



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 10:54 pm BST

PDB ID : 2PEI
Title : Crystal structure of selenomethionine-labeled RbcX
Authors : Saschenbrecker, S.; Bracher, A.; Vasudeva Rao, K.; Vasudeva Rao, B.; Hartl, F.U.; Hayer-Hartl, M.
Deposited on : 2007-04-03
Resolution : 2.70 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 ⓘ symbol.

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

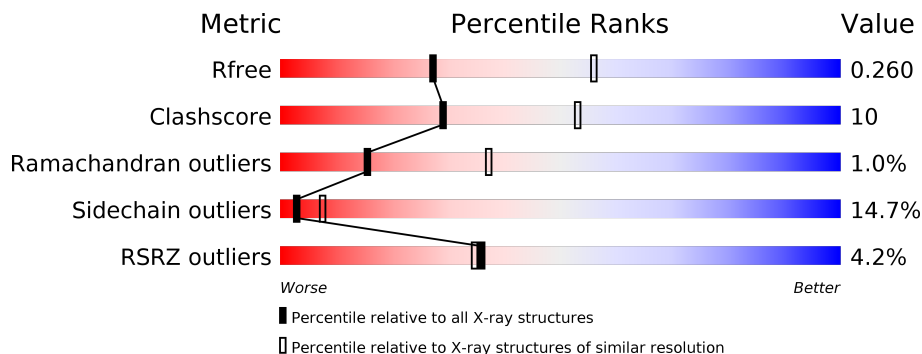
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

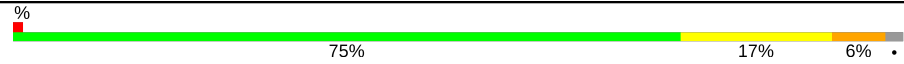


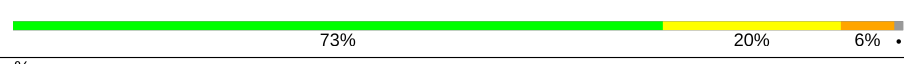
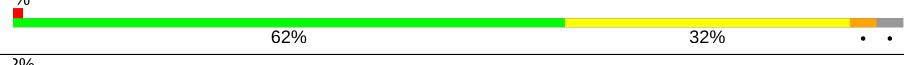

The reported resolution of this entry is 2.70 Å.

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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 $\leq 5\%$. 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	109	
1	B	109	
1	C	109	
1	D	109	
1	E	109	
1	F	109	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	109	<p>74% 23% ..</p>
1	H	109	<p>69% 26% 5% ..</p>
1	I	109	<p>72% 22% ..</p>
1	J	109	<p>79% 17% ..</p>
1	K	109	<p>60% 31% 12% ..</p>
1	L	109	<p>69% 25% 20% ..</p>

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 10068 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 ORF134.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	107	841	534	146	158	3	0	0	0
1	B	107	832	529	144	156	3	0	0	0
1	C	105	835	531	144	157	3	0	0	0
1	D	108	861	546	151	161	3	0	0	0
1	E	106	856	544	148	161	3	0	0	0
1	F	106	850	541	151	155	3	0	0	0
1	G	108	849	541	148	157	3	0	0	0
1	H	108	875	556	153	163	3	0	0	0
1	I	107	845	537	146	159	3	0	0	0
1	J	106	810	518	141	148	3	0	0	0
1	K	105	780	499	138	140	3	0	0	0
1	L	107	792	503	143	143	3	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
A	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
A	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
A	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
B	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
B	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
C	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
C	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
C	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
D	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
D	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
D	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
E	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
E	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
E	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
F	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
F	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
F	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
G	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
G	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
G	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
G	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
H	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
H	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
H	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
H	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
I	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
I	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
I	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
I	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
J	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
J	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
J	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
J	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
K	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
K	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
K	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177
K	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177
L	1	MSE	MET	MODIFIED RESIDUE	UNP Q44177
L	60	MSE	MET	MODIFIED RESIDUE	UNP Q44177
L	61	MSE	MET	MODIFIED RESIDUE	UNP Q44177

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	89	MSE	MET	MODIFIED RESIDUE	UNP Q44177

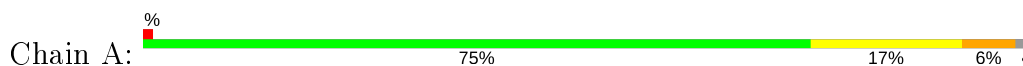
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	6	Total O 6 6	0	0
2	B	4	Total O 4 4	0	0
2	C	5	Total O 5 5	0	0
2	E	3	Total O 3 3	0	0
2	F	4	Total O 4 4	0	0
2	G	6	Total O 6 6	0	0
2	H	1	Total O 1 1	0	0
2	I	5	Total O 5 5	0	0
2	J	5	Total O 5 5	0	0
2	K	1	Total O 1 1	0	0
2	L	2	Total O 2 2	0	0

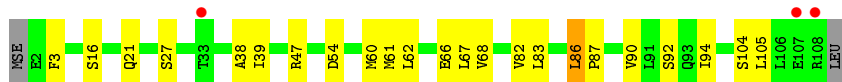
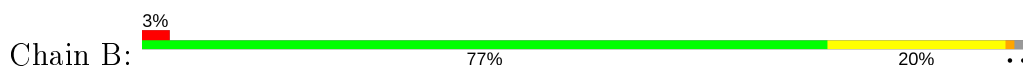
3 Residue-property plots [i](#)

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

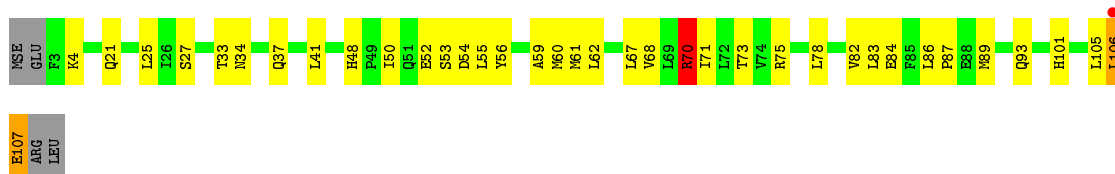
- Molecule 1: ORF134



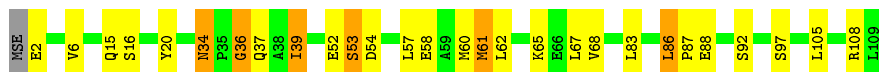
- Molecule 1: ORF134



- Molecule 1: ORF134

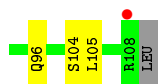


- Molecule 1: ORF134



- Molecule 1: ORF134

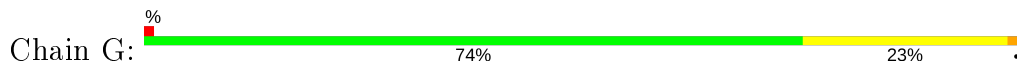




• Molecule 1: ORF134



• Molecule 1: ORF134



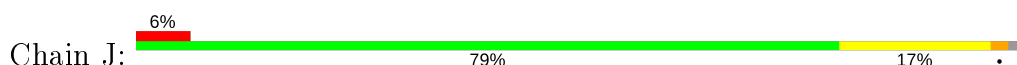
• Molecule 1: ORF134



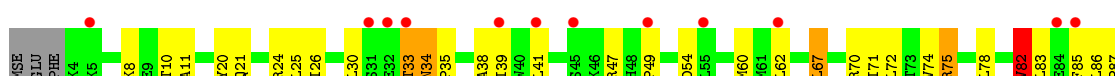
• Molecule 1: ORF134



• Molecule 1: ORF134

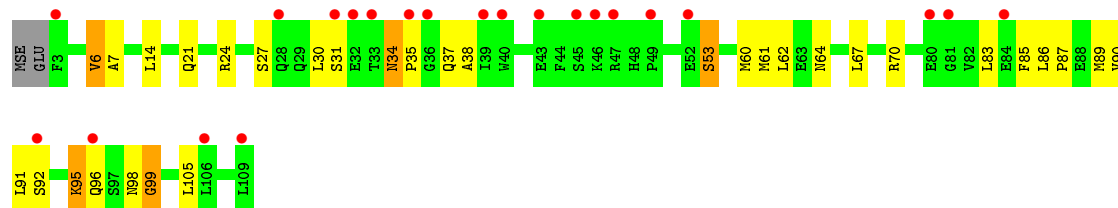


• Molecule 1: ORF134



● Molecule 1: ORF134

Chain L:  20% 69% 25% 5%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	70.36Å 71.10Å 113.16Å 89.58° 81.10° 80.09°	Depositor
Resolution (Å)	20.00 – 2.70 19.96 – 2.70	Depositor EDS
% Data completeness (in resolution range)	96.9 (20.00-2.70) 96.9 (19.96-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.56 (at 2.71Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.212 , 0.262 0.214 , 0.260	Depositor DCC
R_{free} test set	2881 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	38.0	Xtrriage
Anisotropy	0.048	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 40.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	10068	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.45% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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

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	0.86	1/850 (0.1%)	0.90	1/1146 (0.1%)
1	B	0.78	0/841	0.81	1/1134 (0.1%)
1	C	0.81	0/844	0.91	4/1136 (0.4%)
1	D	0.83	0/870	0.86	1/1170 (0.1%)
1	E	0.91	0/865	0.87	1/1161 (0.1%)
1	F	0.91	0/859	0.95	2/1153 (0.2%)
1	G	0.88	0/858	0.90	2/1156 (0.2%)
1	H	0.84	1/884 (0.1%)	0.88	1/1186 (0.1%)
1	I	0.85	0/854	0.92	3/1150 (0.3%)
1	J	0.73	0/819	0.81	1/1107 (0.1%)
1	K	0.74	0/787	0.88	1/1065 (0.1%)
1	L	0.81	0/800	0.84	0/1083
All	All	0.83	2/10131 (0.0%)	0.88	18/13647 (0.1%)

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	D	0	2
1	G	0	1
1	H	0	1
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	66	GLU	CD-OE2	5.34	1.31	1.25
1	A	60	MSE	SE-CE	-5.22	1.64	1.95

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	70	ARG	NE-CZ-NH1	8.23	124.42	120.30
1	F	54	ASP	CB-CG-OD2	7.61	125.15	118.30
1	J	54	ASP	CB-CG-OD2	7.43	124.98	118.30
1	E	54	ASP	CB-CG-OD2	7.42	124.98	118.30
1	A	47	ARG	NE-CZ-NH1	6.70	123.65	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	2	GLU	Peptide
1	D	36	GLY	Peptide
1	G	2	GLU	Peptide
1	H	32	GLU	Peptide

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	841	0	838	23	0
1	B	832	0	825	16	0
1	C	835	0	843	27	0
1	D	861	0	868	16	0
1	E	856	0	877	26	0
1	F	850	0	876	25	0
1	G	849	0	856	13	0
1	H	875	0	901	14	0
1	I	845	0	847	18	0
1	J	810	0	803	13	0
1	K	780	0	765	27	0
1	L	792	0	760	22	0
2	A	6	0	0	0	0
2	B	4	0	0	1	0
2	C	5	0	0	2	0
2	E	3	0	0	0	0
2	F	4	0	0	0	0
2	G	6	0	0	0	0
2	H	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	I	5	0	0	1	0
2	J	5	0	0	1	0
2	K	1	0	0	1	0
2	L	2	0	0	0	0
All	All	10068	0	10059	192	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 192 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:E:89:MSE:SE	1:E:89:MSE:CE	2.16	1.44
1:C:89:MSE:CE	1:C:89:MSE:SE	2.14	1.44
1:F:61:MSE:CE	1:F:61:MSE:SE	2.16	1.42
1:L:61:MSE:CE	1:L:61:MSE:SE	2.18	1.41
1:H:61:MSE:SE	1:H:61:MSE:CE	2.19	1.40

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	105/109 (96%)	102 (97%)	3 (3%)	0	100	100
1	B	105/109 (96%)	95 (90%)	10 (10%)	0	100	100
1	C	103/109 (94%)	98 (95%)	4 (4%)	1 (1%)	15	37
1	D	106/109 (97%)	102 (96%)	4 (4%)	0	100	100
1	E	104/109 (95%)	100 (96%)	4 (4%)	0	100	100
1	F	104/109 (95%)	100 (96%)	4 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	106/109 (97%)	103 (97%)	3 (3%)	0	100	100
1	H	106/109 (97%)	103 (97%)	3 (3%)	0	100	100
1	I	105/109 (96%)	97 (92%)	7 (7%)	1 (1%)	15	37
1	J	104/109 (95%)	95 (91%)	8 (8%)	1 (1%)	15	37
1	K	103/109 (94%)	82 (80%)	15 (15%)	6 (6%)	1	2
1	L	105/109 (96%)	87 (83%)	14 (13%)	4 (4%)	3	7
All	All	1256/1308 (96%)	1164 (93%)	79 (6%)	13 (1%)	15	37

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	84	GLU
1	I	84	GLU
1	K	33	THR
1	L	53	SER
1	L	99	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	90/95 (95%)	79 (88%)	11 (12%)	5	11
1	B	88/95 (93%)	79 (90%)	9 (10%)	7	17
1	C	91/95 (96%)	75 (82%)	16 (18%)	2	4
1	D	93/95 (98%)	79 (85%)	14 (15%)	3	7
1	E	95/95 (100%)	78 (82%)	17 (18%)	2	4
1	F	93/95 (98%)	80 (86%)	13 (14%)	3	8
1	G	91/95 (96%)	76 (84%)	15 (16%)	2	5
1	H	97/95 (102%)	78 (80%)	19 (20%)	1	3
1	I	91/95 (96%)	78 (86%)	13 (14%)	3	8
1	J	84/95 (88%)	76 (90%)	8 (10%)	8	20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	76/95 (80%)	64 (84%)	12 (16%)	2	6
1	L	76/95 (80%)	66 (87%)	10 (13%)	4	9
All	All	1065/1140 (93%)	908 (85%)	157 (15%)	3	7

5 of 157 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	62	LEU
1	G	58	GLU
1	K	93	GLN
1	F	67	LEU
1	F	108	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	34	ASN
1	F	48	HIS
1	K	15	GLN
1	E	37	GLN
1	F	15	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(Å ²)	Q < 0.9
1	A	104/109 (95%)	-0.31	1 (0%) 82 83	20, 29, 48, 65	0
1	B	104/109 (95%)	-0.05	3 (2%) 51 52	19, 39, 61, 64	0
1	C	102/109 (93%)	-0.25	1 (0%) 82 83	19, 35, 53, 66	0
1	D	105/109 (96%)	-0.15	0 100 100	25, 37, 49, 57	0
1	E	103/109 (94%)	-0.17	1 (0%) 82 83	17, 33, 46, 57	0
1	F	103/109 (94%)	-0.23	2 (1%) 66 69	19, 34, 46, 68	0
1	G	105/109 (96%)	-0.25	1 (0%) 82 83	21, 34, 47, 56	0
1	H	105/109 (96%)	-0.17	0 100 100	19, 34, 51, 62	0
1	I	104/109 (95%)	-0.23	2 (1%) 66 69	20, 32, 55, 67	0
1	J	103/109 (94%)	0.07	6 (5%) 23 22	28, 43, 65, 67	0
1	K	102/109 (93%)	0.77	13 (12%) 3 3	38, 64, 75, 78	0
1	L	104/109 (95%)	0.87	22 (21%) 0 0	38, 64, 76, 78	0
All	All	1244/1308 (95%)	-0.01	52 (4%) 36 35	17, 37, 69, 78	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	109	LEU	5.4
1	E	108	ARG	4.7
1	L	32	GLU	4.7
1	L	33	THR	4.4
1	F	108	ARG	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 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.