wwPDB NMR Structure Validation Summary Report (i)

Jun 16, 2024 – 04:15 AM EDT

PDB ID	:	2L3O
Title	:	Solution structure of murine interleukin 3
Authors	:	Yao, S.; Young, I.G.; Norton, R.S.; Murphy, J.M.
Deposited on	:	2010-09-20

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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

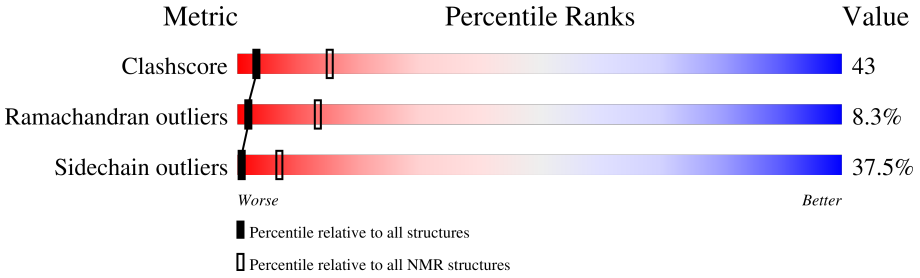
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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	127	16% 44% 15% • 22% •

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 8 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:43-A:138 (96)	1.05	8

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 5 clusters and 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1980 atoms, of which 994 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Interleukin 3.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	124	1980	614	994	177	192	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	GLY	-	EXPRESSION TAG	UNP Q5SX77
A	31	ALA	-	EXPRESSION TAG	UNP Q5SX77
A	32	MET	-	EXPRESSION TAG	UNP Q5SX77
A	105	ALA	CYS	ENGINEERED MUTATION	UNP Q5SX77

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: Interleukin 3

Chain A: 16% 44% 15% • 22% •

GLY ALA MET D33 T34 H35 R36 L37 T38 R39 T40 L41 N42 C43 I46 V47 K48 E49 I50 T51 G52 K53 L54 P55 E56 P57 E58 L59 K60 T61 D62 D63 E64 G65 P66 S67 L68 R69 N70 K71 S72 F73 R74 R75 V76 N77 L78 S79 K80 F81 Q85 G86 E87 V88 D89 P90 E91 D92

R93 Y94 V95 I96 K97 S98 N99 L100 Q101 K102 L103 L107 P108 T109 S110 A111 N112 D113 S114 A115 L116 P117 G118 V119 F120 I121 R122 D123 L124 D125 D126 F127 R128 K129 K130 L131 R132 M135 V136 H137 L138 N139 D140 L141 E142 T143 V144 L145 T146 S147 R148 P149 P150 Q151 P152 A153 S154 G155

S156

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

The representative model is number 8. Colouring as in section 4.1 above.

- Molecule 1: Interleukin 3

Chain A: 24% 30% 19% • 22% •

GLY ALA MET D33 T34 H35 R36 L37 T38 R39 T40 L41 N42 C43 S44 S45 I46 V47 K48 E49 I50 T51 G52 K53 L54 P55 E56 P57 E58 L59 K60 T61 D62 D63 P66 S67 L68 R69 N70 K71 S72 F73 R74 R75 V76 N77 L78 S79 K80 F81 Q85 G86 E87 V88 D89 P90 E91 D92

R93 Y94 V95 I96 K97 S98 N99 L100 Q101 T109 S110 A111 N112 D113 S114 A115 L116 P117 G118 V119 F120 I121 R122 D123 L124 D125 D126 F127 R128 K129 K130 L131 R132 F133 Y134 M135 V136 H137 L138 N139 D140 L141 E142 T143 V144 L145 T146 S147 R148 P149 P150 Q151 P152 A153 S154 G155 S156

5 Refinement protocol and experimental data overview (i)

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *on the basis of stereochemistry and energy considerations*.

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

Software name	Classification	Version
CYANA	structure solution	2.1
X-PLOR NIH	structure solution	2.17.0
X-PLOR NIH	refinement	2.17.0

No chemical shift data was provided.

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	6.8±0.4
All	All	0	135

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 7 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	69	ARG	Sidechain	20
1	A	74	ARG	Sidechain	20
1	A	93	ARG	Sidechain	20
1	A	122	ARG	Sidechain	20
1	A	132	ARG	Sidechain	20

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	770	782	782	66±10
All	All	15400	15640	15640	1320

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 43.

5 of 586 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:78:LEU:HD11	1:A:131:LEU:HD13	1.01	1.33	18	3
1:A:128:ARG:HA	1:A:131:LEU:HD22	0.98	1.32	9	4
1:A:76:VAL:HG11	1:A:113:ASP:HA	0.87	1.47	11	1
1:A:91:GLU:CB	1:A:124:LEU:HD21	0.86	2.00	9	8
1:A:46:ILE:HD13	1:A:103:LEU:HB2	0.86	1.47	6	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	96/127 (76%)	76±2 (79±2%)	12±2 (12±2%)	8±1 (8±1%)	2	13
All	All	1920/2540 (76%)	1523 (79%)	237 (12%)	160 (8%)	2	13

5 of 22 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	95	VAL	20
1	A	60	LYS	19
1	A	94	TYR	18
1	A	66	PRO	18
1	A	118	GLY	11

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	89/116 (77%)	56±4 (63±4%)	33±4 (37±4%)	1	7
All	All	1780/2320 (77%)	1113 (63%)	667 (37%)	1	7

5 of 76 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	81	PHE	20
1	A	127	PHE	20
1	A	131	LEU	20
1	A	46	ILE	18
1	A	49	GLU	18

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 monosaccharides 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