wwPDB NMR Structure Validation Summary Report (i)

May 7, 2024 – 04:29 pm BST

PDB ID	:	1DZC
Title	:	High resolution structure of acidic fibroblast growth factor. Mutant FGF-4-ALA-(24-154), 24 NMR structures
Authors	:	Lozano, R.M.; Pineda-Lucena, A.; Gonzalez, C.; Jimenez, M.A.; Cuevas, P.; Redondo-Horcajo, M.; Sanz, J.M.; Rico, M.; Gimenez-Gallego, G.
Deposited on	:	2000-02-24

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:

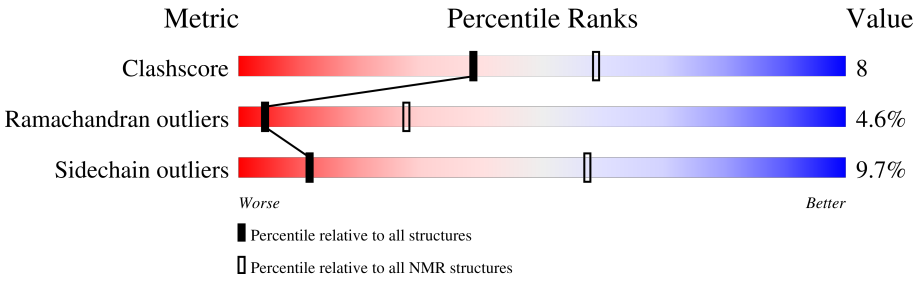
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson 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	131	64% 26% 5% 5%

2 Ensemble composition and analysis (i)

This entry contains 24 models. Model 13 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:26-A:150 (125)	0.82	13

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 3 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1033 atoms, of which 0 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called FIBROBLAST GROWTH FACTOR 1.

Mol	Chain	Residues	Atoms					Trace
			Total	C	N	O	S	
1	A	131	1033	648	180	201	4	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	ALA	LYS	engineered mutation	UNP P05230
A	25	ALA	PRO	engineered mutation	UNP P05230
A	26	ALA	LYS	engineered mutation	UNP P05230

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: FIBROBLAST GROWTH FACTOR 1

Chain A: 64% 26% 5% 5%

A24 A25 A26 L27 L28 Y29 N32 L37 R38 I39 R49 D50 R51 S52 D53 Q54 H55 I56 Q57 L58 Q59 Y69 S72 T73 Y78 G85 L86 L87 Y88 F99 L100 E101 R102 H107 Y108 Y111 K115 F122 K132 R133 G134 P135 R136 T137 H138 Y139 G140

Q141 K142 A143 I144 L149 P150 V151 S152 S153 D154

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

The representative model is number 13. Colouring as in section 4.1 above.

• Molecule 1: FIBROBLAST GROWTH FACTOR 1

Chain A: 62% 24% 6% • 5%

A24 A25 A26 L27 L28 Y29 C30 S31 N32 H35 R38 I39 R51 S52 D53 Q54 H55 I56 Q57 L58 V65 G66 E67 V68 Y69 S72 T73 Y78 Y88 E96 C97 L98 F99 L100 L103 N106 H107 Y108 K115 E118 K119 N120 W121 F122 L125 K126 K127

K132 R133 R136 T137 H138 Y139 G140 Q141 K142 A143 I144 L147 V151 S152 S153 D154

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *RESTRAINED MOLECULAR DYNAMICS*.

Of the 24 calculated structures, 24 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
GROMOS	refinement	
GROMOS	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	0.82±0.01	0±0/1018 (0.0± 0.0%)	1.44±0.08	9±3/1376 (0.6± 0.2%)
All	All	0.82	1/24432 (0.0%)	1.45	206/33024 (0.6%)

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	9.6±2.2
All	All	0	231

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	91	GLN	CD-NE2	-7.78	1.13	1.32	11	1

5 of 43 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	91	GLN	OE1-CD-NE2	-29.10	54.98	121.90	11	1
1	A	91	GLN	CG-CD-NE2	-12.24	87.32	116.70	11	1
1	A	139	TYR	CB-CG-CD2	-10.32	114.81	121.00	2	14
1	A	29	TYR	CB-CG-CD2	-9.92	115.05	121.00	11	23
1	A	102	ARG	NE-CZ-NH1	8.85	124.72	120.30	17	10

There are no chirality outliers.

5 of 43 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	88	TYR	Sidechain	21
1	A	122	PHE	Sidechain	21
1	A	29	TYR	Sidechain	20
1	A	108	TYR	Sidechain,Mainchain	20
1	A	69	TYR	Sidechain	17

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	995	0	971	16±4
All	All	23880	0	23304	374

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 8.

5 of 177 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:133:ARG:HB2	1:A:137:THR:HG21	0.88	1.44	21	1
1:A:85:GLY:HA3	1:A:134:GLY:H	0.77	1.39	10	12
1:A:87:LEU:HD21	1:A:133:ARG:HH11	0.76	1.41	6	2
1:A:27:LEU:HD21	1:A:36:PHE:HE2	0.74	1.43	15	1
1:A:27:LEU:HD21	1:A:53:ASP:HB3	0.72	1.62	1	2

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	125/131 (95%)	95±4 (76±3%)	24±3 (20±2%)	6±2 (5±1%)	4 27
All	All	3000/3144 (95%)	2275 (76%)	586 (20%)	139 (5%)	4 27

5 of 36 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	73	THR	24
1	A	101	GLU	13
1	A	107	HIS	10
1	A	66	GLY	7
1	A	65	VAL	7

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	108/112 (96%)	98±3 (90±3%)	10±3 (10±3%)	12 57
All	All	2592/2688 (96%)	2340 (90%)	252 (10%)	12 57

5 of 58 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	100	LEU	24
1	A	133	ARG	15
1	A	58	LEU	12
1	A	49	ARG	12
1	A	142	LYS	10

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