wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 02:28 PM BST

PDB ID	:	1EQ1
Title	:	NMR STRUCTURE OF AN EXCHANGEABLE APOLIPOPROTEIN-MAN DUCA SEXTA APOLIPOPHORIN-III
Authors	:	Wang, J.; Sykes, B.D.; Ryan, R.O.
Deposited on	:	2000-03-31

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:

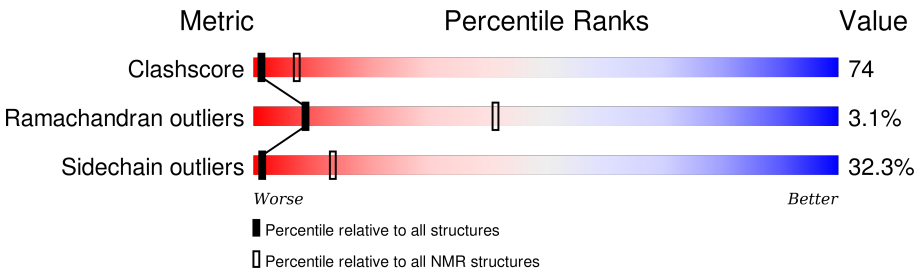
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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	166	8% 75% 5% • 11%

2 Ensemble composition and analysis (i)

This entry contains 25 models. Model 21 is the overall representative, medoid model (most similar to other models). The authors have identified model 2 as representative, based on the following criterion: *closest to the average, 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:10-A:35, A:39-A:128, A:135-A:165 (147)	0.48	21

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 4 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2566 atoms, of which 1273 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called APOLIPOPHORIN-III.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	166	2566	795	1273	228	268	2	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: APOLIPOPHORIN-III

Chain A: 8% 75% 5% • 11%

D1 A2 P3 A4 G5 G6 N7 A8 F9 E10 E11 M12 E13 K14 H15 A16 K17 E18 F19 Q20 K21 T22 F23 S24 E25 Q26 F27 N28 S29 L30 V31 N32 S33 K34 N35 T36 Q37 D38 F39 N40 K41 A42 L43 K44 D45 G46 S47 D48 S49 V50 L51 Q52 Q53 L54 S55 A56 F57 S58 S59 S60

L61 Q62 G63 A64 I65 S66 D67 A68 N69 G70 K71 A72 K73 E74 A75 L76 E77 Q78 A79 R80 Q81 N82 V83 E84 K85 T86 A87 E88 E89 L90 R91 K92 A93 H94 P95 D96 V97 E98 K99 E100 A101 N102 A103 F104 K105 D106 K107 L108 Q109 A110 A111 V112 Q113 T114 T115 V116 Q117 E118 S119 Q120

K121 L122 A123 K124 E125 S128 N129 M130 E131 E132 T133 N134 K135 K136 L137 A138 P139 K140 I141 K142 Q143 A144 Y145 D146 D147 F148 V149 K150 H151 A152 E153 E154 V155 Q156 K157 K158 L159 H160 E161 A162 A163 T164 K165 Q166

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

The representative model is number 21. Colouring as in section 4.1 above.

- Molecule 1: APOLIPOPHORIN-III

Chain A: 19% 51% 17% • 11%

D1 A2 P3 A4 G5 G6 N7 A8 F9 E10 E11 M12 E13 K14 H15 A16 K17 E18 F19 Q20 K21 T22 F23 S24 E25 Q26 F27 N28 S29 L30 V31 N32 S33 K34 N35 T36 Q37 D38 F39 N40 K41 A42 L43 K44 D45 G46 S47 D48 S49 V50 L51 Q52 Q53 L54 S55 A56 F57 S58 S59 S60

L61 Q62 G63 A64 I65 A68 N69 G70 K71 A72 K73 L76 E77 Q78 A79 R80 Q81 N82 K85 T86 A87 E88 E89 L90 R91 K92 A93 H94 P95 D96 V97 E100 A101 N102 A103 F104 K105 D106 K107 L108 Q109 A110 A111 V112 Q113 T114 T115 V116 Q117 E118 S119 L122 A123 K124 E125

V126 A127 S128 N129 M130 E131 E132 T133 N134 K135 K136 L137 A138 P139 K140 I141 K142 Q143 A144 Y145 D146 D147 F148 V149 K150 H151 A152 E153 E154 V155 Q156 K157 K158 L159 H160 E161 A162 A163 T164 K165 Q166

5 Refinement protocol and experimental data overview (i)

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

Of the 40 calculated structures, 25 were deposited, based on the following criterion: *structures with acceptable covalent geometry, structures with favorable non-bond energy, structures with the least restraint violations, structures with the lowest energy, target function*.

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

Software name	Classification	Version
X-PLOR	structure solution	3.1
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	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)
1	A	91	ARG	Sidechain	24
1	A	80	ARG	Sidechain	24

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	1153	1157	1157	172±14
All	All	28825	28925	28925	4288

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 74.

5 of 1729 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:76:LEU:HD11	1:A:137:LEU:HD21	1.14	1.18	14	12
1:A:50:VAL:HG21	1:A:155:VAL:HG13	1.11	1.20	19	4
1:A:54:LEU:HD11	1:A:87:ALA:HB2	1.07	1.21	24	21
1:A:30:LEU:HD12	1:A:159:LEU:HD13	1.06	1.28	10	1
1:A:43:LEU:HD22	1:A:104:PHE:CE2	1.03	1.88	3	1

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	147/166 (89%)	129±2 (88±2%)	13±3 (9±2%)	5±1 (3±1%)	9 41
All	All	3675/4150 (89%)	3237 (88%)	323 (9%)	115 (3%)	9 41

5 of 18 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	103	ALA	19
1	A	94	HIS	16
1	A	31	VAL	13
1	A	165	LYS	10
1	A	68	ALA	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	123/137 (90%)	83±4 (68±4%)	40±4 (32±4%)	1 13
All	All	3075/3425 (90%)	2083 (68%)	992 (32%)	1 13

5 of 110 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	73	LYS	25
1	A	124	LYS	24
1	A	140	LYS	22
1	A	80	ARG	20
1	A	91	ARG	20

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