



Full wwPDB NMR Structure Validation Report ⓘ

Mar 8, 2026 – 02:28 PM UTC

PDB ID : 7MY8 / pdb_00007my8
BMRB ID : 30909
Title : Fusion Peptide of SARS-CoV-2 Spike Rearranges into a Wedge Inserted in Bilayered Micelles
Authors : Koppiseti, R.K.; Fulcher, Y.G.; Van Doren, S.R.
Deposited on : 2021-05-20

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

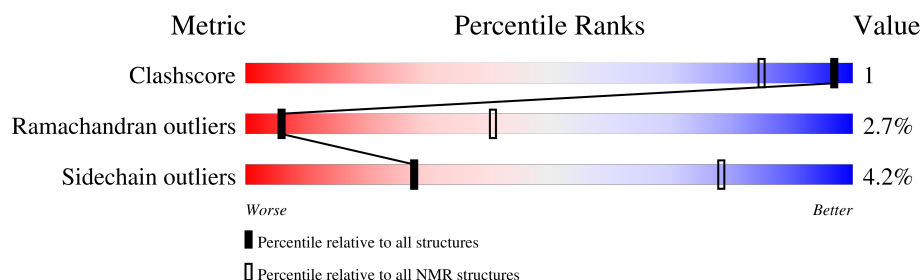
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 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	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	42	43% 52% 5%

2 Ensemble composition and analysis

This entry contains 15 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:816-A:855 (40)	1.14	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 9 single-model clusters were found.

Cluster number	Models
1	1, 2, 7, 10
2	3, 4
Single-model clusters	5; 6; 8; 9; 11; 12; 13; 14; 15

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 637 atoms, of which 315 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Spike protein S2.

Mol	Chain	Residues	Atoms						Trace
1	A	42	Total	C	H	N	O	S	0
			637	206	315	52	62	2	

There is a discrepancy between the modelled and reference sequences:

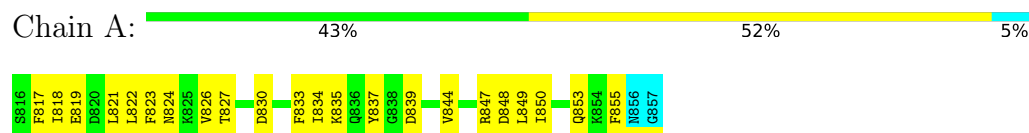
Chain	Residue	Modelled	Actual	Comment	Reference
A	844	VAL	ILE	conflict	UNP P0DTC2

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: Spike protein S2

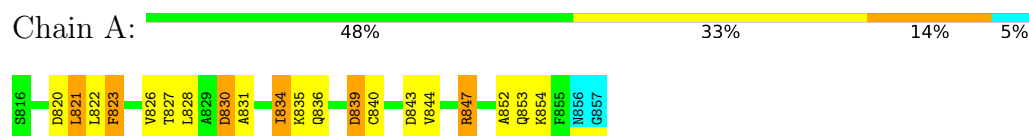


4.2 Scores per residue for each member of the ensemble

Colouring as in section [4.1](#) above.

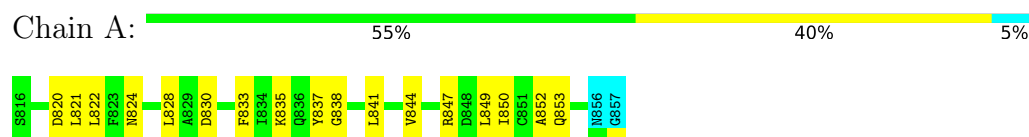
4.2.1 Score per residue for model 1 (medoid)

- Molecule 1: Spike protein S2



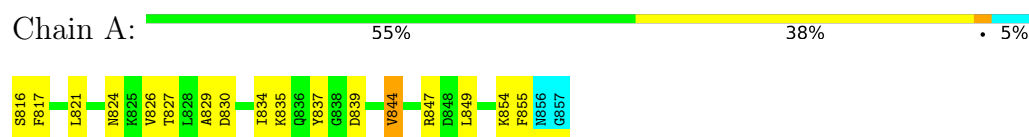
4.2.2 Score per residue for model 2

- Molecule 1: Spike protein S2



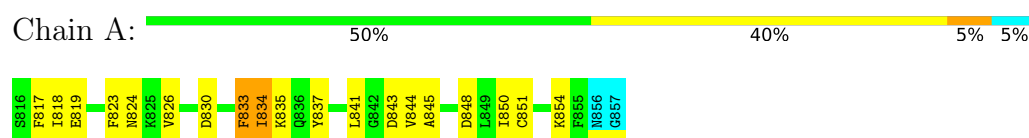
4.2.3 Score per residue for model 3

- Molecule 1: Spike protein S2



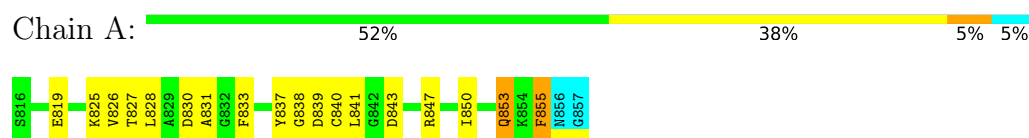
4.2.4 Score per residue for model 4

- Molecule 1: Spike protein S2



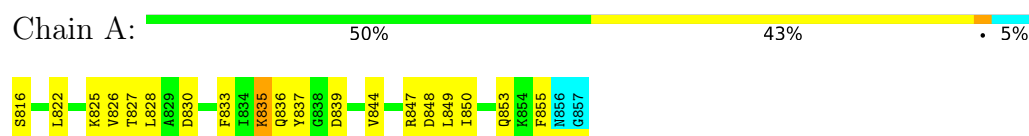
4.2.5 Score per residue for model 5

- Molecule 1: Spike protein S2



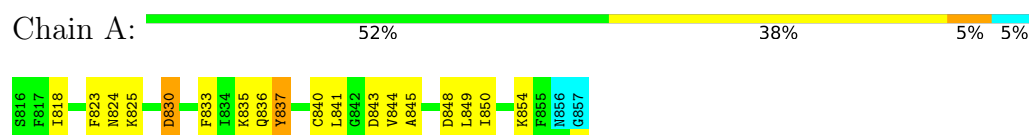
4.2.6 Score per residue for model 6

- Molecule 1: Spike protein S2



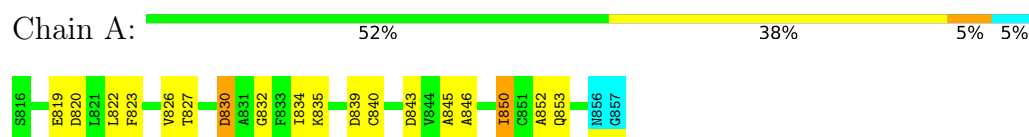
4.2.7 Score per residue for model 7

- Molecule 1: Spike protein S2



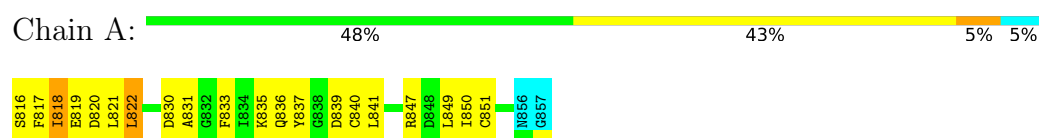
4.2.8 Score per residue for model 8

- Molecule 1: Spike protein S2



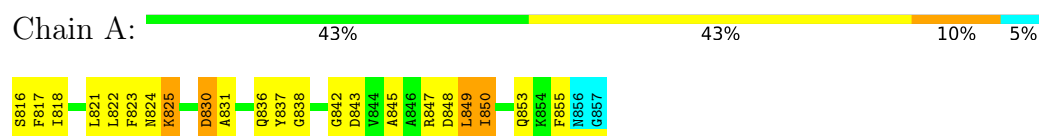
4.2.9 Score per residue for model 9

- Molecule 1: Spike protein S2



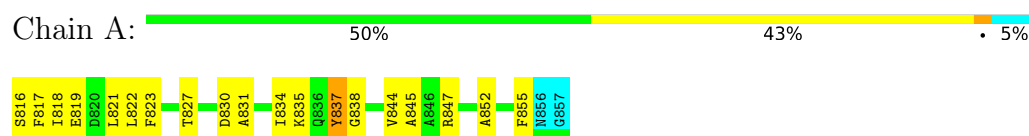
4.2.10 Score per residue for model 10

- Molecule 1: Spike protein S2



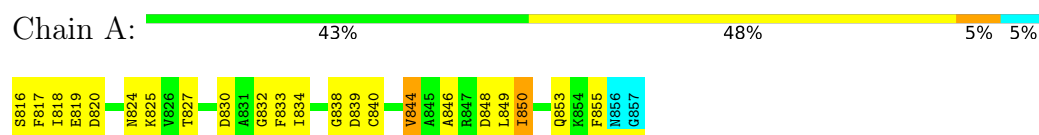
4.2.11 Score per residue for model 11

- Molecule 1: Spike protein S2



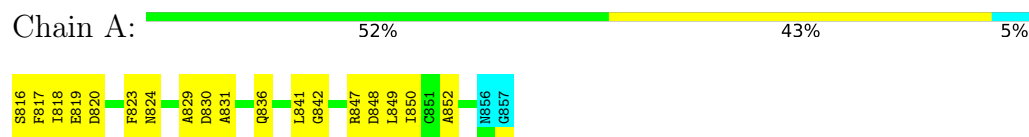
4.2.12 Score per residue for model 12

- Molecule 1: Spike protein S2



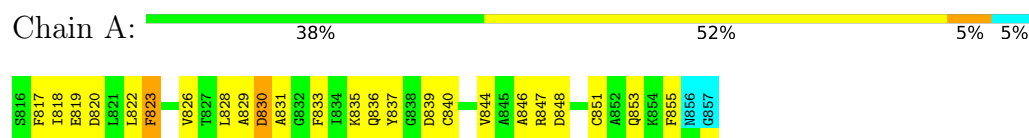
4.2.13 Score per residue for model 13

- Molecule 1: Spike protein S2



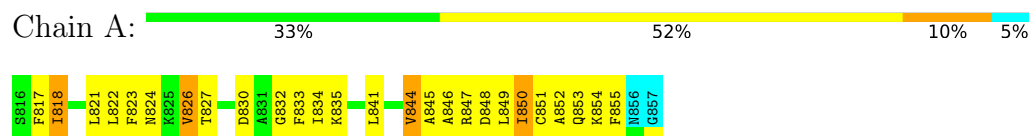
4.2.14 Score per residue for model 14

- Molecule 1: Spike protein S2



4.2.15 Score per residue for model 15

- Molecule 1: Spike protein S2



5 Refinement protocol and experimental data overview

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

Of the 50 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
CYANA	structure calculation	3.0
GROMACS	refinement	

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	488
Number of shifts mapped to atoms	488
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

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	1.99±0.12	6±3/314 (2.0± 0.9%)	2.45±0.15	22±7/423 (5.3± 1.7%)
All	All	1.99	94/4710 (2.0%)	2.45	334/6345 (5.3%)

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	0.7±0.7
All	All	0	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
1	A	833	PHE	CA-C	8.09	1.63	1.52	9	1
1	A	824	ASN	CA-C	7.99	1.63	1.52	3	1
1	A	844	VAL	CA-CB	-7.87	1.43	1.54	4	3
1	A	818	ILE	C-N	7.50	1.43	1.33	7	1
1	A	853	GLN	CA-C	-7.49	1.41	1.52	8	2
1	A	840	CYS	CA-CB	7.18	1.64	1.53	1	2
1	A	844	VAL	N-CA	7.13	1.55	1.46	4	1
1	A	844	VAL	CA-C	6.96	1.60	1.52	12	3
1	A	832	GLY	CA-C	6.93	1.60	1.52	15	1
1	A	843	ASP	CA-C	-6.83	1.43	1.52	10	3
1	A	816	SER	C-N	6.51	1.42	1.33	6	1
1	A	826	VAL	N-CA	6.44	1.54	1.46	3	2
1	A	834	ILE	CA-CB	-6.27	1.47	1.54	12	1
1	A	852	ALA	CA-C	6.16	1.60	1.52	13	2
1	A	820	ASP	N-CA	-6.14	1.38	1.46	1	1
1	A	837	TYR	CA-CB	6.14	1.63	1.53	4	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	843	ASP	C-N	6.11	1.42	1.33	4	3
1	A	819	GLU	CA-CB	6.08	1.63	1.53	12	2
1	A	850	ILE	C-N	6.08	1.42	1.33	13	1
1	A	837	TYR	N-CA	6.07	1.53	1.46	3	3
1	A	837	TYR	CG-CD2	6.04	1.52	1.39	7	1
1	A	846	ALA	CA-C	6.03	1.61	1.52	14	2
1	A	854	LYS	CA-CB	6.02	1.63	1.53	15	1
1	A	833	PHE	CA-CB	5.98	1.62	1.53	9	2
1	A	840	CYS	C-N	5.97	1.41	1.33	9	2
1	A	851	CYS	CA-CB	5.87	1.62	1.53	15	1
1	A	839	ASP	C-N	5.85	1.41	1.33	6	2
1	A	835	LYS	CA-CB	5.82	1.62	1.53	9	2
1	A	818	ILE	CA-CB	-5.82	1.47	1.54	4	2
1	A	850	ILE	CB-CG1	5.79	1.65	1.53	4	1
1	A	824	ASN	C-N	5.79	1.41	1.33	4	1
1	A	840	CYS	CA-C	5.79	1.61	1.53	9	1
1	A	819	GLU	CA-C	5.78	1.60	1.52	4	1
1	A	847	ARG	CD-NE	5.74	1.54	1.46	11	2
1	A	821	LEU	CA-CB	5.62	1.62	1.53	3	1
1	A	832	GLY	C-O	-5.58	1.16	1.23	12	1
1	A	855	PHE	C-N	5.56	1.41	1.33	15	1
1	A	848	ASP	C-N	5.56	1.41	1.34	7	1
1	A	830	ASP	CA-C	5.55	1.60	1.52	10	2
1	A	822	LEU	CA-C	5.53	1.60	1.52	2	2
1	A	841	LEU	C-N	5.52	1.41	1.32	15	2
1	A	816	SER	N-CA	5.51	1.56	1.46	10	1
1	A	847	ARG	CA-C	-5.51	1.45	1.52	1	2
1	A	816	SER	CA-CB	5.47	1.64	1.53	11	2
1	A	827	THR	CB-OG1	-5.44	1.35	1.43	12	1
1	A	825	LYS	N-CA	5.44	1.52	1.46	6	1
1	A	828	LEU	CA-C	5.43	1.59	1.52	5	1
1	A	852	ALA	CA-CB	-5.38	1.45	1.53	2	1
1	A	831	ALA	C-N	5.35	1.40	1.33	14	3
1	A	836	GLN	C-N	5.33	1.41	1.33	9	1
1	A	835	LYS	CA-C	5.30	1.59	1.52	3	1
1	A	830	ASP	C-N	5.29	1.41	1.34	4	1
1	A	833	PHE	C-O	-5.22	1.17	1.24	14	1
1	A	839	ASP	N-CA	-5.21	1.39	1.46	3	1
1	A	834	ILE	CB-CG1	5.20	1.63	1.53	12	1
1	A	821	LEU	CA-C	5.19	1.59	1.52	10	1
1	A	826	VAL	C-N	5.18	1.40	1.33	8	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	841	LEU	CA-CB	5.15	1.61	1.53	9	1
1	A	827	THR	C-N	5.09	1.41	1.33	3	1
1	A	855	PHE	CA-C	5.08	1.59	1.53	11	1
1	A	817	PHE	C-N	5.05	1.40	1.33	3	1
1	A	819	GLU	N-CA	5.02	1.52	1.46	8	1
1	A	834	ILE	CA-C	5.01	1.59	1.52	3	1
1	A	845	ALA	N-CA	-5.00	1.40	1.46	10	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(°)	Ideal(°)	Models	
								Worst	Total
1	A	823	PHE	CA-CB-CG	-11.80	102.00	113.80	4	3
1	A	843	ASP	CA-CB-CG	11.42	124.02	112.60	10	3
1	A	833	PHE	CA-CB-CG	-10.29	103.51	113.80	12	4
1	A	844	VAL	CB-CA-C	9.42	123.23	110.99	6	2
1	A	841	LEU	N-CA-C	-9.22	102.17	113.97	4	2
1	A	818	ILE	N-CA-CB	9.17	121.93	110.47	14	1
1	A	823	PHE	N-CA-C	8.81	120.50	111.07	15	4
1	A	833	PHE	N-CA-C	8.58	120.63	111.28	12	4
1	A	855	PHE	CA-CB-CG	-8.47	105.33	113.80	6	2
1	A	839	ASP	CA-CB-CG	-8.28	104.32	112.60	14	4
1	A	848	ASP	N-CA-C	-8.25	103.12	113.02	15	1
1	A	849	LEU	N-CA-C	8.22	122.27	111.75	7	6
1	A	845	ALA	N-CA-C	8.10	121.14	111.33	15	3
1	A	829	ALA	N-CA-C	8.08	120.09	111.28	14	3
1	A	827	THR	CA-C-O	-8.04	111.02	120.10	6	2
1	A	844	VAL	O-C-N	-8.00	112.57	122.57	15	5
1	A	817	PHE	N-CA-C	7.98	120.99	111.33	10	2
1	A	828	LEU	N-CA-C	-7.97	103.15	113.12	14	1
1	A	820	ASP	CA-C-N	7.75	130.66	120.28	12	2
1	A	820	ASP	C-N-CA	7.75	130.66	120.28	12	2
1	A	819	GLU	N-CA-C	7.67	119.64	111.28	9	1
1	A	834	ILE	CA-C-N	7.49	130.65	120.54	15	3
1	A	834	ILE	C-N-CA	7.49	130.65	120.54	15	3
1	A	855	PHE	CB-CA-C	7.48	122.35	111.73	11	1
1	A	840	CYS	CA-C-N	7.45	130.87	120.29	1	2
1	A	840	CYS	C-N-CA	7.45	130.87	120.29	1	2
1	A	833	PHE	CA-C-N	7.43	129.92	120.56	12	3
1	A	833	PHE	C-N-CA	7.43	129.92	120.56	12	3
1	A	825	LYS	CA-C-O	7.39	128.38	120.55	12	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	849	LEU	CA-C-N	7.34	132.07	120.47	2	5
1	A	849	LEU	C-N-CA	7.34	132.07	120.47	2	5
1	A	853	GLN	CG-CD-OE1	7.33	135.45	120.80	15	1
1	A	838	GLY	O-C-N	-7.30	113.21	122.70	10	2
1	A	816	SER	CA-C-N	7.22	129.95	120.28	11	2
1	A	816	SER	C-N-CA	7.22	129.95	120.28	11	2
1	A	853	GLN	N-CA-C	-7.20	103.05	112.41	14	1
1	A	818	ILE	CA-C-O	7.18	128.27	119.58	15	2
1	A	826	VAL	CA-C-N	7.18	129.90	120.28	5	2
1	A	826	VAL	C-N-CA	7.18	129.90	120.28	5	2
1	A	835	LYS	N-CA-C	7.14	118.72	111.07	6	2
1	A	851	CYS	N-CA-C	7.14	119.14	111.36	9	2
1	A	851	CYS	CB-CA-C	-7.03	98.90	110.85	9	1
1	A	818	ILE	CA-C-N	7.00	130.36	120.28	12	3
1	A	818	ILE	C-N-CA	7.00	130.36	120.28	12	3
1	A	848	ASP	CA-CB-CG	7.00	119.60	112.60	4	5
1	A	817	PHE	CA-C-N	6.99	130.38	120.53	14	2
1	A	817	PHE	C-N-CA	6.99	130.38	120.53	14	2
1	A	855	PHE	CA-C-N	6.97	132.52	122.35	3	3
1	A	855	PHE	C-N-CA	6.97	132.52	122.35	3	3
1	A	844	VAL	CA-C-N	6.87	129.79	120.38	15	2
1	A	844	VAL	C-N-CA	6.87	129.79	120.38	15	2
1	A	835	LYS	CA-C-N	6.84	130.01	120.29	4	5
1	A	835	LYS	C-N-CA	6.84	130.01	120.29	4	5
1	A	831	ALA	CA-C-N	6.84	127.57	119.98	9	1
1	A	831	ALA	C-N-CA	6.84	127.57	119.98	9	1
1	A	850	ILE	N-CA-C	6.83	117.58	110.62	4	2
1	A	817	PHE	CA-CB-CG	-6.81	106.99	113.80	11	3
1	A	845	ALA	CA-C-N	6.79	129.38	120.28	11	2
1	A	845	ALA	C-N-CA	6.79	129.38	120.28	11	2
1	A	847	ARG	NE-CZ-NH2	-6.78	113.10	119.20	13	2
1	A	854	LYS	CA-C-N	6.72	133.09	122.61	7	2
1	A	854	LYS	C-N-CA	6.72	133.09	122.61	7	2
1	A	826	VAL	CB-CA-C	-6.67	104.52	112.19	6	1
1	A	819	GLU	O-C-N	-6.67	115.05	122.12	9	3
1	A	821	LEU	N-CA-C	6.66	118.54	111.28	15	2
1	A	824	ASN	CA-CB-CG	-6.65	105.95	112.60	15	2
1	A	841	LEU	CA-C-N	6.65	133.32	122.09	13	1
1	A	841	LEU	C-N-CA	6.65	133.32	122.09	13	1
1	A	818	ILE	O-C-N	-6.63	115.17	121.87	12	2
1	A	840	CYS	N-CA-C	6.59	120.98	111.56	1	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	827	THR	N-CA-C	6.57	118.44	111.28	11	5
1	A	853	GLN	OE1-CD-NE2	-6.55	116.05	122.60	15	3
1	A	847	ARG	O-C-N	-6.53	113.45	122.46	9	2
1	A	831	ALA	CA-C-O	-6.48	113.68	120.55	5	1
1	A	847	ARG	CA-C-O	6.45	126.00	118.90	14	2
1	A	838	GLY	CA-C-N	6.43	128.90	120.28	11	2
1	A	838	GLY	C-N-CA	6.43	128.90	120.28	11	2
1	A	853	GLN	N-CA-CB	6.42	120.27	110.44	10	1
1	A	820	ASP	CA-CB-CG	-6.34	106.26	112.60	2	3
1	A	818	ILE	N-CA-C	6.33	117.08	110.62	9	3
1	A	842	GLY	N-CA-C	6.33	123.76	115.40	13	2
1	A	837	TYR	N-CA-C	6.33	118.18	111.28	6	1
1	A	853	GLN	CA-C-N	6.33	131.13	122.34	12	3
1	A	853	GLN	C-N-CA	6.33	131.13	122.34	12	3
1	A	834	ILE	CB-CA-C	-6.32	103.61	112.14	3	1
1	A	821	LEU	CA-C-N	6.32	129.38	120.28	9	3
1	A	821	LEU	C-N-CA	6.32	129.38	120.28	9	3
1	A	830	ASP	O-C-N	-6.31	114.20	122.59	8	1
1	A	836	GLN	N-CA-C	6.29	118.22	111.36	13	3
1	A	821	LEU	N-CA-CB	6.28	119.18	110.07	2	1
1	A	848	ASP	CA-C-N	6.28	129.32	120.28	15	4
1	A	848	ASP	C-N-CA	6.28	129.32	120.28	15	4
1	A	826	VAL	CA-C-O	6.26	127.15	120.26	14	1
1	A	822	LEU	N-CA-C	6.25	118.89	111.33	8	1
1	A	830	ASP	CA-C-N	6.20	129.73	120.31	8	2
1	A	830	ASP	C-N-CA	6.20	129.73	120.31	8	2
1	A	816	SER	O-C-N	-6.12	113.20	123.00	11	2
1	A	836	GLN	N-CA-CB	6.03	119.08	110.16	14	1
1	A	837	TYR	CA-CB-CG	-6.02	103.06	113.90	2	1
1	A	847	ARG	NH1-CZ-NH2	-6.02	111.47	119.30	2	1
1	A	829	ALA	CA-C-N	6.02	133.04	121.54	13	1
1	A	829	ALA	C-N-CA	6.02	133.04	121.54	13	1
1	A	847	ARG	N-CA-CB	-6.00	101.84	110.65	10	2
1	A	822	LEU	CB-CA-C	-5.97	101.50	110.88	14	2
1	A	850	ILE	CA-CB-CG1	5.96	120.53	110.40	7	4
1	A	828	LEU	CA-C-O	5.94	125.44	118.90	6	1
1	A	852	ALA	N-CA-C	5.92	117.40	111.07	15	2
1	A	830	ASP	CA-CB-CG	5.91	118.51	112.60	7	3
1	A	832	GLY	CA-C-N	5.91	128.19	120.28	12	1
1	A	832	GLY	C-N-CA	5.91	128.19	120.28	12	1
1	A	854	LYS	N-CA-C	5.90	119.53	111.39	1	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	832	GLY	O-C-N	-5.89	116.29	122.13	8	1
1	A	817	PHE	CA-C-O	5.88	126.98	120.63	10	2
1	A	825	LYS	CB-CG-CD	5.86	124.78	111.30	7	1
1	A	824	ASN	N-CA-CB	5.85	118.72	110.12	12	2
1	A	837	TYR	N-CA-CB	-5.83	101.96	110.53	4	2
1	A	855	PHE	CA-C-O	5.81	127.20	119.59	12	1
1	A	834	ILE	O-C-N	-5.80	116.24	121.87	15	1
1	A	853	GLN	O-C-N	-5.76	115.66	122.58	6	2
1	A	833	PHE	CB-CG-CD2	-5.76	110.90	120.70	15	1
1	A	824	ASN	OD1-CG-ND2	-5.76	116.84	122.60	7	1
1	A	851	CYS	CA-C-N	5.74	128.55	120.28	14	1
1	A	851	CYS	C-N-CA	5.74	128.55	120.28	14	1
1	A	845	ALA	N-CA-CB	-5.74	101.45	110.06	15	1
1	A	847	ARG	NE-CZ-NH1	5.73	127.23	121.50	3	3
1	A	848	ASP	CA-C-O	5.71	125.66	119.15	7	1
1	A	820	ASP	N-CA-C	5.70	118.27	111.71	8	1
1	A	824	ASN	O-C-N	-5.67	116.23	122.07	13	1
1	A	822	LEU	CA-C-N	5.66	127.87	120.28	11	2
1	A	822	LEU	C-N-CA	5.66	127.87	120.28	11	2
1	A	839	ASP	CA-C-O	5.65	125.39	119.51	6	1
1	A	855	PHE	N-CA-CB	-5.62	103.47	111.84	10	1
1	A	830	ASP	N-CA-CB	5.61	119.97	110.49	14	1
1	A	821	LEU	CA-C-O	5.60	126.47	120.70	2	1
1	A	829	ALA	N-CA-CB	-5.59	101.91	110.12	14	1
1	A	831	ALA	N-CA-CB	-5.57	101.65	110.28	11	1
1	A	835	LYS	CA-CB-CG	5.55	125.20	114.10	11	1
1	A	819	GLU	CA-C-N	5.54	127.97	120.38	5	2
1	A	819	GLU	C-N-CA	5.54	127.97	120.38	5	2
1	A	844	VAL	N-CA-C	5.53	120.85	109.34	3	1
1	A	826	VAL	O-C-N	-5.46	115.74	122.57	3	1
1	A	850	ILE	CA-CB-CG2	-5.46	101.22	110.50	10	1
1	A	850	ILE	CA-C-O	-5.44	114.75	120.57	2	1
1	A	839	ASP	CB-CA-C	5.43	122.20	110.41	6	1
1	A	823	PHE	CA-C-N	5.41	127.47	120.44	13	2
1	A	823	PHE	C-N-CA	5.41	127.47	120.44	13	2
1	A	831	ALA	O-C-N	-5.39	115.64	122.27	1	2
1	A	823	PHE	N-CA-CB	5.38	118.03	110.12	8	1
1	A	850	ILE	CA-C-N	5.36	127.91	120.29	15	1
1	A	850	ILE	C-N-CA	5.36	127.91	120.29	15	1
1	A	853	GLN	CB-CA-C	-5.36	104.68	111.86	6	1
1	A	820	ASP	O-C-N	-5.33	116.08	122.15	12	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	847	ARG	CB-CG-CD	5.32	123.54	111.30	5	1
1	A	816	SER	CA-CB-OG	5.30	121.71	111.10	13	1
1	A	847	ARG	CB-CA-C	5.30	118.45	109.55	10	1
1	A	824	ASN	CA-C-O	-5.27	114.93	120.63	10	1
1	A	853	GLN	CG-CD-NE2	-5.27	108.49	116.40	15	1
1	A	836	GLN	CA-C-N	5.25	127.26	120.44	10	1
1	A	836	GLN	C-N-CA	5.25	127.26	120.44	10	1
1	A	849	LEU	CB-CA-C	-5.24	102.09	110.79	3	1
1	A	825	LYS	N-CA-C	5.22	118.81	112.23	10	1
1	A	839	ASP	N-CA-C	-5.22	106.59	113.17	5	1
1	A	850	ILE	N-CA-CB	5.21	116.65	110.55	9	1
1	A	840	CYS	CA-C-O	5.21	124.22	118.65	14	1
1	A	828	LEU	O-C-N	-5.13	116.15	122.20	1	1
1	A	846	ALA	N-CA-C	5.11	118.78	112.54	15	1
1	A	837	TYR	CA-C-N	5.10	126.61	121.35	9	1
1	A	837	TYR	C-N-CA	5.10	126.61	121.35	9	1
1	A	848	ASP	CB-CA-C	5.10	119.26	110.79	12	1
1	A	844	VAL	N-CA-CB	5.09	119.63	111.23	4	1
1	A	835	LYS	CG-CD-CE	5.06	122.94	111.30	2	1
1	A	836	GLN	OE1-CD-NE2	-5.06	117.54	122.60	1	1
1	A	834	ILE	CA-C-O	5.04	125.97	120.57	11	1
1	A	852	ALA	CB-CA-C	-5.02	102.36	110.79	8	1
1	A	817	PHE	O-C-N	-5.01	116.81	122.12	15	1
1	A	832	GLY	N-CA-C	5.00	119.80	113.24	12	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	837	TYR	Sidechain	3
1	A	847	ARG	Sidechain	2
1	A	833	PHE	Sidechain	1
1	A	843	ASP	Mainchain	1
1	A	855	PHE	Sidechain	1
1	A	841	LEU	Mainchain	1
1	A	823	PHE	Sidechain	1

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	309	305	304	0±0
All	All	4635	4575	4560	6

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 1.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:822:LEU:O	1:A:826:VAL:HG22	0.47	2.09	15	1
1:A:844:VAL:HG22	1:A:846:ALA:H	0.46	1.70	12	1
1:A:826:VAL:HG12	1:A:834:ILE:HD12	0.44	1.88	1	2
1:A:850:ILE:HG23	1:A:855:PHE:CZ	0.43	2.48	5	1
1:A:818:ILE:HG12	1:A:822:LEU:HD13	0.43	1.90	9	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	39/42 (93%)	35±2 (89±5%)	3±2 (9±5%)	1±0 (3±1%)	6	41
All	All	585/630 (93%)	519 (89%)	50 (9%)	16 (3%)	6	41

All 4 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	830	ASP	12
1	A	844	VAL	2
1	A	838	GLY	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	840	CYS	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	32/33 (97%)	31±1 (96±4%)	1±1 (4±4%)	28	78
All	All	480/495 (97%)	460 (96%)	20 (4%)	28	78

All 15 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	850	ILE	4
1	A	854	LYS	2
1	A	849	LEU	2
1	A	821	LEU	1
1	A	822	LEU	1
1	A	823	PHE	1
1	A	839	ASP	1
1	A	828	LEU	1
1	A	834	ILE	1
1	A	837	TYR	1
1	A	853	GLN	1
1	A	835	LYS	1
1	A	825	LYS	1
1	A	816	SER	1
1	A	818	ILE	1

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates [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 [i](#)

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

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *protein_nmrstar31_corrected1.str*

7.1.1 Bookkeeping [i](#)

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	488
Number of shifts mapped to atoms	488
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 [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	41	0.02 ± 0.50	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	38	0.34 ± 0.13	None needed (< 0.5 ppm)
$^{13}\text{C}'$	41	-0.40 ± 0.31	None needed (< 0.5 ppm)
^{15}N	40	-0.22 ± 0.21	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments [i](#)

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. 474 atoms were assigned a chemical shift out of a possible 537. 0 out of 7 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	197/203 (97%)	81/83 (98%)	78/80 (98%)	38/40 (95%)
Sidechain	241/285 (85%)	164/186 (88%)	77/90 (86%)	0/9 (0%)

Continued on next page...

Continued from previous page...

	Total	¹ H	¹³ C	¹⁵ N
Aromatic	36/49 (73%)	20/24 (83%)	16/25 (64%)	0/0 (—%)
Overall	474/537 (88%)	265/293 (90%)	171/195 (88%)	38/49 (78%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 88%, i.e. 488 atoms were assigned a chemical shift out of a possible 555. 0 out of 7 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	207/214 (97%)	85/88 (97%)	82/84 (98%)	40/42 (95%)
Sidechain	245/292 (84%)	166/190 (87%)	78/92 (85%)	1/10 (10%)
Aromatic	36/49 (73%)	20/24 (83%)	16/25 (64%)	0/0 (—%)
Overall	488/555 (88%)	271/302 (90%)	176/201 (88%)	41/52 (79%)

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:

