



Full wwPDB NMR Structure Validation Report ⓘ

Mar 26, 2026 – 07:24 AM UTC

PDB ID : 2MSN / pdb_00002msn
BMRB ID : 25127
Title : NMR structure of a putative phosphoglycolate phosphatase (NP_346487.1) from *Streptococcus pneumoniae* TIGR4
Authors : Jaudzems, K.; Serrano, P.; Pedrini, B.; Geralt, M.; Wuthrich, K.; Joint Center for Structural Genomics (JCSG)
Deposited on : 2014-08-04

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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
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) : 2.49

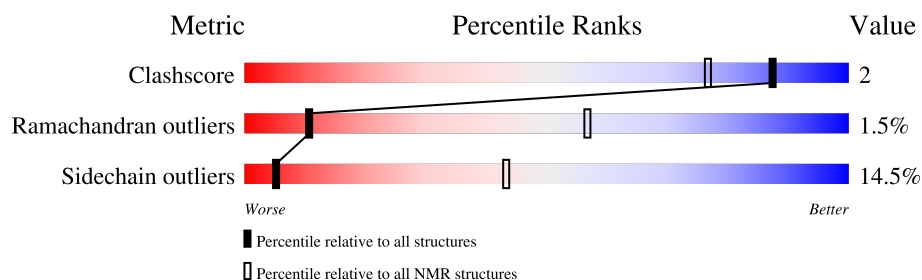
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 77%.

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	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	208	83% 10% 7%

2 Ensemble composition and analysis

This entry contains 20 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:17, A:84-A:204 (138)	0.68	1
2	A:19-A:73 (55)	0.56	18

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 4 clusters and 1 single-model cluster was found.

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3302 atoms, of which 1628 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Hydrolase, haloacid dehalogenase-like family.

Mol	Chain	Residues	Atoms						Trace
1	A	208	Total	C	H	N	O	S	0
			3302	1066	1628	276	330	2	

There are 2 discrepancies between the modelled and reference sequences:

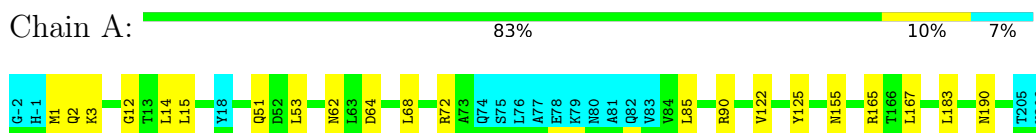
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q97NG6
A	-1	HIS	-	expression tag	UNP Q97NG6

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: Hydrolase, haloacid dehalogenase-like family

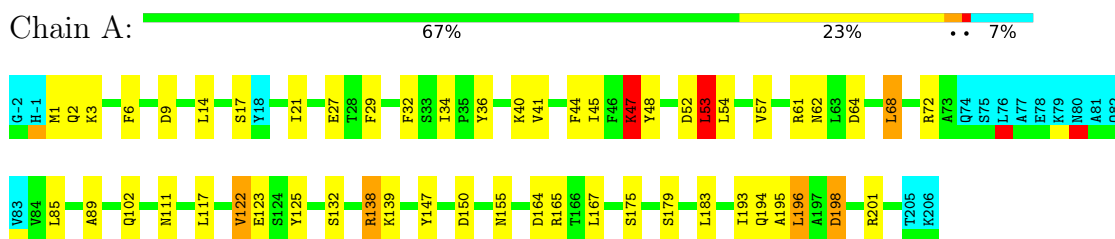


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

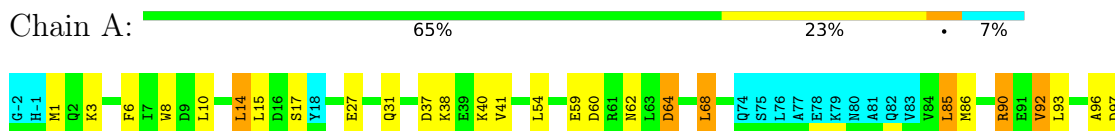
4.2.1 Score per residue for model 1 (medoid)

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



4.2.2 Score per residue for model 2

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

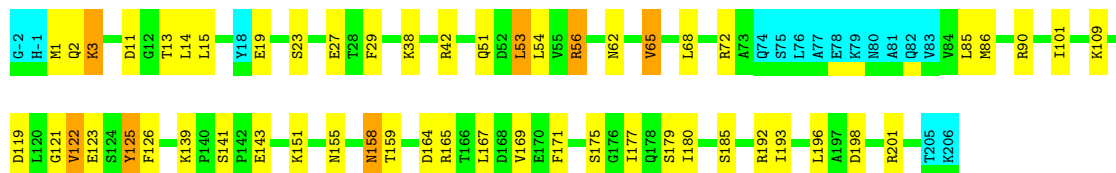




4.2.3 Score per residue for model 3

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

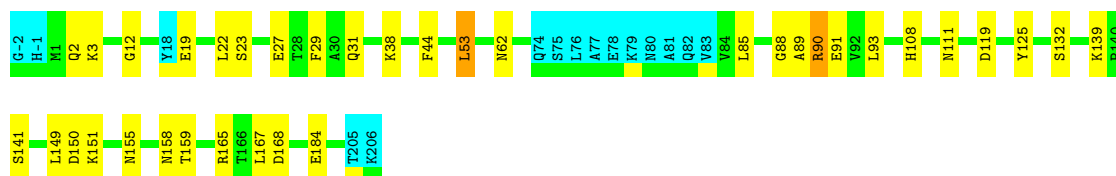
Chain A: 67% 23% 7%



4.2.4 Score per residue for model 4

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

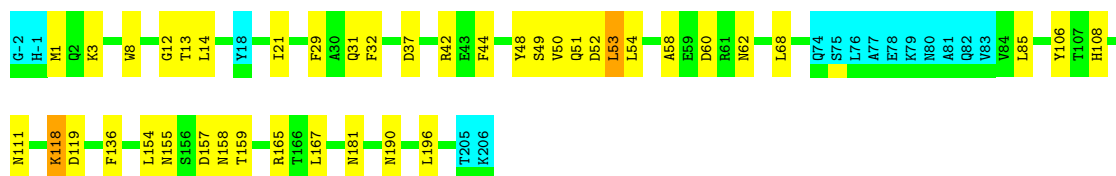
Chain A: 75% 16% 7%



4.2.5 Score per residue for model 5

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

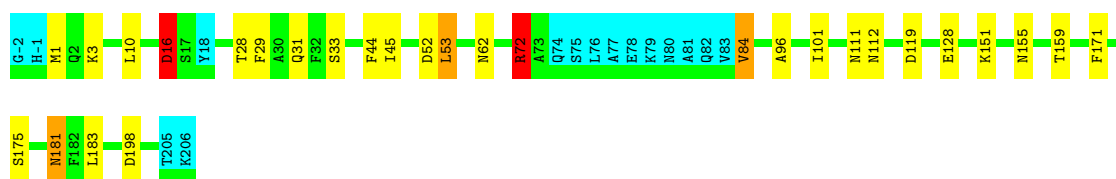
Chain A: 73% 19% 7%



4.2.6 Score per residue for model 6

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

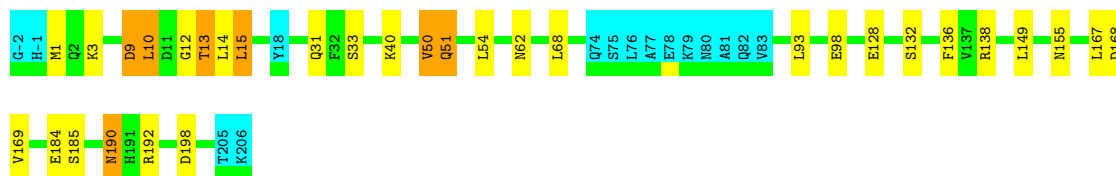
Chain A: 79% 12% 7%



4.2.7 Score per residue for model 7

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

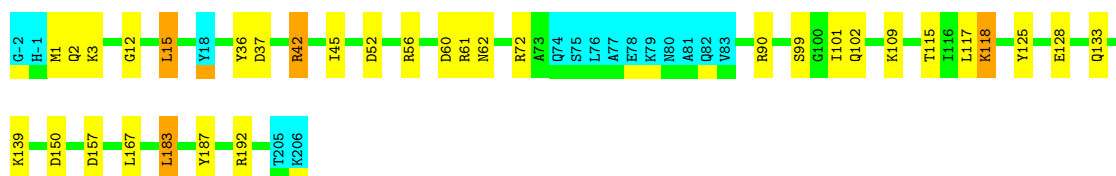
Chain A: 77% 12% 7%



4.2.8 Score per residue for model 8

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

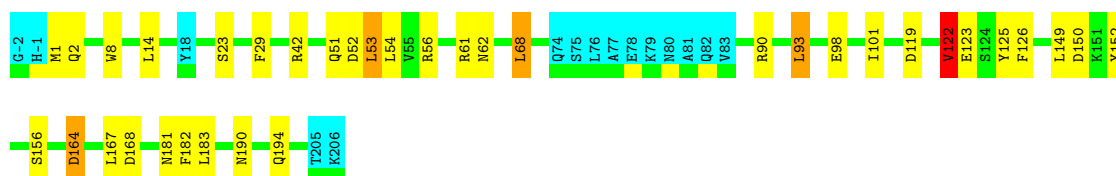
Chain A: 77% 14% 7%



4.2.9 Score per residue for model 9

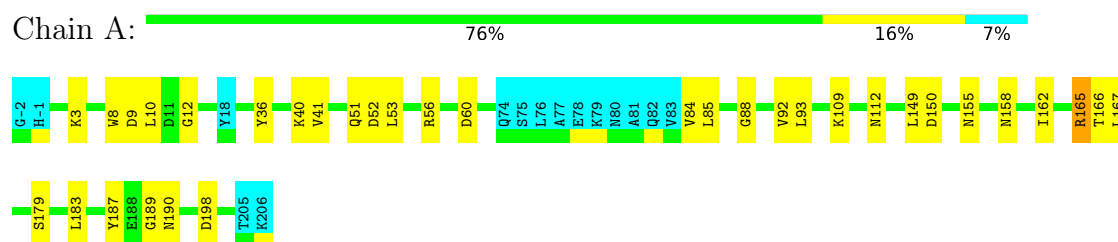
- Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 75% 15% 7%



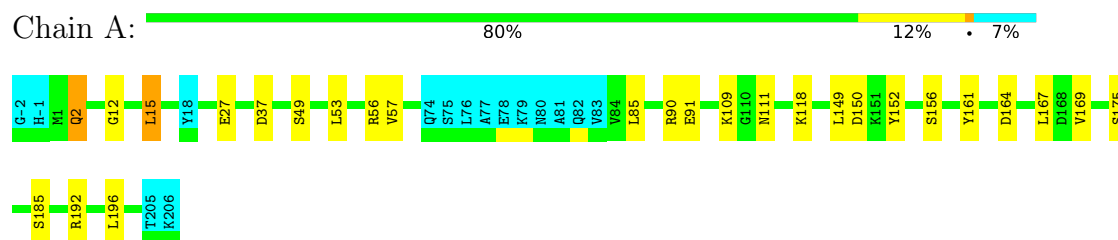
4.2.10 Score per residue for model 10

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



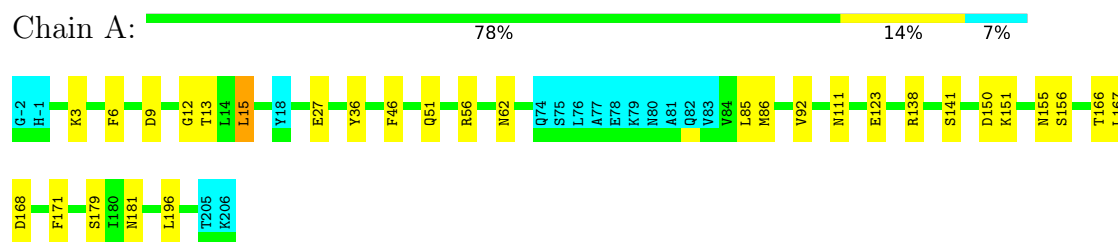
4.2.11 Score per residue for model 11

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



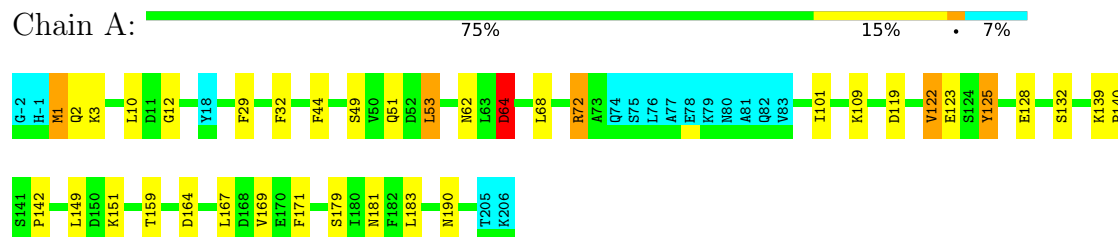
4.2.12 Score per residue for model 12

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



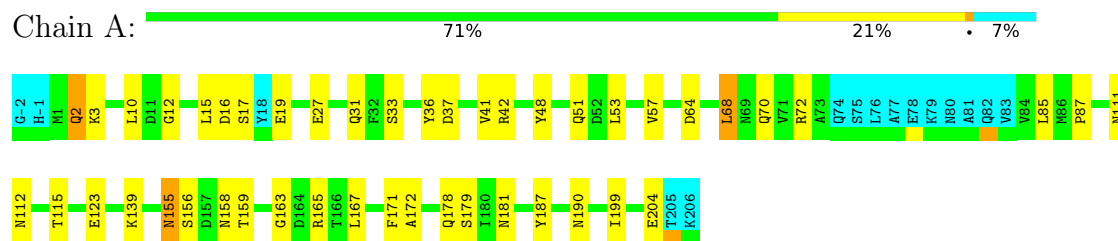
4.2.13 Score per residue for model 13

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



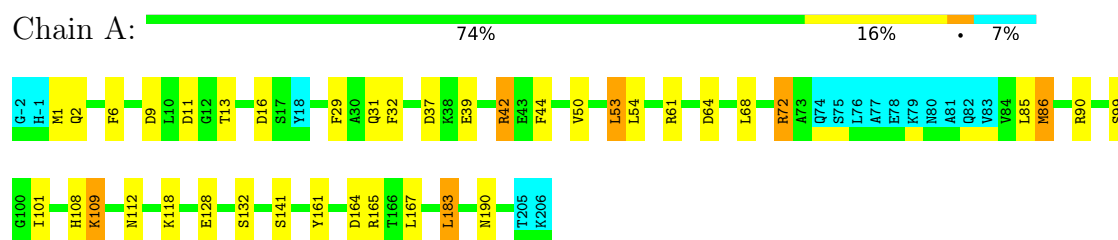
4.2.14 Score per residue for model 14

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



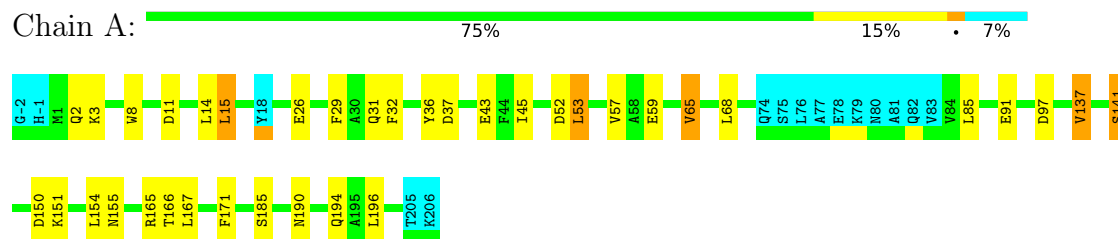
4.2.15 Score per residue for model 15

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



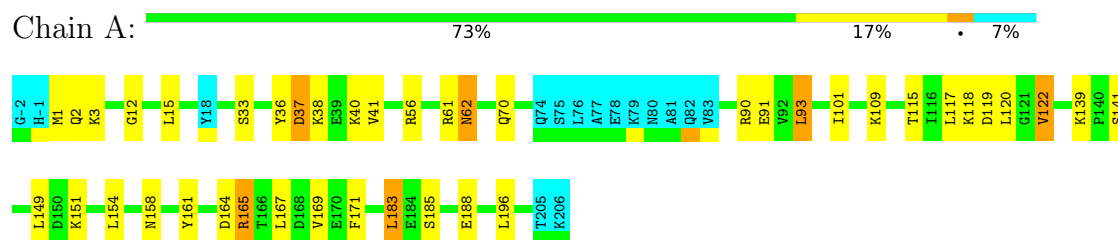
4.2.16 Score per residue for model 16

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



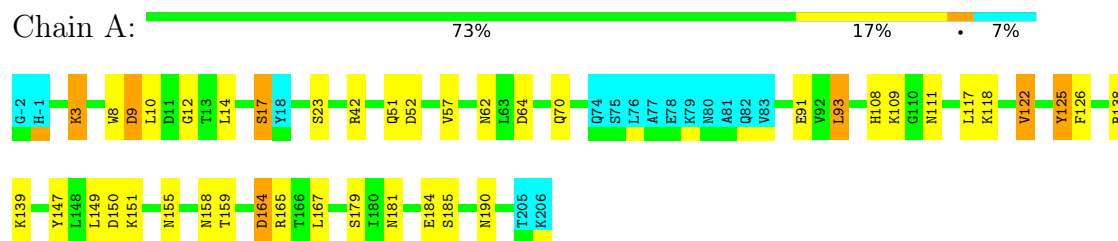
4.2.17 Score per residue for model 17

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



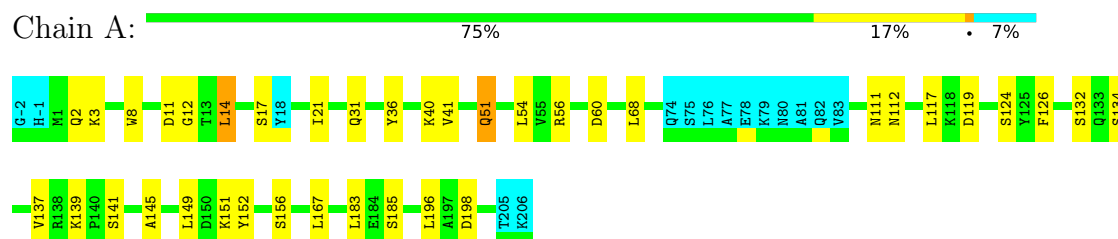
4.2.18 Score per residue for model 18

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



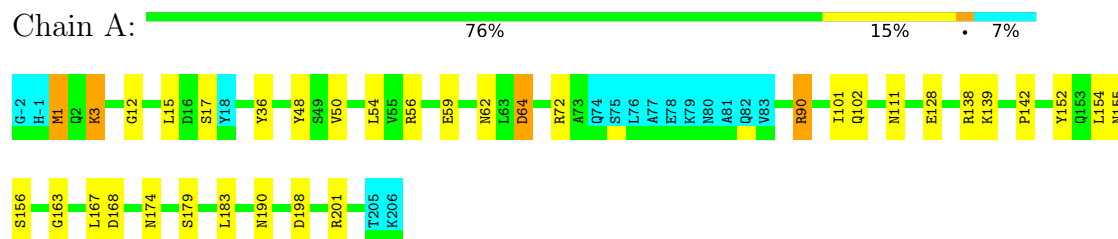
4.2.19 Score per residue for model 19

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



4.2.20 Score per residue for model 20

- Molecule 1: Hydrolase, haloacid dehalogenase-like family



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	refinement	3.0
CYANA	structure solution	3.0
OPAL	refinement	
UNIO-ATNOS/CANDID	structure solution	

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	2213
Number of shifts mapped to atoms	2213
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	77%

6 Model quality i

6.1 Standard geometry i

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	0.65±0.02	0±0/1588 (0.0± 0.0%)	1.40±0.04	7±3/2153 (0.3± 0.1%)
All	All	0.65	0/31760 (0.0%)	1.40	134/43060 (0.3%)

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	2.2±1.5
All	All	0	44

There are no bond-length outliers.

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(°)	Ideal(°)	Models	
								Worst	Total
1	A	16	ASP	CA-CB-CG	9.46	122.06	112.60	6	2
1	A	57	VAL	CB-CA-C	9.37	120.34	111.44	16	2
1	A	122	VAL	CA-CB-CG2	9.10	125.86	110.40	1	4
1	A	181	ASN	CA-CB-CG	8.49	121.09	112.60	2	5
1	A	61	ARG	N-CA-CB	-8.29	98.22	110.49	1	2
1	A	64	ASP	CA-CB-CG	8.01	120.61	112.60	20	4
1	A	158	ASN	CA-CB-CG	7.89	120.49	112.60	18	7
1	A	168	ASP	CA-CB-CG	7.81	120.41	112.60	9	2
1	A	97	ASP	CA-CB-CG	7.49	120.09	112.60	2	2
1	A	60	ASP	CA-CB-CG	7.33	119.93	112.60	2	3
1	A	65	VAL	CG1-CB-CG2	7.32	126.90	110.80	16	1
1	A	171	PHE	CA-CB-CG	7.24	121.04	113.80	12	5
1	A	9	ASP	CA-CB-CG	6.88	119.48	112.60	7	5
1	A	164	ASP	CA-CB-CG	6.80	119.40	112.60	9	5
1	A	190	ASN	CA-CB-CG	-6.67	105.93	112.60	14	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	198	ASP	CA-CB-CG	6.61	119.21	112.60	1	4
1	A	108	HIS	CA-C-N	6.23	133.44	121.54	15	1
1	A	108	HIS	C-N-CA	6.23	133.44	121.54	15	1
1	A	91	GLU	CB-CA-C	6.13	120.59	110.90	17	1
1	A	13	THR	CA-C-N	6.04	132.36	121.92	7	1
1	A	13	THR	C-N-CA	6.04	132.36	121.92	7	1
1	A	169	VAL	CB-CA-C	6.04	117.57	110.93	3	3
1	A	119	ASP	CA-CB-CG	5.94	118.54	112.60	5	1
1	A	118	LYS	CA-C-N	5.83	128.02	120.44	8	4
1	A	118	LYS	C-N-CA	5.83	128.02	120.44	8	4
1	A	62	ASN	CA-CB-CG	5.73	118.33	112.60	17	1
1	A	145	ALA	CA-C-N	5.64	128.11	120.38	19	1
1	A	145	ALA	C-N-CA	5.64	128.11	120.38	19	1
1	A	122	VAL	CB-CA-C	5.63	120.52	111.29	17	2
1	A	150	ASP	CA-CB-CG	-5.58	107.02	112.60	4	1
1	A	155	ASN	CA-CB-CG	5.57	118.17	112.60	14	1
1	A	126	PHE	CA-C-N	5.54	128.01	120.54	19	1
1	A	126	PHE	C-N-CA	5.54	128.01	120.54	19	1
1	A	52	ASP	CA-CB-CG	5.53	118.13	112.60	18	2
1	A	204	GLU	CA-C-N	5.47	131.99	121.54	14	1
1	A	204	GLU	C-N-CA	5.47	131.99	121.54	14	1
1	A	59	GLU	CA-C-N	5.44	127.56	120.28	20	2
1	A	59	GLU	C-N-CA	5.44	127.56	120.28	20	2
1	A	133	GLN	CA-C-N	5.43	128.09	120.28	8	1
1	A	133	GLN	C-N-CA	5.43	128.09	120.28	8	1
1	A	90	ARG	CA-C-N	5.42	127.86	120.54	20	1
1	A	90	ARG	C-N-CA	5.42	127.86	120.54	20	1
1	A	47	LYS	CB-CG-CD	5.42	123.76	111.30	1	1
1	A	93	LEU	N-CA-CB	-5.38	100.57	109.72	17	1
1	A	134	SER	N-CA-C	-5.33	104.05	110.88	19	1
1	A	96	ALA	CA-C-N	5.33	128.20	120.79	2	2
1	A	96	ALA	C-N-CA	5.33	128.20	120.79	2	2
1	A	110	GLY	CA-C-N	5.33	129.21	120.63	2	1
1	A	110	GLY	C-N-CA	5.33	129.21	120.63	2	1
1	A	136	PHE	CA-CB-CG	5.28	119.08	113.80	5	1
1	A	122	VAL	CG1-CB-CG2	-5.27	99.21	110.80	17	1
1	A	14	LEU	N-CA-CB	-5.22	102.86	110.53	7	1
1	A	124	SER	CA-C-N	5.19	127.96	120.38	2	1
1	A	124	SER	C-N-CA	5.19	127.96	120.38	2	1
1	A	108	HIS	CB-CG-CD2	-5.19	124.45	131.20	5	2
1	A	72	ARG	CD-NE-CZ	5.16	131.62	124.40	15	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	172	ALA	CA-C-N	5.15	127.14	120.44	14	1
1	A	172	ALA	C-N-CA	5.15	127.14	120.44	14	1
1	A	185	SER	CA-C-N	5.14	129.68	122.79	11	1
1	A	185	SER	C-N-CA	5.14	129.68	122.79	11	1
1	A	93	LEU	CB-CG-CD1	5.13	126.09	110.70	9	1
1	A	21	ILE	CA-C-N	5.12	127.14	120.28	1	2
1	A	21	ILE	C-N-CA	5.12	127.14	120.28	1	2
1	A	122	VAL	N-CA-CB	-5.12	102.78	111.23	1	1
1	A	53	LEU	CD1-CG-CD2	5.11	122.05	110.80	1	1
1	A	13	THR	CA-CB-CG2	5.11	119.18	110.50	7	1
1	A	50	VAL	CA-C-O	-5.08	114.44	120.78	7	1
1	A	17	SER	CA-C-N	5.05	127.75	120.38	19	1
1	A	17	SER	C-N-CA	5.05	127.75	120.38	19	1
1	A	68	LEU	CB-CG-CD1	5.03	125.80	110.70	1	1
1	A	58	ALA	CA-C-N	5.03	127.02	120.28	5	1
1	A	58	ALA	C-N-CA	5.03	127.02	120.28	5	1
1	A	142	PRO	CA-C-N	5.02	127.32	120.54	13	2
1	A	142	PRO	C-N-CA	5.02	127.32	120.54	13	2
1	A	141	SER	CA-CB-OG	5.01	121.13	111.10	16	1

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	56	ARG	Sidechain	6
1	A	125	TYR	Sidechain	5
1	A	90	ARG	Sidechain	4
1	A	36	TYR	Sidechain	4
1	A	165	ARG	Sidechain	3
1	A	72	ARG	Sidechain	3
1	A	42	ARG	Sidechain	3
1	A	187	TYR	Sidechain	3
1	A	161	TYR	Sidechain	3
1	A	147	TYR	Sidechain	2
1	A	61	ARG	Sidechain	2
1	A	48	TYR	Sidechain	1
1	A	201	ARG	Sidechain	1
1	A	106	TYR	Sidechain	1
1	A	60	ASP	Peptide	1
1	A	139	LYS	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	138	ARG	Sidechain	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	1557	1513	1513	6±3
All	All	31140	30260	30260	120

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 2.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:93:LEU:HD23	1:A:125:TYR:CD2	0.83	2.08	4	1
1:A:1:MET:HE2	1:A:101:ILE:HD11	0.76	1.57	2	9
1:A:90:ARG:HD3	1:A:93:LEU:HD22	0.72	1.59	4	1
1:A:85:LEU:HD21	1:A:122:VAL:HG11	0.69	1.63	2	1
1:A:36:TYR:CE2	1:A:41:VAL:HG21	0.67	2.25	1	1
1:A:36:TYR:CE1	1:A:41:VAL:HG21	0.66	2.25	19	4
1:A:28:THR:HG21	1:A:72:ARG:NH2	0.65	2.06	6	1
1:A:22:LEU:HD13	1:A:38:LYS:HD2	0.64	1.70	4	1
1:A:54:LEU:HD12	1:A:54:LEU:H	0.63	1.54	7	4
1:A:169:VAL:HG11	1:A:185:SER:HB3	0.60	1.72	7	1
1:A:85:LEU:CD2	1:A:122:VAL:HG11	0.60	2.27	2	1
1:A:15:LEU:N	1:A:15:LEU:HD13	0.59	2.12	11	3
1:A:122:VAL:HG23	1:A:125:TYR:CD2	0.59	2.32	3	6
1:A:32:PHE:CZ	1:A:68:LEU:HD12	0.59	2.32	1	3
1:A:89:ALA:HB3	1:A:90:ARG:NH2	0.55	2.16	4	1
1:A:15:LEU:HD13	1:A:15:LEU:H	0.53	1.63	11	2
1:A:29:PHE:CZ	1:A:53:LEU:HD11	0.53	2.39	1	9
1:A:37:ASP:O	1:A:41:VAL:HG23	0.53	2.03	14	3
1:A:44:PHE:HA	1:A:47:LYS:CG	0.52	2.34	1	1
1:A:28:THR:HG21	1:A:72:ARG:HH22	0.51	1.64	6	1
1:A:32:PHE:CE1	1:A:68:LEU:HD23	0.51	2.41	5	2
1:A:85:LEU:HD13	1:A:86:MET:N	0.51	2.21	15	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:122:VAL:HG22	1:A:126:PHE:CE2	0.51	2.41	2	4
1:A:64:ASP:O	1:A:68:LEU:HD13	0.50	2.06	1	2
1:A:44:PHE:HA	1:A:47:LYS:HG3	0.50	1.83	1	1
1:A:193:ILE:HD12	1:A:195:ALA:O	0.49	2.07	2	2
1:A:8:TRP:CZ3	1:A:162:ILE:HG21	0.49	2.42	2	2
1:A:8:TRP:CE3	1:A:14:LEU:HD13	0.49	2.42	9	4
1:A:15:LEU:CD1	1:A:15:LEU:N	0.48	2.76	16	1
1:A:53:LEU:O	1:A:57:VAL:HG23	0.48	2.08	11	2
1:A:15:LEU:N	1:A:15:LEU:CD1	0.48	2.77	11	2
1:A:64:ASP:O	1:A:68:LEU:HD22	0.48	2.07	2	2
1:A:16:ASP:CG	1:A:84:VAL:HG23	0.48	2.34	6	1
1:A:14:LEU:C	1:A:15:LEU:HD12	0.48	2.34	3	2
1:A:50:VAL:O	1:A:54:LEU:HD12	0.47	2.09	5	3
1:A:54:LEU:HD22	1:A:68:LEU:HB3	0.47	1.86	2	2
1:A:36:TYR:CZ	1:A:41:VAL:HG21	0.47	2.44	1	1
1:A:15:LEU:HD22	1:A:120:LEU:CD1	0.46	2.41	17	1
1:A:90:ARG:CD	1:A:93:LEU:HD22	0.46	2.36	4	1
1:A:90:ARG:HA	1:A:93:LEU:HD12	0.46	1.86	2	1
1:A:29:PHE:CZ	1:A:57:VAL:HG21	0.44	2.48	1	1
1:A:164:ASP:HB2	1:A:182:PHE:CD2	0.44	2.48	9	1
1:A:44:PHE:CD1	1:A:53:LEU:HB2	0.43	2.47	13	4
1:A:8:TRP:CZ3	1:A:14:LEU:HD13	0.43	2.48	5	1
1:A:34:ILE:HG21	1:A:57:VAL:HG11	0.43	1.89	1	1
1:A:183:LEU:HD12	1:A:183:LEU:H	0.43	1.72	2	1
1:A:89:ALA:HB3	1:A:90:ARG:CZ	0.43	2.44	4	1
1:A:183:LEU:HD13	1:A:183:LEU:N	0.43	2.29	15	1
1:A:51:GLN:HE21	1:A:51:GLN:HA	0.42	1.74	19	1
1:A:12:GLY:H	1:A:15:LEU:HD11	0.42	1.74	12	2
1:A:10:LEU:HA	1:A:15:LEU:HD21	0.42	1.90	7	1
1:A:51:GLN:HA	1:A:54:LEU:HD13	0.42	1.90	7	1
1:A:93:LEU:HB3	1:A:125:TYR:CZ	0.42	2.50	18	1
1:A:137:VAL:HG23	1:A:141:SER:CB	0.42	2.45	16	1
1:A:15:LEU:HD22	1:A:15:LEU:N	0.42	2.30	8	1
1:A:169:VAL:HG11	1:A:185:SER:CB	0.41	2.45	17	1
1:A:180:ILE:HG22	1:A:193:ILE:HD13	0.41	1.91	3	1
1:A:50:VAL:HG23	1:A:51:GLN:N	0.41	2.30	5	1
1:A:183:LEU:N	1:A:183:LEU:HD13	0.41	2.31	17	2
1:A:6:PHE:CD1	1:A:101:ILE:HG21	0.41	2.51	2	1
1:A:89:ALA:HB2	1:A:196:LEU:HD23	0.40	1.92	1	1
1:A:8:TRP:CH2	1:A:92:VAL:HG11	0.40	2.51	2	1
1:A:44:PHE:CE1	1:A:48:TYR:HB2	0.40	2.52	5	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:54:LEU:HB2	1:A:65:VAL:HG23	0.40	1.93	3	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	193/208 (93%)	171±4 (88±2%)	20±4 (10±2%)	3±2 (1±1%)	11	57
All	All	3860/4160 (93%)	3413 (88%)	391 (10%)	56 (1%)	11	57

All 19 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	3	LYS	13
1	A	12	GLY	11
1	A	122	VAL	5
1	A	1	MET	4
1	A	139	LYS	4
1	A	138	ARG	2
1	A	121	GLY	2
1	A	88	GLY	2
1	A	2	GLN	2
1	A	163	GLY	2
1	A	136	PHE	1
1	A	189	GLY	1
1	A	140	PRO	1
1	A	87	PRO	1
1	A	109	LYS	1
1	A	164	ASP	1
1	A	137	VAL	1
1	A	17	SER	1
1	A	50	VAL	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	169/181 (93%)	144±5 (86±3%)	24±5 (14±3%)	5	43
All	All	3380/3620 (93%)	2890 (86%)	490 (14%)	5	43

All 106 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	167	LEU	19
1	A	62	ASN	14
1	A	155	ASN	13
1	A	2	GLN	12
1	A	183	LEU	11
1	A	53	LEU	10
1	A	85	LEU	10
1	A	111	ASN	10
1	A	151	LYS	10
1	A	165	ARG	10
1	A	150	ASP	9
1	A	179	SER	9
1	A	31	GLN	9
1	A	109	LYS	9
1	A	51	GLN	9
1	A	149	LEU	9
1	A	196	LEU	8
1	A	159	THR	8
1	A	190	ASN	8
1	A	3	LYS	7
1	A	27	GLU	7
1	A	72	ARG	7
1	A	10	LEU	7
1	A	128	GLU	7
1	A	141	SER	7
1	A	119	ASP	7
1	A	15	LEU	7
1	A	40	LYS	6
1	A	117	LEU	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	123	GLU	6
1	A	132	SER	6
1	A	68	LEU	6
1	A	42	ARG	6
1	A	37	ASP	6
1	A	52	ASP	6
1	A	156	SER	6
1	A	139	LYS	5
1	A	175	SER	5
1	A	90	ARG	5
1	A	185	SER	5
1	A	13	THR	5
1	A	112	ASN	5
1	A	93	LEU	5
1	A	198	ASP	4
1	A	86	MET	4
1	A	11	ASP	4
1	A	23	SER	4
1	A	56	ARG	4
1	A	192	ARG	4
1	A	91	GLU	4
1	A	168	ASP	4
1	A	118	LYS	4
1	A	154	LEU	4
1	A	181	ASN	4
1	A	33	SER	4
1	A	152	TYR	4
1	A	6	PHE	3
1	A	9	ASP	3
1	A	14	LEU	3
1	A	102	GLN	3
1	A	138	ARG	3
1	A	194	GLN	3
1	A	17	SER	3
1	A	38	LYS	3
1	A	92	VAL	3
1	A	19	GLU	3
1	A	164	ASP	3
1	A	184	GLU	3
1	A	49	SER	3
1	A	115	THR	3
1	A	166	THR	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	64	ASP	3
1	A	70	GLN	3
1	A	45	ILE	2
1	A	201	ARG	2
1	A	124	SER	2
1	A	65	VAL	2
1	A	158	ASN	2
1	A	157	ASP	2
1	A	16	ASP	2
1	A	98	GLU	2
1	A	99	SER	2
1	A	122	VAL	2
1	A	48	TYR	2
1	A	171	PHE	2
1	A	47	LYS	1
1	A	131	THR	1
1	A	146	THR	1
1	A	170	GLU	1
1	A	143	GLU	1
1	A	177	ILE	1
1	A	108	HIS	1
1	A	60	ASP	1
1	A	84	VAL	1
1	A	1	MET	1
1	A	46	PHE	1
1	A	178	GLN	1
1	A	199	ILE	1
1	A	39	GLU	1
1	A	26	GLU	1
1	A	43	GLU	1
1	A	59	GLU	1
1	A	137	VAL	1
1	A	61	ARG	1
1	A	188	GLU	1
1	A	174	ASN	1

6.3.3 RNA ⓘ

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 oligosaccharides 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

The completeness of assignment taking into account all chemical shift lists is 77% for the well-defined parts and 77% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

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	2213
Number of shifts mapped to atoms	2213
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	7

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	205	-0.09 ± 0.08	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	196	0.44 ± 0.11	None needed (< 0.5 ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
^{15}N	196	0.24 ± 0.26	None needed (< 0.5 ppm)

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 77%, i.e. 2065 atoms were assigned a chemical shift out of a possible 2670. 0 out of 31 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	564/967 (58%)	191/392 (49%)	190/386 (49%)	183/189 (97%)
Sidechain	1331/1464 (91%)	905/949 (95%)	397/460 (86%)	29/55 (53%)

Continued on next page...

Continued from previous page...

	Total	¹H	¹³C	¹⁵N
Aromatic	170/239 (71%)	105/116 (91%)	63/119 (53%)	2/4 (50%)
Overall	2065/2670 (77%)	1201/1457 (82%)	650/965 (67%)	214/248 (86%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 77%, i.e. 2213 atoms were assigned a chemical shift out of a possible 2868. 0 out of 33 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹H	¹³C	¹⁵N
Backbone	608/1043 (58%)	207/423 (49%)	205/416 (49%)	196/204 (96%)
Sidechain	1425/1570 (91%)	969/1017 (95%)	424/493 (86%)	32/60 (53%)
Aromatic	180/255 (71%)	111/124 (90%)	67/126 (53%)	2/5 (40%)
Overall	2213/2868 (77%)	1287/1564 (82%)	696/1035 (67%)	230/269 (86%)

7.1.4 Statistically unusual chemical shifts [i](#)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	162	ILE	HD11	-1.47	-0.72 – 2.09	-7.7
1	A	162	ILE	HD12	-1.47	-0.72 – 2.09	-7.7
1	A	162	ILE	HD13	-1.47	-0.72 – 2.09	-7.7
1	A	151	LYS	HG2	-0.37	0.13 – 2.61	-7.0
1	A	200	SER	HB2	2.12	2.61 – 5.13	-7.0
1	A	173	GLN	HB2	0.59	0.80 – 3.29	-5.9
1	A	93	LEU	HB2	-0.14	-0.07 – 3.30	-5.2

7.1.5 Random Coil Index (RCI) plots [i](#)

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:

