



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:34 PM GMT

PDB ID : 1PGR
Title : 2:2 COMPLEX OF G-CSF WITH ITS RECEPTOR
Authors : Aritomi, M.; Kunishima, N.; Okamoto, T.; Kuroki, R.; Ota, Y.; Morikawa, K.
Deposited on : 1999-03-08
Resolution : 3.50 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

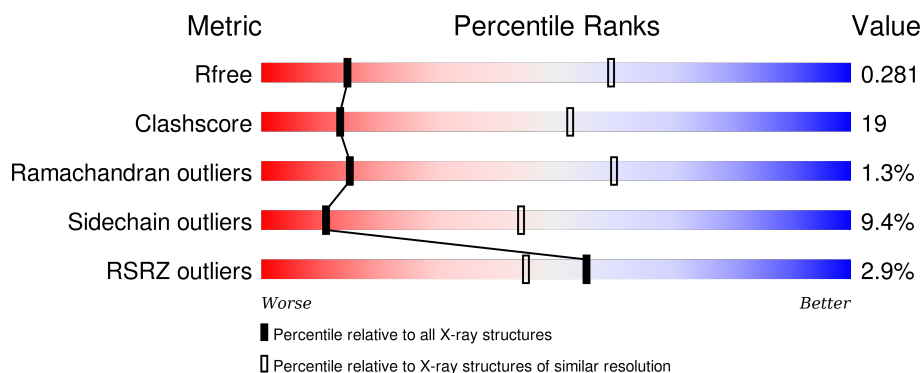
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION




The reported resolution of this entry is 3.50 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	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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 $\leq 5\%$. 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	175	 57% 34% 8%
1	C	175	 59% 33% 6%
1	E	175	 56% 36% 7%
1	G	175	 62% 30% 6%
2	B	215	 60% 33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	215	<div><div></div><div>7%</div><div>46%</div><div>40%</div><div>6%</div><div>7%</div></div>
2	F	215	<div><div></div><div>4%</div><div>59%</div><div>35%</div><div></div><div>•</div><div>•</div></div>
2	H	215	<div><div></div><div>5%</div><div>46%</div><div>40%</div><div>7%</div><div>7%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11424 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 PROTEIN (GRANULOCYTE COLONY-STIMULATING FACTOR).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	161	Total	C	N	O	S	0	0	0
			1219	781	206	224	8			
1	C	165	Total	C	N	O	S	0	0	0
			1248	799	211	230	8			
1	E	163	Total	C	N	O	S	0	0	0
			1230	788	208	226	8			
1	G	165	Total	C	N	O	S	0	0	0
			1248	799	211	230	8			

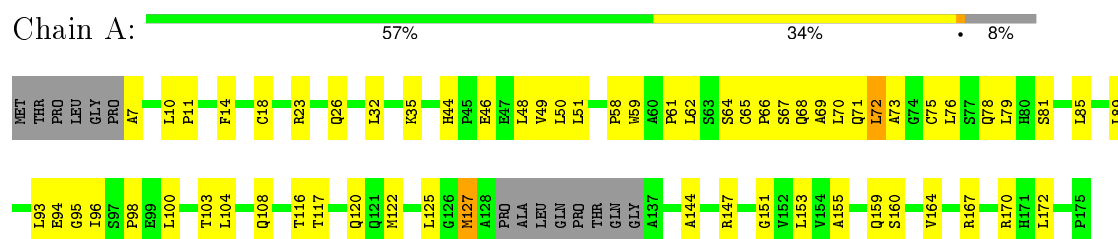
- Molecule 2 is a protein called PROTEIN (G-CSF RECEPTOR).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	208	Total	C	N	O	S	0	0	0
			1650	1056	276	300	18			
2	D	200	Total	C	N	O	S	0	0	0
			1587	1016	265	289	17			
2	F	209	Total	C	N	O	S	0	0	0
			1655	1059	277	301	18			
2	H	200	Total	C	N	O	S	0	0	0
			1587	1016	265	289	17			

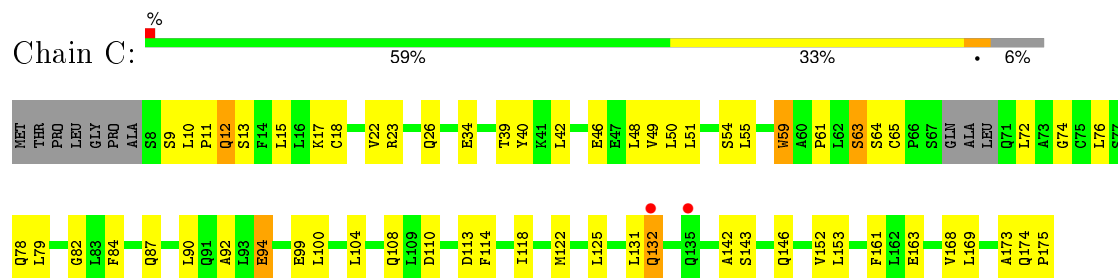
3 Residue-property plots

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

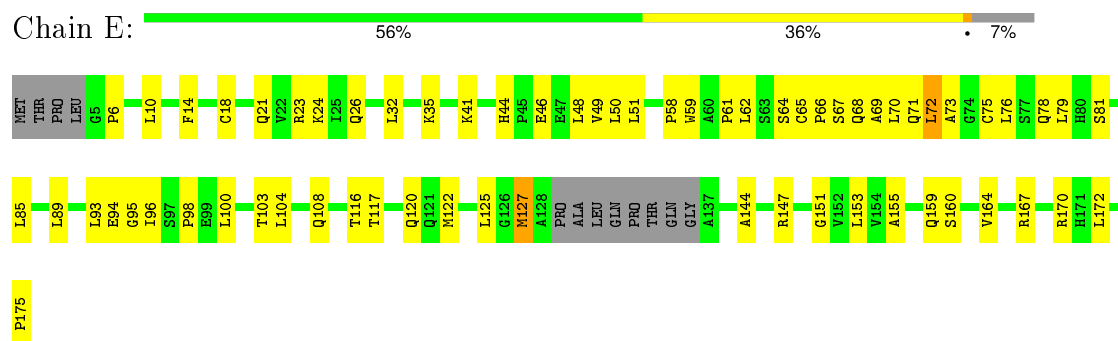
- Molecule 1: PROTEIN (GRANULOCYTE COLONY-STIMULATING FACTOR)



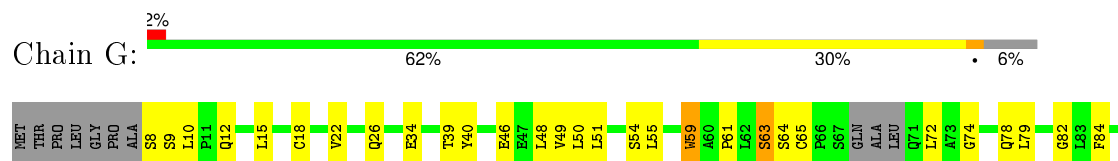
- Molecule 1: PROTEIN (GRANULOCYTE COLONY-STIMULATING FACTOR)



- Molecule 1: PROTEIN (GRANULOCYTE COLONY-STIMULATING FACTOR)

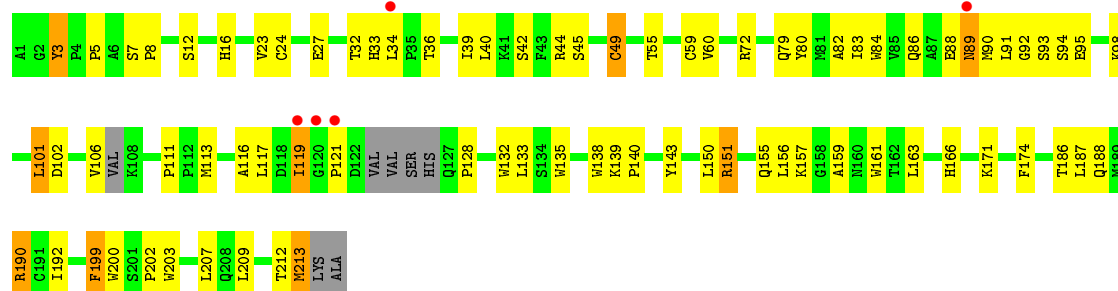


- Molecule 1: PROTEIN (GRANULOCYTE COLONY-STIMULATING FACTOR)

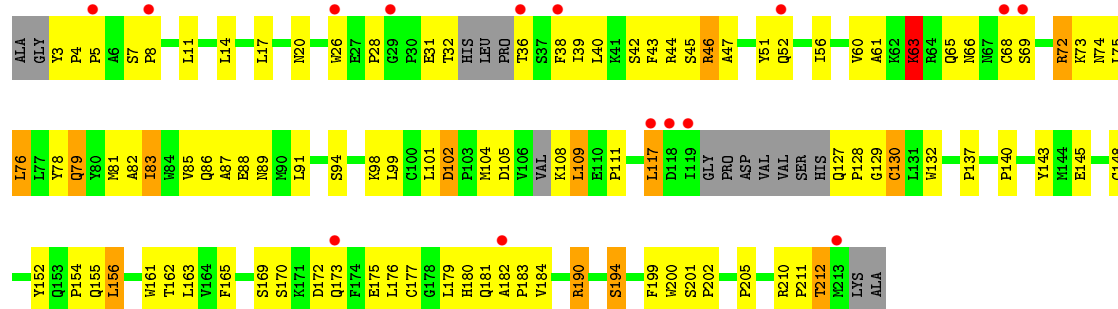




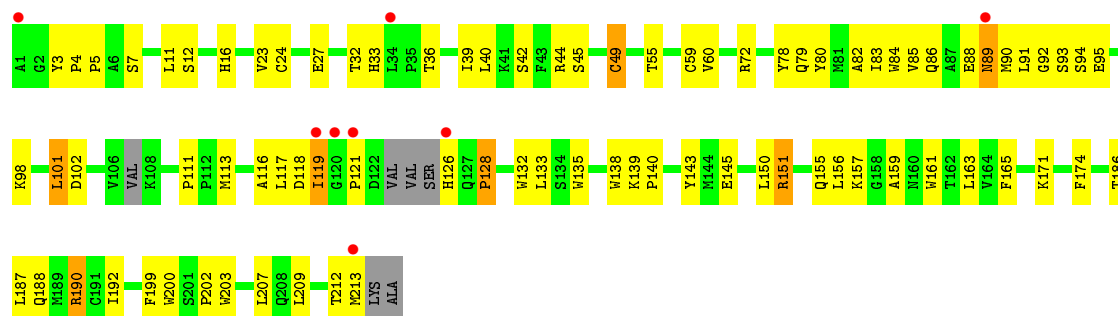
• Molecule 2: PROTEIN (G-CSF RECEPTOR)



• Molecule 2: PROTEIN (G-CSF RECEPTOR)

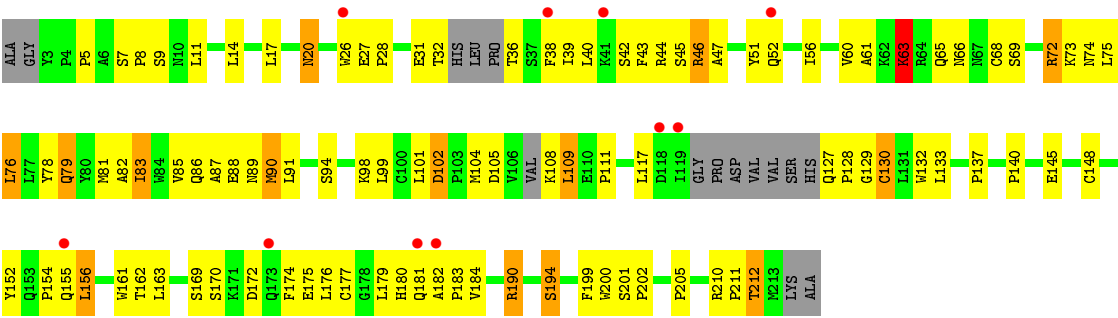


• Molecule 2: PROTEIN (G-CSF RECEPTOR)



• Molecule 2: PROTEIN (G-CSF RECEPTOR)





4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	125.72Å 125.72Å 373.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 3.50 48.16 – 3.50	Depositor EDS
% Data completeness (in resolution range)	69.8 (8.00-3.50) 68.9 (48.16-3.50)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.50 (at 3.48Å)	Xtriage
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.317 , (Not available) 0.290 , 0.281	Depositor DCC
R_{free} test set	504 reflections (2.05%)	DCC
Wilson B-factor (Å ²)	67.9	Xtriage
Anisotropy	0.247	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 68.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 26680 reflections	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	11424	wwPDB-VP
Average B, all atoms (Å ²)	36.0	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 68.54 % 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 4.2931e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.53	0/1244	0.71	0/1687
1	C	0.56	0/1275	0.77	0/1731
1	E	0.53	0/1256	0.71	0/1704
1	G	0.56	0/1275	0.77	0/1731
2	B	0.57	0/1702	0.86	1/2321 (0.0%)
2	D	0.54	0/1635	0.81	2/2228 (0.1%)
2	F	0.56	0/1707	0.85	1/2328 (0.0%)
2	H	0.54	0/1635	0.81	2/2228 (0.1%)
All	All	0.55	0/11729	0.80	6/15958 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	102	ASP	N-CA-C	-5.69	95.64	111.00
2	D	102	ASP	N-CA-C	-5.68	95.65	111.00
2	B	102	ASP	N-CA-C	-5.31	96.65	111.00
2	D	76	LEU	N-CA-C	-5.31	96.67	111.00
2	H	76	LEU	N-CA-C	-5.31	96.67	111.00

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	1219	0	1225	42	0
1	C	1248	0	1252	46	0
1	E	1230	0	1235	43	0
1	G	1248	0	1252	40	0
2	B	1650	0	1598	56	1
2	D	1587	0	1525	75	1
2	F	1655	0	1600	61	0
2	H	1587	0	1525	77	0
All	All	11424	0	11212	425	1

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 19.

The worst 5 of 425 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:H:14:LEU:HD13	2:H:108:LYS:HB2	1.28	1.13
1:E:76:LEU:HD12	1:E:127:MET:HE1	1.58	0.85
2:F:121:PRO:HB2	2:F:126:HIS:HA	1.60	0.84
1:E:76:LEU:HD12	1:E:127:MET:CE	2.09	0.83
1:E:6:PRO:HA	1:E:175:PRO:HG2	1.59	0.82

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:213:MET:CE	2:D:91:LEU:O[6_555]	1.99	0.21

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	157/175 (90%)	142 (90%)	14 (9%)	1 (1%)	30	75
1	C	161/175 (92%)	146 (91%)	12 (8%)	3 (2%)	10	51
1	E	159/175 (91%)	144 (91%)	14 (9%)	1 (1%)	30	75
1	G	161/175 (92%)	146 (91%)	12 (8%)	3 (2%)	10	51
2	B	202/215 (94%)	190 (94%)	11 (5%)	1 (0%)	34	78
2	D	192/215 (89%)	179 (93%)	9 (5%)	4 (2%)	9	50
2	F	203/215 (94%)	191 (94%)	11 (5%)	1 (0%)	34	78
2	H	192/215 (89%)	179 (93%)	9 (5%)	4 (2%)	9	50
All	All	1427/1560 (92%)	1317 (92%)	92 (6%)	18 (1%)	15	60

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	69	ALA
1	C	63	SER
2	D	63	LYS
1	E	69	ALA
1	G	63	SER

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	130/142 (92%)	121 (93%)	9 (7%)	19	59
1	C	134/142 (94%)	122 (91%)	12 (9%)	12	46
1	E	131/142 (92%)	120 (92%)	11 (8%)	14	50
1	G	134/142 (94%)	122 (91%)	12 (9%)	12	46
2	B	187/195 (96%)	172 (92%)	15 (8%)	15	52
2	D	179/195 (92%)	156 (87%)	23 (13%)	5	28
2	F	187/195 (96%)	175 (94%)	12 (6%)	22	62
2	H	179/195 (92%)	155 (87%)	24 (13%)	5	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	1261/1348 (94%)	1143 (91%)	118 (9%)	11	44

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	177	CYS
1	E	108	GLN
2	H	161	TRP
2	D	180	HIS
1	E	24	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 61 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	153	GLN
1	E	78	GLN
2	H	74	ASN
2	D	155	GLN
2	D	181	GLN

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 [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(Å ²)	Q<0.9
1	A	161/175 (92%)	-0.25	0 100 100	8, 29, 72, 97	0
1	C	165/175 (94%)	0.01	2 (1%) 81 72	10, 32, 85, 101	0
1	E	163/175 (93%)	-0.23	0 100 100	8, 29, 77, 97	0
1	G	165/175 (94%)	0.02	3 (1%) 71 62	10, 32, 85, 101	0
2	B	208/215 (96%)	0.26	5 (2%) 62 52	6, 28, 57, 109	0
2	D	200/215 (93%)	0.58	15 (7%) 17 14	12, 43, 76, 92	0
2	F	209/215 (97%)	0.29	8 (3%) 44 36	6, 28, 58, 109	0
2	H	200/215 (93%)	0.50	10 (5%) 32 25	12, 43, 76, 94	0
All	All	1471/1560 (94%)	0.18	43 (2%) 55 45	6, 32, 78, 109	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	120	GLY	3.5
1	G	135	GLN	3.3
2	D	38	PHE	3.3
2	F	121	PRO	3.2
2	B	120	GLY	3.1

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 ⓘ

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.