Full wwPDB NMR Structure Validation Report

Nov 7, 2023 – 04:48 pm GMT

PDB ID	:	7ZAX
BMRB ID	:	34716
Title	:	Solution structure of thanatin-like derivative 7 in complex with K. pneumoniae LptA
Authors	:	Oi, K.K.; Moehle, K.; Zerbe, O.
Deposited on	:	2022-03-22

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:

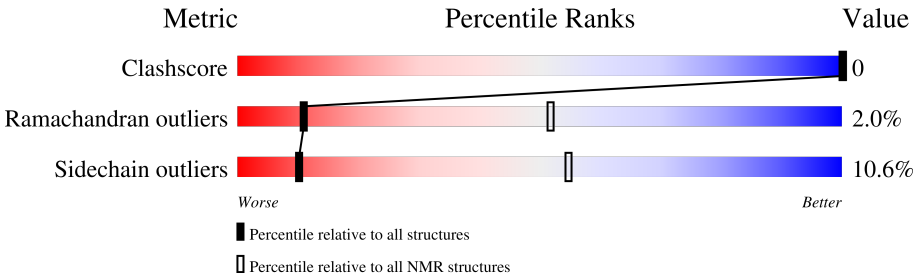
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	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 is 88%.

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	133	86% · 11%
2	B	16	50% 12% 38%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA and RNA chains that are outliers for geometric criteria:

Mol	Chain	Compound	Res	Total models with violations	
				Chirality	Geometry
2	B	EU0	206	-	2

2 Ensemble composition and analysis

This entry contains 20 models. Model 5 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *target function*.

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:35-A:152, B:208-B:210, B:212-B:213, B:215-B:215, B:217-B:218, B:220-B:221 (128)	1.28	5

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 and 1 single-model cluster was found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2357 atoms, of which 1181 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Lipopolysaccharide export system protein LptA.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	133	2071	642	1034	183	208	4	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	GLY	-	expression tag	UNP A0A2X3CC60

- Molecule 2 is a protein called Thanatin-like derivative.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	16	286	85	147	29	23	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	209	THR	ILE	engineered mutation	UNP P55788
B	211	LE1	CYS	modified residue	UNP P55788
B	214	DAB	ARG	modified residue	UNP P55788
B	216	4FO	GLY	modified residue	UNP P55788
B	219	DAB	GLN	modified residue	UNP P55788
B	221	TYR	MET	engineered mutation	UNP P55788

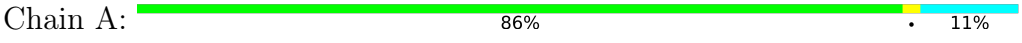
4 Residue-property plots (i)

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: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

Chain B:

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

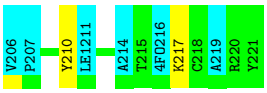
4.2.1 Score per residue for model 1

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:

• Molecule 2: Thanatin-like derivative

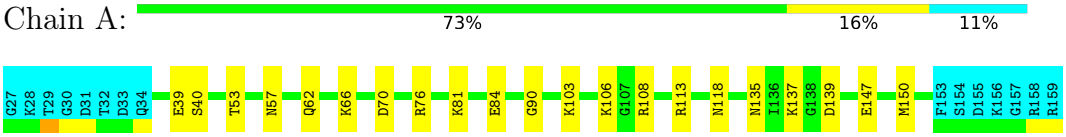
Chain B:



4.2.2 Score per residue for model 2

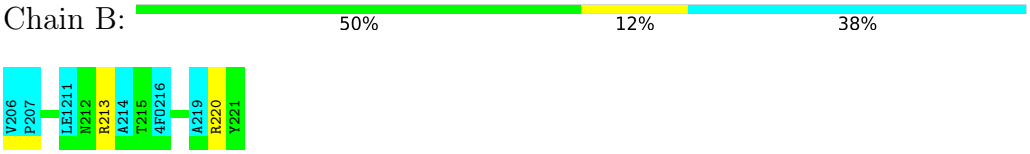
• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

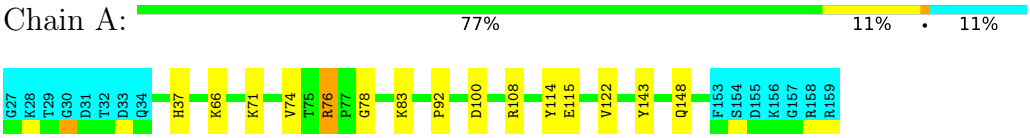
Chain B:



4.2.3 Score per residue for model 3

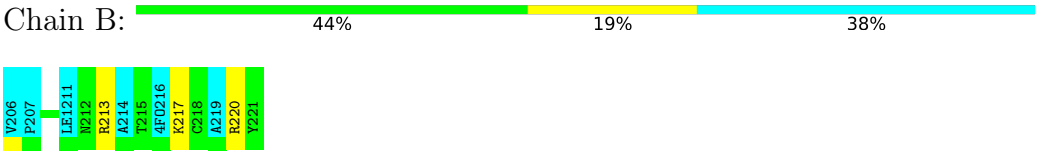
• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

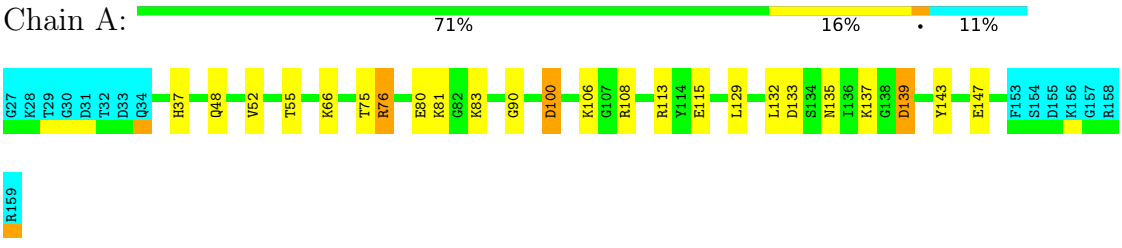
Chain B:



4.2.4 Score per residue for model 4

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

Chain B:

4.2.5 Score per residue for model 5 (medoid)

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:

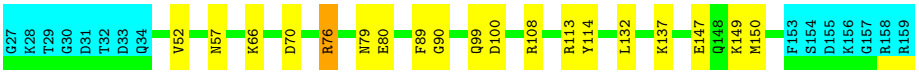
• Molecule 2: Thanatin-like derivative

Chain B:

4.2.6 Score per residue for model 6

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

Chain B:

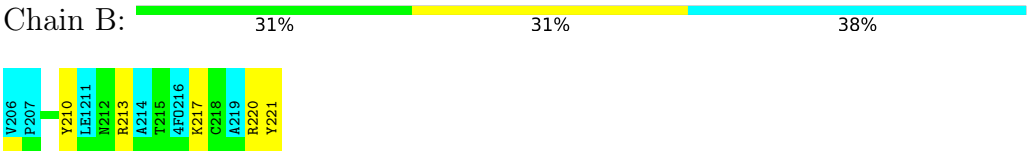
4.2.7 Score per residue for model 7

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:

• Molecule 2: Thanatin-like derivative

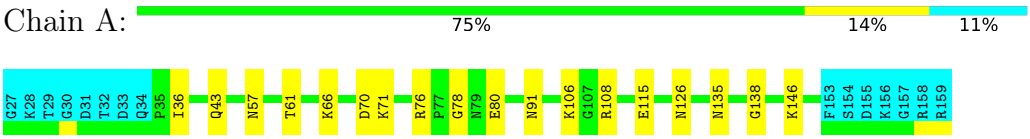
Chain B:



4.2.8 Score per residue for model 8

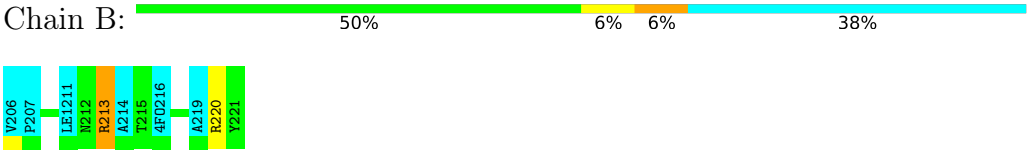
• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

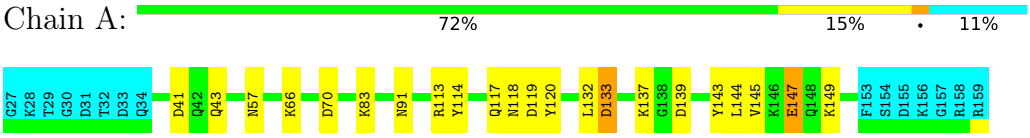
Chain B:



4.2.9 Score per residue for model 9

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



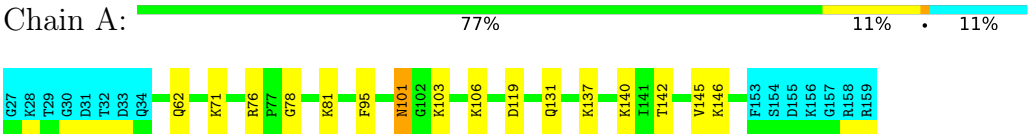
• Molecule 2: Thanatin-like derivative

Chain B:

4.2.10 Score per residue for model 10

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



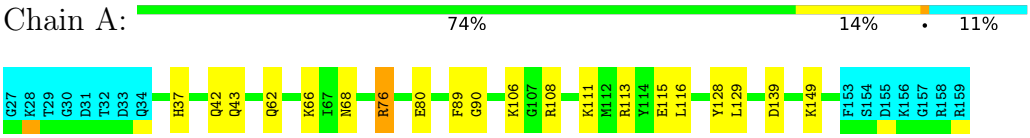
• Molecule 2: Thanatin-like derivative

Chain B:

4.2.11 Score per residue for model 11

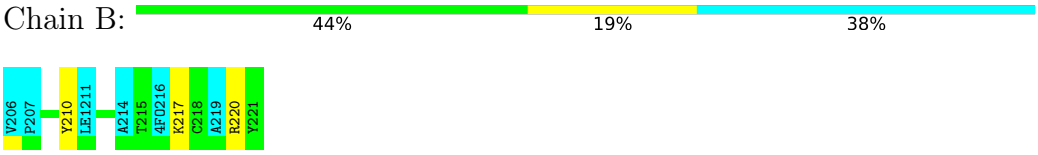
• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

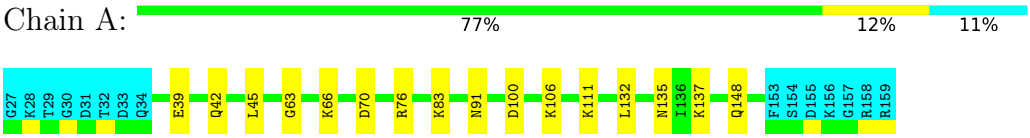
Chain B:



4.2.12 Score per residue for model 12

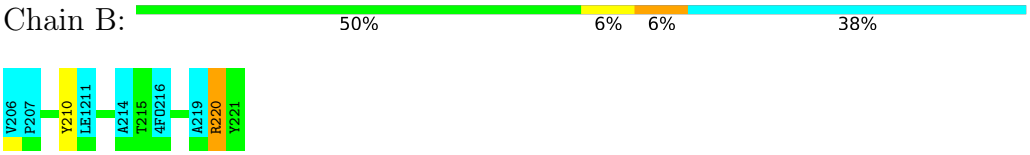
• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

Chain B:



4.2.13 Score per residue for model 13

• Molecule 1: Lipopolysaccharide export system protein LptA

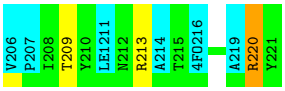
Chain A:

• Molecule 2: Thanatin-like derivative

4.2.14 Score per residue for model 14

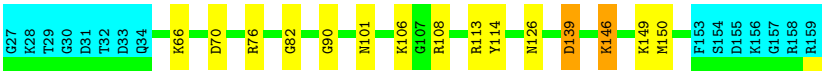
• Molecule 1: Lipopolysaccharide export system protein LptA

• Molecule 2: Thanatin-like derivative

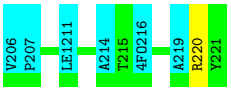


4.2.15 Score per residue for model 15

• Molecule 1: Lipopolysaccharide export system protein LptA



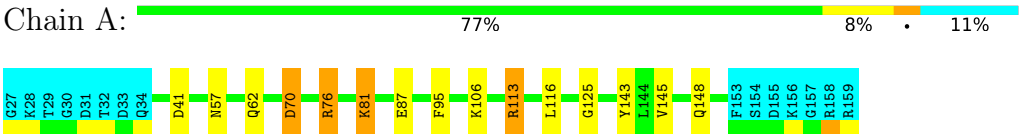
• Molecule 2: Thanatin-like derivative



4.2.16 Score per residue for model 16

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



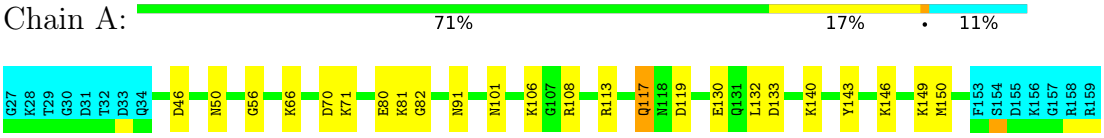
• Molecule 2: Thanatin-like derivative

Chain B:

4.2.17 Score per residue for model 17

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



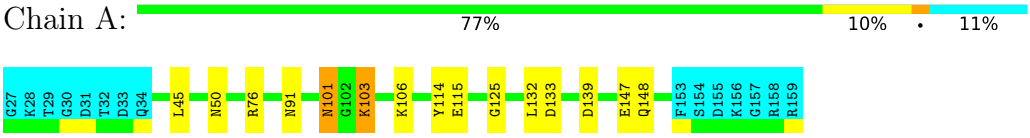
• Molecule 2: Thanatin-like derivative

Chain B:

4.2.18 Score per residue for model 18

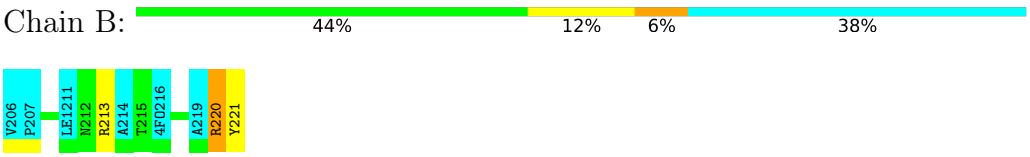
• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

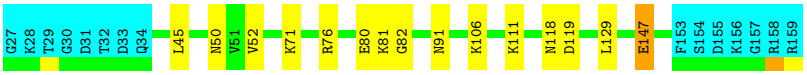
Chain B:



4.2.19 Score per residue for model 19

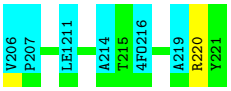
• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:



• Molecule 2: Thanatin-like derivative

Chain B:



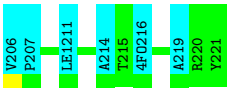
4.2.20 Score per residue for model 20

• Molecule 1: Lipopolysaccharide export system protein LptA

Chain A:

• Molecule 2: Thanatin-like derivative

Chain B:



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 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
CYANA	structure calculation	3.98
MOE	refinement	2020.09

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	1713
Number of shifts mapped to atoms	1713
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	88%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DAB, HYP, 4FO, EU0, LE1

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	0.66±0.01	0±0/933 (0.0± 0.0%)	1.07±0.04	3±1/1258 (0.2± 0.1%)
2	B	0.94±0.03	0±0/89 (0.0± 0.0%)	1.39±0.15	1±1/110 (1.0± 0.9%)
All	All	0.69	0/20440 (0.0%)	1.10	78/27360 (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	2.6±1.2
2	B	0.0±0.0	0.7±1.1
All	All	0	67

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	113	ARG	NE-CZ-NH1	12.21	126.41	120.30	1	13
1	A	76	ARG	NE-CZ-NH1	11.35	125.97	120.30	6	18
1	A	108	ARG	NE-CZ-NH1	9.34	124.97	120.30	11	9
2	B	213	ARG	NE-CZ-NH1	9.08	124.84	120.30	2	11
2	B	220	ARG	NE-CZ-NH1	9.06	124.83	120.30	3	9
1	A	139	ASP	CB-CG-OD2	-7.49	111.56	118.30	2	1
1	A	114	TYR	CB-CG-CD2	-7.28	116.64	121.00	15	2
1	A	113	ARG	NE-CZ-NH2	-7.11	116.75	120.30	2	4
1	A	113	ARG	CD-NE-CZ	6.63	132.88	123.60	1	1
1	A	143	TYR	CB-CG-CD2	-6.12	117.33	121.00	9	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	220	ARG	CD-NE-CZ	5.83	131.76	123.60	17	1
1	A	142	THR	CA-CB-CG2	5.68	120.36	112.40	13	1
1	A	101	ASN	N-CA-CB	-5.64	100.44	110.60	10	1
2	B	220	ARG	NE-CZ-NH2	-5.57	117.52	120.30	18	1
1	A	70	ASP	CB-CG-OD1	-5.30	113.53	118.30	16	1
1	A	76	ARG	NH1-CZ-NH2	-5.05	113.84	119.40	20	1
2	B	213	ARG	NE-CZ-NH2	-5.03	117.78	120.30	2	1
1	A	108	ARG	NH1-CZ-NH2	-5.03	113.87	119.40	11	1
1	A	76	ARG	CD-NE-CZ	5.02	130.63	123.60	6	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	143	TYR	Sidechain	5
1	A	139	ASP	Peptide	5
2	B	220	ARG	Sidechain,Peptide,Mainchain	5
1	A	114	TYR	Sidechain	4
2	B	210	TYR	Sidechain,Peptide	3
1	A	76	ARG	Peptide,Sidechain	3
1	A	35	PRO	Peptide	2
2	B	217	LYS	Peptide	2
1	A	70	ASP	Peptide	2
1	A	78	GLY	Peptide	2
1	A	92	PRO	Peptide	2
1	A	100	ASP	Peptide	2
1	A	80	GLU	Peptide	2
1	A	147	GLU	Peptide	2
1	A	146	LYS	Peptide	2
1	A	116	LEU	Peptide	2
1	A	149	LYS	Peptide	2
1	A	119	ASP	Peptide	1
1	A	40	SER	Peptide	1
1	A	52	VAL	Peptide	1
1	A	75	THR	Peptide	1
1	A	99	GLN	Peptide	1
2	B	213	ARG	Sidechain	1
1	A	91	ASN	Peptide	1
1	A	138	GLY	Peptide	1
1	A	42	GLN	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	83	LYS	Peptide	1
1	A	79	ASN	Peptide	1
1	A	126	ASN	Peptide	1
1	A	81	LYS	Peptide	1
1	A	113	ARG	Sidechain	1
1	A	56	GLY	Peptide	1
1	A	118	ASN	Peptide	1
1	A	129	LEU	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	920	923	923	0±0
All	All	20240	20300	20300	2

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:148:GLN:CD	1:A:148:GLN:H	0.47	2.13	14	1
1:A:89:PHE:CD2	1:A:111:LYS:HE2	0.42	2.48	11	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	118/133 (89%)	105±2 (89±2%)	10±3 (9±2%)	2±1 (2±1%)	10 50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	9/16 (56%)	8±1 (89±10%)	1±1 (10±9%)	0±0 (1±2%)	29 74
All	All	2540/2980 (85%)	2262 (89%)	227 (9%)	51 (2%)	11 52

All 24 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	90	GLY	7
1	A	101	ASN	7
1	A	133	ASP	6
1	A	63	GLY	4
1	A	125	GLY	3
1	A	82	GLY	3
1	A	148	GLN	2
1	A	36	ILE	2
1	A	145	VAL	2
1	A	100	ASP	1
1	A	81	LYS	1
1	A	118	ASN	1
1	A	78	GLY	1
1	A	76	ARG	1
1	A	132	LEU	1
1	A	126	ASN	1
2	B	220	ARG	1
1	A	35	PRO	1
1	A	60	VAL	1
1	A	146	LYS	1
1	A	91	ASN	1
1	A	117	GLN	1
1	A	103	LYS	1
1	A	147	GLU	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	102/114 (89%)	91±3 (89±3%)	11±3 (11±3%)	10 54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	10/10 (100%)	9±1 (94±7%)	1±1 (6±7%)	23 72
All	All	2240/2480 (90%)	2003 (89%)	237 (11%)	10 55

All 74 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	106	LYS	15
1	A	66	LYS	10
1	A	91	ASN	8
1	A	80	GLU	7
1	A	137	LYS	7
1	A	70	ASP	7
1	A	43	GLN	6
1	A	45	LEU	6
1	A	62	GLN	6
1	A	81	LYS	6
1	A	71	LYS	6
1	A	115	GLU	6
1	A	148	GLN	6
1	A	132	LEU	6
1	A	146	LYS	5
1	A	57	ASN	5
1	A	135	ASN	5
1	A	147	GLU	5
1	A	119	ASP	5
1	A	103	LYS	4
1	A	118	ASN	4
1	A	150	MET	4
1	A	83	LYS	4
1	A	149	LYS	4
2	B	220	ARG	4
1	A	39	GLU	3
1	A	53	THR	3
1	A	37	HIS	3
1	A	76	ARG	3
1	A	129	LEU	3
1	A	139	ASP	3
1	A	52	VAL	3
1	A	133	ASP	3
2	B	221	TYR	3
1	A	111	LYS	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	50	ASN	3
1	A	101	ASN	2
1	A	74	VAL	2
2	B	217	LYS	2
1	A	55	THR	2
1	A	100	ASP	2
1	A	42	GLN	2
1	A	130	GLU	2
1	A	144	LEU	2
1	A	151	GLN	2
1	A	126	ASN	2
1	A	41	ASP	2
1	A	117	GLN	2
1	A	95	PHE	2
1	A	140	LYS	2
1	A	46	ASP	2
1	A	113	ARG	1
1	A	84	GLU	1
1	A	122	VAL	1
1	A	48	GLN	1
1	A	79	ASN	1
1	A	89	PHE	1
1	A	61	THR	1
2	B	213	ARG	1
1	A	120	TYR	1
1	A	131	GLN	1
1	A	142	THR	1
1	A	68	ASN	1
1	A	128	TYR	1
2	B	210	TYR	1
1	A	99	GLN	1
1	A	47	MET	1
1	A	58	VAL	1
1	A	108	ARG	1
2	B	209	THR	1
1	A	87	GLU	1
1	A	116	LEU	1
1	A	145	VAL	1
1	A	124	THR	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)

6 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
2	4FO	B	216	2	5,6,7	0.63±0.09	0±0 (0±0%)
2	DAB	B	219	2	5,6,7	0.65±0.13	0±0 (0±0%)
2	HYP	B	207	2	6,8,9	0.51±0.04	0±0 (0±0%)
2	DAB	B	214	2	5,6,7	0.62±0.07	0±0 (0±0%)
2	LE1	B	211	2	3,7,8	0.58±0.03	0±0 (0±0%)
2	EU0	B	206	2	8,9,10	1.34±0.13	1±0 (8±5%)

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
2	4FO	B	216	2	1,6,8	0.14±0.14	0±0 (0±0%)
2	DAB	B	219	2	1,6,8	0.25±0.17	0±0 (0±0%)
2	HYP	B	207	2	5,10,12	1.01±0.24	0±1 (5±10%)
2	DAB	B	214	2	1,6,8	0.31±0.18	0±0 (0±0%)
2	LE1	B	211	2	3,10,12	0.88±0.19	0±0 (0±0%)
2	EU0	B	206	2	10,11,13	1.39±0.36	2±1 (16±12%)

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
2	DAB	B	219	2	-	0±0,4,5,7	-
2	EU0	B	206	2	-	0±0,9,10,12	-
2	DAB	B	214	2	-	0±0,4,5,7	-
2	4FO	B	216	2	-	0±0,4,5,7	-
2	HYP	B	207	2	-	0±0,0,11,13	0±0,1,1,1
2	LE1	B	211	2	-	0±0,4,8,10	-

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
2	B	206	EU0	CA-N	4.04	1.52	1.46	2	12
2	B	206	EU0	CT-N	2.09	1.29	1.33	13	2

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
2	B	206	EU0	CB-CA-N	4.44	116.95	111.17	8	5
2	B	206	EU0	CB-CA-C	4.16	107.82	113.04	8	5
2	B	206	EU0	CA-N-CT	3.23	132.58	123.19	8	3
2	B	206	EU0	CG1-CB-CA	3.15	116.03	111.21	8	3
2	B	207	HYP	CG-CB-CA	2.62	100.66	103.96	10	2
2	B	206	EU0	CG2-CB-CA	2.61	115.20	111.21	13	2
2	B	206	EU0	O-C-CA	2.57	117.66	124.83	7	12
2	B	207	HYP	CB-CG-CD	2.44	100.27	103.27	10	2
2	B	206	EU0	N-CT-NT1	2.20	124.43	120.59	4	2
2	B	207	HYP	O-C-CA	2.07	119.35	124.78	2	1

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates (i)

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

The completeness of assignment taking into account all chemical shift lists is 88% for the well-defined parts and 86% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_output*

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	1713
Number of shifts mapped to atoms	1713
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	131	-0.13 ± 0.14	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	118	-0.12 ± 0.19	None needed (< 0.5 ppm)
$^{13}C'$	123	0.44 ± 0.08	None needed (< 0.5 ppm)
^{15}N	123	-0.70 ± 0.37	None needed (imprecise)

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 88%, i.e. 1543 atoms were assigned a chemical shift out of a possible 1752. 0 out of 19 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	598/643 (93%)	258/263 (98%)	229/256 (89%)	111/124 (90%)
Sidechain	875/1009 (87%)	608/649 (94%)	248/313 (79%)	19/47 (40%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	70/100 (70%)	43/47 (91%)	27/52 (52%)	0/1 (0%)
Overall	1543/1752 (88%)	909/959 (95%)	504/621 (81%)	130/172 (76%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 86%, i.e. 1675 atoms were assigned a chemical shift out of a possible 1942. 0 out of 19 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	663/721 (92%)	286/296 (97%)	254/286 (89%)	123/139 (88%)
Sidechain	942/1111 (85%)	651/711 (92%)	271/344 (79%)	20/56 (36%)
Aromatic	70/110 (64%)	43/52 (83%)	27/57 (47%)	0/1 (0%)
Overall	1675/1942 (86%)	980/1059 (93%)	552/687 (80%)	143/196 (73%)

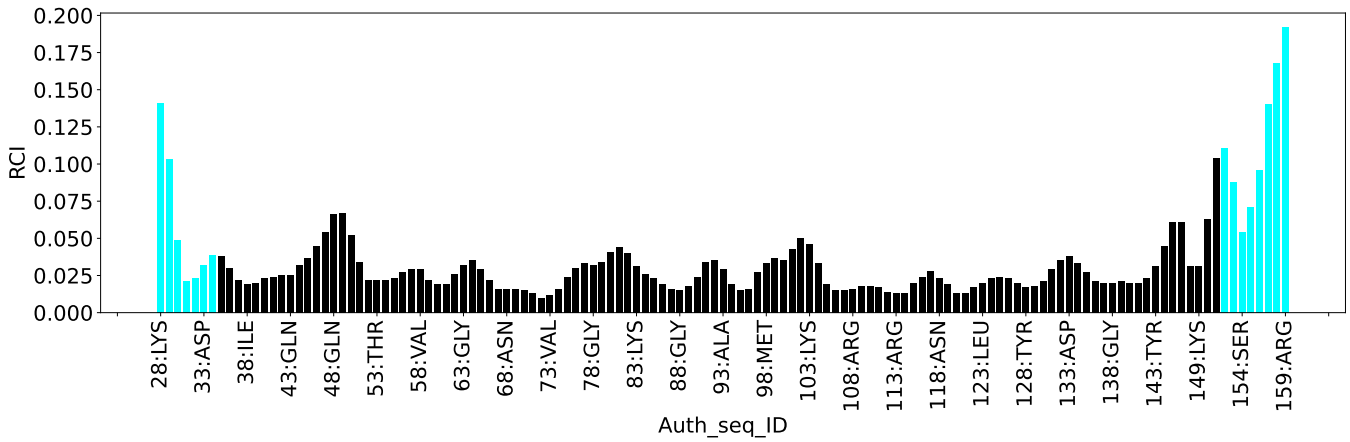
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

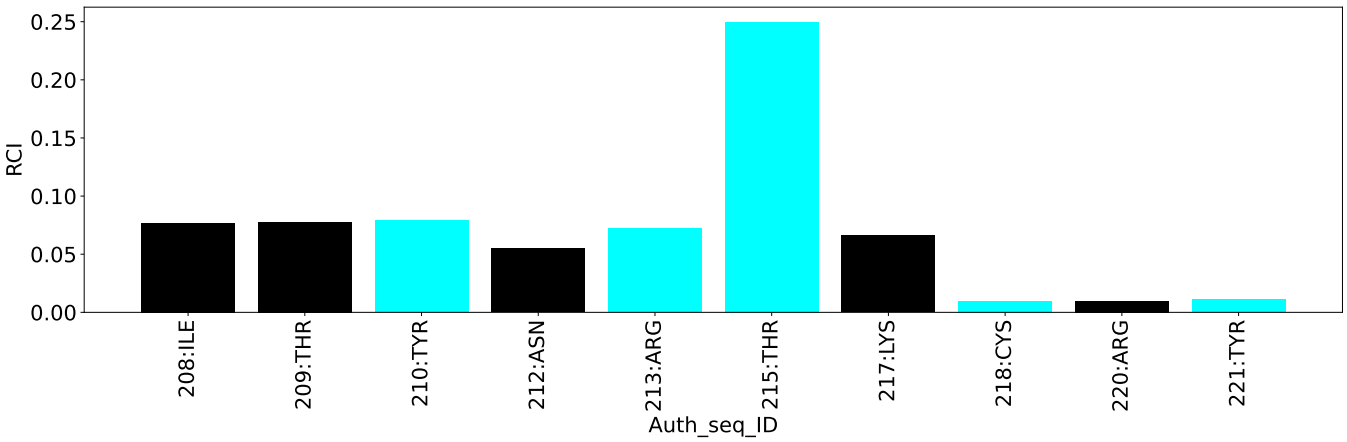
7.1.5 Random Coil Index (RCI) plots (i)

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:



Random coil index (RCI) for chain B:


RCI
0.25
0.20
0.15
0.10
0.05
0.00
208:ILE
209:THR
210:TYR
212:ASN
213:ARG
215:THR
217:LYS
218:CYS
220:ARG
221:TYR
Auth_seq_ID