wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 10, 2023 – 06:25 PM JST

PDB ID	:	8HBR
Title	:	The C-terminal domain of Spiral2
Authors	:	Hayashi, I.
Deposited on	:	2022-10-30
Resolution	:	2.20 Å(reported)

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	:	2.31.3
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)	:	2.31.3

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 2.20 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.
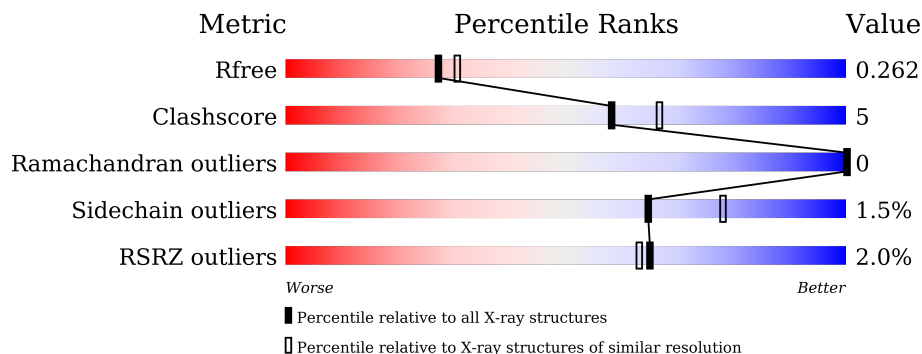
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	135	4% 84% 13% ••
1	B	135	90% 7% ••
1	C	135	.% 91% 6% ••
1	D	135	.% 92% 7% ••
1	E	135	2% 84% 13% ••
1	F	135	4% 82% 14% ••

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6271 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 TOG domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	133	Total 1030	C 654	N 173	O 196	S 1	Se 6	0	0	0
1	B	134	Total 1040	C 660	N 176	O 197	S 1	Se 6	0	0	0
1	C	134	Total 1041	C 660	N 176	O 198	S 1	Se 6	0	0	0
1	D	134	Total 1040	C 660	N 176	O 197	S 1	Se 6	0	0	0
1	E	133	Total 1030	C 654	N 173	O 196	S 1	Se 6	0	0	0
1	F	132	Total 1025	C 651	N 172	O 195	S 1	Se 6	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	771	GLY	-	expression tag	UNP A0A2K1IUB4
A	772	SER	-	expression tag	UNP A0A2K1IUB4
A	773	HIS	-	expression tag	UNP A0A2K1IUB4
B	771	GLY	-	expression tag	UNP A0A2K1IUB4
B	772	SER	-	expression tag	UNP A0A2K1IUB4
B	773	HIS	-	expression tag	UNP A0A2K1IUB4
C	771	GLY	-	expression tag	UNP A0A2K1IUB4
C	772	SER	-	expression tag	UNP A0A2K1IUB4
C	773	HIS	-	expression tag	UNP A0A2K1IUB4
D	771	GLY	-	expression tag	UNP A0A2K1IUB4
D	772	SER	-	expression tag	UNP A0A2K1IUB4
D	773	HIS	-	expression tag	UNP A0A2K1IUB4
E	771	GLY	-	expression tag	UNP A0A2K1IUB4
E	772	SER	-	expression tag	UNP A0A2K1IUB4
E	773	HIS	-	expression tag	UNP A0A2K1IUB4
F	771	GLY	-	expression tag	UNP A0A2K1IUB4
F	772	SER	-	expression tag	UNP A0A2K1IUB4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	773	HIS	-	expression tag	UNP A0A2K1IUB4

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	12	Total 12	O 12	0	0
2	B	10	Total 10	O 10	0	0
2	C	19	Total 19	O 19	0	0
2	D	8	Total 8	O 8	0	0
2	E	13	Total 13	O 13	0	0
2	F	3	Total 3	O 3	0	0

3 Residue-property plots ⓘ

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: TOG domain-containing protein
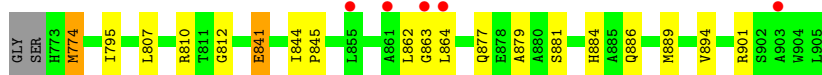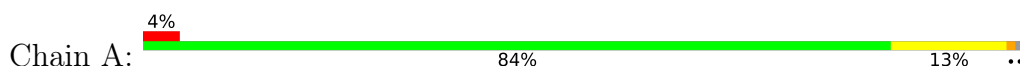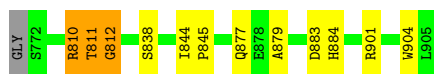
- Molecule 1: TOG domain-containing protein

- Molecule 1: TOG domain-containing protein

- Molecule 1: TOG domain-containing protein
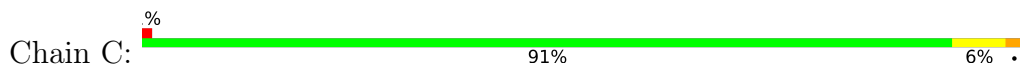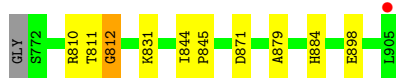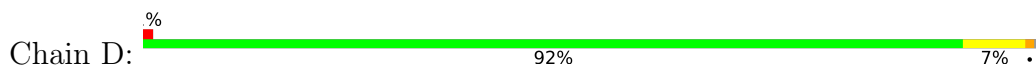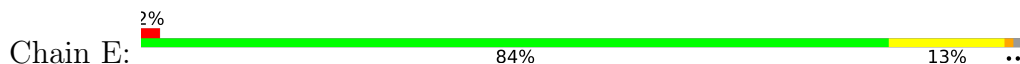

- Molecule 1: TOG domain-containing protein
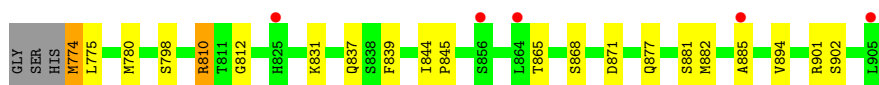

- Molecule 1: TOG domain-containing protein

Chain F:

4%

82%

14%

GLY SER HIS M774 L775 M780 S798 R810 T811 G812 H825 K831 Q837 S838 F839 I844 P845 S856 L864 T865 S868 D871 Q877 S881 M882 A885 V894 R901 S902 L905

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	55.51Å 109.56Å 69.79Å 90.00° 113.20° 90.00°	Depositor
Resolution (Å)	40.00 – 2.20 46.14 – 2.16	Depositor EDS
% Data completeness (in resolution range)	99.0 (40.00-2.20) 97.6 (46.14-2.16)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.79 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.202 , 0.256 0.215 , 0.262	Depositor DCC
R_{free} test set	2079 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	48.0	Xtriage
Anisotropy	0.211	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 41.0	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6271	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.31% 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 ⓘ

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	0.47	0/1042	0.73	0/1404
1	B	0.52	0/1053	0.80	1/1419 (0.1%)
1	C	0.57	1/1054 (0.1%)	0.79	1/1420 (0.1%)
1	D	0.54	1/1053 (0.1%)	0.79	0/1419
1	E	0.49	0/1042	0.74	0/1404
1	F	0.47	0/1038	0.74	0/1400
All	All	0.51	2/6282 (0.0%)	0.77	2/8466 (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	A	0	2
1	B	0	2
1	C	0	3
1	D	0	2
1	E	0	2
1	F	0	2
All	All	0	13

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	895	GLU	CD-OE2	-5.88	1.19	1.25
1	D	898	GLU	CD-OE2	5.09	1.31	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	811	THR	C-N-CA	-5.76	110.20	122.30
1	C	811	THR	C-N-CA	-5.11	111.58	122.30

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	810	ARG	Sidechain
1	A	812	GLY	Peptide
1	B	810	ARG	Sidechain
1	B	812	GLY	Peptide
1	C	806	ARG	Sidechain

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	1030	0	1022	16	0
1	B	1040	0	1029	4	0
1	C	1041	0	1032	4	0
1	D	1040	0	1029	6	0
1	E	1030	0	1022	17	0
1	F	1025	0	1020	18	0
2	A	12	0	0	3	0
2	B	10	0	0	0	0
2	C	19	0	0	0	0
2	D	8	0	0	0	0
2	E	13	0	0	2	0
2	F	3	0	0	0	0
All	All	6271	0	6154	59	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 59 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:A:841:GLU:HG2	2:A:1001:HOH:O	1.12	1.27
1:D:879:ALA:O	1:D:884:HIS:HE1	1.67	0.76
1:E:896:LEU:O	1:E:900:LEU:HD13	1.91	0.71
1:F:837:GLN:HB2	1:F:882:MSE:SE	2.43	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:881:SER:OG	1:F:881:SER:HB3	1.93	0.67

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	131/135 (97%)	127 (97%)	4 (3%)	0	100	100
1	B	132/135 (98%)	128 (97%)	4 (3%)	0	100	100
1	C	132/135 (98%)	128 (97%)	4 (3%)	0	100	100
1	D	132/135 (98%)	129 (98%)	3 (2%)	0	100	100
1	E	131/135 (97%)	128 (98%)	3 (2%)	0	100	100
1	F	130/135 (96%)	128 (98%)	2 (2%)	0	100	100
All	All	788/810 (97%)	768 (98%)	20 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	111/107 (104%)	109 (98%)	2 (2%)	59	72
1	B	112/107 (105%)	108 (96%)	4 (4%)	35	45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	113/107 (106%)	111 (98%)	2 (2%)	59	72
1	D	112/107 (105%)	112 (100%)	0	100	100
1	E	111/107 (104%)	111 (100%)	0	100	100
1	F	111/107 (104%)	109 (98%)	2 (2%)	59	72
All	All	670/642 (104%)	660 (98%)	10 (2%)	65	78

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	774	MSE
1	F	774	MSE
1	F	902	SER
1	B	838	SER
1	B	883	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	832	HIS
1	F	848	GLN
1	D	837	GLN
1	F	860	ASN
1	D	899	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	127/135 (94%)	0.07	5 (3%)	39	37	38, 59, 99, 115	0
1	B	128/135 (94%)	-0.13	0	100	100	37, 50, 72, 87	0
1	C	128/135 (94%)	-0.20	1 (0%)	86	85	29, 46, 67, 79	0
1	D	128/135 (94%)	-0.06	1 (0%)	86	85	32, 51, 79, 99	0
1	E	127/135 (94%)	-0.03	3 (2%)	59	56	37, 56, 94, 113	0
1	F	126/135 (93%)	0.22	5 (3%)	38	36	44, 67, 99, 116	0
All	All	764/810 (94%)	-0.02	15 (1%)	65	63	29, 54, 91, 116	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	905	LEU	8.6
1	A	861	ALA	4.8
1	E	863	GLY	3.1
1	F	885	ALA	3.0
1	A	903	ALA	2.8

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.