



Full wwPDB NMR Structure Validation Report ⓘ

Mar 25, 2026 – 10:15 AM UTC

PDB ID : 2K7X / pdb_00002k7x
Title : solution structure of C-terminal domain of SARS-CoV main protease
Authors : Xia, B.; Zhong, N.
Deposited on : 2008-08-28

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

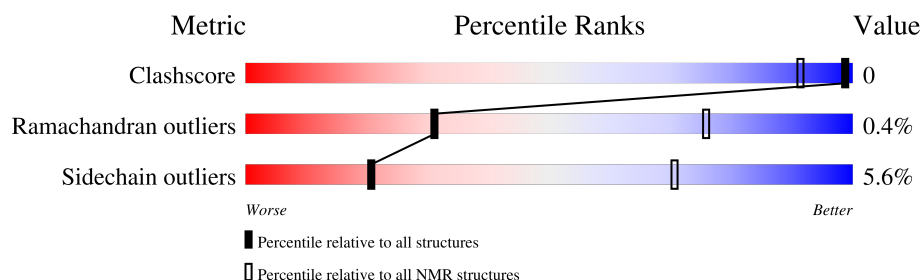
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment 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	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	120	

2 Ensemble composition and analysis

This entry contains 21 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:198-A:273, A:281-A:299 (95)	0.27	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. No single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 6, 8, 9, 10, 11, 12, 13, 14, 16, 18, 19, 20, 21
2	4, 5, 7, 15, 17

3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1851 atoms, of which 918 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called SARS-CoV main protease.

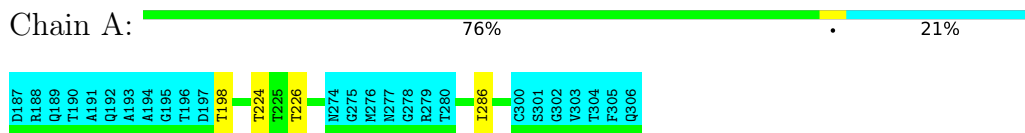
Mol	Chain	Residues	Atoms						Trace
1	A	120	Total	C	H	N	O	S	0
			1851	590	918	156	182	5	

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: SARS-CoV main protease

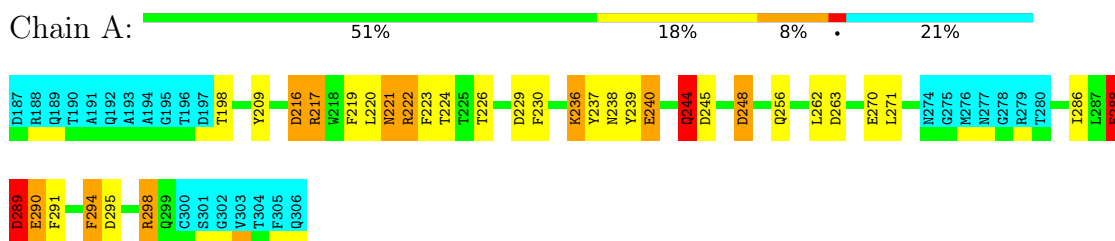


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

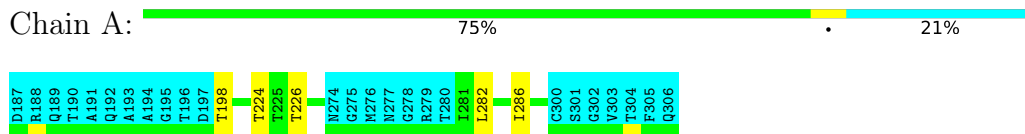
4.2.1 Score per residue for model 1 (medoid)

- Molecule 1: SARS-CoV main protease



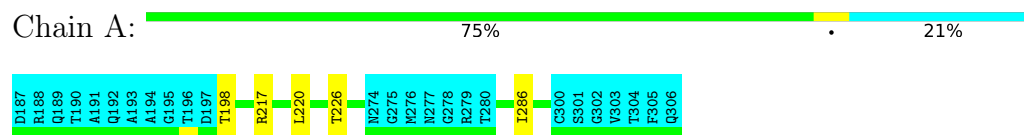
4.2.2 Score per residue for model 2

- Molecule 1: SARS-CoV main protease



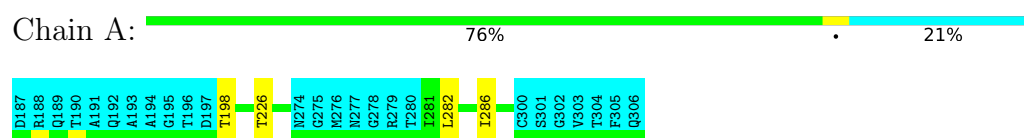
4.2.3 Score per residue for model 3

- Molecule 1: SARS-CoV main protease



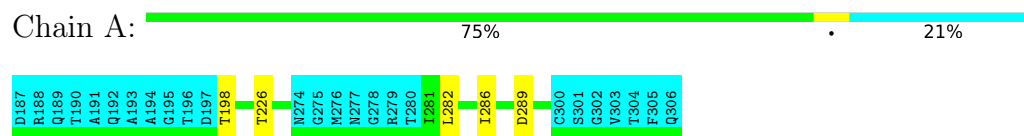
4.2.4 Score per residue for model 4

- Molecule 1: SARS-CoV main protease



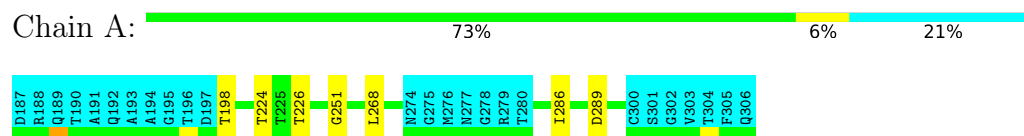
4.2.5 Score per residue for model 5

- Molecule 1: SARS-CoV main protease



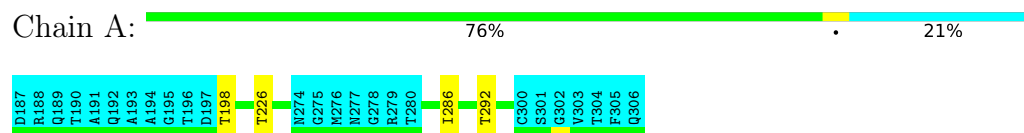
4.2.6 Score per residue for model 6

- Molecule 1: SARS-CoV main protease



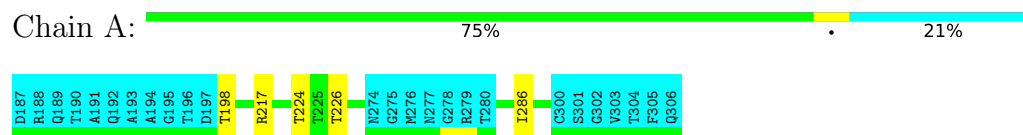
4.2.7 Score per residue for model 7

- Molecule 1: SARS-CoV main protease



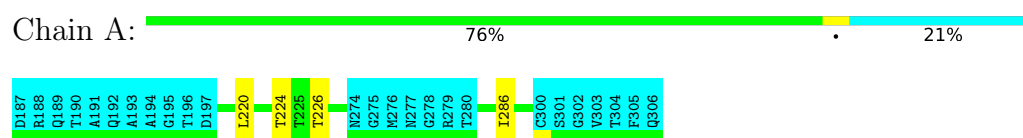
4.2.8 Score per residue for model 8

- Molecule 1: SARS-CoV main protease



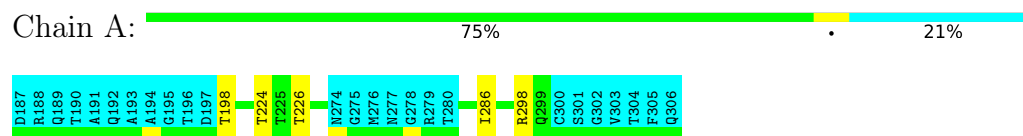
4.2.9 Score per residue for model 9

- Molecule 1: SARS-CoV main protease



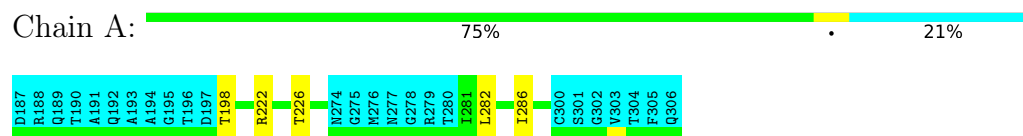
4.2.10 Score per residue for model 10

- Molecule 1: SARS-CoV main protease



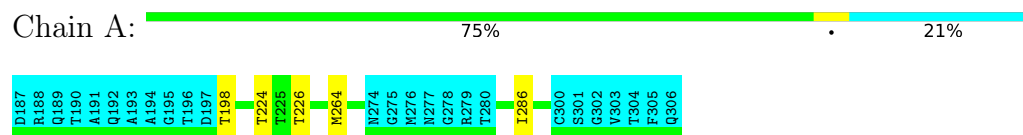
4.2.11 Score per residue for model 11

- Molecule 1: SARS-CoV main protease



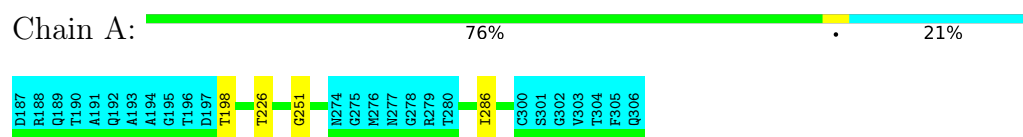
4.2.12 Score per residue for model 12

- Molecule 1: SARS-CoV main protease



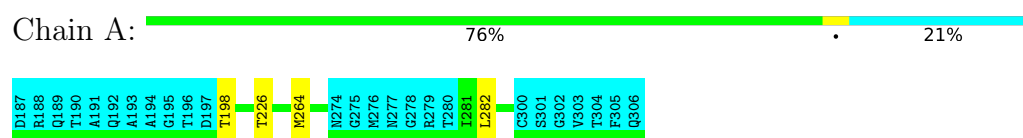
4.2.13 Score per residue for model 13

- Molecule 1: SARS-CoV main protease



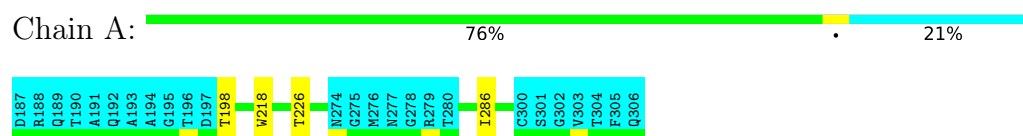
4.2.14 Score per residue for model 14

- Molecule 1: SARS-CoV main protease



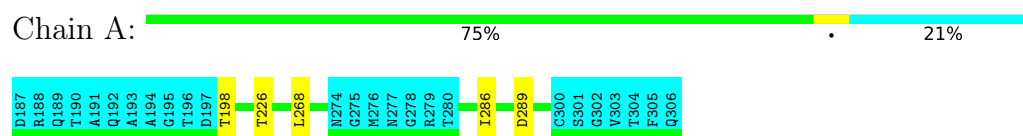
4.2.15 Score per residue for model 15

- Molecule 1: SARS-CoV main protease



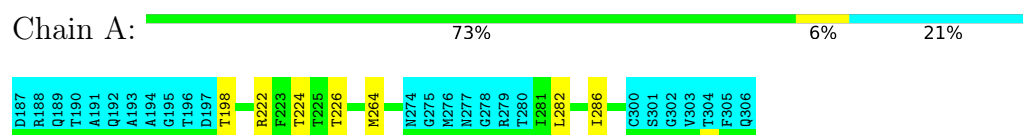
4.2.16 Score per residue for model 16

- Molecule 1: SARS-CoV main protease



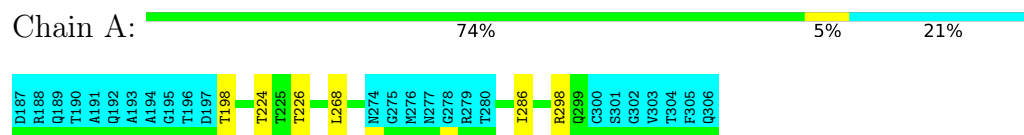
4.2.17 Score per residue for model 17

- Molecule 1: SARS-CoV main protease



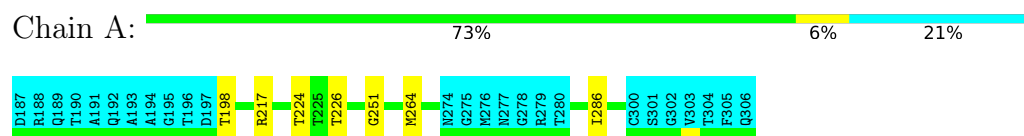
4.2.18 Score per residue for model 18

- Molecule 1: SARS-CoV main protease



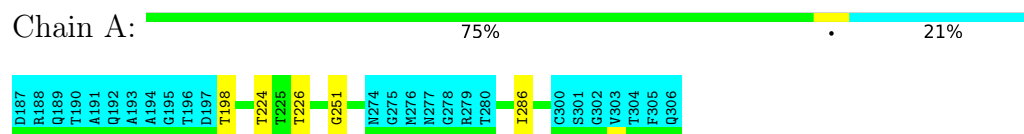
4.2.19 Score per residue for model 19

- Molecule 1: SARS-CoV main protease



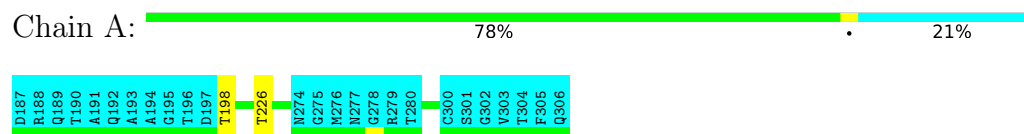
4.2.20 Score per residue for model 20

- Molecule 1: SARS-CoV main protease



4.2.21 Score per residue for model 21

- Molecule 1: SARS-CoV main protease



5 Refinement protocol and experimental data overview ⓘ

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

Of the 100 calculated structures, 21 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
Amber	refinement	
CYANA	structure solution	

No chemical shift data was provided.

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	0.68±0.25	1±3/770 (0.1± 0.4%)	1.29±0.45	5±21/1055 (0.5± 2.0%)
All	All	0.73	13/16170 (0.1%)	1.36	100/22155 (0.5%)

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	0.3±1.5
All	All	0	7

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	298	ARG	CD-NE	-15.13	1.25	1.46	1	1
1	A	217	ARG	CD-NE	-13.00	1.28	1.46	1	1
1	A	217	ARG	NE-CZ	-11.43	1.20	1.33	1	1
1	A	298	ARG	NE-CZ	-11.31	1.20	1.33	1	1
1	A	294	PHE	CB-CG	-8.66	1.30	1.50	1	1
1	A	298	ARG	CG-CD	-7.10	1.31	1.52	1	1
1	A	298	ARG	CZ-NH2	-6.96	1.24	1.33	1	1
1	A	221	ASN	CA-CB	-6.33	1.43	1.53	1	1
1	A	294	PHE	CG-CD2	-6.02	1.26	1.38	1	1
1	A	222	ARG	CD-NE	-5.95	1.38	1.46	1	1
1	A	270	GLU	CG-CD	-5.88	1.37	1.52	1	1
1	A	294	PHE	CG-CD1	-5.70	1.26	1.38	1	1
1	A	236	LYS	CD-CE	-5.70	1.35	1.52	1	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	221	ASN	CA-CB-CG	25.11	137.71	112.60	1	1
1	A	244	GLN	OE1-CD-NE2	-22.65	99.95	122.60	1	1
1	A	240	GLU	OE1-CD-OE2	-18.84	77.67	122.90	1	1
1	A	289	ASP	OD1-CG-OD2	-17.50	80.89	122.90	1	1
1	A	290	GLU	OE1-CD-OE2	-15.02	86.86	122.90	1	1
1	A	238	ASN	OD1-CG-ND2	-14.92	107.68	122.60	1	1
1	A	248	ASP	OD1-CG-OD2	-14.63	87.79	122.90	1	1
1	A	216	ASP	OD1-CG-OD2	-14.58	87.91	122.90	1	1
1	A	298	ARG	NE-CZ-NH2	-14.33	106.30	119.20	1	1
1	A	237	TYR	CD1-CG-CD2	-14.22	96.77	118.10	1	1
1	A	230	PHE	CD1-CG-CD2	-14.07	97.49	118.60	1	1
1	A	230	PHE	CE1-CZ-CE2	-13.93	94.92	120.00	1	1
1	A	239	TYR	CE1-CZ-CE2	-13.88	92.54	120.30	1	1
1	A	239	TYR	CD1-CG-CD2	-13.79	97.42	118.10	1	1
1	A	244	GLN	CG-CD-NE2	13.72	136.98	116.40	1	1
1	A	288	GLU	OE1-CD-OE2	-13.47	90.56	122.90	1	1
1	A	237	TYR	CE1-CZ-CE2	-13.19	93.91	120.30	1	1
1	A	291	PHE	CE1-CZ-CE2	-12.91	96.76	120.00	1	1
1	A	291	PHE	CD1-CG-CD2	-12.65	99.62	118.60	1	1
1	A	219	PHE	CE1-CZ-CE2	-12.47	97.56	120.00	1	1
1	A	217	ARG	NE-CZ-NH2	-12.10	108.31	119.20	1	1
1	A	219	PHE	CD1-CG-CD2	-12.02	100.57	118.60	1	1
1	A	209	TYR	CE1-CZ-CE2	-11.90	96.49	120.30	1	1
1	A	245	ASP	OD1-CG-OD2	-11.61	95.03	122.90	1	1
1	A	209	TYR	CD1-CG-CD2	-11.49	100.86	118.10	1	1
1	A	223	PHE	CE1-CZ-CE2	-10.46	101.16	120.00	1	1
1	A	223	PHE	CD1-CG-CD2	-9.92	103.72	118.60	1	1
1	A	239	TYR	CE1-CZ-OH	9.80	149.31	119.90	1	1
1	A	237	TYR	CE1-CZ-OH	9.74	149.11	119.90	1	1
1	A	229	ASP	OD1-CG-OD2	-9.68	99.67	122.90	1	1
1	A	239	TYR	CZ-CE2-CD2	9.63	136.93	119.60	1	1
1	A	248	ASP	CA-CB-CG	-8.80	103.80	112.60	1	1
1	A	262	LEU	CD1-CG-CD2	-8.59	91.90	110.80	1	1
1	A	237	TYR	CZ-CE2-CD2	8.58	135.04	119.60	1	1
1	A	298	ARG	CD-NE-CZ	-8.45	112.57	124.40	1	1
1	A	270	GLU	CG-CD-OE1	8.42	137.77	118.40	1	1
1	A	295	ASP	OD1-CG-OD2	-8.35	102.87	122.90	1	1
1	A	263	ASP	OD1-CG-OD2	-8.29	102.99	122.90	1	1
1	A	209	TYR	CD1-CE1-CZ	7.86	133.75	119.60	1	1
1	A	240	GLU	CG-CD-OE2	7.79	136.31	118.40	1	1
1	A	209	TYR	OH-CZ-CE2	7.76	143.20	119.90	1	1
1	A	237	TYR	CB-CG-CD1	7.61	132.21	120.80	1	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	240	GLU	CB-CG-CD	-7.49	99.87	112.60	1	1
1	A	221	ASN	CB-CG-ND2	-7.42	105.27	116.40	1	1
1	A	248	ASP	CB-CG-OD2	7.41	135.45	118.40	1	1
1	A	290	GLU	CG-CD-OE1	7.36	135.32	118.40	1	1
1	A	221	ASN	N-CA-CB	-7.30	99.11	111.31	1	1
1	A	294	PHE	CB-CA-C	-7.23	98.78	110.79	1	1
1	A	298	ARG	CA-CB-CG	-7.22	99.65	114.10	1	1
1	A	239	TYR	CB-CG-CD1	7.16	131.54	120.80	1	1
1	A	217	ARG	NH1-CZ-NH2	7.06	128.48	119.30	1	1
1	A	230	PHE	CD1-CE1-CZ	6.87	132.36	120.00	1	1
1	A	219	PHE	CB-CG-CD2	6.86	132.36	120.70	1	1
1	A	288	GLU	CB-CG-CD	-6.86	100.94	112.60	1	1
1	A	230	PHE	CZ-CE2-CD2	6.86	132.34	120.00	1	1
1	A	239	TYR	CD1-CE1-CZ	6.84	131.91	119.60	1	1
1	A	291	PHE	CZ-CE2-CD2	6.63	131.94	120.00	1	1
1	A	291	PHE	CD1-CE1-CZ	6.62	131.91	120.00	1	1
1	A	217	ARG	O-C-N	-6.53	111.70	121.63	1	1
1	A	256	GLN	OE1-CD-NE2	-6.48	116.12	122.60	1	1
1	A	219	PHE	CD1-CE1-CZ	6.45	131.61	120.00	1	1
1	A	209	TYR	CZ-CE2-CD2	6.36	131.04	119.60	1	1
1	A	222	ARG	O-C-N	-6.32	114.58	122.35	1	1
1	A	290	GLU	CB-CG-CD	-6.30	101.89	112.60	1	1
1	A	221	ASN	OD1-CG-ND2	6.24	128.84	122.60	1	1
1	A	289	ASP	CA-CB-CG	-6.19	106.41	112.60	1	1
1	A	239	TYR	CG-CD2-CE2	6.17	130.46	121.20	1	1
1	A	219	PHE	CZ-CE2-CD2	6.10	130.99	120.00	1	1
1	A	216	ASP	CA-CB-CG	-6.04	106.56	112.60	1	1
1	A	230	PHE	CB-CG-CD2	5.95	130.82	120.70	1	1
1	A	239	TYR	CB-CG-CD2	5.95	129.73	120.80	1	1
1	A	237	TYR	CG-CD2-CE2	5.94	130.11	121.20	1	1
1	A	245	ASP	CA-CB-CG	-5.92	106.68	112.60	1	1
1	A	245	ASP	CB-CG-OD1	5.91	131.98	118.40	1	1
1	A	209	TYR	CB-CG-CD1	5.89	129.63	120.80	1	1
1	A	239	TYR	CG-CD1-CE1	5.88	130.02	121.20	1	1
1	A	237	TYR	CD1-CE1-CZ	5.88	130.18	119.60	1	1
1	A	230	PHE	CG-CD1-CE1	5.81	130.58	120.70	1	1
1	A	209	TYR	CB-CG-CD2	5.74	129.41	120.80	1	1
1	A	291	PHE	CB-CG-CD1	5.70	130.39	120.70	1	1
1	A	237	TYR	CG-CD1-CE1	5.69	129.73	121.20	1	1
1	A	222	ARG	NE-CZ-NH2	-5.63	114.13	119.20	1	1
1	A	298	ARG	NH1-CZ-NH2	5.63	126.62	119.30	1	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	294	PHE	CE1-CZ-CE2	-5.62	109.89	120.00	1	1
1	A	217	ARG	CG-CD-NE	-5.56	99.78	112.00	1	1
1	A	230	PHE	CG-CD2-CE2	5.55	130.13	120.70	1	1
1	A	220	LEU	CD1-CG-CD2	-5.53	98.63	110.80	1	1
1	A	298	ARG	NE-CZ-NH1	5.50	127.00	121.50	1	1
1	A	291	PHE	CB-CG-CD2	5.43	129.94	120.70	1	1
1	A	291	PHE	CG-CD1-CE1	5.41	129.89	120.70	1	1
1	A	223	PHE	CD1-CE1-CZ	5.40	129.72	120.00	1	1
1	A	291	PHE	CG-CD2-CE2	5.39	129.86	120.70	1	1
1	A	223	PHE	CZ-CE2-CD2	5.36	129.65	120.00	1	1
1	A	229	ASP	CB-CG-OD2	5.34	130.69	118.40	1	1
1	A	209	TYR	CG-CD2-CE2	5.32	129.18	121.20	1	1
1	A	244	GLN	CB-CG-CD	5.29	121.60	112.60	1	1
1	A	294	PHE	N-CA-C	5.23	116.98	111.28	1	1
1	A	216	ASP	CB-CG-OD1	5.21	130.39	118.40	1	1
1	A	219	PHE	CG-CD1-CE1	5.17	129.50	120.70	1	1
1	A	224	THR	CB-CA-C	5.01	118.27	111.70	10	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	216	ASP	Sidechain	1
1	A	240	GLU	Sidechain	1
1	A	248	ASP	Sidechain	1
1	A	288	GLU	Sidechain	1
1	A	289	ASP	Sidechain	1
1	A	290	GLU	Sidechain	1
1	A	294	PHE	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	754	754	754	0±1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes
All	All	15834	15834	15834	4

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:222:ARG:HD3	1:A:222:ARG:O	0.60	1.96	1	1
1:A:217:ARG:HG3	1:A:217:ARG:O	0.52	2.04	1	1
1:A:244:GLN:HE21	1:A:244:GLN:HA	0.50	1.66	1	1
1:A:220:LEU:N	1:A:220:LEU:HD22	0.41	2.30	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	95/120 (79%)	90±1 (94±1%)	5±1 (5±1%)	0±1 (0±1%)	31	76
All	All	1995/2520 (79%)	1880 (94%)	108 (5%)	7 (0%)	31	76

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	251	GLY	4
1	A	217	ARG	3

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	83/101 (82%)	78±2 (94±2%)	5±2 (6±2%)	21	70
All	All	1743/2121 (82%)	1645 (94%)	98 (6%)	21	70

All 16 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	226	THR	21
1	A	198	THR	20
1	A	286	ILE	19
1	A	224	THR	10
1	A	282	LEU	6
1	A	289	ASP	4
1	A	264	MET	4
1	A	298	ARG	3
1	A	268	LEU	3
1	A	222	ARG	2
1	A	221	ASN	1
1	A	236	LYS	1
1	A	244	GLN	1
1	A	288	GLU	1
1	A	220	LEU	1
1	A	292	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](#)

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

No chemical shift data were provided