WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report ⓘ

Nov 7, 2023 – 07:05 pm GMT

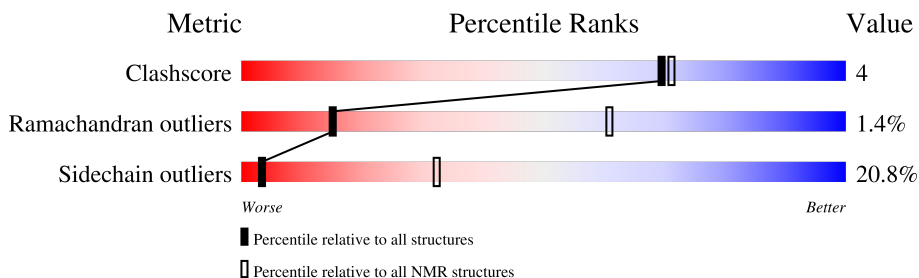|		
PDB ID	:	8BWW
BMRB ID	:	34778
Title	:	Targeting Toll-like receptor-driven systemic inflammation by engineering an innate structural fold into drugs
Authors	:	Diehl, C.
Deposited on	:	2022-12-07

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 is 87%.

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		1.4%
Sidechain outliers		20.8%

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	18	39%　28%　11%　22%

2 Ensemble composition and analysis ⓘ

This entry contains 10 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:2-A:9, A:11-A:13, A:15-A:17 (14)	0.73	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 2 clusters and 4 single-model clusters were found.

Cluster number	Models
1	1, 3, 5, 9
2	6, 10
Single-model clusters	2; 4; 7; 8

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 342 atoms, of which 179 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Thrombin light chain.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	18	342	113	179	28	22	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	MK8	GLN	engineered mutation	UNP P00734
A	14	MK8	ASP	engineered mutation	UNP P00734

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.

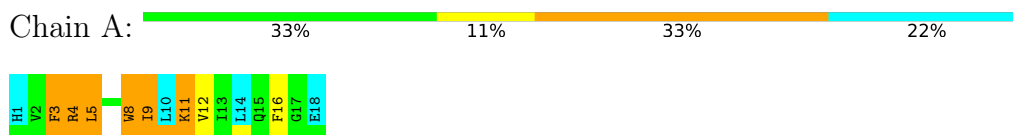• Molecule 1: Thrombin light chain
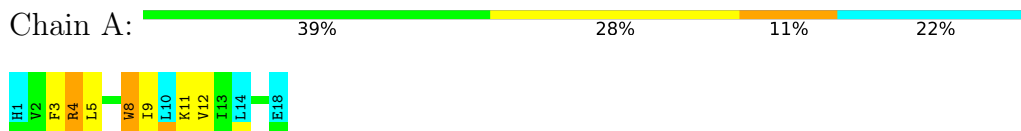
Chain A:

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: Thrombin light chain

Chain A:

5   Refinement protocol and experimental data overview ⓘ

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

Of the 60 calculated structures, 10 were deposited, based on the following criterion: *structures with the least restraint violations.*

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

Software name	Classification	Version
GROMACS	structure calculation	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	255
Number of shifts mapped to atoms	251
Number of unparsed shifts	0
Number of shifts with mapping errors	4
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MK8

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.61±0.02	0±0/127 (0.0± 0.0%)	2.25±0.20	6±2/166 (3.3± 1.4%)
All	All	0.61	0/1270 (0.0%)	2.26	55/1660 (3.3%)

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±0.5
All	All	0	19

There are no bond-length outliers.

5 of 30 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
1	A	4	ARG	NE-CZ-NH1	12.85	126.72	120.30	3	3
1	A	3	PHE	CB-CG-CD1	-12.63	111.96	120.80	3	3
1	A	16	PHE	CB-CG-CD2	-10.61	113.38	120.80	6	2
1	A	4	ARG	NE-CZ-NH2	-10.44	115.08	120.30	5	3
1	A	8	TRP	CB-CG-CD1	-9.37	114.82	127.00	2	3

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	9	ILE	Peptide	9

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	11	LYS	Peptide	5
1	A	4	ARG	Sidechain,Mainchain	2
1	A	6	LYS	Peptide	2
1	A	7	LYS	Peptide	1

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	125	142	142	1±1
All	All	1250	1420	1420	10

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 6 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:8:TRP:CZ2	1:A:9:ILE:CG1	0.49	2.96	7	1
1:A:5:LEU:HA	1:A:8:TRP:CD1	0.45	2.47	3	5
1:A:8:TRP:CD1	1:A:9:ILE:N	0.45	2.84	6	1
1:A:8:TRP:N	1:A:8:TRP:CD1	0.44	2.86	7	1
1:A:8:TRP:CZ2	1:A:9:ILE:HD11	0.44	2.48	9	1

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	14/18 (78%)	12±2 (82±11%)	2±1 (16±10%)	0±0 (1±3%)	15	61
All	All	140/180 (78%)	115 (82%)	23 (16%)	2 (1%)	15	61

All 2 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	12	VAL	1
1	A	4	ARG	1

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	13/15 (87%)	10±1 (79±7%)	3±1 (21±7%)	3	32
All	All	130/150 (87%)	103 (79%)	27 (21%)	3	32

All 5 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	4	ARG	9
1	A	8	TRP	9
1	A	5	LEU	5
1	A	12	VAL	3
1	A	3	PHE	1

6.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
1	MK8	A	10	1	5,8,9	0.88±0.12	0±0 (2±6%)
1	MK8	A	14	1	5,8,9	0.91±0.15	0±0 (4±8%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
1	MK8	A	10	1	4,10,12	2.36±0.78	1±1 (30±15%)
1	MK8	A	14	1	4,10,12	1.42±0.45	1±0 (17±11%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MK8	A	10	1	-	0±0,6,8,11	-
1	MK8	A	14	1	-	0±0,6,8,11	-

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	14	MK8	CB-CA	2.48	1.52	1.55	3	2
1	A	10	MK8	CB-CA	2.04	1.52	1.55	9	1

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	10	MK8	CB1-CA-CB	6.92	98.81	111.00	2	9
1	A	14	MK8	CB1-CA-CB	4.27	103.48	111.00	1	5
1	A	10	MK8	CB1-CA-C	3.88	123.59	108.99	8	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
1	A	10	MK8	CB-CG-CD	2.41	105.67	113.14	3	1
1	A	14	MK8	CB-CG-CD	2.28	120.17	113.14	6	2

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

7 Chemical shift validation ⓘ

The completeness of assignment taking into account all chemical shift lists is 87% for the well-defined parts and 85% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list*

7.1.1 Bookkeeping ⓘ

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	255
Number of shifts mapped to atoms	251
Number of unparsed shifts	0
Number of shifts with mapping errors	4
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. All 4 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	10	MK8	HB2	2.189	0.004	2
1	A	10	MK8	HB3	2.12	0.013	2
1	A	14	MK8	HB2	2.113	0.009	2
1	A	14	MK8	HB3	1.696	0.007	2

7.1.2 Chemical shift referencing ⓘ

No chemical shift referencing corrections were calculated (not enough data).

7.1.3 Completeness of resonance assignments ⓘ

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 87%, i.e. 206 atoms were assigned a chemical

shift out of a possible 238. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	55/71 (77%)	28/29 (97%)	14/28 (50%)	13/14 (93%)
Sidechain	119/135 (88%)	81/88 (92%)	37/40 (92%)	1/7 (14%)
Aromatic	32/32 (100%)	16/16 (100%)	15/15 (100%)	1/1 (100%)
Overall	206/238 (87%)	125/133 (94%)	66/83 (80%)	15/22 (68%)

7.1.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots ⓘ

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: