wwPDB NMR Structure Validation Summary Report

Feb 7, 2022 – 04:45 PM EST

PDB ID	:	1A0N
Title	:	NMR STUDY OF THE SH3 DOMAIN FROM FYN PROTO-ONCOGENE TYROSINE KINASE COMPLEXED WITH THE SYNTHETIC PEPTIDE P2L CORRESPONDING TO RESIDUES 91-104 OF THE P85 SUBUNIT OF PI3-KINASE, FAMILY OF 25 STRUCTURES
Authors	:	Renzoni, D.A.; Pugh, D.J.R.; Siligardi, G.; Das, P.; Morton, C.J.; Rossi, C.; Waterfield, M.D.; Campbell, I.D.; Ladbury, J.E.
Deposited on	:	1997-12-05

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

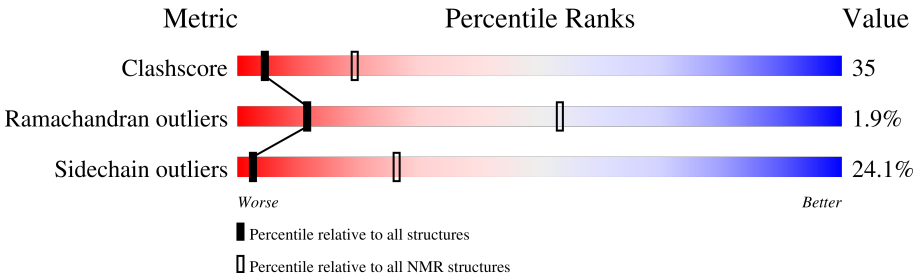
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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	14	7% 7% 86%
2	B	69	17% 58% • 7% 16%

2 Ensemble composition and analysis (i)

This entry contains 25 models. Model 5 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:95-A:96, B:86-B:89, B:100-B:124, B:128-B:151 (55)	0.30	5

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 6 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 4, 8, 11, 13, 14, 20, 21, 25
2	5, 6, 7, 17, 18, 23
3	2, 3
4	15, 16
5	12, 24
6	10, 22
Single-model clusters	9; 19

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1108 atoms, of which 540 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PRO-PRO-ARG-PRO-LEU-PRO-VAL-ALA-PRO-GLY-SER -SER-LYS-THR.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	14	207	63	108	18	18	0

- Molecule 2 is a protein called FYN.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	58	901	301	432	73	95	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	81	SER	GLY	conflict	UNP P06241

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: PRO-PRO-ARG-PRO-LEU-PRO-VAL-ALA-PRO-GLY-SER-SER-LYS-THR

Chain A: 7% 7% 86%

• Molecule 2: FYN

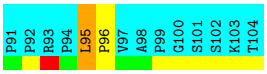
Chain B: 17% 58% 7% 16%

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

The representative model is number 5. Colouring as in section 4.1 above.

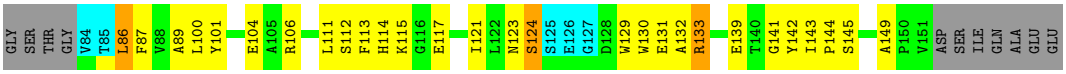
• Molecule 1: PRO-PRO-ARG-PRO-LEU-PRO-VAL-ALA-PRO-GLY-SER-SER-LYS-THR

Chain A: 7% 7% 86%



• Molecule 2: FYN

Chain B: 36% 36% 7% 16%


GLY
SER
THR
GLY
V84
T85
L86
F87
V88
A89
L100
Y101
E104
A105
R106
L111
S112
F113
H114
K115
G116
E117
I121
L122
N123
S124
S125
E126
G127
D128
W129
W130
E131
A132
R133
E139
T140
G141
Y142
I143
P144
S145
A149
P150
V151
ASP
SER
ILE
GLN
ALA
GLU
GLU

5 Refinement protocol and experimental data overview

Of the 50 calculated structures, 25 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

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

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
2	B	0.0±0.0	1.9±0.3
All	All	0	48

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

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

Mol	Chain	Res	Type	Group	Models (Total)
2	B	106	ARG	Sidechain	24
2	B	133	ARG	Sidechain	24

6.2 Too-close contacts

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	15	18	18	1±1
2	B	436	403	403	30±5
All	All	11275	10525	10525	755

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 35.

5 of 262 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:121:ILE:HD13	2:B:132:ALA:HB2	0.94	1.35	9	11
2:B:111:LEU:HD21	2:B:140:THR:C	0.89	1.87	16	10
2:B:111:LEU:HD12	2:B:143:ILE:HG23	0.87	1.47	2	24
2:B:143:ILE:CD1	2:B:148:VAL:HG11	0.85	2.02	21	9
2:B:105:ALA:HB2	2:B:111:LEU:O	0.83	1.72	13	8

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	2/14 (14%)	1±1 (36±27%)	1±1 (26±32%)	1±1 (38±32%)	0	0
2	B	52/69 (75%)	49±1 (94±3%)	3±1 (6±2%)	0±1 (0±1%)	32	76
All	All	1350/2075 (65%)	1234 (91%)	91 (7%)	25 (2%)	11	53

5 of 7 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	95	LEU	15
1	A	96	PRO	4
2	B	110	ASP	2
2	B	124	SER	1
2	B	113	PHE	1

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	2/12 (17%)	1±0 (64±22%)	1±0 (36±22%)	1	8
2	B	46/58 (79%)	35±3 (76±6%)	11±3 (24±6%)	3	27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1200/1750 (69%)	911 (76%)	289 (24%)	2 26

5 of 34 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	95	LEU	18
2	B	117	GLU	16
2	B	100	LEU	16
2	B	115	LYS	15
2	B	133	ARG	14

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

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 (i)

No chemical shift data were provided