wwPDB X-ray Structure Validation Summary Report (i)

Dec 10, 2023 – 04:16 am GMT

PDB ID	:	2BJY
Title	:	The X-ray crystal structure of Listeria innocua Dps H31G-H43G mutant.
Authors	:	Ilari, A.; Stefanini, S.; Chiancone, E.
Deposited on	:	2005-02-08
Resolution	:	2.60 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 (1)) were used in the production of this report:

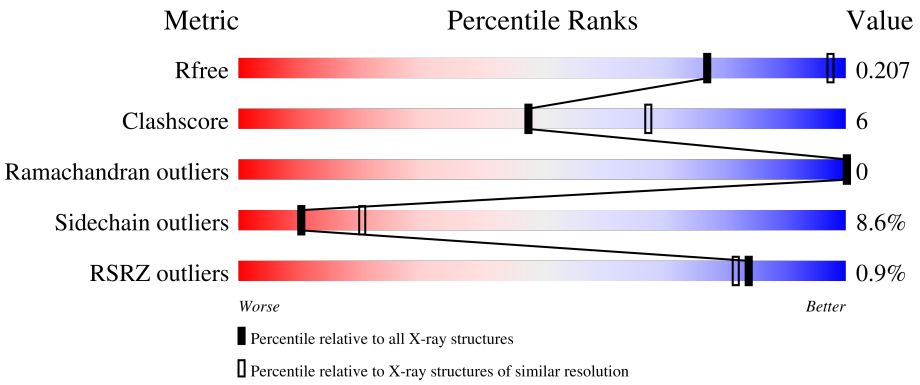
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.60 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	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 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	156	3% 79% 13%
1	B	156	% 74% 19%
1	C	156	80% 12%
1	D	156	2% 81% 12%
1	E	156	% 81% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	156	77% 14%
1	G	156	81% 12%
1	H	156	72% 18% 5%
1	I	156	2% 76% 16%
1	J	156	72% 19%
1	K	156	% 74% 18%
1	L	156	% 79% 13%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 14696 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 NON-HEME IRON-CONTAINING FERRITIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	150	1210	776	192	235	7	0	0	0
1	B	151	1216	779	193	237	7	0	0	0
1	C	150	1210	776	192	235	7	0	0	0
1	D	150	1214	778	192	237	7	0	1	0
1	E	150	1210	776	192	235	7	0	0	0
1	F	150	1210	776	192	235	7	0	0	0
1	G	150	1210	776	192	235	7	0	0	0
1	H	149	1207	773	191	236	7	0	1	0
1	I	150	1210	776	192	235	7	0	0	0
1	J	150	1210	776	192	235	7	0	0	0
1	K	149	1203	771	191	234	7	0	0	0
1	L	152	1224	783	195	239	7	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	31	GLY	HIS	engineered mutation	UNP P80725
A	43	GLY	HIS	engineered mutation	UNP P80725
B	31	GLY	HIS	engineered mutation	UNP P80725
B	43	GLY	HIS	engineered mutation	UNP P80725
C	31	GLY	HIS	engineered mutation	UNP P80725

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	43	GLY	HIS	engineered mutation	UNP P80725
D	31	GLY	HIS	engineered mutation	UNP P80725
D	43	GLY	HIS	engineered mutation	UNP P80725
E	31	GLY	HIS	engineered mutation	UNP P80725
E	43	GLY	HIS	engineered mutation	UNP P80725
F	31	GLY	HIS	engineered mutation	UNP P80725
F	43	GLY	HIS	engineered mutation	UNP P80725
G	31	GLY	HIS	engineered mutation	UNP P80725
G	43	GLY	HIS	engineered mutation	UNP P80725
H	31	GLY	HIS	engineered mutation	UNP P80725
H	43	GLY	HIS	engineered mutation	UNP P80725
I	31	GLY	HIS	engineered mutation	UNP P80725
I	43	GLY	HIS	engineered mutation	UNP P80725
J	31	GLY	HIS	engineered mutation	UNP P80725
J	43	GLY	HIS	engineered mutation	UNP P80725
K	31	GLY	HIS	engineered mutation	UNP P80725
K	43	GLY	HIS	engineered mutation	UNP P80725
L	31	GLY	HIS	engineered mutation	UNP P80725
L	43	GLY	HIS	engineered mutation	UNP P80725

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	7	7	7	0	0
			Total	O		
2	B	14	14	14	0	0
			Total	O		
2	C	13	13	13	0	0
			Total	O		
2	D	17	17	17	0	0
			Total	O		
2	E	16	16	16	0	0
			Total	O		
2	F	8	8	8	0	0
			Total	O		
2	G	11	11	11	0	0
			Total	O		
2	H	25	25	25	0	0
			Total	O		
2	I	11	11	11	0	0
			Total	O		
2	J	17	17	17	0	0

Continued on next page...

Continued from previous page...

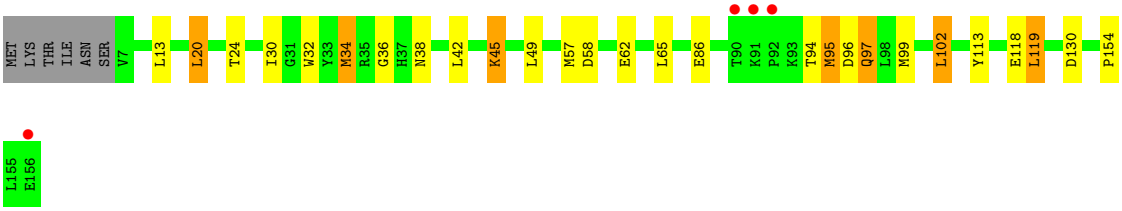
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	K	9	9	9	0	0
			Total	O		
2	L	14	14	14	0	0

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.

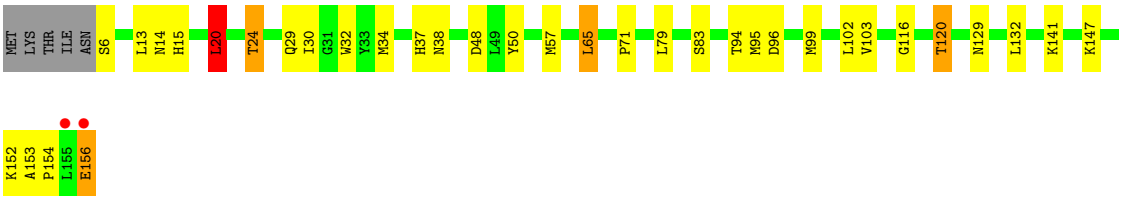
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain A:



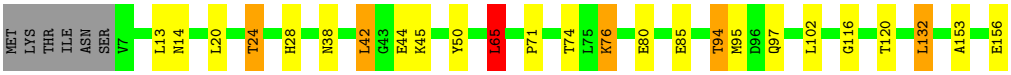
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain B:



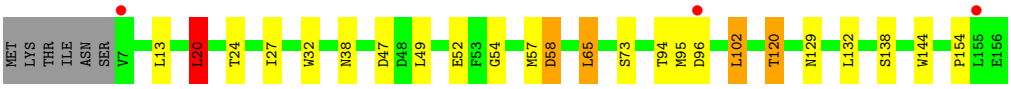
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain C:



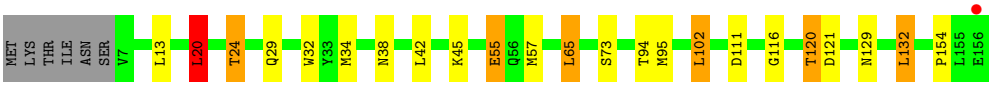
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain D:



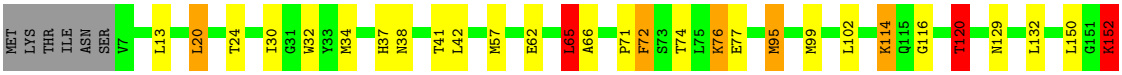
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain E:



• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain F:

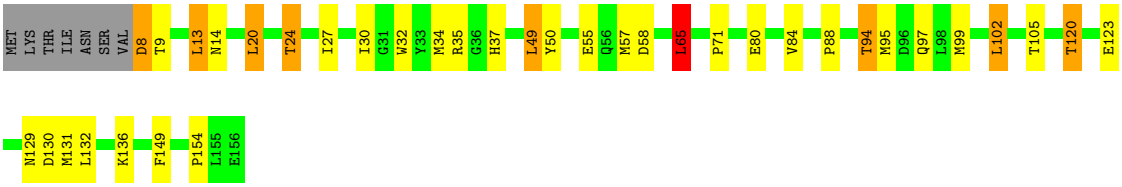


• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain G:

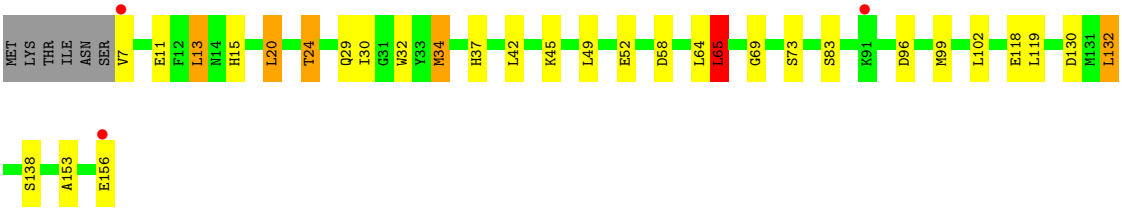
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain H:



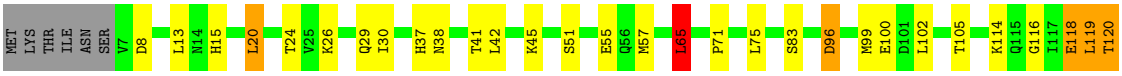
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

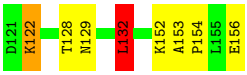
Chain I:



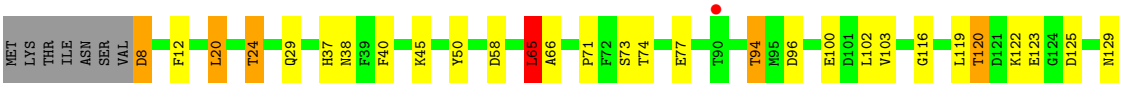
• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN

Chain J:

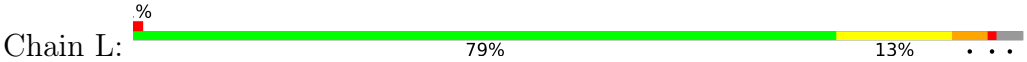


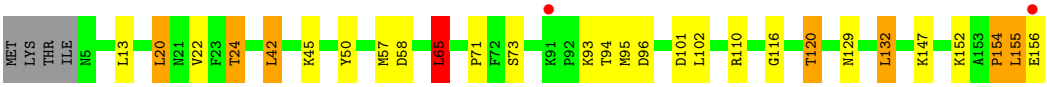


• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN



• Molecule 1: NON-HEME IRON-CONTAINING FERRITIN





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.62Å 132.21Å 168.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.60 19.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	93.7 (30.00-2.60) 93.9 (19.99-2.60)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.91 (at 2.59Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, R_{free}	0.207 , 0.252 0.208 , 0.207	Depositor DCC
R_{free} test set	2863 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	21.0	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 26.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	14696	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% 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 (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	0.93	0/1235	1.08	3/1664 (0.2%)
1	B	0.98	2/1241 (0.2%)	1.11	5/1672 (0.3%)
1	C	0.94	1/1235 (0.1%)	1.04	3/1664 (0.2%)
1	D	0.97	0/1243	1.03	3/1675 (0.2%)
1	E	0.95	1/1235 (0.1%)	1.06	6/1664 (0.4%)
1	F	0.96	4/1235 (0.3%)	1.08	3/1664 (0.2%)
1	G	0.91	0/1235	1.05	5/1664 (0.3%)
1	H	0.93	1/1236 (0.1%)	1.14	7/1665 (0.4%)
1	I	0.93	0/1235	1.14	8/1664 (0.5%)
1	J	0.91	0/1235	1.06	4/1664 (0.2%)
1	K	0.97	1/1228 (0.1%)	1.09	5/1654 (0.3%)
1	L	0.92	1/1249 (0.1%)	1.07	7/1683 (0.4%)
All	All	0.94	11/14842 (0.1%)	1.08	59/19997 (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	K	1	0

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	55	GLU	CG-CD	6.96	1.62	1.51
1	L	154	PRO	N-CD	-6.02	1.39	1.47
1	F	120	THR	CB-CG2	-5.71	1.33	1.52
1	B	156	GLU	CG-CD	5.59	1.60	1.51
1	F	156	GLU	CB-CG	5.52	1.62	1.52

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	I	119	LEU	CA-CB-CG	8.38	134.57	115.30
1	A	119	LEU	CA-CB-CG	8.01	133.71	115.30
1	A	34	MET	CG-SD-CE	-7.97	87.45	100.20
1	L	155	LEU	O-C-N	-7.44	110.80	122.70
1	B	96	ASP	CB-CG-OD1	7.38	124.95	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	K	8	ASP	CA

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	1210	0	1178	16	0
1	B	1216	0	1183	29	0
1	C	1210	0	1178	12	0
1	D	1214	0	1178	13	0
1	E	1210	0	1178	11	0
1	F	1210	0	1178	23	0
1	G	1210	0	1178	11	0
1	H	1207	0	1169	23	0
1	I	1210	0	1178	16	0
1	J	1210	0	1178	20	0
1	K	1203	0	1169	21	0
1	L	1224	0	1189	17	0
2	A	7	0	0	0	0
2	B	14	0	0	1	0
2	C	13	0	0	0	0
2	D	17	0	0	0	0
2	E	16	0	0	0	0
2	F	8	0	0	0	0
2	G	11	0	0	0	0
2	H	25	0	0	0	0
2	I	11	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	J	17	0	0	0	0
2	K	9	0	0	0	0
2	L	14	0	0	0	0
All	All	14696	0	14134	174	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 6.

The worst 5 of 174 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:D:54:GLY:O	1:D:58[A]:ASP:OD2	1.91	0.87
1:I:153:ALA:HB3	1:I:156:GLU:HG2	1.63	0.80
1:J:38:ASN:HB3	1:J:42:LEU:HD23	1.69	0.73
1:L:147:LYS:NZ	1:L:156:GLU:C	2.41	0.73
1:G:116:GLY:O	1:G:120:THR:HB	1.92	0.68

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	148/156 (95%)	147 (99%)	1 (1%)	0	100	100
1	B	149/156 (96%)	148 (99%)	1 (1%)	0	100	100
1	C	148/156 (95%)	145 (98%)	3 (2%)	0	100	100
1	D	149/156 (96%)	147 (99%)	2 (1%)	0	100	100
1	E	148/156 (95%)	145 (98%)	3 (2%)	0	100	100
1	F	148/156 (95%)	146 (99%)	2 (1%)	0	100	100
1	G	148/156 (95%)	146 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	148/156 (95%)	146 (99%)	2 (1%)	0	100	100
1	I	148/156 (95%)	147 (99%)	1 (1%)	0	100	100
1	J	148/156 (95%)	147 (99%)	1 (1%)	0	100	100
1	K	147/156 (94%)	145 (99%)	2 (1%)	0	100	100
1	L	150/156 (96%)	148 (99%)	2 (1%)	0	100	100
All	All	1779/1872 (95%)	1757 (99%)	22 (1%)	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	131/137 (96%)	120 (92%)	11 (8%)	11	21
1	B	132/137 (96%)	122 (92%)	10 (8%)	13	26
1	C	131/137 (96%)	121 (92%)	10 (8%)	13	26
1	D	132/137 (96%)	121 (92%)	11 (8%)	11	22
1	E	131/137 (96%)	120 (92%)	11 (8%)	11	21
1	F	131/137 (96%)	118 (90%)	13 (10%)	8	15
1	G	131/137 (96%)	120 (92%)	11 (8%)	11	21
1	H	131/137 (96%)	117 (89%)	14 (11%)	6	12
1	I	131/137 (96%)	119 (91%)	12 (9%)	9	17
1	J	131/137 (96%)	117 (89%)	14 (11%)	6	12
1	K	130/137 (95%)	120 (92%)	10 (8%)	13	25
1	L	133/137 (97%)	123 (92%)	10 (8%)	13	27
All	All	1575/1644 (96%)	1438 (91%)	137 (9%)	10	20

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	132	LEU
1	K	45	LYS
1	L	42	LEU
1	E	94	THR
1	E	65	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	38	ASN
1	K	15	HIS
1	I	15	HIS
1	J	15	HIS
1	L	15	HIS

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)

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	150/156 (96%)	-0.31	4 (2%)	54	48	10, 18, 33, 41	0
1	B	151/156 (96%)	-0.27	2 (1%)	77	73	12, 18, 33, 42	0
1	C	150/156 (96%)	-0.29	0	100	100	12, 18, 33, 39	0
1	D	150/156 (96%)	-0.27	3 (2%)	65	60	12, 18, 33, 39	0
1	E	150/156 (96%)	-0.31	1 (0%)	87	86	12, 18, 32, 41	0
1	F	150/156 (96%)	-0.31	1 (0%)	87	86	11, 18, 32, 43	0
1	G	150/156 (96%)	-0.38	0	100	100	10, 18, 32, 40	0
1	H	149/156 (95%)	-0.37	0	100	100	10, 18, 31, 40	0
1	I	150/156 (96%)	-0.22	3 (2%)	65	60	10, 18, 32, 42	0
1	J	150/156 (96%)	-0.34	0	100	100	10, 18, 32, 40	0
1	K	149/156 (95%)	-0.33	1 (0%)	87	86	11, 18, 31, 40	0
1	L	152/156 (97%)	-0.30	2 (1%)	77	73	11, 18, 34, 43	0
All	All	1801/1872 (96%)	-0.31	17 (0%)	84	82	10, 18, 33, 43	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	7	VAL	3.7
1	F	156	GLU	3.6
1	B	156	GLU	3.4
1	E	156	GLU	3.4
1	A	156	GLU	3.2

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.