wwPDB NMR Structure Validation Summary Report

Jun 15, 2024 – 07:03 AM EDT

PDB ID	:	2CKN
BMRB ID	:	6885
Title	:	NMR Structure of the First Ig Module of mouse FGFR1
Authors	:	Kiselyov, V.V.; Bock, E.; Berezin, V.; Poulsen, F.M.
Deposited on	:	2006-04-20

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:

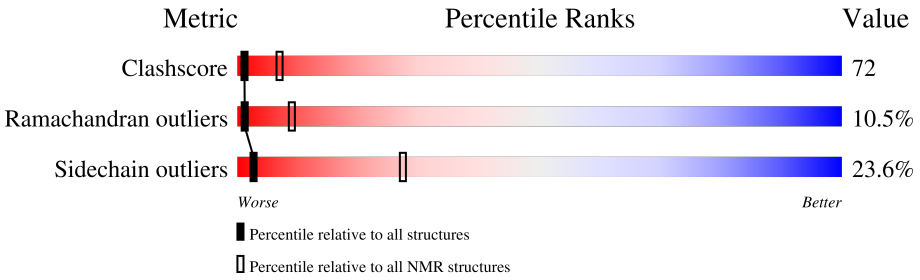
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 is 51%.

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	95	17% 41% 24% 18%

2 Ensemble composition and analysis

This entry contains 20 models. Model 14 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:28-A:63, A:67-A:108 (78)	0.69	14

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	1, 2, 3, 6, 8, 10, 13, 14, 15, 16, 17, 18, 19
2	4, 5, 11
3	7, 12
Single-model clusters	9; 20

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1445 atoms, of which 711 are hydrogens and 0 are deuteriums.

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

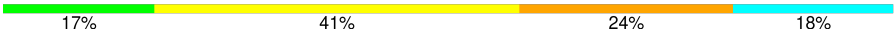
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	95	1445	454	711	128	150	2	0

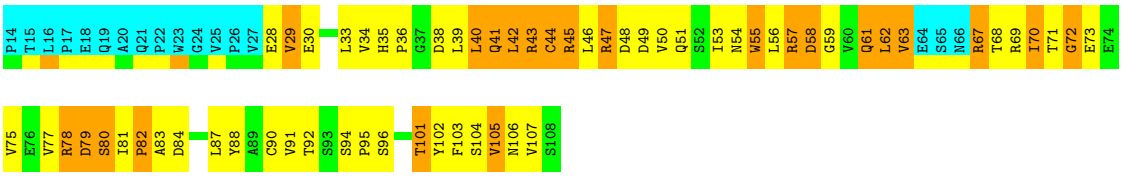
4 Residue-property plots

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: BASIC FIBROBLAST GROWTH FACTOR RECEPTOR 1

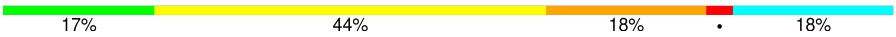
Chain A: 

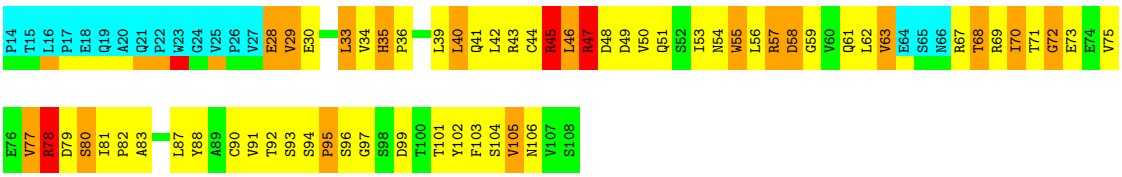


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 14. Colouring as in section 4.1 above.

• Molecule 1: BASIC FIBROBLAST GROWTH FACTOR RECEPTOR 1

Chain A: 



5 Refinement protocol and experimental data overview

Of the ? calculated structures, 20 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	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	702
Number of shifts mapped to atoms	634
Number of unparsed shifts	0
Number of shifts with mapping errors	68
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	51%

6 Model quality

6.1 Standard geometry

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	1.11±0.00	1±0/612 (0.2± 0.0%)	1.42±0.02	5±1/833 (0.6± 0.1%)
All	All	1.11	19/12240 (0.2%)	1.42	96/16660 (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	5.9±1.0
All	All	0	118

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	55	TRP	CG-CD2	-5.32	1.34	1.43	20	19

5 of 12 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	55	TRP	CD1-NE1-CE2	8.99	117.09	109.00	9	20
1	A	55	TRP	NE1-CE2-CZ2	7.89	139.08	130.40	9	20
1	A	55	TRP	CG-CD1-NE1	-7.45	102.66	110.10	10	20
1	A	47	ARG	NE-CZ-NH2	-7.37	116.61	120.30	10	4
1	A	57	ARG	NE-CZ-NH2	-7.17	116.72	120.30	20	1

There are no chirality outliers.

5 of 9 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	67	ARG	Sidechain	18
1	A	69	ARG	Sidechain	18
1	A	45	ARG	Sidechain	16
1	A	78	ARG	Sidechain	16
1	A	43	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	604	590	590	85±4
All	All	12080	11800	11800	1709

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 72.

5 of 479 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:29:VAL:HG21	1:A:101:THR:HB	1.04	1.27	5	18
1:A:29:VAL:HG21	1:A:101:THR:CB	0.95	1.91	5	16
1:A:29:VAL:HG11	1:A:103:PHE:CE2	0.93	1.98	10	17
1:A:63:VAL:HB	1:A:70:ILE:HD13	0.89	1.45	2	7
1:A:71:THR:O	1:A:73:GLU:N	0.84	2.10	12	18

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	77/95 (81%)	57±2 (74±2%)	12±2 (15±3%)	8±2 (11±2%)	1	9
All	All	1540/1900 (81%)	1142 (74%)	236 (15%)	162 (11%)	1	9

5 of 19 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	29	VAL	20
1	A	63	VAL	20
1	A	72	GLY	19
1	A	58	ASP	16
1	A	79	ASP	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	71/86 (83%)	54±2 (76±2%)	17±2 (24±2%)	3	27
All	All	1420/1720 (83%)	1085 (76%)	335 (24%)	3	27

5 of 51 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	34	VAL	18
1	A	101	THR	18
1	A	42	LEU	17
1	A	54	ASN	17
1	A	78	ARG	15

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

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 51% for the well-defined parts and 50% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	702
Number of shifts mapped to atoms	634
Number of unparsed shifts	0
Number of shifts with mapping errors	68
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 68) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	GLU	H	8.32	0.005	1
1	A	1	GLU	HA	4.22	0.005	1
1	A	1	GLU	HB2	1.99	0.005	2
1	A	1	GLU	HB3	1.86	0.005	2
1	A	1	GLU	HG2	2.18	0.005	2
1	A	1	GLU	N	122.54	0.02	1
1	A	2	ALA	H	8.33	0.005	1
1	A	2	ALA	HA	4.28	0.005	1
1	A	2	ALA	HB1	1.34	0.005	1
1	A	2	ALA	HB2	1.34	0.005	1
1	A	2	ALA	HB3	1.34	0.005	1
1	A	2	ALA	N	125.6	0.02	1
1	A	3	ARG	H	8.26	0.005	1
1	A	3	ARG	HA	4.59	0.005	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	ARG	HB2	1.68	0.005	2
1	A	3	ARG	HB3	1.81	0.005	2
1	A	3	ARG	HG2	1.63	0.005	2
1	A	3	ARG	HD2	3.16	0.005	2
1	A	3	ARG	N	121.63	0.02	1
1	A	4	PRO	HA	4.38	0.005	1
1	A	4	PRO	HB2	1.86	0.005	2
1	A	4	PRO	HG2	1.99	0.005	2
1	A	4	PRO	HD2	3.75	0.005	2
1	A	5	ALA	H	8.38	0.005	1
1	A	5	ALA	HA	4.52	0.005	1
1	A	5	ALA	HB1	1.34	0.005	1
1	A	5	ALA	HB2	1.34	0.005	1
1	A	5	ALA	HB3	1.34	0.005	1
1	A	5	ALA	N	125.84	0.02	1
1	A	101	ASP	H	8.54	0.005	1
1	A	101	ASP	HA	4.56	0.005	1
1	A	101	ASP	HB2	2.64	0.005	2
1	A	101	ASP	HB3	2.54	0.005	2
1	A	101	ASP	N	123.69	0.02	1
1	A	102	ALA	H	8.23	0.005	1
1	A	102	ALA	HA	4.26	0.005	1
1	A	102	ALA	HB1	1.26	0.005	1
1	A	102	ALA	HB2	1.26	0.005	1
1	A	102	ALA	HB3	1.26	0.005	1
1	A	102	ALA	N	123.38	0.02	1
1	A	103	LEU	H	8.22	0.005	1
1	A	103	LEU	HA	4.54	0.005	1
1	A	103	LEU	HB2	1.57	0.005	2
1	A	103	LEU	HB3	1.52	0.005	2
1	A	103	LEU	HG	1.66	0.005	1
1	A	103	LEU	HD11	0.88	0.005	2
1	A	103	LEU	HD12	0.88	0.005	2
1	A	103	LEU	HD13	0.88	0.005	2
1	A	103	LEU	N	123.11	0.02	1
1	A	104	PRO	HA	4.43	0.005	1
1	A	104	PRO	HB2	1.9	0.005	2
1	A	104	PRO	HB3	2.28	0.005	2
1	A	105	SER	H	8.54	0.005	1
1	A	105	SER	HA	4.61	0.005	1
1	A	105	SER	HB2	3.94	0.005	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	105	SER	HB3	3.8	0.005	2
1	A	105	SER	N	116.39	0.02	1
1	A	106	SER	H	8.31	0.005	1
1	A	106	SER	HA	4.48	0.005	1
1	A	106	SER	HB2	3.83	0.005	2
1	A	106	SER	HB3	3.86	0.005	2
1	A	106	SER	N	118.35	0.02	1
1	A	107	GLU	H	8.1	0.005	1
1	A	107	GLU	HA	4.14	0.005	1
1	A	107	GLU	HB2	2.04	0.005	2
1	A	107	GLU	HB3	1.88	0.005	2
1	A	107	GLU	HG2	2.2	0.005	2
1	A	107	GLU	N	127.79	0.02	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	0	—	None (insufficient data)
$^{13}C_{\beta}$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	96	-1.19 ± 0.46	Should be applied

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 51%, i.e. 531 atoms were assigned a chemical shift out of a possible 1041. 0 out of 19 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	228/389 (59%)	154/158 (97%)	0/156 (0%)	74/75 (99%)
Sidechain	288/605 (48%)	283/393 (72%)	0/186 (0%)	5/26 (19%)
Aromatic	15/47 (32%)	14/23 (61%)	0/22 (0%)	1/2 (50%)
Overall	531/1041 (51%)	451/574 (79%)	0/364 (0%)	80/103 (78%)

7.1.4 Statistically unusual chemical shifts (i)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	75	VAL	HG21	-0.70	-0.58 – 2.19	-5.4
1	A	75	VAL	HG22	-0.70	-0.58 – 2.19	-5.4
1	A	75	VAL	HG23	-0.70	-0.58 – 2.19	-5.4

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: