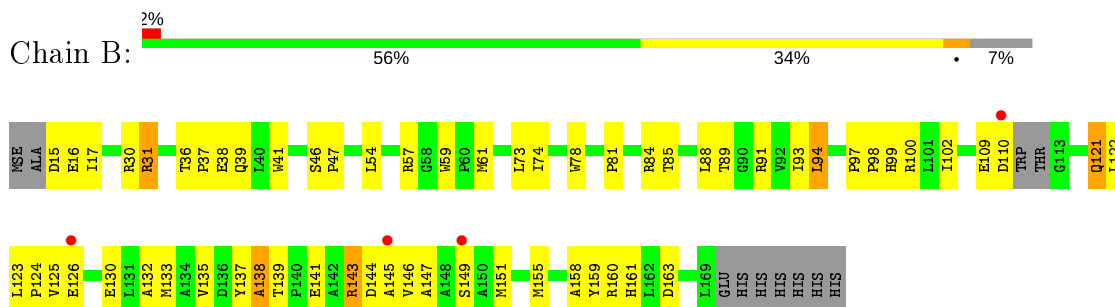
- Molecule 1: uncharacterized protein
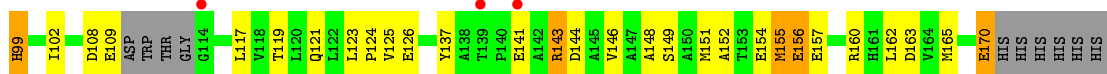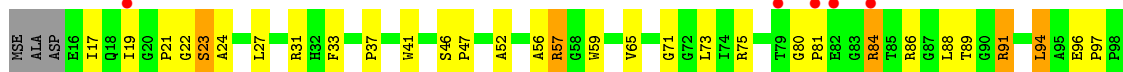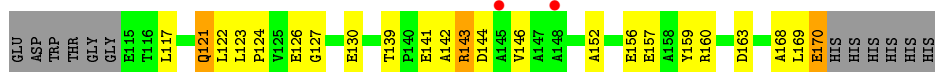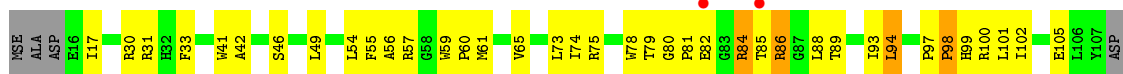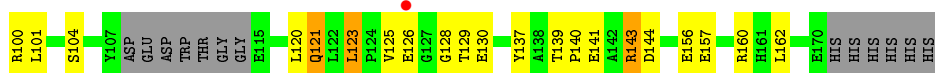

- Molecule 1: uncharacterized protein

- Molecule 1: uncharacterized protein

● Molecule 1: uncharacterized protein

Chain E:

● Molecule 1: uncharacterized protein

Chain F:

● Molecule 1: uncharacterized protein

Chain G:

● Molecule 1: uncharacterized protein

Chain H:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	75.21Å　35.86Å　230.47Å 90.00°　98.98°　90.00°	Depositor
Resolution (Å)	29.10　–　2.61 29.10　–　2.61	Depositor EDS
% Data completeness (in resolution range)	93.3 (29.10-2.61) 95.7 (29.10-2.61)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.80 (at 2.61Å)	Xtriage
Refinement program	PHENIX 1.6_289	Depositor
R, R_{free}	0.230　,　0.263 0.225　,　0.255	Depositor DCC
R_{free} test set	1833 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	47.0	Xtriage
Anisotropy	0.679	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 50.6	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9371	wwPDB-VP
Average B, all atoms (Å^2)	66.0	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 70.39 % 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 3.1464e-06. 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 ⓘ

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	0.55	0/1194	0.68	0/1619
1	B	0.50	0/1208	0.69	0/1638
1	C	0.55	0/1209	0.71	0/1639
1	D	0.50	0/1184	0.70	2/1606 (0.1%)
1	E	0.40	0/1193	0.72	0/1618
1	F	0.38	0/1166	0.69	1/1583 (0.1%)
1	G	0.36	0/1178	0.65	0/1598
1	H	0.43	0/1172	0.70	0/1590
All	All	0.47	0/9504	0.69	3/12891 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	127	GLY	N-CA-C	-5.86	98.46	113.10
1	D	108	ASP	N-CA-C	5.14	124.88	111.00
1	D	107	TYR	N-CA-C	-5.04	97.40	111.00

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	1167	0	1157	50	0
1	B	1181	0	1160	65	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1182	0	1163	53	0
1	D	1157	0	1146	52	0
1	E	1169	0	1150	55	0
1	F	1142	0	1126	58	0
1	G	1153	0	1134	41	0
1	H	1148	0	1137	47	0
2	A	14	0	0	3	0
2	B	14	0	0	0	0
2	C	10	0	0	2	0
2	D	8	0	0	0	0
2	E	7	0	0	2	0
2	F	10	0	0	1	0
2	G	4	0	0	0	0
2	H	5	0	0	0	0
All	All	9371	0	9173	404	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 22.

The worst 5 of 404 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:133:MSE:HE2	1:B:155[B]:MSE:HE1	1.28	1.13
1:B:133:MSE:CE	1:B:155[B]:MSE:CE	2.29	1.10
1:B:133:MSE:HE2	1:B:155[B]:MSE:CE	1.82	1.08
1:F:17:ILE:HG22	1:F:31:ARG:HB2	1.37	1.03
1:B:155[A]:MSE:HE3	1:B:159:TYR:CZ	1.94	1.01

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	148/164 (90%)	136 (92%)	11 (7%)	1 (1%)	22	43
1	B	150/164 (92%)	141 (94%)	8 (5%)	1 (1%)	22	43
1	C	149/164 (91%)	141 (95%)	8 (5%)	0	100	100
1	D	146/164 (89%)	136 (93%)	9 (6%)	1 (1%)	22	43
1	E	147/164 (90%)	140 (95%)	6 (4%)	1 (1%)	22	43
1	F	144/164 (88%)	130 (90%)	13 (9%)	1 (1%)	22	43
1	G	145/164 (88%)	133 (92%)	11 (8%)	1 (1%)	22	43
1	H	144/164 (88%)	130 (90%)	13 (9%)	1 (1%)	22	43
All	All	1173/1312 (89%)	1087 (93%)	79 (7%)	7 (1%)	25	47

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	148	ALA
1	F	169	LEU
1	A	150	ALA
1	D	150	ALA
1	G	150	ALA

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	118/123 (96%)	108 (92%)	10 (8%)	10	21
1	B	119/123 (97%)	111 (93%)	8 (7%)	16	33
1	C	120/123 (98%)	105 (88%)	15 (12%)	4	8
1	D	117/123 (95%)	105 (90%)	12 (10%)	7	13
1	E	118/123 (96%)	104 (88%)	14 (12%)	5	9
1	F	115/123 (94%)	103 (90%)	12 (10%)	7	13
1	G	116/123 (94%)	108 (93%)	8 (7%)	15	31
1	H	116/123 (94%)	108 (93%)	8 (7%)	15	31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	939/984 (95%)	852 (91%)	87 (9%)	9 17

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	123	LEU
1	E	89	THR
1	H	31	ARG
1	D	136	ASP
1	D	155[B]	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	18	GLN
1	E	18	GLN
1	G	18	GLN
1	C	103	HIS
1	C	121	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 carbohydrates in this entry.

5.6   Ligand geometry (i)

There are no ligands in this entry.

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(Å²)	Q<0.9
1	A	146/164 (89%)	-0.22	1 (0%)	87	86	22, 57, 95, 118	0
1	B	148/164 (90%)	-0.16	4 (2%)	54	48	30, 55, 112, 161	0
1	C	147/164 (89%)	-0.30	3 (2%)	65	60	22, 50, 100, 140	0
1	D	144/164 (87%)	-0.06	6 (4%)	36	29	28, 64, 125, 188	0
1	E	146/164 (89%)	-0.08	8 (5%)	25	19	42, 63, 98, 138	0
1	F	143/164 (87%)	-0.00	4 (2%)	53	46	35, 64, 120, 156	0
1	G	144/164 (87%)	0.11	10 (6%)	16	12	34, 71, 138, 180	0
1	H	143/164 (87%)	0.05	7 (4%)	29	23	34, 70, 114, 164	0
All	All	1161/1312 (88%)	-0.08	43 (3%)	41	34	22, 62, 117, 188	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	82	GLU	6.2
1	C	170	GLU	4.7
1	G	83	GLY	4.6
1	D	148	ALA	4.5
1	H	85	THR	4.1

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.