Full wwPDB NMR Structure Validation Report (i)

Dec 13, 2023 – 08:47 PM EST

PDB ID	:	2JOB
Title	:	Solution structure of an antilipopolysaccharide factor from shrimp and its possible Lipid A binding site
Authors	:	Yang, Y.; Boze, H.; Chemardin, P.; Padilla, A.; Moulin, G.; Tassanakajon, A.; Pugniere, M.; Roquet, F.; Gueguen, Y.; Bachere, E.; Aumelas, A.
Deposited on	:	2007-03-02

This is a Full wwPDB NMR Structure Validation 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:

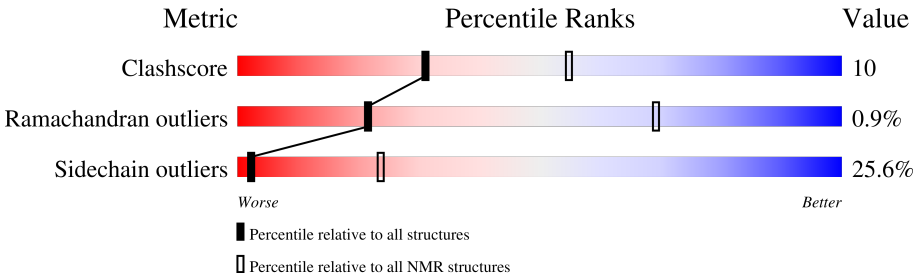
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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	102	46% 33% • 17%

2 Ensemble composition and analysis (i)

This entry contains 15 models. Model 3 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 energy*.

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:9-A:40, A:48-A:100 (85)	0.57	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. No single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 5, 6, 9, 10, 11, 12
2	8, 14, 15
3	4, 7, 13

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1626 atoms, of which 812 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called antilipopolysaccharide factor.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	102	1626	521	812	146	144	3	0

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: antilipopolysaccharide factor

Chain A: 46% 33% • 17%

E1 A2 Y3 V4 Q5 G6 W7 E8 V14 K17 I18 L21 W22 R23 N24 E25 L29 L30 C34 K35 F36 K39 P40 Y41 L42 K43 R44 F45 Q46 V47 Y48 Y49 K50 G51 R52 M53 W54 R62 G63 E64 A65 S66 T67 R68 S69 Q70 V73 K76 T77 A78 K79

D80 F81 V82 R83 K84 A85 F86 G89 L90 I91 E95 Q98 W99 L100 S101 S102

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: antilipopolysaccharide factor

Chain A: 51% 25% 7% 17%

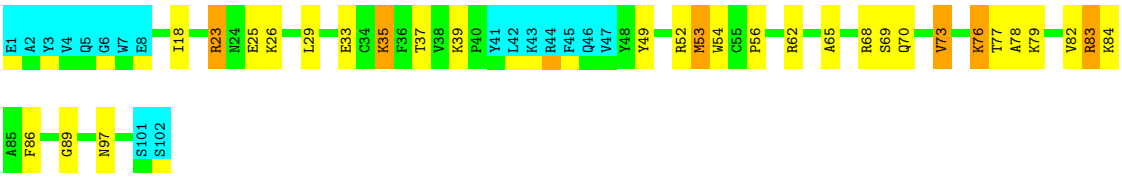
E1 A2 Y3 V4 Q5 G6 W7 E8 A12 S16 K17 I18 V19 G20 L21 W22 R23 K26 F36 T37 Y41 L42 K43 R44 F45 Q46 V47 Y48 Y49 K50 G51 R52 M53 W54 C55 P56 R62 A65 R68 S69 V73 A74 G75 K76 K79 D80 R83 K84 A85

F86 Q87 K88 G89 L90 I91 S92 Q93 Q98 W99 L100 S101 S102

4.2.2 Score per residue for model 2

• Molecule 1: antilipopolysaccharide factor

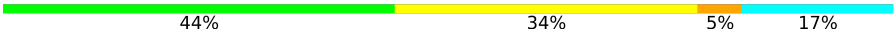
Chain A: 54% 24% 6% 17%

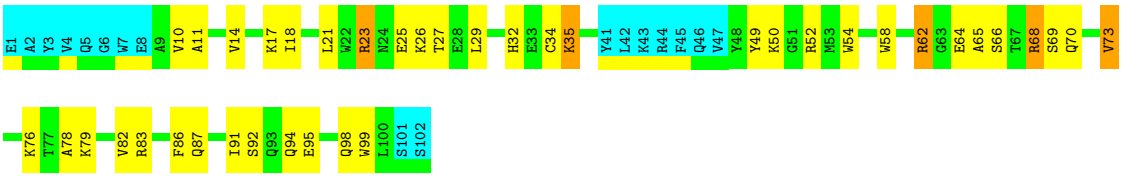


4.2.3 Score per residue for model 3 (medoid)

- Molecule 1: antilipopolysaccharide factor

Chain A:

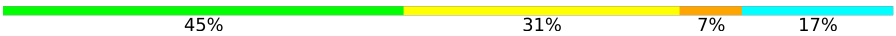


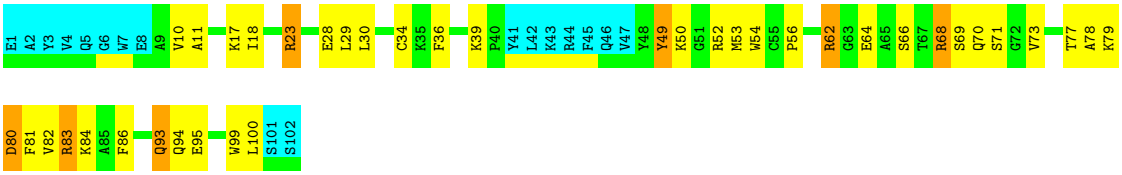


4.2.4 Score per residue for model 4

- Molecule 1: antilipopolysaccharide factor

Chain A:

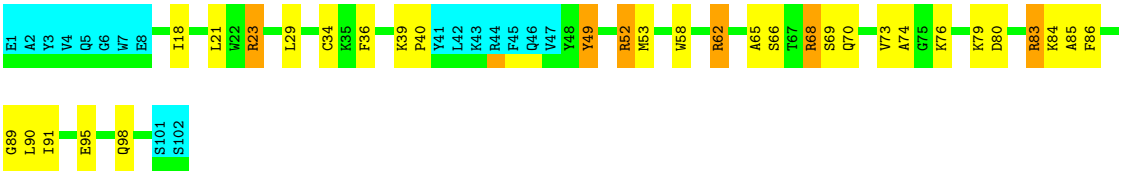




4.2.5 Score per residue for model 5

- Molecule 1: antilipopolysaccharide factor

Chain A:



4.2.6 Score per residue for model 6

- Molecule 1: antilipopolysaccharide factor

Chain A: 45% 30% 7% • 17%

E1 A2 Y3 V4 Q5 G6 W7 E8 A9 V10 V14 A15 S16 K17 I18 L21 W22 R23 K26 L29 L30 F36 T37 P40 Y41 L42 K43 R44 F45 Q46 V47 Y48 Y49 R52 M53 W54 C55 R62 G63 E64 A65 S66 T67 R68 S69 Q70 V73 A74 G75 K76 T77

A78 K79 D80 R83 K84 A85 F86 G89 L90 I91 A96 N97 Q98 W99 L100 S101 S102

4.2.7 Score per residue for model 7

- Molecule 1: antilipopolysaccharide factor

Chain A: 51% 24% 8% • 17%

E1 A2 Y3 V4 Q5 G6 W7 E8 A15 S16 K17 L21 W22 R23 N24 E25 E33 C34 K35 F36 T37 V38 K39 P40 Y41 L42 K43 R44 F45 Q46 V47 Y48 Y49 K50 G51 R52 M53 W54 C55 P56 I61 R62 G63 E64 R68 S69 V73 K76 T77 A78 K79 R83

K84 A85 F86 Q87 K88 G89 L90 I91 W99 L100 S101 S102

4.2.8 Score per residue for model 8

- Molecule 1: antilipopolysaccharide factor

Chain A: 45% 33% • • 17%

E1 A2 Y3 V4 Q5 G6 W7 E8 A9 V10 V14 A15 S16 K17 I18 L21 W22 R23 N24 E25 K26 T27 E28 L29 L30 E33 C34 K35 F36 K39 P40 Y41 L42 K43 R44 F45 Q46 V47 K50 G51 R52 M53 W54 R62 G63 E64 A65 S66 T67 R68 S69 V73

K76 T77 A78 K79 D80 R83 F86 Q87 K88 Q93 Q94 E95 W99 L100 S101 S102

4.2.9 Score per residue for model 9

- Molecule 1: antilipopolysaccharide factor

Chain A: 48% 31% • • 17%

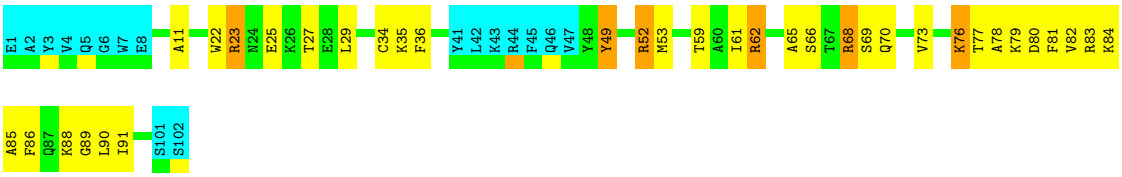
E1 A2 Y3 V4 Q5 G6 W7 E8 S16 K17 R23 L29 L30 E33 C34 K35 F36 T37 V38 K39 P40 Y41 L42 K43 R44 F45 Q46 V47 R52 M53 W54 C55 P56 T59 A60 I61 R62 G63 E64 T67 R68 V73 A74 G75 K76 T77 A78 K79 V82 R83 K84

A85 F86 Q87 K88 G89 L90 E95 Q98 W99 L100 S101 S102

4.2.10 Score per residue for model 10

- Molecule 1: antilipopolysaccharide factor

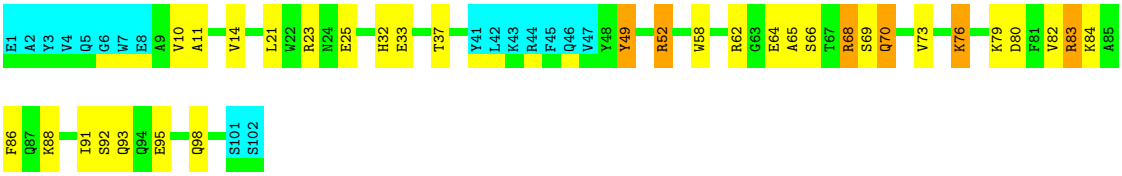
Chain A: 48% 29% 6% 17%



4.2.11 Score per residue for model 11

- Molecule 1: antilipopolysaccharide factor

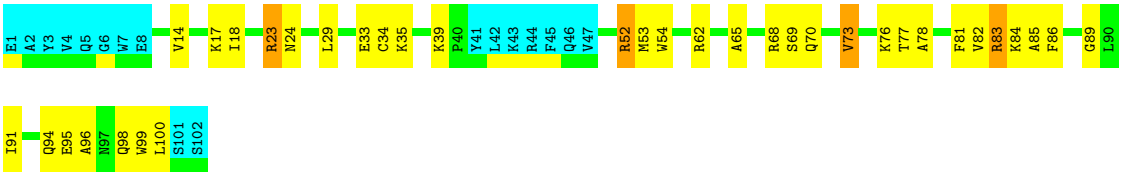
Chain A: 51% 26% 6% 17%



4.2.12 Score per residue for model 12

- Molecule 1: antilipopolysaccharide factor

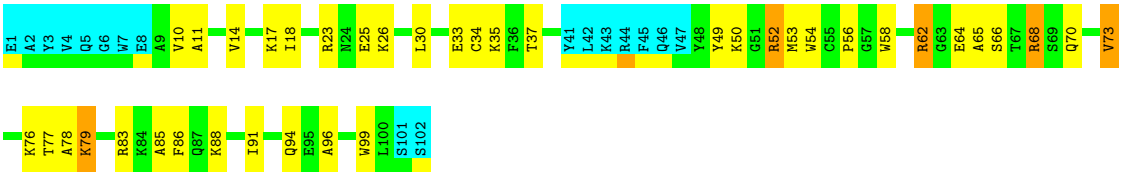
Chain A: 48% 31% • 17%



4.2.13 Score per residue for model 13

- Molecule 1: antilipopolysaccharide factor

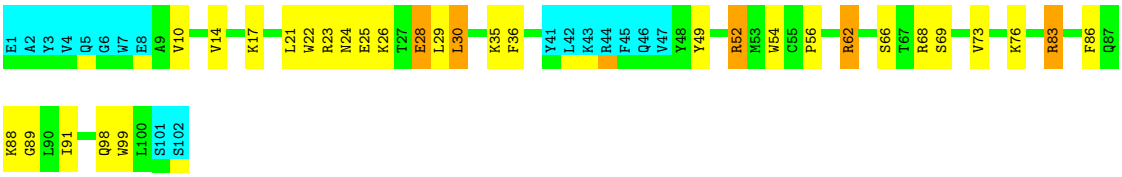
Chain A: 45% 33% 5% 17%



4.2.14 Score per residue for model 14

- Molecule 1: antilipopolysaccharide factor

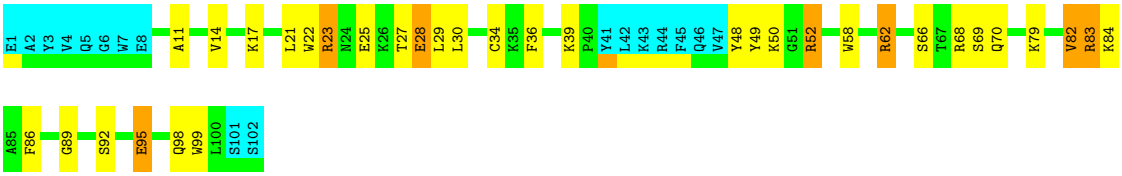
Chain A: 53% 25% 5% 17%



4.2.15 Score per residue for model 15

- Molecule 1: antilipopolysaccharide factor

Chain A: 50% 26% 7% 17%



5 Refinement protocol and experimental data overview (i)

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

Of the 60 calculated structures, 15 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
X-PLOR	structure solution	
X-PLOR	refinement	

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	5.0±0.0
All	All	0	75

There are no bond-length outliers.

There are no bond-angle outliers.

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	23	ARG	Sidechain	15
1	A	52	ARG	Sidechain	15
1	A	62	ARG	Sidechain	15
1	A	68	ARG	Sidechain	15
1	A	83	ARG	Sidechain	15

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	665	672	672	13±3
All	All	9975	10080	10080	194

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 10.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:15:ALA:HB2	1:A:79:LYS:HE2	0.88	1.45	7	1
1:A:15:ALA:HB2	1:A:79:LYS:CE	0.84	2.02	7	1
1:A:53:MET:HG2	1:A:77:THR:HG22	0.68	1.63	2	6
1:A:18:ILE:HD12	1:A:78:ALA:HB1	0.68	1.63	2	4
1:A:85:ALA:HB3	1:A:91:ILE:HD12	0.66	1.67	1	7
1:A:30:LEU:HD12	1:A:99:TRP:HD1	0.65	1.51	8	2
1:A:17:LYS:HD3	1:A:100:LEU:HD23	0.65	1.67	12	2
1:A:30:LEU:HD12	1:A:99:TRP:CD1	0.64	2.27	8	2
1:A:29:LEU:HD12	1:A:55:CYS:SG	0.63	2.32	6	1
1:A:82:VAL:HG13	1:A:91:ILE:HG21	0.63	1.68	11	1
1:A:11:ALA:HA	1:A:82:VAL:HG11	0.61	1.73	11	4
1:A:28:GLU:C	1:A:29:LEU:HD23	0.60	2.15	14	3
1:A:14:VAL:CG2	1:A:96:ALA:HB1	0.59	2.28	13	3
1:A:17:LYS:HD2	1:A:100:LEU:HD23	0.59	1.74	1	1
1:A:21:LEU:HD21	1:A:28:GLU:O	0.59	1.98	8	3
1:A:79:LYS:CE	1:A:79:LYS:HA	0.58	2.28	7	1
1:A:61:ILE:HD12	1:A:90:LEU:HD12	0.58	1.76	10	1
1:A:17:LYS:CD	1:A:100:LEU:HD23	0.58	2.27	12	1
1:A:18:ILE:HD11	1:A:29:LEU:HG	0.58	1.75	5	1
1:A:61:ILE:HD11	1:A:84:LYS:HB3	0.58	1.75	9	1
1:A:65:ALA:HB1	1:A:76:LYS:HB3	0.57	1.75	10	4
1:A:30:LEU:HD21	1:A:95:GLU:HG2	0.57	1.75	4	1
1:A:61:ILE:HD12	1:A:88:LYS:HD2	0.55	1.76	9	1
1:A:52:ARG:HG2	1:A:64:GLU:HB3	0.55	1.77	7	1
1:A:30:LEU:HD11	1:A:95:GLU:HG3	0.54	1.79	4	2
1:A:30:LEU:HD21	1:A:95:GLU:CG	0.54	2.32	4	1
1:A:17:LYS:HB3	1:A:99:TRP:CZ3	0.54	2.37	7	8
1:A:85:ALA:CB	1:A:91:ILE:HD12	0.53	2.34	7	1
1:A:65:ALA:HB1	1:A:76:LYS:HG2	0.53	1.81	2	1
1:A:29:LEU:HD12	1:A:34:CYS:SG	0.53	2.44	9	4
1:A:10:VAL:HG21	1:A:93:GLN:HG2	0.52	1.82	4	1
1:A:79:LYS:HA	1:A:79:LYS:NZ	0.51	2.20	7	1
1:A:19:VAL:HG22	1:A:74:ALA:HB1	0.51	1.82	1	1
1:A:78:ALA:O	1:A:82:VAL:HG23	0.51	2.05	9	4
1:A:54:TRP:CD1	1:A:56:PRO:HD3	0.51	2.41	4	7
1:A:14:VAL:HG13	1:A:99:TRP:CZ3	0.51	2.39	8	3
1:A:29:LEU:HB2	1:A:34:CYS:SG	0.50	2.46	4	3
1:A:14:VAL:HG21	1:A:91:ILE:CG2	0.50	2.36	14	1
1:A:35:LYS:HB3	1:A:54:TRP:CE2	0.50	2.41	7	3
1:A:10:VAL:O	1:A:14:VAL:HG23	0.50	2.07	8	6
1:A:65:ALA:HB1	1:A:76:LYS:CG	0.50	2.37	1	1
1:A:30:LEU:HD12	1:A:99:TRP:HB2	0.49	1.84	6	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:36:PHE:CD1	1:A:53:MET:HB3	0.49	2.43	6	5
1:A:18:ILE:HG21	1:A:78:ALA:HB2	0.49	1.85	4	3
1:A:29:LEU:HD13	1:A:81:PHE:CE2	0.49	2.43	10	1
1:A:22:TRP:CD1	1:A:36:PHE:HB3	0.48	2.44	14	4
1:A:17:LYS:HE3	1:A:100:LEU:HD23	0.48	1.84	9	1
1:A:49:TYR:CD2	1:A:70:GLN:HA	0.48	2.43	15	2
1:A:66:SER:HA	1:A:73:VAL:HG23	0.47	1.86	5	2
1:A:65:ALA:CB	1:A:76:LYS:HB3	0.47	2.39	11	3
1:A:14:VAL:HG21	1:A:91:ILE:HG21	0.47	1.86	14	1
1:A:10:VAL:HG21	1:A:93:GLN:CG	0.47	2.39	4	1
1:A:70:GLN:O	1:A:73:VAL:HG12	0.46	2.11	12	4
1:A:85:ALA:HB3	1:A:91:ILE:CD1	0.46	2.38	1	1
1:A:63:GLY:O	1:A:77:THR:HG23	0.46	2.11	8	1
1:A:28:GLU:O	1:A:29:LEU:HD23	0.45	2.12	15	2
1:A:53:MET:CG	1:A:77:THR:HG22	0.45	2.41	12	1
1:A:34:CYS:HA	1:A:56:PRO:HD2	0.45	1.88	4	3
1:A:92:SER:HB3	1:A:95:GLU:HB2	0.45	1.89	15	1
1:A:18:ILE:HD11	1:A:29:LEU:CG	0.45	2.42	5	1
1:A:79:LYS:O	1:A:82:VAL:HG12	0.45	2.12	15	1
1:A:32:HIS:CD2	1:A:58:TRP:CD1	0.45	3.05	3	2
1:A:49:TYR:CE2	1:A:70:GLN:HA	0.44	2.48	11	1
1:A:66:SER:HA	1:A:73:VAL:CG2	0.44	2.42	5	2
1:A:64:GLU:HG2	1:A:65:ALA:N	0.44	2.27	3	1
1:A:59:THR:HG21	1:A:90:LEU:HD13	0.44	1.90	10	1
1:A:11:ALA:HB1	1:A:82:VAL:CG1	0.44	2.43	15	1
1:A:82:VAL:HA	1:A:91:ILE:CD1	0.44	2.42	3	1
1:A:27:THR:HG23	1:A:34:CYS:HB2	0.44	1.90	15	2
1:A:52:ARG:HD3	1:A:64:GLU:HB3	0.44	1.89	9	1
1:A:27:THR:HG23	1:A:34:CYS:CB	0.43	2.44	3	2
1:A:18:ILE:HD11	1:A:29:LEU:HD21	0.43	1.89	2	1
1:A:15:ALA:HA	1:A:78:ALA:HB1	0.43	1.89	8	2
1:A:67:THR:HG22	1:A:73:VAL:HG23	0.43	1.90	6	1
1:A:61:ILE:HD12	1:A:88:LYS:CD	0.43	2.43	9	1
1:A:12:ALA:HA	1:A:79:LYS:HE3	0.43	1.91	1	1
1:A:68:ARG:HG3	1:A:69:SER:N	0.43	2.28	4	1
1:A:53:MET:SD	1:A:81:PHE:CD2	0.43	3.11	12	2
1:A:30:LEU:HD22	1:A:58:TRP:CZ2	0.43	2.48	13	2
1:A:40:PRO:HB3	1:A:49:TYR:CE1	0.43	2.49	5	1
1:A:36:PHE:CZ	1:A:77:THR:HB	0.43	2.48	10	1
1:A:52:ARG:CD	1:A:64:GLU:HB3	0.42	2.44	9	1
1:A:80:ASP:O	1:A:83:ARG:HG2	0.42	2.14	4	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:80:ASP:O	1:A:83:ARG:HG3	0.42	2.14	6	1
1:A:18:ILE:HD12	1:A:78:ALA:CB	0.42	2.41	8	1
1:A:59:THR:CB	1:A:90:LEU:HD11	0.42	2.44	9	1
1:A:11:ALA:HB1	1:A:79:LYS:HE2	0.42	1.91	13	1
1:A:35:LYS:HG2	1:A:54:TRP:CZ2	0.42	2.50	3	3
1:A:61:ILE:HD11	1:A:88:LYS:HG3	0.42	1.91	7	1
1:A:39:LYS:CD	1:A:39:LYS:N	0.41	2.83	8	1
1:A:29:LEU:HD23	1:A:29:LEU:N	0.41	2.30	14	1
1:A:49:TYR:CZ	1:A:70:GLN:HG3	0.41	2.50	10	1
1:A:50:LYS:HA	1:A:73:VAL:HG21	0.41	1.92	7	1
1:A:59:THR:HG21	1:A:90:LEU:CD1	0.41	2.45	10	1
1:A:49:TYR:O	1:A:73:VAL:HG21	0.41	2.15	1	1
1:A:52:ARG:HG3	1:A:64:GLU:HG3	0.41	1.92	8	1
1:A:36:PHE:CE1	1:A:77:THR:HG21	0.41	2.51	4	1
1:A:61:ILE:CD1	1:A:90:LEU:HD12	0.41	2.44	10	1
1:A:40:PRO:HB3	1:A:49:TYR:CD1	0.40	2.51	6	1
1:A:73:VAL:CG1	1:A:74:ALA:N	0.40	2.85	5	3
1:A:27:THR:OG1	1:A:29:LEU:HD21	0.40	2.16	15	1
1:A:58:TRP:CZ2	1:A:90:LEU:HD22	0.40	2.51	5	1
1:A:79:LYS:HZ2	1:A:79:LYS:HA	0.40	1.77	13	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	85/102 (83%)	77±1 (91±2%)	7±1 (8±2%)	1±1 (1±1%)	21	69
All	All	1275/1530 (83%)	1161 (91%)	103 (8%)	11 (1%)	21	69

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	89	GLY	10
1	A	67	THR	1

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	66/81 (81%)	49±3 (74±4%)	17±3 (26±4%)	2 24
All	All	990/1215 (81%)	737 (74%)	253 (26%)	2 24

All 41 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	86	PHE	14
1	A	52	ARG	12
1	A	69	SER	12
1	A	73	VAL	12
1	A	23	ARG	11
1	A	49	TYR	11
1	A	62	ARG	11
1	A	79	LYS	11
1	A	25	GLU	9
1	A	83	ARG	9
1	A	84	LYS	9
1	A	98	GLN	8
1	A	39	LYS	8
1	A	76	LYS	8
1	A	66	SER	8
1	A	37	THR	7
1	A	33	GLU	7
1	A	95	GLU	7
1	A	21	LEU	6
1	A	26	LYS	6
1	A	80	ASP	6
1	A	68	ARG	6
1	A	35	LYS	5
1	A	88	LYS	5
1	A	17	LYS	4
1	A	50	LYS	4
1	A	93	GLN	4
1	A	94	GLN	4
1	A	64	GLU	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	53	MET	3
1	A	70	GLN	3
1	A	28	GLU	3
1	A	30	LEU	3
1	A	16	SER	2
1	A	87	GLN	2
1	A	97	ASN	2
1	A	92	SER	2
1	A	24	ASN	2
1	A	71	SER	1
1	A	77	THR	1
1	A	82	VAL	1

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