



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 08:43 pm BST

PDB ID : 1E50
Title : AML1/CBFbeta complex
Authors : Warren, A.J.; Bravo, J.; Williams, R.L.; Rabbits, T.H.
Deposited on : 2000-07-13
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 ⓘ symbol.

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

MolProbity : 4.02b-467
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

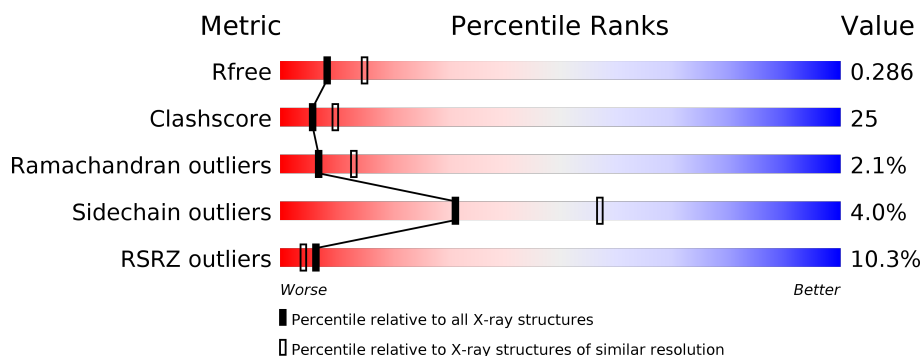
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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 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	134	
1	C	134	
1	E	134	
1	G	134	
1	Q	134	
1	R	134	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	134	<div><div></div><div>7%</div><div>57%</div><div>37%</div><div></div><div></div></div>
2	D	134	<div><div></div><div>12%</div><div>53%</div><div>36%</div><div></div><div>7%</div></div>
2	F	134	<div><div></div><div>12%</div><div>46%</div><div>40%</div><div>7%</div><div>7%</div></div>
2	H	134	<div><div></div><div>9%</div><div>55%</div><div>37%</div><div></div><div></div></div>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9696 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 CORE-BINDING FACTOR ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	115	Total	C	N	O	S	0	0	0
			872	551	155	162	4			
1	C	125	Total	C	N	O	S	0	0	0
			945	593	174	174	4			
1	E	119	Total	C	N	O	S	0	0	0
			908	571	165	168	4			
1	G	119	Total	C	N	O	S	0	0	0
			908	571	165	168	4			
1	Q	110	Total	C	N	O	S	0	0	0
			847	535	153	155	4			
1	R	110	Total	C	N	O	S	0	0	0
			847	535	153	155	4			

- Molecule 2 is a protein called CORE-BINDING FACTOR CBF-BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	130	Total	C	N	O	S	0	0	0
			1070	672	194	198	6			
2	D	125	Total	C	N	O	S	0	0	0
			1036	649	188	193	6			
2	F	125	Total	C	N	O	S	0	0	0
			1036	649	188	193	6			
2	H	130	Total	C	N	O	S	0	0	0
			1070	672	194	198	6			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	23	Total	O	0	0
			23	23		
3	B	8	Total	O	0	0
			8	8		

Continued on next page...

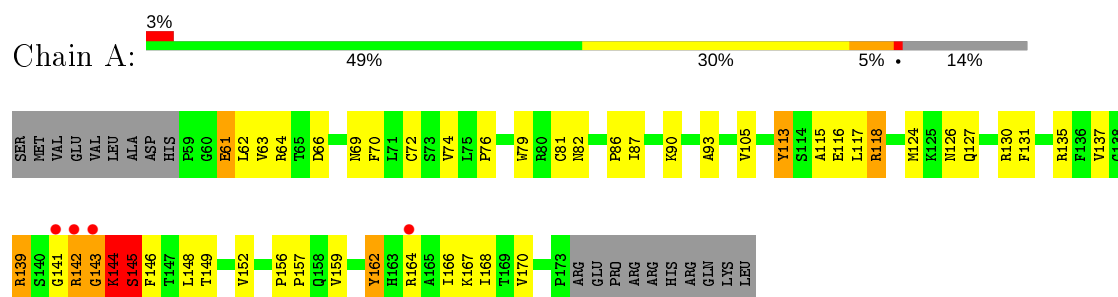
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	25	Total 25	O 25	0	0
3	D	24	Total 24	O 24	0	0
3	E	21	Total 21	O 21	0	0
3	F	19	Total 19	O 19	0	0
3	G	21	Total 21	O 21	0	0
3	H	14	Total 14	O 14	0	0
3	Q	1	Total 1	O 1	0	0
3	R	1	Total 1	O 1	0	0

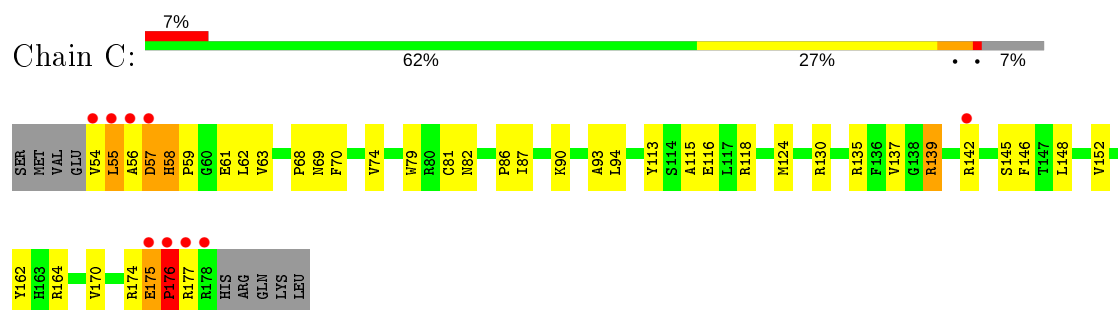
3 Residue-property plots

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.

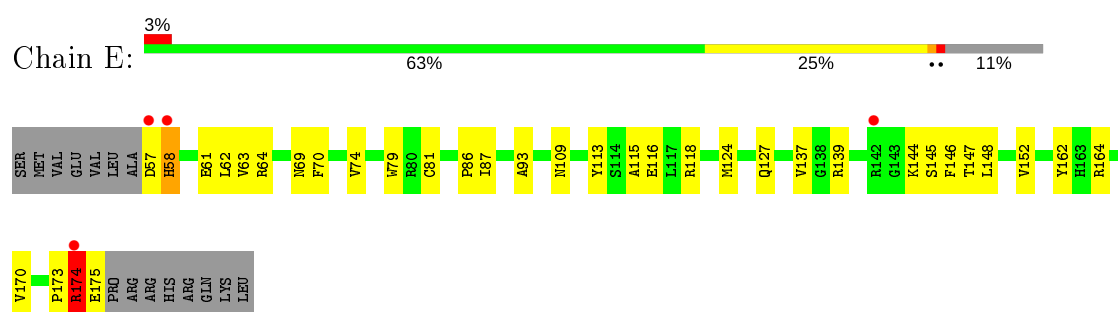
• Molecule 1: CORE-BINDING FACTOR ALPHA SUBUNIT



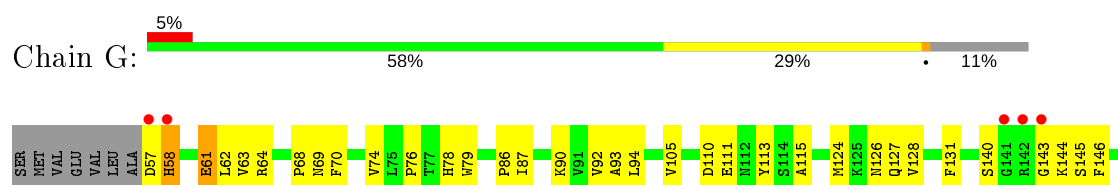
• Molecule 1: CORE-BINDING FACTOR ALPHA SUBUNIT

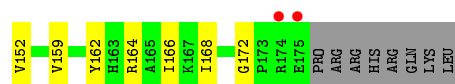


• Molecule 1: CORE-BINDING FACTOR ALPHA SUBUNIT



• Molecule 1: CORE-BINDING FACTOR ALPHA SUBUNIT

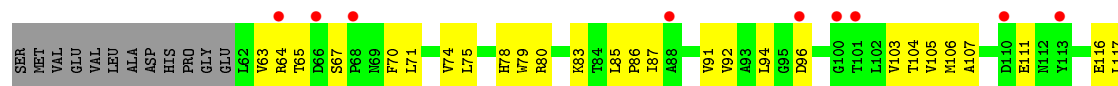




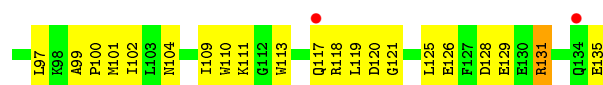
• Molecule 1: CORE-BINDING FACTOR ALPHA SUBUNIT



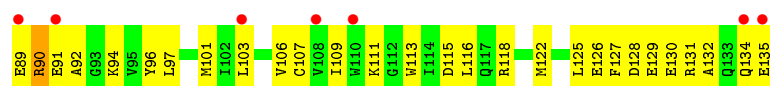
• Molecule 1: CORE-BINDING FACTOR ALPHA SUBUNIT



• Molecule 2: CORE-BINDING FACTOR CBF-BETA



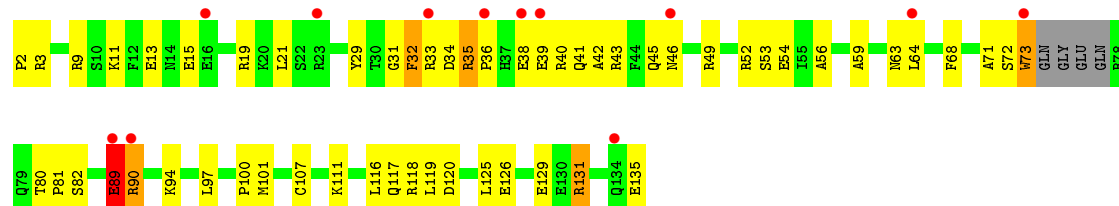
• Molecule 2: CORE-BINDING FACTOR CBF-BETA



• Molecule 2: CORE-BINDING FACTOR CBF-BETA



- Molecule 2: CORE-BINDING FACTOR CBF-BETA



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	103.26Å 79.38Å 130.10Å 90.00° 101.39° 90.00°	Depositor
Resolution (Å)	23.82 – 2.60 23.82 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.3 (23.82-2.60) 99.4 (23.82-2.60)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.65 (at 2.60Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.268 , 0.287 0.259 , 0.286	Depositor DCC
R_{free} test set	1905 reflections (3.01%)	wwPDB-VP
Wilson B-factor (Å ²)	51.4	Xtriage
Anisotropy	0.320	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 54.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	9696	wwPDB-VP
Average B, all atoms (Å ²)	55.0	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 20.76 % 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 8.1096e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

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	0.97	4/891 (0.4%)	0.84	4/1216 (0.3%)
1	C	0.44	0/965	0.74	0/1316
1	E	0.70	4/927 (0.4%)	1.32	6/1264 (0.5%)
1	G	0.42	0/927	0.71	0/1264
1	Q	0.32	0/864	0.56	0/1178
1	R	0.32	0/864	0.56	0/1178
2	B	0.44	0/1093	0.63	0/1467
2	D	0.44	0/1057	0.65	0/1416
2	F	0.46	0/1057	0.68	0/1416
2	H	0.71	3/1093 (0.3%)	0.80	4/1467 (0.3%)
All	All	0.56	11/9738 (0.1%)	0.78	14/13182 (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	A	0	2
1	E	0	2
2	F	0	1
2	H	0	1
All	All	0	6

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	144	LYS	CA-C	14.18	1.89	1.52
1	A	144	LYS	C-O	-12.96	0.98	1.23
1	A	145	SER	N-CA	12.54	1.71	1.46
1	E	174	ARG	CA-C	12.06	1.84	1.52
2	H	89	GLU	C-O	-9.37	1.05	1.23

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	174	ARG	O-C-N	-27.34	78.95	122.70
1	E	174	ARG	CA-C-N	20.83	163.03	117.20
1	E	174	ARG	CA-C-O	-13.46	91.84	120.10
1	E	175	GLU	CA-CB-CG	8.25	131.54	113.40
2	H	89	GLU	CA-C-N	7.62	133.96	117.20

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	144	LYS	Mainchain,Peptide
1	E	174	ARG	Mainchain,Peptide
2	F	129	GLU	Mainchain
2	H	35	ARG	Mainchain

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	872	0	867	65	0
1	C	945	0	930	52	0
1	E	908	0	900	41	0
1	G	908	0	900	46	0
1	Q	847	0	851	35	0
1	R	847	0	851	45	0
2	B	1070	0	1022	62	0
2	D	1036	0	1001	70	0
2	F	1036	0	1001	70	0
2	H	1070	0	1022	57	0
3	A	23	0	0	3	0
3	B	8	0	0	1	0
3	C	25	0	0	2	0
3	D	24	0	0	5	0
3	E	21	0	0	2	0
3	F	19	0	0	5	0
3	G	21	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	14	0	0	3	0
3	Q	1	0	0	0	0
3	R	1	0	0	0	0
All	All	9696	0	9345	477	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 25.

The worst 5 of 477 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:D:83:ARG:CB	2:D:83:ARG:HH11	1.13	1.56
1:A:145:SER:N	1:A:145:SER:CA	1.71	1.54
2:H:89:GLU:C	2:H:89:GLU:CA	1.74	1.52
1:E:174:ARG:C	1:E:174:ARG:CA	1.84	1.45
2:D:83:ARG:NH1	2:D:83:ARG:HB2	1.16	1.43

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	113/134 (84%)	104 (92%)	6 (5%)	3 (3%)	5 8
1	C	123/134 (92%)	109 (89%)	8 (6%)	6 (5%)	2 2
1	E	117/134 (87%)	110 (94%)	5 (4%)	2 (2%)	9 18
1	G	117/134 (87%)	110 (94%)	6 (5%)	1 (1%)	17 35
1	Q	108/134 (81%)	97 (90%)	9 (8%)	2 (2%)	8 15
1	R	108/134 (81%)	89 (82%)	16 (15%)	3 (3%)	5 7
2	B	126/134 (94%)	118 (94%)	6 (5%)	2 (2%)	9 19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	121/134 (90%)	112 (93%)	7 (6%)	2 (2%)	9	18
2	F	121/134 (90%)	107 (88%)	12 (10%)	2 (2%)	9	18
2	H	126/134 (94%)	115 (91%)	9 (7%)	2 (2%)	9	19
All	All	1180/1340 (88%)	1071 (91%)	84 (7%)	25 (2%)	7	13

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	142	ARG
1	A	145	SER
2	B	90	ARG
1	C	58	HIS
1	C	175	GLU

5.3.2 Protein sidechains ⓘ

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	95/115 (83%)	90 (95%)	5 (5%)	22	45
1	C	100/115 (87%)	97 (97%)	3 (3%)	41	67
1	E	98/115 (85%)	97 (99%)	1 (1%)	76	90
1	G	98/115 (85%)	96 (98%)	2 (2%)	55	78
1	Q	93/115 (81%)	92 (99%)	1 (1%)	73	88
1	R	93/115 (81%)	91 (98%)	2 (2%)	52	76
2	B	111/117 (95%)	106 (96%)	5 (4%)	27	52
2	D	110/117 (94%)	103 (94%)	7 (6%)	17	35
2	F	110/117 (94%)	101 (92%)	9 (8%)	11	22
2	H	111/117 (95%)	105 (95%)	6 (5%)	22	44
All	All	1019/1158 (88%)	978 (96%)	41 (4%)	31	57

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	106	VAL
2	F	21	LEU
2	H	131	ARG
2	D	125	LEU
1	E	81	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	14	ASN
2	D	133	GLN
1	Q	78	HIS
2	B	134	GLN
1	G	155	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 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 ⓘ

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, 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	115/134 (85%)	0.22	4 (3%)	44	36	24, 39, 72, 83	0
1	C	125/134 (93%)	0.43	9 (7%)	15	11	25, 38, 86, 97	0
1	E	119/134 (88%)	0.24	4 (3%)	45	38	24, 38, 67, 89	0
1	G	119/134 (88%)	0.35	7 (5%)	22	17	24, 41, 86, 112	0
1	Q	110/134 (82%)	1.08	23 (20%)	1	0	58, 84, 100, 104	0
1	R	110/134 (82%)	1.21	24 (21%)	0	0	61, 85, 100, 105	0
2	B	130/134 (97%)	0.60	10 (7%)	13	10	26, 53, 81, 90	0
2	D	125/134 (93%)	0.63	16 (12%)	3	2	25, 51, 75, 89	0
2	F	125/134 (93%)	0.84	16 (12%)	3	2	25, 51, 78, 98	0
2	H	130/134 (97%)	0.64	12 (9%)	9	6	25, 53, 81, 88	0
All	All	1208/1340 (90%)	0.62	125 (10%)	6	4	24, 51, 93, 112	0

The worst 5 of 125 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	57	ASP	7.4
1	R	113	TYR	7.1
1	G	142	ARG	6.9
1	Q	110	ASP	6.5
1	C	56	ALA	6.4

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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.