



# wwPDB NMR Structure Validation Summary Report ⓘ

Jun 15, 2020 – 11:00 pm BST

PDB ID : 2KI6  
Title : The FGF1-S100A13-C2A hetero-hexameric complex structure: A component in the non-classical pathway for FGF1 secretion  
Authors : Krishna, S.M.; Rani, S.G.; Yu, C.  
Deposited on : 2009-04-28

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

<https://www.wwpdb.org/validation/2017/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 : 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 : 2.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

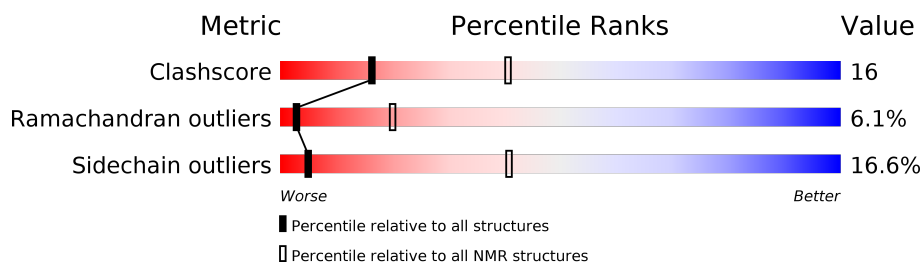
# 1 Overall quality at a glance

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	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  $\geq 3$ , 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	 45% 41% 14%
1	F	128	 53% 38% 9%
2	B	133	 50% 40% 9% .
2	E	133	 53% 38% 8% .
3	C	98	 65% 19% . 12%
4	D	98	 64% 31% . .

## 2 Ensemble composition and analysis ⓘ

This entry contains 18 models. Model 6 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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, C:11-C:96, D:1-D:93, D:95-D:98, E:1-E:133, F:1-F:128 (705)	0.20	6

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

NmrClust was unable to cluster the ensemble.

Error message: Inconsistent models

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

There are 4 unique types of molecules in this entry. The entry contains 11675 atoms, of which 5848 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
			2085	677	1042	169	194	3	
1	F	128	Total	C	H	N	O	S	0
			2085	677	1042	169	194	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
			2113	671	1049	185	204	4	
2	E	133	Total	C	H	N	O	S	0
			2113	671	1049	185	204	4	

- Molecule 3 is a protein called Protein S100-A13.

Mol	Chain	Residues	Atoms						Trace
3	C	98	Total	C	H	N	O	S	0
			1640	512	834	136	156	2	

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	97	DLY	LYS	SEE REMARK 999	UNP Q99584
C	98	DLY	LYS	SEE REMARK 999	UNP Q99584

- Molecule 4 is a protein called Protein S100-A13.

Mol	Chain	Residues	Atoms						Trace
4	D	98	Total	C	H	N	O	S	0
			1639	512	832	136	157	2	

There is a discrepancy between the modelled and reference sequences:

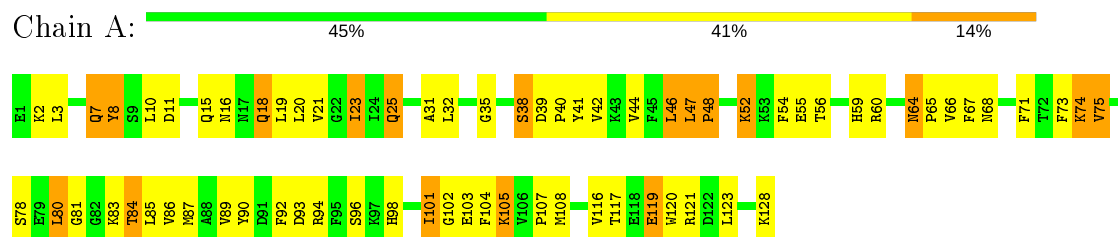
Chain	Residue	Modelled	Actual	Comment	Reference
D	94	DLY	LYS	SEE REMARK 999	UNP Q99584

## 4 Residue-property plots [i](#)

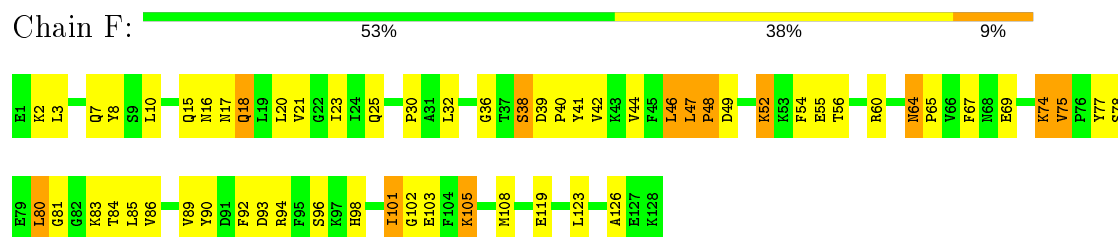
### 4.1 Average score per residue in the NMR ensemble

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.

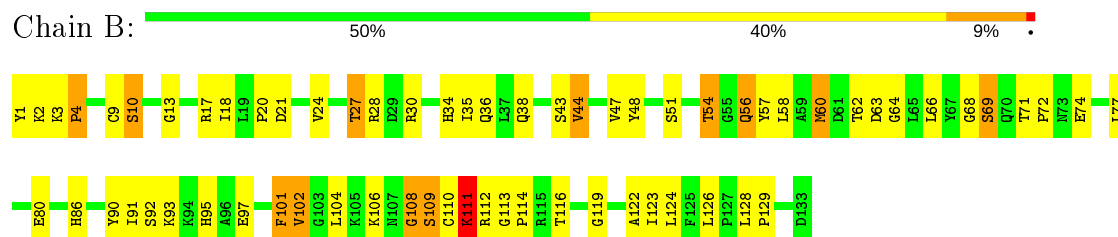
- Molecule 1: Synaptotagmin-1



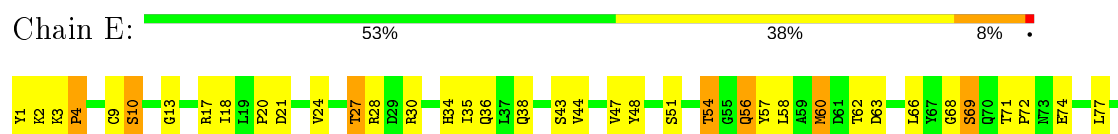
- Molecule 1: Synaptotagmin-1



- Molecule 2: Heparin-binding growth factor 1



- Molecule 2: Heparin-binding growth factor 1





• Molecule 3: Protein S100-A13



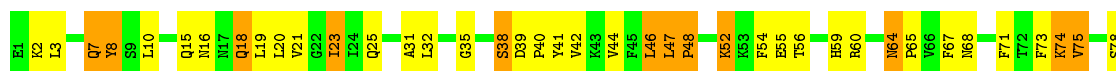
• Molecule 4: Protein S100-A13



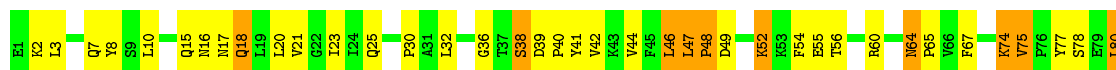
## 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 6. Colouring as in section 4.1 above.

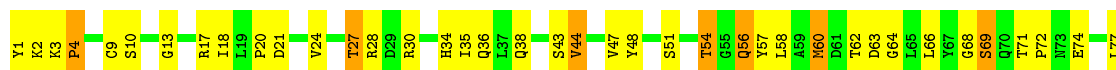
• Molecule 1: Synaptotagmin-1



• Molecule 1: Synaptotagmin-1



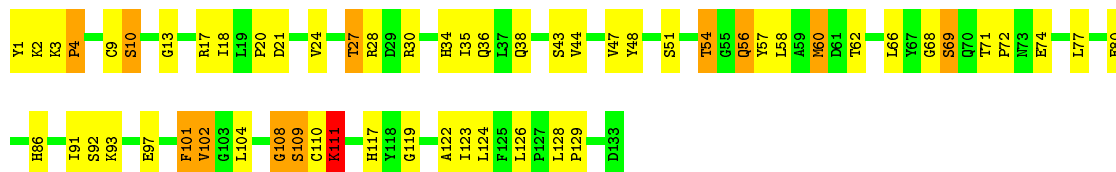
• Molecule 2: Heparin-binding growth factor 1





• Molecule 2: Heparin-binding growth factor 1

Chain E: 56% 35% 8%



• Molecule 3: Protein S100-A13

Chain C: 65% 20% 12%



• Molecule 4: Protein S100-A13

Chain D: 65% 30%



## 5 Refinement protocol and experimental data overview

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

Of the 2000 calculated structures, 18 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
ARIA	structure solution	1.2 & 2.2
CNSSOLVE	refinement	1.1 & 1.2
CNSSOLVE	structure solution	1.1 & 1.2

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

Bond lengths and bond angles in the following residue types are not validated in this section: DLY

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	0.30±0.00	0±0/1068 ( 0.0± 0.0%)	0.38±0.00	0±0/1440 ( 0.0± 0.0%)
1	F	0.30±0.00	0±0/1068 ( 0.0± 0.0%)	0.39±0.00	0±0/1440 ( 0.0± 0.0%)
2	B	0.30±0.00	0±0/1088 ( 0.0± 0.0%)	0.39±0.00	0±0/1466 ( 0.0± 0.0%)
2	E	0.30±0.00	0±0/1088 ( 0.0± 0.0%)	0.39±0.00	0±0/1466 ( 0.0± 0.0%)
3	C	0.51±0.00	0±0/722 ( 0.0± 0.0%)	0.93±0.00	2±0/965 ( 0.2± 0.0%)
4	D	0.51±0.00	0±0/807 ( 0.0± 0.0%)	0.98±0.01	0±0/1077 ( 0.0± 0.0%)
All	All	0.36	0/105138 ( 0.0%)	0.59	36/141372 ( 0.0%)

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
3	C	0.0±0.0	4.0±0.0
4	D	0.0±0.0	9.0±0.0
All	All	0	234

There are no bond-length outliers.

All 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(°)	Ideal(°)	Models	
								Worst	Total
3	C	37	GLU	OE1-CD-OE2	-5.92	116.19	123.30	17	18
3	C	68	ASP	CB-CG-OD2	5.34	123.10	118.30	18	18

There are no chirality outliers.

5 of 13 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
4	D	2	ALA	Peptide	18
4	D	76	TYR	Sidechain	18
4	D	78	ARG	Sidechain	18
3	C	38	PHE	Sidechain	18
3	C	78	ARG	Sidechain	18

## 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	1043	1042	1039	47±2
1	F	1043	1042	1039	31±1
2	B	1064	1049	1046	47±3
2	E	1064	1049	1046	46±1
3	C	713	733	732	5±1
4	D	798	819	820	8±2
All	All	103050	103212	102996	3249

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 16.

5 of 253 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:66:LEU:HD12	2:B:110:CYS:HB2	0.73	1.60	3	18
1:F:3:LEU:HD23	1:F:123:LEU:HB3	0.73	1.60	18	18
4:D:4:GLU:CG	4:D:5:PRO:HD2	0.73	2.13	1	18
2:E:66:LEU:HD12	2:E:110:CYS:HB2	0.70	1.62	17	18
1:A:3:LEU:HD23	1:A:123:LEU:HB3	0.69	1.64	6	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	126/128 (98%)	97±1 (77±1%)	23±1 (18±1%)	6±1 (5±0%)	4	28
1	F	126/128 (98%)	96±1 (76±1%)	24±1 (19±1%)	6±0 (5±0%)	4	26
2	B	131/133 (98%)	96±1 (73±1%)	26±1 (20±1%)	10±0 (7±0%)	2	15
2	E	131/133 (98%)	96±0 (73±0%)	26±1 (19±1%)	10±0 (7±0%)	2	15
3	C	86/98 (88%)	65±1 (75±1%)	17±1 (20±1%)	4±0 (5±0%)	4	26
4	D	95/98 (97%)	70±1 (74±1%)	18±1 (19±1%)	7±1 (7±1%)	2	16
All	All	12510/12924 (97%)	9347 (75%)	2401 (19%)	762 (6%)	3	20

5 of 47 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	10	SER	18
4	D	95	ILE	18
2	E	102	VAL	18
2	B	111	LYS	18
2	E	44	VAL	18

### 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%)	94±0 (82±0%)	21±0 (18±0%)	4	37
1	F	115/115 (100%)	94±0 (82±0%)	21±0 (18±0%)	4	37
2	B	117/117 (100%)	98±0 (84±0%)	19±0 (16±0%)	5	41
2	E	117/117 (100%)	97±1 (83±1%)	20±1 (17±1%)	5	41
3	C	81/89 (91%)	72±0 (89±1%)	9±0 (11±1%)	10	55
4	D	90/90 (100%)	74±2 (82±2%)	16±2 (18±2%)	4	38
All	All	11430/11574 (99%)	9528 (83%)	1902 (17%)	5	41

5 of 116 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)
4	D	57	ASP	18
2	B	30	ARG	18
4	D	4	GLU	18
2	B	51	SER	18
1	F	105	LYS	18

### 6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

3 non-standard protein/DNA/RNA residues are modelled in this entry.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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