wwPDB NMR Structure Validation Summary Report (i)

Feb 13, 2022 – 11:06 PM EST

PDB ID	:	1ITF
Title	:	INTERFERON ALPHA-2A, NMR, 24 STRUCTURES
Authors	:	Klaus, W.; Gsell, B.; Labhardt, A.M.; Wipf, B.; Senn, H.
Deposited on	:	1997-08-22

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:

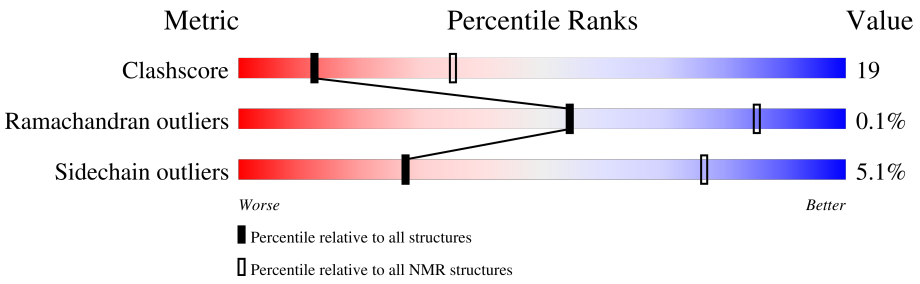
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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	165	41% 32% • 27%

2 Ensemble composition and analysis (i)

This entry contains 24 models. Model 15 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:11-A:43, A:52-A:95, A:114-A:156 (120)	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 2 clusters. No single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2698 atoms, of which 1347 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called INTERFERON ALPHA-2A.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	165	2698	860	1347	227	255	9	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: INTERFERON ALPHA-2A

Chain A: 41% 32% • 27%

C1 D2 L3 P4 Q5 T6 H7 S8 L9 G10 S11 R12 R13 T14 L15 M16 L17 L18 M21 R22 K23 I24 C29 L30 K31 D32 R33 F36 G37 F38 P39 Q40 F43 G44 N45 Q46 F47 Q48 K49 A50 E51 T52 I53 P54 V55 L56 H57 E58 M59 I60 I63 F64 N65 L66 F67

S68 T69 S72 W76 F84 L88 L95 E96 A97 C98 V99 I100 Q101 G102 V103 G104 V105 T106 E107 T108 P109 L110 M111 K112 E113 D114 S115 I116 L117 A118 V119 R120 K121 Y122 R125 I126 W140 E141 V142 V143 R144 R149 S150 F151 S152 L153 S154 T155 N156 L157 Q158 E159

S160 L161 R162 S163 K164 E165

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

The representative model is number 15. Colouring as in section 4.1 above.

- Molecule 1: INTERFERON ALPHA-2A

Chain A: 44% 27% •• 27%

C1 D2 L3 P4 Q5 T6 H7 S8 L9 G10 S11 R12 R13 T14 L15 M16 L17 R22 K23 I24 C29 D32 R33 F36 P39 Q40 E41 E42 F43 G44 N45 Q46 F47 Q48 K49 A50 E51 T52 I53 P54 V55 L56 I60 I63 F64 N65 L66 F67 S68 T69 K70 D71

W76 F84 L88 L95 E96 A97 C98 V99 I100 Q101 G102 V103 G104 V105 T106 E107 T108 P109 L110 M111 K112 E113 D114 S115 I116 L117 A118 V119 R120 K121 Y122 R125 I126 L130 Y135 W140 E141 V142 V143 R144 R149 S150 F151 S152 L153 S154 T155 N156 L157 Q158 E159 S160

L161 R162 S163 K164 E165

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 24 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	3.1
X-PLOR	structure solution	3.1

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

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	7.9±0.3
All	All	0	190

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 8 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	12	ARG	Sidechain	24
1	A	13	ARG	Sidechain	24
1	A	22	ARG	Sidechain	24
1	A	120	ARG	Sidechain	24
1	A	144	ARG	Sidechain	24

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	1008	1009	1009	38±7
All	All	24192	24216	24216	917

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.

5 of 217 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:24:ILE:HD12	1:A:29:CYS:SG	0.90	2.07	1	24
1:A:17:LEU:HD13	1:A:84:PHE:HA	0.81	1.52	6	21
1:A:58:GLU:HG3	1:A:116:ILE:HD13	0.78	1.56	14	8
1:A:15:LEU:HD23	1:A:18:LEU:HD12	0.77	1.55	1	11
1:A:15:LEU:HD21	1:A:151:PHE:HB3	0.77	1.54	5	24

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	120/165 (73%)	116±1 (97±1%)	3±1 (3±1%)	0±0 (0±0%)	54	85
All	All	2880/3960 (73%)	2795 (97%)	83 (3%)	2 (0%)	54	85

All 2 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	22	ARG	1
1	A	37	GLY	1

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	113/152 (74%)	107±2 (95±2%)	6±2 (5±2%)	27	77
All	All	2712/3648 (74%)	2573 (95%)	139 (5%)	27	77

5 of 45 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	23	LYS	11
1	A	144	ARG	9
1	A	30	LEU	8
1	A	149	ARG	7
1	A	40	GLN	7

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