



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 10:15 PM GMT

PDB ID : 5BPU
Title : Crystal structure of Norrin, a Wnt signalling activator, Crystal Form I
Authors : Chang, T.-H.; Hsieh, F.-L.; Harlos, K.; Jones, E.Y.
Deposited on : 2015-05-28
Resolution : 2.40 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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026982
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) : rb-20026982

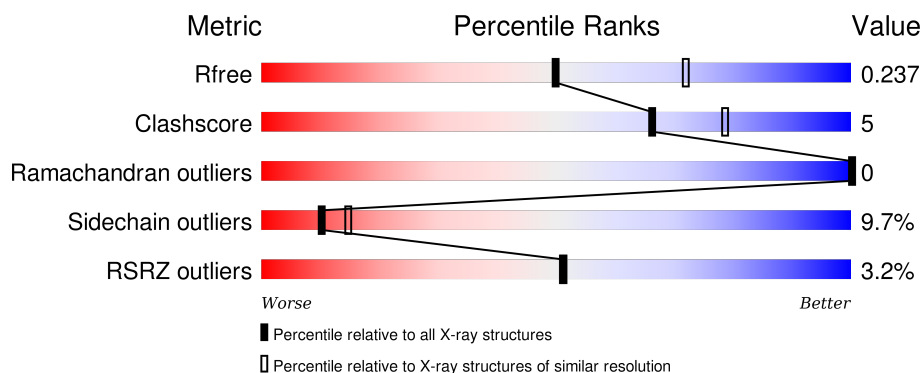
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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.



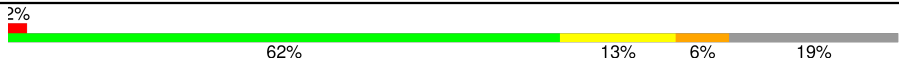


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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	122	<div> <div>8%</div> <div> <div></div> <div>74%</div> <div>16%</div> <div>•</div> <div>9%</div> </div> </div>
1	B	122	<div> <div>%</div> <div> <div></div> <div>72%</div> <div>9%</div> <div></div> <div>19%</div> </div> </div>
1	C	122	<div> <div>%</div> <div> <div></div> <div>67%</div> <div>16%</div> <div>•</div> <div>16%</div> </div> </div>
1	D	122	<div> <div>2%</div> <div> <div></div> <div>76%</div> <div>7%</div> <div>•</div> <div>16%</div> </div> </div>
1	E	122	<div> <div>2%</div> <div> <div></div> <div>59%</div> <div>19%</div> <div>•</div> <div>18%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	122	
2	H	7	
3	I	4	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GGL	H	1	X	-	-	-
3	GGL	I	1	X	-	-	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5087 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 Norrin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	111	Total	C	N	O	S	0	0	0
			866	526	166	159	15			
1	B	99	Total	C	N	O	S	0	0	0
			782	477	153	138	14			
1	C	103	Total	C	N	O	S	0	0	0
			807	490	157	146	14			
1	D	102	Total	C	N	O	S	0	0	0
			804	489	156	144	15			
1	E	100	Total	C	N	O	S	0	0	0
			788	480	154	140	14			
1	F	99	Total	C	N	O	S	0	0	0
			782	477	153	138	14			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	GLY	-	expression tag	UNP Q00604
A	23	PRO	-	expression tag	UNP Q00604
A	24	GLY	-	expression tag	UNP Q00604
A	134	GLY	-	expression tag	UNP Q00604
A	135	THR	-	expression tag	UNP Q00604
A	136	GLU	-	expression tag	UNP Q00604
A	137	THR	-	expression tag	UNP Q00604
A	138	SER	-	expression tag	UNP Q00604
A	139	GLN	-	expression tag	UNP Q00604
A	140	VAL	-	expression tag	UNP Q00604
A	141	ALA	-	expression tag	UNP Q00604
A	142	PRO	-	expression tag	UNP Q00604
A	143	ALA	-	expression tag	UNP Q00604
B	22	GLY	-	expression tag	UNP Q00604
B	23	PRO	-	expression tag	UNP Q00604
B	24	GLY	-	expression tag	UNP Q00604
B	134	GLY	-	expression tag	UNP Q00604

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	135	THR	-	expression tag	UNP Q00604
B	136	GLU	-	expression tag	UNP Q00604
B	137	THR	-	expression tag	UNP Q00604
B	138	SER	-	expression tag	UNP Q00604
B	139	GLN	-	expression tag	UNP Q00604
B	140	VAL	-	expression tag	UNP Q00604
B	141	ALA	-	expression tag	UNP Q00604
B	142	PRO	-	expression tag	UNP Q00604
B	143	ALA	-	expression tag	UNP Q00604
C	22	GLY	-	expression tag	UNP Q00604
C	23	PRO	-	expression tag	UNP Q00604
C	24	GLY	-	expression tag	UNP Q00604
C	134	GLY	-	expression tag	UNP Q00604
C	135	THR	-	expression tag	UNP Q00604
C	136	GLU	-	expression tag	UNP Q00604
C	137	THR	-	expression tag	UNP Q00604
C	138	SER	-	expression tag	UNP Q00604
C	139	GLN	-	expression tag	UNP Q00604
C	140	VAL	-	expression tag	UNP Q00604
C	141	ALA	-	expression tag	UNP Q00604
C	142	PRO	-	expression tag	UNP Q00604
C	143	ALA	-	expression tag	UNP Q00604
D	22	GLY	-	expression tag	UNP Q00604
D	23	PRO	-	expression tag	UNP Q00604
D	24	GLY	-	expression tag	UNP Q00604
D	134	GLY	-	expression tag	UNP Q00604
D	135	THR	-	expression tag	UNP Q00604
D	136	GLU	-	expression tag	UNP Q00604
D	137	THR	-	expression tag	UNP Q00604
D	138	SER	-	expression tag	UNP Q00604
D	139	GLN	-	expression tag	UNP Q00604
D	140	VAL	-	expression tag	UNP Q00604
D	141	ALA	-	expression tag	UNP Q00604
D	142	PRO	-	expression tag	UNP Q00604
D	143	ALA	-	expression tag	UNP Q00604
E	22	GLY	-	expression tag	UNP Q00604
E	23	PRO	-	expression tag	UNP Q00604
E	24	GLY	-	expression tag	UNP Q00604
E	134	GLY	-	expression tag	UNP Q00604
E	135	THR	-	expression tag	UNP Q00604
E	136	GLU	-	expression tag	UNP Q00604
E	137	THR	-	expression tag	UNP Q00604

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	138	SER	-	expression tag	UNP Q00604
E	139	GLN	-	expression tag	UNP Q00604
E	140	VAL	-	expression tag	UNP Q00604
E	141	ALA	-	expression tag	UNP Q00604
E	142	PRO	-	expression tag	UNP Q00604
E	143	ALA	-	expression tag	UNP Q00604
F	22	GLY	-	expression tag	UNP Q00604
F	23	PRO	-	expression tag	UNP Q00604
F	24	GLY	-	expression tag	UNP Q00604
F	134	GLY	-	expression tag	UNP Q00604
F	135	THR	-	expression tag	UNP Q00604
F	136	GLU	-	expression tag	UNP Q00604
F	137	THR	-	expression tag	UNP Q00604
F	138	SER	-	expression tag	UNP Q00604
F	139	GLN	-	expression tag	UNP Q00604
F	140	VAL	-	expression tag	UNP Q00604
F	141	ALA	-	expression tag	UNP Q00604
F	142	PRO	-	expression tag	UNP Q00604
F	143	ALA	-	expression tag	UNP Q00604

- Molecule 2 is a protein called (GGL)EEEEEE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	H	7	Total	C	N	O	0	0	0
			59	35	7	17			

- Molecule 3 is a protein called (GGL)EEE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	I	4	Total	C	N	O	0	0	0
			35	20	4	11			

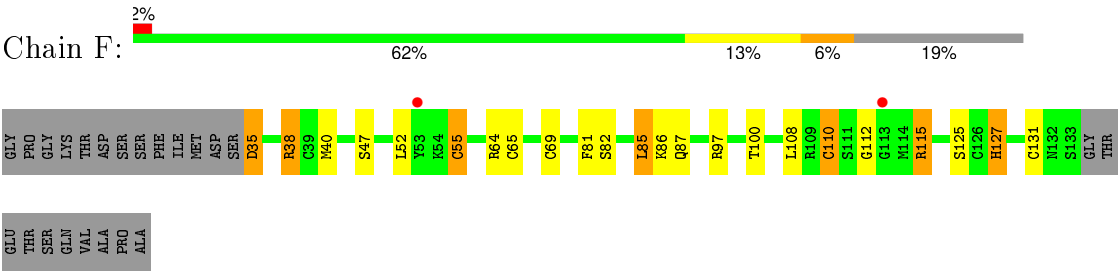
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	28	Total	O	0	0
			28	28		
4	B	29	Total	O	0	0
			29	29		
4	C	27	Total	O	0	0
			27	27		

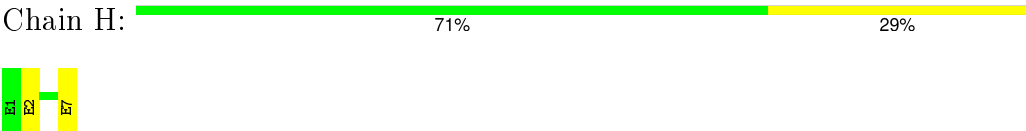
Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	22	Total 22	O 22	0	0
4	E	31	Total 31	O 31	0	0
4	F	25	Total 25	O 25	0	0
4	H	2	Total 2	O 2	0	0



● Molecule 2: (GGL)EEEEEE



● Molecule 3: (GGL)EEE



There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	46.37Å 79.10Å 234.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	65.56 – 2.40 65.56 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (65.56-2.40) 99.8 (65.56-2.40)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.71 (at 2.40Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.217 , 0.263 0.244 , 0.237	Depositor DCC
R_{free} test set	1717 reflections (5.21%)	DCC
Wilson B-factor (Å ²)	47.3	Xtriage
Anisotropy	0.214	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 46.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 34700 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5087	wwPDB-VP
Average B, all atoms (Å ²)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.02% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GGL

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	0.51	2/882 (0.2%)	0.56	1/1184 (0.1%)
1	B	0.27	0/798	0.48	0/1070
1	C	0.50	0/823	0.55	0/1104
1	D	0.28	0/820	0.46	0/1099
1	E	0.41	0/804	0.54	0/1078
1	F	0.37	0/798	0.57	1/1070 (0.1%)
2	H	1.03	0/44	1.28	0/44
3	I	1.01	0/23	1.19	0/23
All	All	0.42	2/4992 (0.0%)	0.54	2/6672 (0.0%)

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	E	0	2
1	F	0	1
2	H	1	0
3	I	1	0
All	All	2	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	88	PRO	N-CD	5.19	1.55	1.47
1	A	36	PRO	N-CD	5.18	1.55	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	35	ASP	C-N-CD	6.22	141.46	128.40
1	A	35	ASP	C-N-CD	5.29	139.50	128.40

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	H	1	GGL	CA
3	I	1	GGL	CA

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	82	SER	Peptide
1	E	84	VAL	Peptide
1	F	85	LEU	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	866	0	850	9	0
1	B	782	0	771	3	0
1	C	807	0	790	15	0
1	D	804	0	789	3	0
1	E	788	0	776	14	0
1	F	782	0	771	9	0
2	H	59	0	37	1	0
3	I	35	0	22	0	0
4	A	28	0	0	1	0
4	B	29	0	0	1	0
4	C	27	0	0	1	0
4	D	22	0	0	0	0
4	E	31	0	0	1	0
4	F	25	0	0	1	0
4	H	2	0	0	0	0
All	All	5087	0	4806	48	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 5.

The worst 5 of 48 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:F:35:ASP:O	1:F:38:ARG:NH1	1.87	1.06
1:C:110:CYS:HB2	1:C:116:LEU:HD13	1.56	0.86
1:A:87:GLN:N	1:A:87:GLN:OE1	2.30	0.65
1:E:76:GLU:OE1	1:E:90:ARG:NH1	2.29	0.65
1:C:116:LEU:HD12	1:C:116:LEU:N	2.14	0.63

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	109/122 (89%)	107 (98%)	2 (2%)	0	100	100
1	B	97/122 (80%)	95 (98%)	2 (2%)	0	100	100
1	C	101/122 (83%)	99 (98%)	2 (2%)	0	100	100
1	D	100/122 (82%)	99 (99%)	1 (1%)	0	100	100
1	E	98/122 (80%)	94 (96%)	4 (4%)	0	100	100
1	F	97/122 (80%)	92 (95%)	5 (5%)	0	100	100
All	All	602/732 (82%)	586 (97%)	16 (3%)	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	102/110 (93%)	94 (92%)	8 (8%)	16	24
1	B	92/110 (84%)	85 (92%)	7 (8%)	16	25
1	C	95/110 (86%)	86 (90%)	9 (10%)	11	15
1	D	95/110 (86%)	89 (94%)	6 (6%)	22	35
1	E	93/110 (84%)	83 (89%)	10 (11%)	8	11
1	F	92/110 (84%)	77 (84%)	15 (16%)	3	3
2	H	6/6 (100%)	5 (83%)	1 (17%)	3	3
3	I	3/3 (100%)	3 (100%)	0	100	100
All	All	578/669 (86%)	522 (90%)	56 (10%)	10	15

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	74	ARG
1	E	58	LYS
1	F	108	LEU
1	D	84	VAL
1	D	119	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	GGL	H	1	2	5,8,9	0.84	0	3,9,11	0.75	0
3	GGL	I	1	3	5,8,9	0.71	0	3,9,11	0.86	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GGL	H	1	2	1/1/2/3	0/3/8/9	0/0/0/0
3	GGL	I	1	3	1/1/2/3	0/3/8/9	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	H	1	GGL	CA
3	I	1	GGL	CA

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	H	6
3	I	3

The worst 5 of 9 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	H	5:GLU	C	6:GLU	N	6.13
1	H	2:GLU	C	3:GLU	N	5.99
1	H	1:GGL	C	2:GLU	N	5.93
1	H	6:GLU	C	7:GLU	N	5.81
1	H	4:GLU	C	5:GLU	N	5.62

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	111/122 (90%)	0.32	10 (9%) 12 11	29, 56, 116, 141	0
1	B	99/122 (81%)	-0.07	1 (1%) 84 83	32, 58, 103, 132	0
1	C	103/122 (84%)	0.18	1 (0%) 84 83	35, 66, 123, 147	0
1	D	102/122 (83%)	0.04	3 (2%) 55 54	36, 61, 96, 116	0
1	E	100/122 (81%)	0.15	3 (3%) 54 53	35, 62, 128, 144	0
1	F	99/122 (81%)	0.08	2 (2%) 68 68	40, 69, 143, 189	0
2	H	6/7 (85%)	0.36	0 100 100	60, 72, 97, 130	0
3	I	3/4 (75%)	0.65	0 100 100	78, 78, 78, 98	0
All	All	623/743 (83%)	0.13	20 (3%) 51 51	29, 64, 121, 189	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	113	GLY	7.9
1	A	84	VAL	5.5
1	A	83	THR	3.6
1	A	31	ILE	3.1
1	A	111	SER	2.8

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	GGL	H	1	9/10	0.80	0.31	-	98,109,154,156	0
3	GGL	I	1	9/10	0.57	0.30	-	76,98,127,140	0

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.