Full wwPDB NMR Structure Validation Report (i)

Nov 6, 2023 – 11:54 pm GMT

PDB ID	:	1HD9
Title	:	The Bowman-Birk Inhibitor Reactive Site Loop Sequence Represents an Independent Structural Beta-Hairpin Motif
Authors	:	Brauer, A.B.E.; Kelly, G.; McBride, J.D.; Cooke, R.M.; Matthews, S.J.; Leatherbarrow, R.J.
Deposited on	:	2000-11-13

This is a Full wwPDB NMR Structure Validation 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/NMRValidationReportHelp
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 (i)) were used in the production of this report:

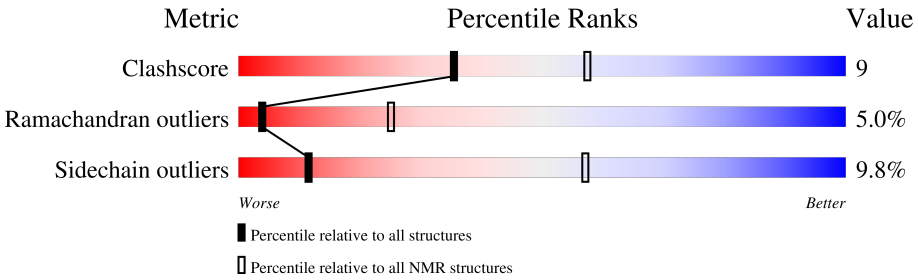
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
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

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

Molprobity failed to run

2 Ensemble composition and analysis (i)

This entry contains 30 models. Model 15 is the overall representative, medoid model (most similar to other models). The authors have identified model 22 as representative.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:3-A:10 (8)	0.23	15

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 6 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 2, 4, 5, 7, 8, 9, 11, 14, 15, 22, 23, 24, 28, 29, 30
2	10, 19, 20, 21, 25
3	3, 6, 26
4	17, 27
5	16, 18
6	12, 13

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 162 atoms, of which 80 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called BOWMAN-BIRK INHIBITOR DERIVED PEPTIDE.

Mol	Chain	Residues	Atoms						Trace
1	A	11	Total	C	H	N	O	S	0
			162	52	80	12	16	2	

SEQUENCE-PLOTS INFOmissingINFO

4 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing*.

Of the 200 calculated structures, 30 were deposited, based on the following criterion: *LOWEST ENERGY*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR	refinement	3.851

No chemical shift data was provided.

5 Model quality

5.1 Standard geometry

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

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	55	54	54	1±1
All	All	1650	1620	1620	28

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 9.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:5:SER:C	1:A:6:ILE:HD12	0.67	2.09	25	2
1:A:6:ILE:CD1	1:A:6:ILE:N	0.57	2.67	25	1
1:A:6:ILE:HD12	1:A:6:ILE:N	0.54	2.16	25	2
1:A:3:THR:O	1:A:5:SER:N	0.51	2.43	28	13
1:A:5:SER:OG	1:A:9:GLN:OE1	0.46	2.34	19	1
1:A:5:SER:O	1:A:6:ILE:HD13	0.46	2.10	26	2
1:A:5:SER:C	1:A:6:ILE:HD13	0.44	2.32	26	1
1:A:3:THR:N	1:A:9:GLN:O	0.43	2.51	24	1
1:A:5:SER:O	1:A:8:PRO:N	0.43	2.51	12	1
1:A:9:GLN:HG3	1:A:10:CYS:N	0.41	2.29	18	1
1:A:5:SER:OG	1:A:9:GLN:HG2	0.41	2.15	24	1
1:A:9:GLN:CG	1:A:10:CYS:N	0.41	2.83	30	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:6:ILE:N	1:A:6:ILE:HD12	0.40	2.31	8	1

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 PDB entries followed by that with respect to all NMR entries. 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	8/11 (73%)	7±1 (90±7%)	0±1 (5±7%)	0±0 (5±6%)	4	25
All	All	240/330 (73%)	215 (90%)	13 (5%)	12 (5%)	4	25

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	4	ALA	12

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 PDB entries followed by that with respect to all NMR entries. 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	0	-	-	-	
All	All	133/270 (49%)	120 (90%)	13 (10%)	11	57

All 4 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	9	GLN	10
1	A	6	ILE	1
1	A	3	THR	1
1	A	5	SER	1

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
1	NLE	A	1	1	6,7,8	0.41±0.01	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
1	NLE	A	1	1	2,7,9	0.39±0.00	0±0 (0±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
1	NLE	A	1	1	-	0±0,5,6,8	-

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 molecules in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Chemical shift validation

No chemical shift data were provided