wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 02:26 PM BST

PDB ID	:	1FU5
Title	:	NMR STRUCTURE OF THE N-SH2 DOMAIN OF THE P85 SUBUNIT OF PI3-KINASE COMPLEXED TO A DOUBLY PHOSPHORYLATED PEPTIDE DERIVED FROM POLYOMAVIRUS MIDDLE T ANTIGEN
Authors	:	Weber, T.; Schaffhausen, B.; Liu, Y.; Guenther, U.L.
Deposited on	:	2000-09-14

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
http://wwpdb.org/validation/2016/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:

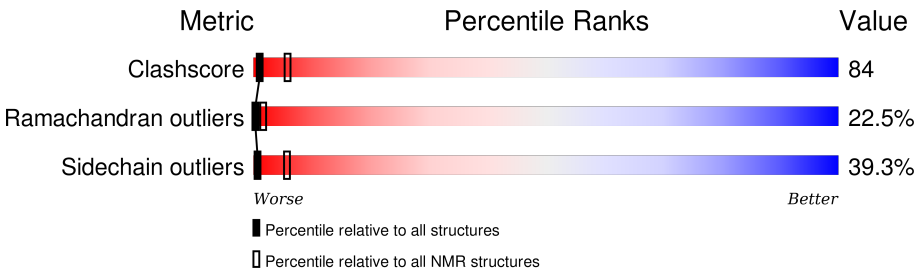
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	111	20% 41% 32% 7%
2	B	15	13% 33% 40% 13%

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2066 atoms, of which 1020 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PHOSPHATIDYLINOSITOL 3-KINASE REGULATORY ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	111	1799	571	893	158	174	3	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	60	SER	LEU	CONFLICT	UNP Q63787

- Molecule 2 is a protein called DOUBLY PHOSPHORYLATED MIDDLE T ANTIGEN.

Mol	Chain	Residues	Atoms							Trace
			Total	C	H	N	O	P	S	
2	B	15	267	85	127	15	36	2	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	4	PTR	TYR	modified residue	UNP P03076
B	11	PTR	TYR	modified residue	UNP P03076

4 Residue-property plots (i)

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: PHOSPHATIDYLINOSITOL 3-KINASE REGULATORY ALPHA SUBUNIT

Chain A: 20% 41% 32% 7%

G1 M6 S7 L8 Q9 D10 A11 E12 W13 Y14 W15 G16 D17 I18 S19 R20 E21 E22 V23 N24 E25 K26 L27 R28 D29 T30 A31 D32 G33 T34 F35 L36 V37 R38 D39 A40 S41 T42 K43 N44 H45 Y48 T49 L50 T51 L52 R53 K54 N57 N58 K59 S60 I61 K62 I63 F64 H65

R66 K69 Y70 G71 F72 S73 L76 T77 F78 N79 S80 V81 V82 E83 L84 I85 N86 H87 Y88 R89 N90 E91 S92 L93 A94 Q95 Y96 K99 L100 D101 V102 K103 L104 L105 Y106 K110 Y111

• Molecule 2: DOUBLY PHOSPHORYLATED MIDDLE T ANTIGEN

Chain B: 13% 33% 40% 13%

E1 E2 E3 Y4 M5 P6 M7 E8 D9 L10 Y11 L12 D13 I14 L15

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *The structures were energy minimized with MSI DISCOVER..*

Of the 110 calculated structures, 1 were deposited, based on the following criterion: *structures with favorable non-bond energy.*

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

Software name	Classification	Version
DYANA	structure solution	
DISCOVER	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)

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

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	0.81	0/926 (0.0%)	1.01	0/1249 (0.0%)
2	B	1.16	1/106 (0.9%)	1.00	0/137 (0.0%)
All	All	0.85	1/1032 (0.1%)	1.01	0/1386 (0.0%)

All bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	15	LEU	C-OXT	7.95	1.38	1.23

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	906	893	878	159
2	B	140	127	117	45
All	All	1046	1020	995	172

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 84.

5 of 172 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
2:B:12:LEU:HD23	2:B:15:LEU:OXT	0.93	1.64
1:A:20:ARG:CZ	2:B:4:PTR:O1P	0.83	2.26
1:A:50:LEU:CD2	1:A:61:ILE:HG22	0.82	2.05
1:A:49:THR:HG23	1:A:62:LYS:HD3	0.81	1.52
1:A:49:THR:HG23	1:A:62:LYS:CD	0.81	2.06

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	109/111 (98%)	49 (45%)	38 (35%)	22 (20%)	0 2
2	B	11/15 (73%)	4 (36%)	2 (18%)	5 (45%)	0 0
All	All	120/126 (95%)	53 (44%)	40 (33%)	27 (22%)	0 2

5 of 27 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	77	THR
1	A	30	THR
2	B	8	GLU
1	A	72	PHE
1	A	81	VAL

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	99/99 (100%)	60 (61%)	39 (39%)	1 6
2	B	13/13 (100%)	8 (62%)	5 (38%)	1 6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	112/112 (100%)	68 (61%)	44 (39%)	1 6

5 of 44 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	23	VAL
1	A	57	ASN
1	A	66	ARG
1	A	15	TRP
1	A	45	HIS

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	PTR	B	11	2	13,16,17	1.51	0 (0%)
2	PTR	B	4	2	13,16,17	1.52	0 (0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	PTR	B	11	2	19,22,24	3.04	1 (5%)
2	PTR	B	4	2	19,22,24	3.04	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PTR	B	11	2	-	0,9,11,13	0,1,1,1
2	PTR	B	4	2	-	0,9,11,13	0,1,1,1

There are no bond-length outliers.

All angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	4	PTR	P-OH-CZ	12.58	159.54	123.85
2	B	11	PTR	P-OH-CZ	12.58	159.54	123.85

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 (i)

No chemical shift data were provided