wwPDB NMR Structure Validation Summary Report (i)

May 28, 2020 – 10:53 pm BST

PDB ID	:	2L6L
Title	:	Solution structure of human J-protein co-chaperone, Dph4
Authors	:	Thakur, A.; Chitoor, B.S.; Atreya, H.S.; Silva, P.D.
Deposited on	:	2010-11-23

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 (1)) were used in the production of this report:

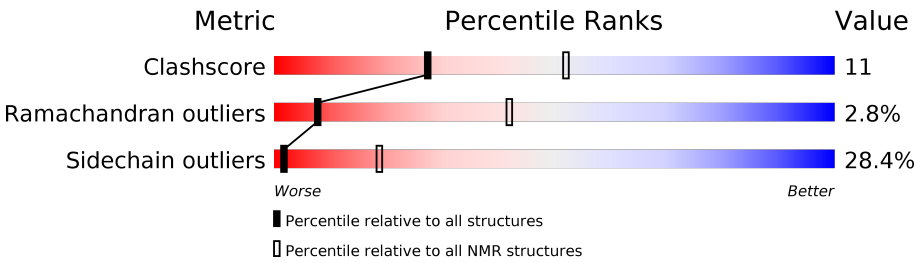
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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	155	52% 30% • 17%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 14 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest target function*.

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:73 (64)	1.10	14
2	A:78-A:88 (11)	0.06	11
3	A:94-A:147 (54)	0.57	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 4 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2238 atoms, of which 980 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DnaJ homolog subfamily C member 24.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	155	2237	786	980	212	248	11	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	SEE REMARK 999	UNP Q6P3W2
A	150	HIS	-	EXPRESSION TAG	UNP Q6P3W2
A	151	HIS	-	EXPRESSION TAG	UNP Q6P3W2
A	152	HIS	-	EXPRESSION TAG	UNP Q6P3W2
A	153	HIS	-	EXPRESSION TAG	UNP Q6P3W2
A	154	HIS	-	EXPRESSION TAG	UNP Q6P3W2
A	155	HIS	-	EXPRESSION TAG	UNP Q6P3W2

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	1	1	1

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: DnaJ homolog subfamily C member 24

Chain A: 52% 30% • 17%

M1 M2 A3 V4 E5 Q6 M7 P8 K9 K10 D11 W12 I15 L16 G17 A18 D19 D26 L27 K30 Y31 L34 I35 Y38 K42 Q43 D46 V47 P48 E61 I62 D63 Q64 I68 L69 G70 N71 T74 K75 R76 E77 L80 Q81 R82 L87 R88 N89 V90

G91 P92 V93 D94 L99 M102 S103 Y112 L113 K120 Y121 S122 Y123 S124 L133 I134 S135 T138 C139 S140 L141 I142 I143 E144 L145 L146 H147 Y148 N149 H150 H151 H152 H153 H154 H155

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: DnaJ homolog subfamily C member 24

Chain A: 52% 24% 7% 17%

M1 M2 A3 V4 E5 Q6 M7 P8 K9 K10 D11 W12 Y13 L16 G17 A18 N23 I24 S25 D26 L27 Y31 Q32 K33 L34 I35 Y38 Q43 S44 T45 D46 V47 P48 V52 Q57 I62 I68 L69 E72 E73 T74 K75 R76 E77 Y78 D79 L80 Q81 L87

R88 N89 V90 G91 P92 V93 D94 A95 Q96 M102 S103 L113 R116 Y121 S122 V123 S124 S132 L133 I134 S135 C139 S140 L141 I142 I143 E144 L145 L146 H147 Y148 N149 H150 H151 H152 H153 H154 H155

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
CYANA	structure solution	
CYANA	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: ZN

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	1038	813	986	22±6
All	All	20780	16260	19715	437

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 11.

5 of 266 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:35:ILE:HG21	1:A:62:ILE:HD12	0.95	1.36	17	1
1:A:35:ILE:HD12	1:A:62:ILE:HG23	0.95	1.34	18	1
1:A:35:ILE:HD12	1:A:62:ILE:HG22	0.93	1.40	3	3
1:A:133:LEU:HD11	1:A:142:ILE:HD13	0.82	1.51	11	2
1:A:18:ALA:HB2	1:A:27:LEU:HD21	0.81	1.52	4	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	129/155 (83%)	106±2 (82±2%)	20±3 (15±2%)	4±2 (3±1%)	8 42
All	All	2580/3100 (83%)	2115 (82%)	392 (15%)	73 (3%)	8 42

5 of 19 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	138	THR	15
1	A	68	ILE	9
1	A	71	ASN	8
1	A	46	ASP	7
1	A	140	SER	4

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	117/141 (83%)	84±4 (72±3%)	33±4 (28±3%)	2 19
All	All	2340/2820 (83%)	1675 (72%)	665 (28%)	2 19

5 of 99 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	143	ILE	15
1	A	38	TYR	15
1	A	138	THR	14
1	A	19	ASP	14
1	A	133	LEU	13

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)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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