wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 03:50 PM BST

PDB ID	:	1MUT
Title	:	NMR STUDY OF MUTT ENZYME, A NUCLEOSIDE TRIPHOSPHATE PYROPHOSPHOHYDROLASE
Authors	:	Abeygunawardana, C.; Weber, D.J.; Gittis, A.G.; Frick, D.N.; Lin, J.; Miller, A.-F.; Bessman, M.J.; Mildvan, A.S.
Deposited on	:	1995-09-14

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:

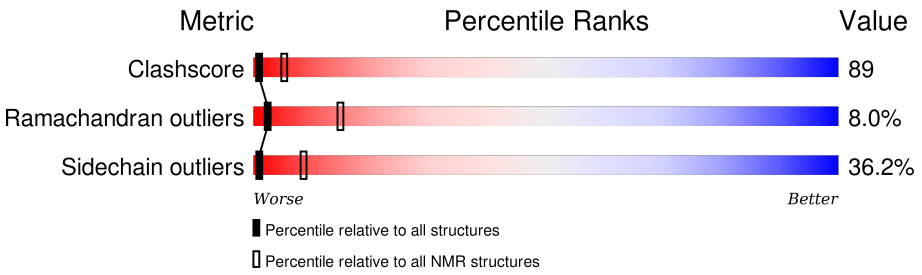
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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	129	7% 63% 20% • 9%

2 Ensemble composition and analysis (i)

This entry contains 15 models. Model 11 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:4-A:22, A:31-A:129 (118)	0.90	11

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 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 5, 8, 9, 11, 13, 14
2	6, 10, 12, 15
3	4, 7
Single-model clusters	3

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2090 atoms, of which 1035 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called NUCLEOSIDE TRIPHOSPHATE PYROPHOSPHOHYDROLASE.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	129	2090	679	1035	180	192	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: NUCLEOSIDE TRIPHOSPHATE PYROPHOSPHOHYDROLASE

Chain A: 7% 63% 20% • 9%

M1 K2 K3 L4 Q5 I6 A7 V8 G9 I10 I11 R12 N13 E14 N15 N16 E17 I18 F19 I20 T21 R22 R23 A24 A25 D26 A27 H28 M29 A30 N31 K32 L33 E34 F35 P36 G37 G38 K39 I40 E41 M42 G43 E44 T45 P46 E47 Q48 A49 V50 V51 R52 E53 L54 Q55 E56 E57 V58 G59 I60

T61 P62 Q63 H64 F65 S66 L67 F68 E69 K70 L71 E72 Y73 E74 F75 P76 D77 R78 H79 I80 T81 L82 W83 F84 W85 L86 V87 E88 R89 W90 E91 G92 E93 P94 W95 G96 K97 E98 G99 Q100 P101 G102 E103 W104 M105 S106 L107 V108 G109 L110 N111 A112 D113 D114 F115 P116 P117 A118 N119 E120

P121 V122 I123 A124 K125 L126 K127 R128 L129

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

The representative model is number 11. Colouring as in section 4.1 above.

- Molecule 1: NUCLEOSIDE TRIPHOSPHATE PYROPHOSPHOHYDROLASE

Chain A: 15% 50% 22% 5% 9%

M1 K2 K3 L4 Q5 I6 A7 V8 G9 I10 I11 R12 N13 E14 N15 N16 E17 I18 F19 I20 T21 R22 R23 A24 A25 D26 A27 H28 M29 A30 N31 K32 L33 E34 F35 P36 G37 I40 E41 M42 G43 E44 T45 P46 E47 Q48 A49 V50 V51 R52 E53 L54 Q55 E56 E57 V58 G59 I60 T61

P62 Q63 H64 F65 S66 L67 F68 E69 K70 L71 E72 Y73 E74 F75 P76 D77 R78 H79 I80 T81 L82 W83 F84 W85 L86 V87 E88 R89 W90 E91 G92 E93 P94 W95 G96 K97 Q100 P101 G102 E103 W104 M105 S106 L107 V108 G109 L110 N111 A112 D113 D114 F115 P116 P117 A118 N119 E120 P121 V122

I123 A124 K125 L126 K127 R128 L129

5 Refinement protocol and experimental data overview (i)

Of the ? calculated structures, 15 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. 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.

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
1	A	0.0±0.0	5.3±0.9
All	All	0	79

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 6 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	52	ARG	Sidechain	14
1	A	89	ARG	Sidechain	14
1	A	78	ARG	Sidechain	14
1	A	22	ARG	Sidechain	13
1	A	128	ARG	Sidechain	12

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	972	945	945	171±11
All	All	14580	14175	14175	2566

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 89.

5 of 1281 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:ILE:HD12	1:A:86:LEU:HD21	1.06	1.14	1	1
1:A:40:ILE:HD11	1:A:83:TRP:CZ2	1.05	1.85	8	1
1:A:8:VAL:HG12	1:A:82:LEU:HD11	1.05	1.21	3	1
1:A:11:ILE:HG23	1:A:87:VAL:HG21	1.05	1.26	2	11
1:A:10:ILE:HD12	1:A:86:LEU:HD11	1.04	1.24	5	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	117/129 (91%)	78±4 (66±3%)	30±4 (25±3%)	9±2 (8±2%)	2 15
All	All	1755/1935 (91%)	1167 (66%)	447 (25%)	141 (8%)	2 15

5 of 43 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	36	PRO	13
1	A	118	ALA	10
1	A	68	PHE	8
1	A	42	MET	8
1	A	61	THR	7

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	103/110 (94%)	66±4 (64±3%)	37±4 (36±3%)	1 8
All	All	1545/1650 (94%)	986 (64%)	559 (36%)	1 8

5 of 91 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	89	ARG	15
1	A	21	THR	15
1	A	60	ILE	14
1	A	52	ARG	14
1	A	45	THR	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