WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report ⓘ

Mar 6, 2022 – 11:29 AM EST

PDB ID	:	2KH2
Title	:	Solution structure of a scFv-IL-1B complex
Authors	:	Wilkinson, I.C.; Hall, C.J.; Veverka, V.; Muskett, F.W.; Stephens, P.E.; Taylor, R.J.; Henry, A.J.; Carr, M.D.
Deposited on	:	2009-03-23

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 ⓘ symbol.

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

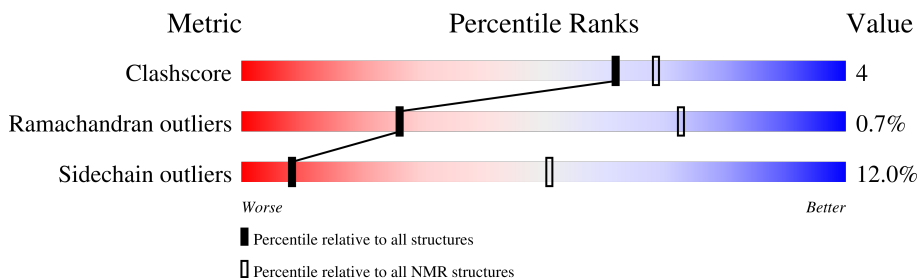|		
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

1 Overall quality at a glance ⓘ

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	Percentile Ranks	Value
Clashscore		4
Ramachandran outliers		0.7%
Sidechain outliers		12.0%

Worse   *Better*

▊ Percentile relative to all structures

◻ Percentile relative to all NMR structures

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	153	88% 12%
2	B	254	75% 17% 7%

2   Ensemble composition and analysis (i)

This entry contains 77 models. Model 1 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:1-A:153, B:1-B:110, B:129-B:254 (389)	0.56	1

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 11 clusters and 9 single-model clusters were found.

Cluster number	Models
1	1, 3, 7, 8, 15, 16, 17, 19, 20, 21, 24, 25, 29, 33, 42, 55, 58, 60
2	23, 31, 35, 37, 39, 40, 48, 56, 62, 66
3	45, 57, 64, 67, 68, 71, 74, 75, 77
4	5, 12, 30, 34, 43, 46, 59, 70, 72
5	6, 11, 22, 38, 54, 69
6	10, 14, 49, 53
7	36, 50, 61
8	2, 4, 41
9	47, 65
10	9, 27
11	44, 63
Single-model clusters	13; 18; 26; 28; 32; 51; 52; 73; 76

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6250 atoms, of which 3099 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Interleukin-1 beta.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	153	2457	773	1238	201	237	8	0

- Molecule 2 is a protein called scFv.

| Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	254	3793	1208	1861	345	372	7	0

4 Residue-property plots ⓘ

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.

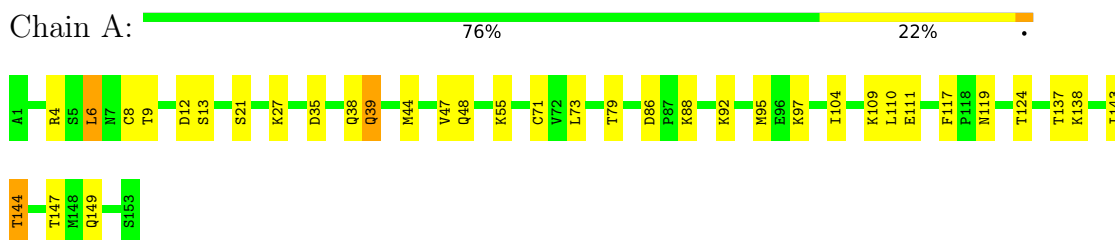• Molecule 1: Interleukin-1 beta

Chain A:

• Molecule 2: scFv

Chain B:
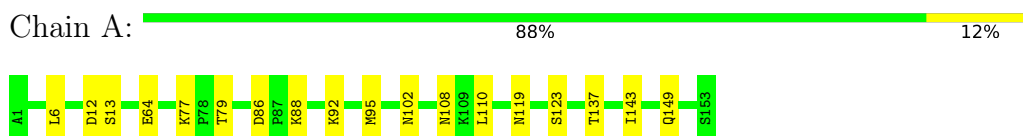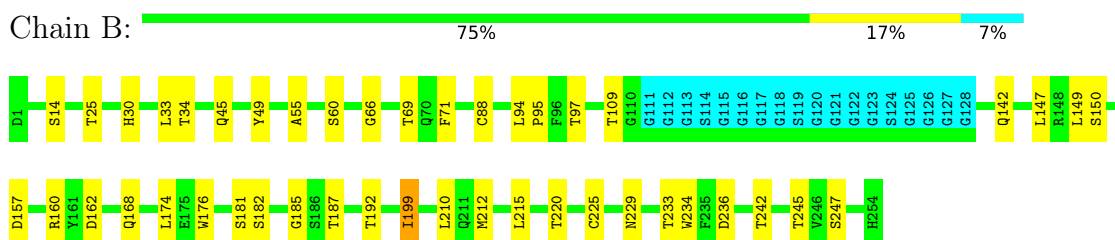
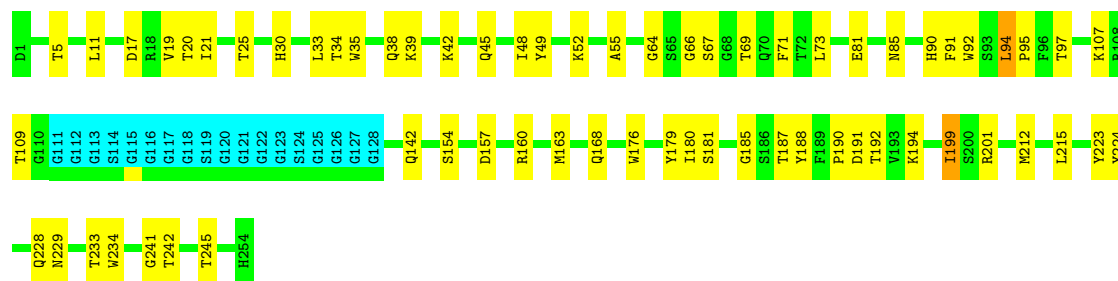## 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

• Molecule 1: Interleukin-1 beta

Chain A:

• Molecule 2: scFv

Chain B:

D1 T5 L11 D17 R18 V19 T20 I21 T25 H30 L33 T34 W35 Q38 K39 K42 Q45 I48 Y49 K52 A55 G64 S65 G66 S67 G68 T69 Q70 F71 I72 L73 E81 N85 H90 F91 W92 S93 L94 P95 F96 T97 K107 R108

T109 G110 G111 G112 G113 S114 G115 G116 G117 G118 S119 G120 G121 G122 G123 S124 G125 G126 G127 G128 Q142 S154 D157 R160 M163 Q168 W176 Y179 I180 S181 G185 S186 T187 Y188 F189 P190 D191 T192 V193 K194 I199 S200 R201 M212 L215 Y223 Y224

Q228 N229 T233 W234 G241 T242 T245 H254

5 Refinement protocol and experimental data overview (i)

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

Of the 77 calculated structures, 77 were deposited, based on the following criterion: *all calculated structures submitted.*

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

Software name	Classification	Version
HADDOCK	structure solution	2.0
HADDOCK	refinement	2.0

No chemical shift data was provided.

6 Model quality ⓘ

6.1 Standard geometry ⓘ

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	0.1±0.5	0.0±0.0
2	B	0.0±0.2	0.0±0.0
All	All	8	0

There are no bond-length outliers.

There are no bond-angle outliers.

5 of 8 unique chiral outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Models (Total)
2	B	76	SER	CA	1
1	A	58	VAL	CA	1
1	A	59	ALA	CA	1
1	A	1	ALA	CA	1
1	A	4	ARG	CA	1

There are no planarity outliers.

6.2 Too-close contacts ⓘ

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	1219	1238	1218	9±3
2	B	1854	1801	1773	20±5
All	All	236621	234003	230307	2095

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 4.

5 of 431 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
2:B:94:LEU:HG	2:B:95:PRO:HA	0.94	1.36	58	32
1:A:12:ASP:HB2	1:A:143:ILE:HG21	0.84	1.48	55	51
1:A:108:ASN:HB3	2:B:185:GLY:HA3	0.83	1.49	23	10
2:B:142:GLN:HG2	2:B:248:SER:HA	0.81	1.49	54	5
2:B:168:GLN:HB2	2:B:174:LEU:HG	0.80	1.54	31	9

6.3   Torsion angles   ⓘ

6.3.1   Protein backbone   ⓘ

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	151/153 (99%)	142±2 (94±2%)	9±2 (6±2%)	1±1 (1±1%)	32	76
2	B	234/254 (92%)	217±3 (93±1%)	15±3 (6±1%)	2±1 (1±0%)	24	71
All	All	29645/31339 (95%)	27623 (93%)	1820 (6%)	202 (1%)	26	73

5 of 35 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	229	ASN	42
2	B	30	HIS	29
2	B	185	GLY	27
1	A	106	ILE	15
1	A	21	SER	11

6.3.2   Protein sidechains   ⓘ

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	140/140 (100%)	128±3 (91±2%)	12±3 (9±2%)	14	61
2	B	200/203 (99%)	171±4 (86±2%)	29±4 (14±2%)	6	46
All	All	26180/26411 (99%)	23050 (88%)	3130 (12%)	8	51

5 of 197 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)
2	B	97	THR	73
2	B	192	THR	73
2	B	199	ILE	73
2	B	33	LEU	72
2	B	234	TRP	70

6.3.3   RNA ⓘ

There are no RNA molecules in this entry.

6.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6   Ligand geometry ⓘ

There are no ligands in this entry.

6.7   Other polymers ⓘ

There are no such molecules in this entry.

6.8   Polymer linkage issues ⓘ

There are no chain breaks in this entry.

7 Chemical shift validation (i)

No chemical shift data were provided