wwPDB NMR Structure Validation Summary Report (i)

Jun 12, 2024 – 04:44 PM EDT

PDB ID	:	1KLD
Title	:	SOLUTION STRUCTURE OF TGF-B1, NMR, MODELS 18-33 OF 33 STRUCTURES
Authors	:	Hinck, A.P.; Archer, S.J.; Qian, S.W.; Roberts, A.B.; Sporn, M.B.; Weatherbee, J.A.; Tsang, M.L.-S.; Lucas, R.; Zhang, B.-L.; Wenker, J.; Torchia, D.A.
Deposited on	:	1996-01-16

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 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:

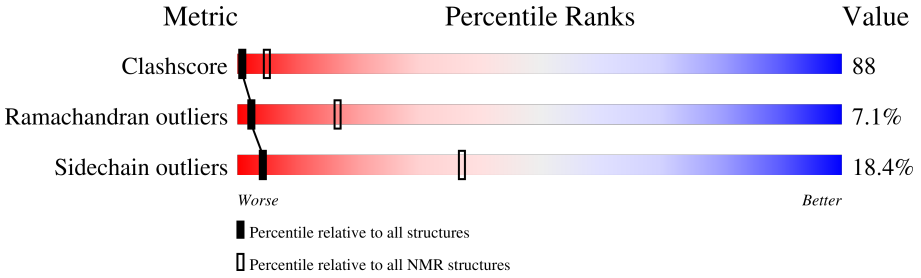
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	112	21% 54% 16% 9%
1	B	112	22% 55% 16% 6%

2 Ensemble composition and analysis (i)

This entry contains 16 models. Model 23 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:1-A:7, A:16-A:45, A:78-A:110, B:48-B:77, B:111-B:112 (102)	0.49	23
2	A:48-A:77, A:111-A:112, B:1-B:7, B:15-B:47, B:78-B:110 (105)	0.53	23

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 4 clusters and 1 single-model cluster was found.

Cluster number	Models
1	21, 23, 25, 31, 33
2	18, 24, 27, 30
3	19, 20, 29, 32
4	22, 28
Single-model clusters	26

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3536 atoms, of which 1742 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called TRANSFORMING GROWTH FACTOR-BETA 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	112	1768	576	871	152	159	10	0
1	B	112	1768	576	871	152	159	10	0

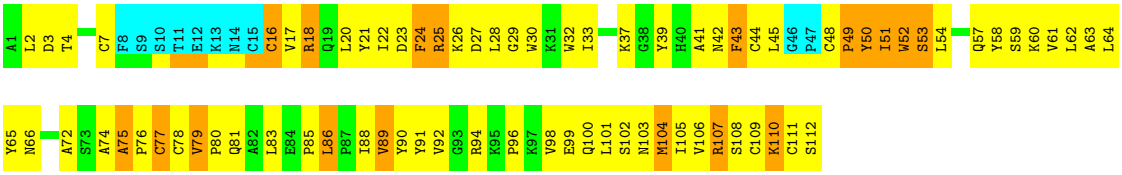
4 Residue-property plots (i)

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.

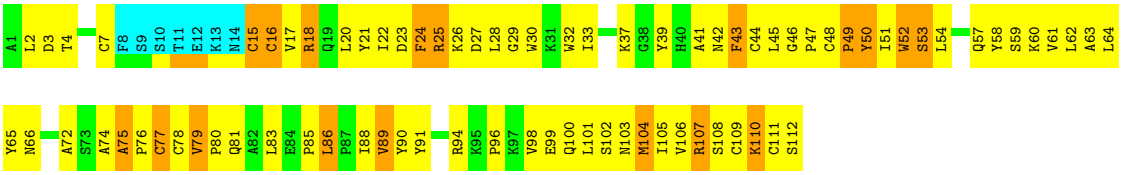
• Molecule 1: TRANSFORMING GROWTH FACTOR-BETA 1

Chain A: 21% 54% 16% 9%



• Molecule 1: TRANSFORMING GROWTH FACTOR-BETA 1

Chain B: 22% 55% 16% 6%

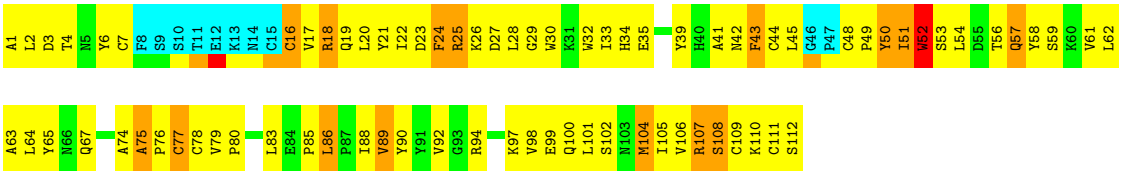


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 23. Colouring as in section 4.1 above.

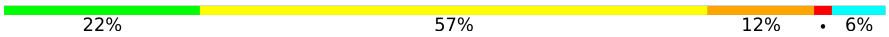
• Molecule 1: TRANSFORMING GROWTH FACTOR-BETA 1

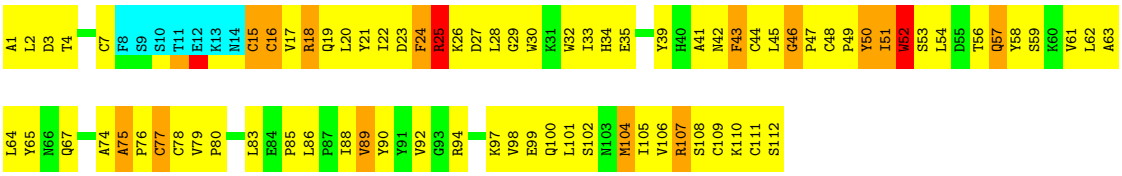
Chain A: 21% 55% 13% • 9%



• Molecule 1: TRANSFORMING GROWTH FACTOR-BETA 1

Chain B:


22%
57%
12%
6%


A1 L2 D3 T4 C7 F8 S9 S10 T11 E12 K13 N14 C15 C16 V17 R18 Q19 L20 Y21 I22 D23 F24 R25 K26 D27 L28 G29 W30 K31 W32 I33 H34 E35 Y39 H40 A41 N42 F43 C44 L45 G46 P47 C48 P49 Y50 I51 W52 S53 L54 D55 T56 Q57 Y58 S59 K60 V61 L62 A63
L64 Y65 N66 Q67 A74 A75 P76 C77 C78 V79 P80 L83 E84 P85 L86 P87 I88 V89 Y90 Y91 V92 G93 R94 K97 V98 E99 Q100 L101 S102 N103 M104 I105 V106 R107 S108 C109 K110 C111 S112

5 Refinement protocol and experimental data overview (i)

Of the 33 calculated structures, 16 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
X-PLOR	refinement	3.1
XPLOR	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	3.9±0.3
1	B	0.0±0.0	3.9±0.2
All	All	0	125

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	18	ARG	Sidechain	16
1	A	25	ARG	Sidechain	16
1	B	18	ARG	Sidechain	16
1	B	94	ARG	Sidechain	16
1	B	107	ARG	Sidechain	16

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	824	806	806	168±10
1	B	841	820	820	169±11
All	All	26640	26016	26016	4628

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 88.

5 of 908 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:61:VAL:HG13	1:B:28:LEU:HD11	1.13	1.15	23	16
1:A:28:LEU:HD11	1:B:61:VAL:HG13	1.11	1.15	32	16
1:B:2:LEU:HD11	1:B:52:TRP:CD1	1.02	1.88	29	12
1:A:2:LEU:HD11	1:A:52:TRP:CD1	1.02	1.88	29	14
1:B:58:TYR:OH	1:B:62:LEU:HD22	0.97	1.60	25	16

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	100/112 (89%)	68±2 (68±2%)	25±3 (25±3%)	7±2 (7±2%)	2	18
1	B	103/112 (92%)	69±2 (67±2%)	27±3 (26±2%)	8±2 (7±2%)	2	15
All	All	3248/3584 (91%)	2188 (67%)	830 (26%)	230 (7%)	2	16

5 of 31 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	89	VAL	16
1	B	89	VAL	16
1	A	24	PHE	15
1	B	24	PHE	15
1	A	49	PRO	14

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	91/100 (91%)	74±2 (81±3%)	17±2 (19±3%)	4 37
1	B	93/100 (93%)	76±2 (82±3%)	17±2 (18±3%)	4 37
All	All	2944/3200 (92%)	2401 (82%)	543 (18%)	4 37

5 of 85 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	3	ASP	16
1	A	50	TYR	16
1	A	77	CYS	16
1	A	99	GLU	16
1	B	3	ASP	16

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