



# wwPDB NMR Structure Validation Summary Report ⓘ

Feb 12, 2017 – 10:40 pm GMT

PDB ID : 2K4A  
Title : FGF-1-C2A binary complex structure: a key component in the fibroblast growthfactor non-classical pathway  
Authors : Yu, C.; Mohan, S.K.  
Deposited on : 2008-06-02

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk28760
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949



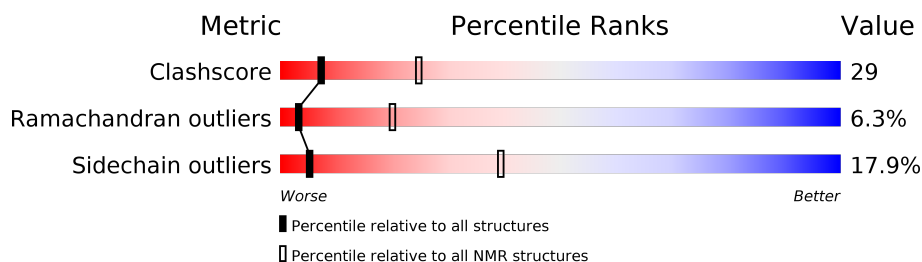
# 1 Overall quality at a glance

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	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	128	
2	B	133	



## 2 Ensemble composition and analysis

This entry contains 20 models. Model 1 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:128, B:1-B:133 (261)	0.35	1

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 2 single-model clusters were found.

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



### 3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 4201 atoms, of which 2093 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Synaptotagmin-1.

Mol	Chain	Residues	Atoms						Trace
1	A	128	Total	C	H	N	O	S	0
			2086	677	1042	169	195	3	

- Molecule 2 is a protein called Heparin-binding growth factor 1.

Mol	Chain	Residues	Atoms						Trace
2	B	133	Total	C	H	N	O	S	0
			2115	671	1051	185	204	4	



## 4 Residue-property plots

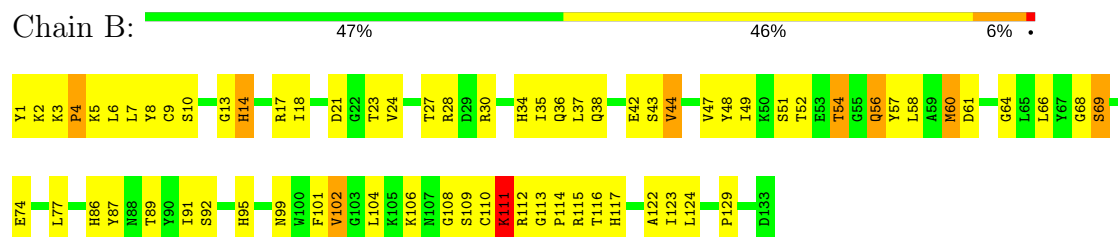
### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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: Synaptotagmin-1



- Molecule 2: Heparin-binding growth factor 1



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

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Synaptotagmin-1





Chain B:

45% 44% 10%

Y1 K2 K3 P4 K5 C9 S10 G13 R17 I18 L19 P20 D21 G22 T23 V24 T27 R28 P29 R30 H34 I35 Q36 L37 Q38 S43 V44 V47 Y48 S51 T54 G55 Q56 L57 L58 A59 M60 D61 T62 D63 G64 L65 L66 V67 G68 S69 Q70 T71 P72 S73 W74



## 5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
ARIA	structure solution	1.2
HADDOCK	structure solution	2.0
HADDOCK	refinement	2.0
ARIA/CNS	refinement	

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.



## 6 Model quality [i](#)

### 6.1 Standard geometry [i](#)

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

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	1044	1042	1039	64±8
2	B	1064	1051	1046	64±8
All	All	42160	41860	41700	2459

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 29.

5 of 662 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:19:LEU:HD11	1:A:80:LEU:HD12	0.87	1.45	5	14
1:A:3:LEU:HD22	1:A:102:GLY:HA3	0.83	1.49	11	19
1:A:56:THR:HG22	1:A:69:GLU:HG2	0.81	1.53	18	8
1:A:80:LEU:HD23	1:A:81:GLY:N	0.79	1.92	3	12
2:B:101:PHE:CD2	2:B:123:ILE:HG21	0.79	2.13	9	14

### 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	126/128 (98%)	96±3 (76±3%)	24±4 (19±3%)	6±1 (5±1%)	5	28
2	B	131/133 (98%)	91±4 (70±3%)	29±4 (22±3%)	10±2 (8±1%)	2	15
All	All	5140/5220 (98%)	3755 (73%)	1062 (21%)	323 (6%)	3	20

5 of 46 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	4	PRO	20
2	B	102	VAL	20
2	B	69	SER	20
2	B	111	LYS	20
1	A	47	LEU	20

### 6.3.2 Protein sidechains ⓘ

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	115/115 (100%)	93±3 (80±2%)	22±3 (20±2%)	4	36
2	B	117/117 (100%)	98±2 (84±2%)	19±2 (16±2%)	6	42
All	All	4640/4640 (100%)	3808 (82%)	832 (18%)	5	39

5 of 107 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)
2	B	2	LYS	20
1	A	8	TYR	20
2	B	35	ILE	20
2	B	111	LYS	20
2	B	56	GLN	20



### 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 carbohydrates 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

No chemical shift data were provided