wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 03:00 PM BST

PDB ID	:	1HX7
Title	:	SOLUTION STRUCTURE OF THE CATALYTIC DOMAIN OF GAMMA DELTA RESOLVASE
Authors	:	Pan, B.; Mullen, G.P.
Deposited on	:	2001-01-11

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:

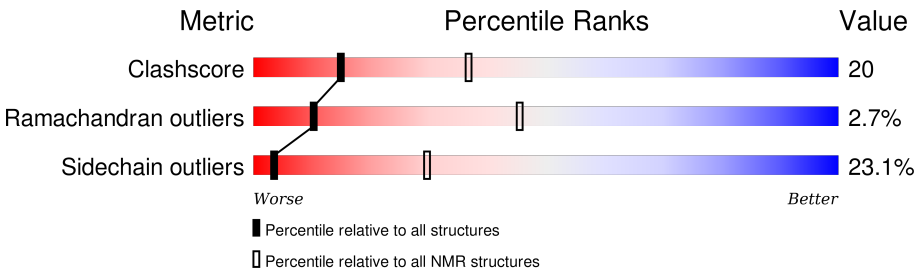
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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	105	37% 31% • 30%

2 Ensemble composition and analysis (i)

This entry contains 19 models. Model 14 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:2-A:38, A:44-A:63, A:77-A:93 (74)	0.28	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, 3, 4, 5, 7, 8, 9, 12, 14, 15, 16, 19
2	2, 10, 11
3	6, 17
Single-model clusters	13; 18

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1647 atoms, of which 836 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called TRANSPOSON GAMMA-DELTA RESOLVASE.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	105	1647	501	836	147	159	4	0

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: TRANSPOSON GAMMA-DELTA RESOLVASE

Chain A: 37% 31% • 30%

M1 R2 L3 A7 R8 V9 S10 T11 S12 Q13 Q14 S15 L16 D17 I18 Q19 V20 R21 A22 L23 K24 V28 N31 R32 I33 F34 T35 D36 K37 A38 S39 G40 S41 S42 S43 L48 D49 L50 L51 V55 V60 I61 L62 V63 K64 K65 L66 D67 R68 L69 G70 R71 D72 T73

A74 D75 M76 I77 Q78 L79 I80 K81 E82 F83 I93 D94 D95 G96 I97 S98 T99 D100 G101 E102 M103 G104 K105

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: TRANSPOSON GAMMA-DELTA RESOLVASE

Chain A: 38% 27% 6% 30%

M1 A7 R8 V9 Q13 Q14 S15 L16 D17 I18 Q19 V20 R21 A22 L23 K24 N31 R32 I33 F34 T35 D36 S39 G40 S41 S42 S43 D44 R45 L48 D49 L50 L51 K54 V55 E56 D59 V60 K64 K65 L66 D67 R68 L69 G70 R71 D72 T73 A74 D75 M76 I77

Q78 L79 I80 K81 E82 V88 S89 D94 D95 G96 I97 S98 T99 D100 G101 E102 M103 G104 K105

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics and simulated annealing.*

Of the 200 calculated structures, 19 were deposited, based on the following criterion: *lowest energy and agreement with experimental NOESY.*

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

Software name	Classification	Version
DYANA	structure solution	1.5
XPLOR	refinement	4.0

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	588	622	622	24±3
All	All	11172	11818	11818	452

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 20.

5 of 155 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:20:VAL:HG22	1:A:33:ILE:HG21	0.98	1.34	4	18
1:A:7:ALA:HB2	1:A:23:LEU:HD11	0.96	1.33	5	13
1:A:7:ALA:HB3	1:A:35:THR:HG23	0.88	1.44	18	1
1:A:80:ILE:HD13	1:A:90:ILE:HD12	0.80	1.51	19	1
1:A:24:LYS:NZ	1:A:33:ILE:HD11	0.80	1.91	6	3

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	74/105 (70%)	62±2 (84±3%)	10±2 (13±2%)	2±1 (3±2%)	10 45
All	All	1406/1995 (70%)	1181 (84%)	187 (13%)	38 (3%)	10 45

5 of 10 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	13	GLN	15
1	A	57	GLU	4
1	A	14	GLN	4
1	A	9	VAL	3
1	A	45	ARG	3

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	63/88 (72%)	48±2 (77±4%)	15±2 (23±4%)	3 30
All	All	1197/1672 (72%)	920 (77%)	277 (23%)	3 30

5 of 45 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	31	ASN	19
1	A	24	LYS	18
1	A	60	VAL	16
1	A	78	GLN	15
1	A	48	LEU	14

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 (i)

No chemical shift data were provided