wwPDB NMR Structure Validation Summary Report (i)

Feb 8, 2022 – 01:34 PM EST

PDB ID	:	1BDC
Title	:	STAPHYLOCOCCUS AUREUS PROTEIN A, IMMUNOGLOBULIN-BIND ING B DOMAIN, NMR, 10 STRUCTURES
Authors	:	Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arata, Y.; Shimada, I.
Deposited on	:	1996-06-28

This is a wwPDB NMR 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

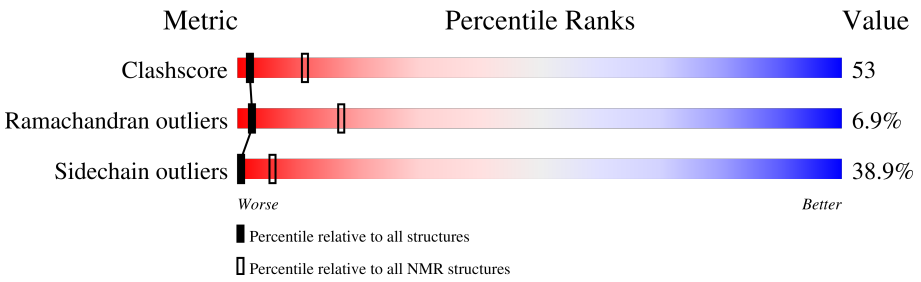
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

1 Overall quality at a glance (i)

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

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	60	15% 48% 23% 13%

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 6 is the overall representative, medoid model (most similar to other models).

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:6-A:57 (52)	0.41	6

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 3 clusters and 2 single-model clusters were found.

Cluster number	Models
1	3, 7, 10
2	1, 5, 6
3	2, 9
Single-model clusters	4; 8

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 941 atoms, of which 463 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called STAPHYLOCOCCUS AUREUS PROTEIN A.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	60	941	295	463	85	98	0

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: STAPHYLOCOCCUS AUREUS PROTEIN A

Chain A: 15% 48% 23% 13%

T1 A2 D3 N4 K5 F6 N7 K8 E9 Q10 Q11 N12 A13 F14 Y15 E16 I17 L18 H19 L20 P21 N22 L23 N24 E25 E26 Q27 R28 N29 G30 F31 I32 Q33 S34 L35 K36 D37 D38 P39 S40 Q41 S42 A43 N44 L45 L46 A49 K50 K51 L52 N53 D54 A55 Q56 A57 P58 K59 A60

4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 6. Colouring as in section 4.1 above.

• Molecule 1: STAPHYLOCOCCUS AUREUS PROTEIN A

Chain A: 18% 38% 23% 7% 13%

T1 A2 D3 N4 K5 F6 N7 K8 E9 Q10 Q11 N12 A13 F14 Y15 E16 I17 L18 H19 L20 L23 N24 Q27 R28 N29 G30 F31 I32 Q33 S34 L35 K36 D37 D38 P39 S40 Q41 S42 A43 N44 L45 L46 A49 K50 K51 L52 N53 D54 A55 Q56 A57 P58 K59 A60

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *HYBRID DISTANCE GEOMETRY-DYNAMICAL SIMULATED ANNEALING METHOD.*

Of the 55 calculated structures, 10 were deposited, based on the following criterion: *AT FIRST, THE DEPOSITORS CARRIED OUT THE DISTANCE GEOMETRY CALCULATION BY STARTING FROM 55 INITIAL STRUCTURES. THIS CALCULATION RESULTED IN 41 SOLUTIONS, WHICH HAD CORRECT POLYPEPTIDE FOLDS EXCLUDING 14 MIRROR-IMAGE SUBSTRUCTURES. NEXT, THE DYNAMICAL SIMULATED ANNEALING CALCULATIONS WERE PERFORMED BY USING THESE 41 SUBSTRUCTURES. THE DISTANCE AND TORSION ANGLE VIOLATIONS OF THE 41 SOLUTIONS OBTAINED BY THE DYNAMICAL SIMULATED ANNEALING CALCULATIONS WERE SMALLER THAN 0.6 ANGSTROMS AND 27 DEGREES, RESPECTIVELY. THE DEPOSITORS SELECTED 10 SOLUTIONS THAT HAD THE DISTANCE AND TORSION ANGLE VIOLATIONS OF SMALLER THAN 0.5 ANGSTROMS AND 10 DEGREES, RESPECTIVELY..*

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

Software name	Classification	Version
X-PLOR	refinement	
EMBOSS	structure solution	
X-PLOR	structure solution	

No chemical shift data was provided.

6 Model quality (i)

6.1 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 (average) 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	1.09±0.01	1±0/426 (0.2± 0.0%)	1.08±0.01	0±0/576 (0.0± 0.0%)
All	All	1.09	10/4260 (0.2%)	1.08	0/5760 (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	Planarity
1	A	0.0±0.0	1.0±0.4
All	All	0	10

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	19	HIS	CG-ND1	-6.61	1.24	1.38	8	10

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	28	ARG	Sidechain	9
1	A	31	PHE	Sidechain	1

6.2 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 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	419	401	401	44±6
All	All	4190	4010	4010	438

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 53.

5 of 236 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:13:ALA:HB3	1:A:46:LEU:HD23	0.92	1.40	2	1
1:A:7:ASN:ND2	1:A:7:ASN:N	0.80	2.29	10	8
1:A:52:LEU:HD12	1:A:53:ASN:N	0.69	2.03	7	1
1:A:35:LEU:HD23	1:A:36:LYS:N	0.69	2.03	9	1
1:A:12:ASN:ND2	1:A:13:ALA:N	0.68	2.41	4	9

6.3 Torsion angles (i)

6.3.1 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	52/60 (87%)	34±3 (65±5%)	15±3 (28±6%)	4±1 (7±2%)	2 17
All	All	520/600 (87%)	337 (65%)	147 (28%)	36 (7%)	2 17

5 of 14 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	8	LYS	8
1	A	9	GLU	7
1	A	20	LEU	3
1	A	29	ASN	3
1	A	17	ILE	2

6.3.2 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	45/51 (88%)	28±2 (61±4%)	18±2 (39±4%)	0 6
All	All	450/510 (88%)	275 (61%)	175 (39%)	0 6

5 of 40 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	7	ASN	10
1	A	12	ASN	10
1	A	34	SER	9
1	A	56	GLN	9
1	A	32	ILE	8

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided