wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 05:16 PM BST

PDB ID	:	1T8D
Title	:	Structure of the C-type lectin domain of CD23
Authors	:	Hibbert, R.G.; Teriete, P.; Grundy, G.J.; Sutton, B.J.; Gould, H.J.; McDonnell, J.M.
Deposited on	:	2004-05-12

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:

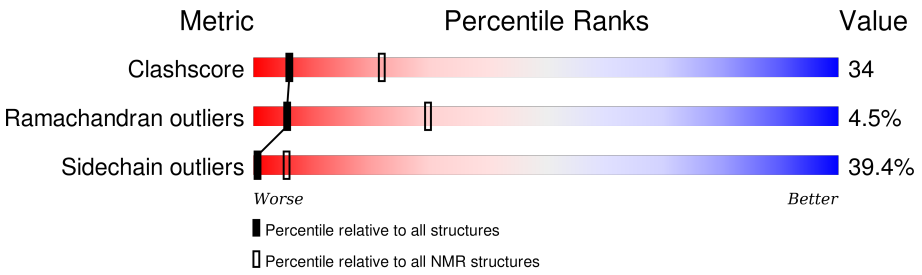
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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	143	18% 25% 11% • 44%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 3 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:11-A:24, A:28-A:58, A:63-A:70, A:77-A:89, A:104-A:106, A:113-A:115, A:123-A:130 (80)	0.14	3

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 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2168 atoms, of which 1035 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Low affinity immunoglobulin epsilon Fc receptor.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	143	2168	706	1035	201	215	11	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: Low affinity immunoglobulin epsilon Fc receptor

Chain A: 18% 25% 11% • 44%

S1 G2 F3 V4 C5 N6 T7 C8 P9 E10 K11 W12 I13 N14 R17 K18 C19 Y20 Y21 F22 G23 K24 G25 T26 K27 V30 H31 A32 R33 Y34 A35 C36 D37 D38 Q42 L43 V44 S45 I46 H47 Q52 D53 F54 L55 T56 K57 H58 A59 S60 H61 T62 G63 S64 W65 I66 G67

L68 R69 N70 L71 D72 L73 K74 G75 E76 F77 I78 W79 V80 V85 S88 N89 W90 A91 P92 G93 E94 P95 T96 S97 R98 S99 Q100 G101 E102 D103 C104 V105 M106 M107 R108 G109 S110 G111 R112 W113 N114 D115 A116 F117 C118 D119 R120 K121 L122 G123 A124 W125 V126 C127 D128 R129 L130 A131

T132 C133 T134 P135 P136 A137 S138 E139 G140 S141 A142 E143

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

The representative model is number 3. Colouring as in section 4.1 above.

- Molecule 1: Low affinity immunoglobulin epsilon Fc receptor

Chain A: 19% 24% 11% • 44%

S1 G2 F3 V4 C5 N6 T7 C8 P9 E10 K11 W12 I13 N14 F15 Q16 R17 K18 C19 Y20 Y21 F22 G23 K24 G25 T26 K27 V30 R33 Y34 A35 C36 D37 D38 M39 E40 G41 Q42 L43 V44 S45 I46 H47 Q52 D53 F54 L55 T56 K57 H58 A59 S60 H61 T62 G63 S64 W65

I66 G67 L68 R69 N70 L71 D72 L73 K74 G75 E76 F77 I78 W79 V80 V87 S88 N89 W90 A91 P92 G93 E94 P95 T96 S97 R98 S99 Q100 G101 E102 D103 C104 V105 M106 M107 R108 G109 S110 G111 R112 W113 N114 D115 A116 F117 C118 D119 R120 K121 L122 G123 A124 W125 V126 C127 D128 R129 L130

A131 T132 C133 T134 P135 P136 A137 S138 E139 G140 S141 A142 E143

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *STANDARD RECOMMENDED PROTOCOL*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *STRUCTURES WITH THE LOWEST ENERGY*.

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

Software name	Classification	Version
CNS	refinement	1.1
CNS	structure solution	1.1

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.

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	676	614	614	43±3
All	All	13520	12280	12280	866

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 34.

5 of 111 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:43:LEU:HB3	1:A:80:VAL:HG11	0.84	1.49	7	20
1:A:68:LEU:HD11	1:A:77:PHE:HB3	0.78	1.54	9	13
1:A:12:TRP:HA	1:A:21:TYR:HA	0.74	1.60	6	3
1:A:68:LEU:HD12	1:A:105:VAL:HG23	0.69	1.65	18	16
1:A:24:LYS:HA	1:A:124:ALA:HA	0.68	1.63	6	17

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	80/143 (56%)	68±1 (85±1%)	8±2 (10±2%)	4±1 (4±2%)	6 30
All	All	1600/2860 (56%)	1365 (85%)	163 (10%)	72 (4%)	6 30

5 of 7 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	17	ARG	20
1	A	44	VAL	20
1	A	88	SER	16
1	A	123	GLY	8
1	A	12	TRP	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	71/120 (59%)	43±3 (61±4%)	28±3 (39±4%)	1 6
All	All	1420/2400 (59%)	861 (61%)	559 (39%)	1 6

5 of 48 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	52	GLN	20
1	A	54	PHE	20
1	A	55	LEU	20
1	A	13	ILE	20
1	A	88	SER	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