Full wwPDB NMR Structure Validation Report (i)

Mar 29, 2026 – 02:43 PM UTC

PDB ID	:	1UKX / pdb_00001ukx
Title	:	Solution structure of the RWD domain of mouse GCN2
Authors	:	Nameki, N.; Yoneyama, M.; Koshiba, S.; Inoue, M.; Kigawa, T.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2003-09-03

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

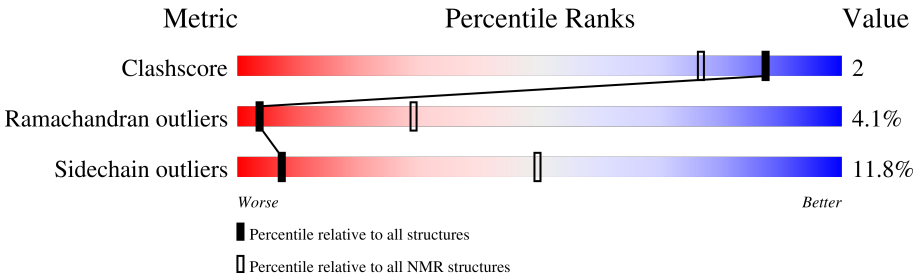
MolProbity	:	4-5-2 with Phenix2.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	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 <=5%

Mol	Chain	Length	Quality of chain
1	A	137	66% 9% 25%

2 Ensemble composition and analysis

This entry contains 20 models. Model 16 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:11-A:33, A:44-A:53, A:61-A:130 (103)	0.50	16

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 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2096 atoms, of which 1033 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called GCN2 eIF2alpha kinase.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	137	2096	661	1033	185	213	4	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP Q9QZ05
A	2	SER	-	cloning artifact	UNP Q9QZ05
A	3	SER	-	cloning artifact	UNP Q9QZ05
A	4	GLY	-	cloning artifact	UNP Q9QZ05
A	5	SER	-	cloning artifact	UNP Q9QZ05
A	6	SER	-	cloning artifact	UNP Q9QZ05
A	7	GLY	-	cloning artifact	UNP Q9QZ05
A	8	MET	-	cloning artifact	UNP Q9QZ05
A	132	SER	-	cloning artifact	UNP Q9QZ05
A	133	GLY	-	cloning artifact	UNP Q9QZ05
A	134	PRO	-	cloning artifact	UNP Q9QZ05
A	135	SER	-	cloning artifact	UNP Q9QZ05
A	136	SER	-	cloning artifact	UNP Q9QZ05
A	137	GLY	-	cloning artifact	UNP Q9QZ05

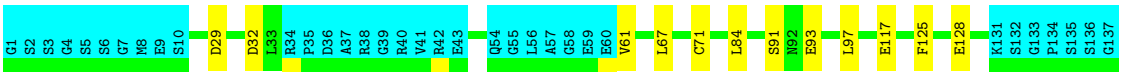
4 Residue-property plots

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: GCN2 eIF2alpha kinase

Chain A:



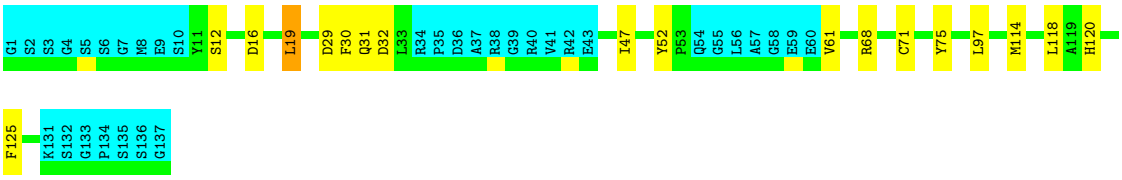
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: GCN2 eIF2alpha kinase

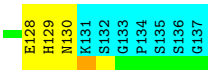
Chain A:



4.2.2 Score per residue for model 2

• Molecule 1: GCN2 eIF2alpha kinase

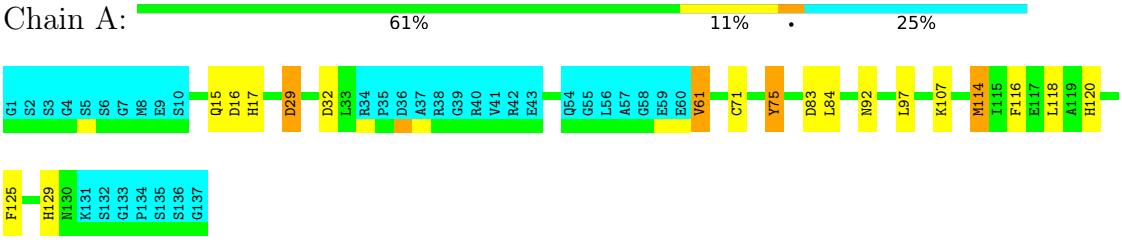
Chain A:



4.2.3 Score per residue for model 3

• Molecule 1: GCN2 eIF2alpha kinase

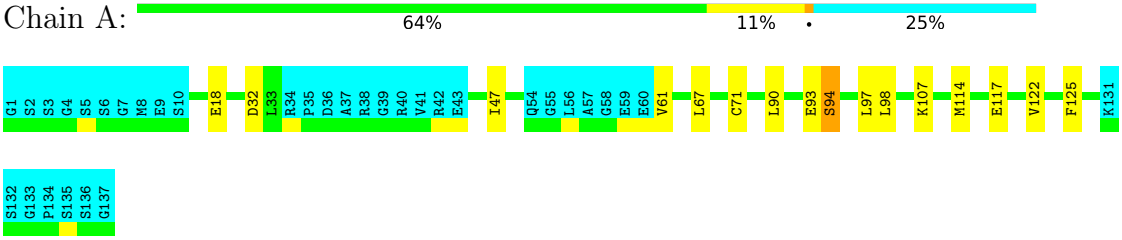
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: GCN2 eIF2alpha kinase

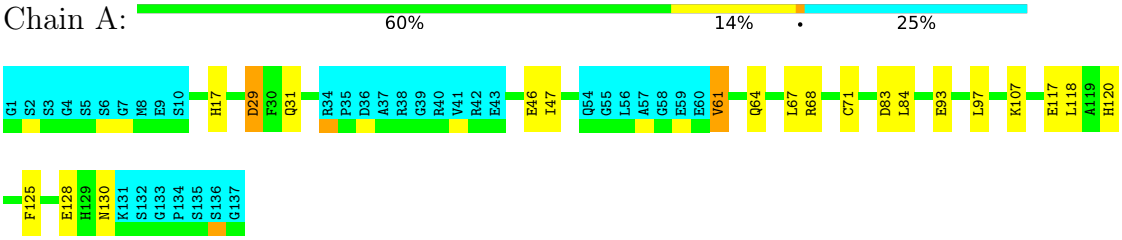
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: GCN2 eIF2alpha kinase

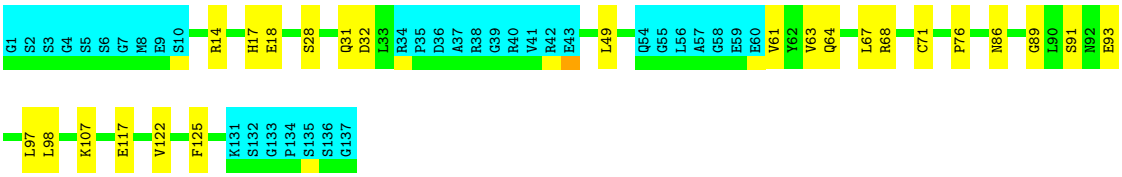
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: GCN2 eIF2alpha kinase

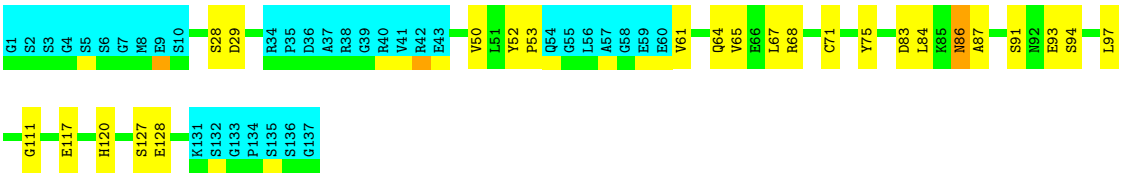
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: GCN2 eIF2alpha kinase

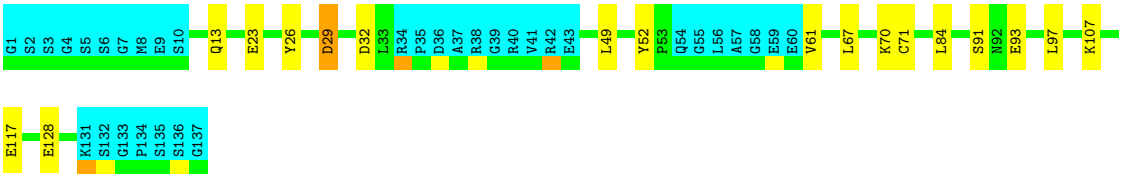
Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: GCN2 eIF2alpha kinase

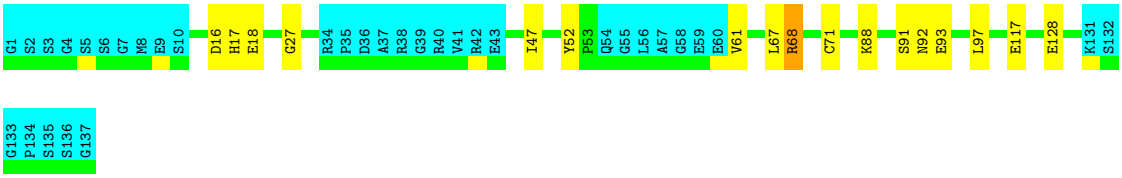
Chain A:



4.2.9 Score per residue for model 9

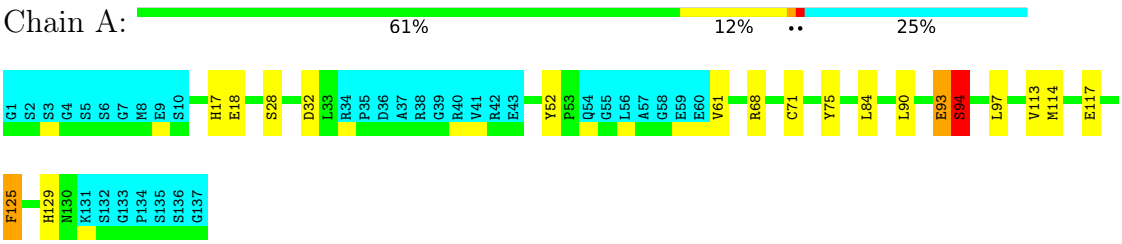
• Molecule 1: GCN2 eIF2alpha kinase

Chain A:



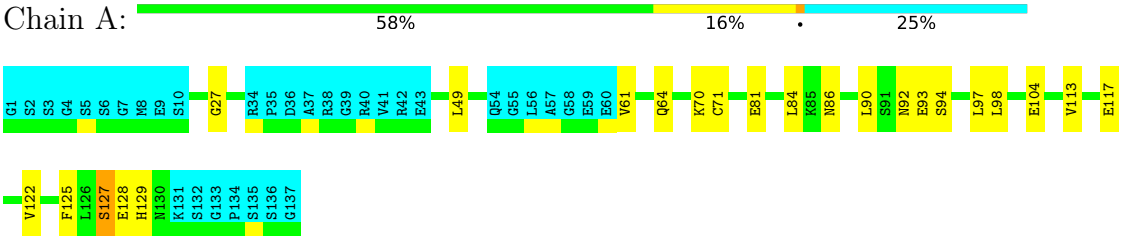
4.2.10 Score per residue for model 10

• Molecule 1: GCN2 eIF2alpha kinase



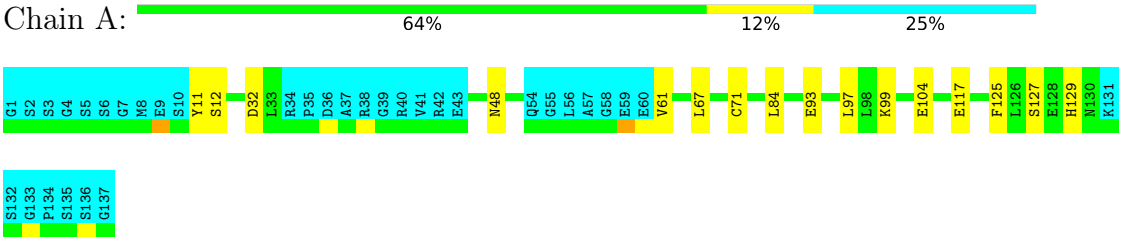
4.2.11 Score per residue for model 11

- Molecule 1: GCN2 eIF2alpha kinase



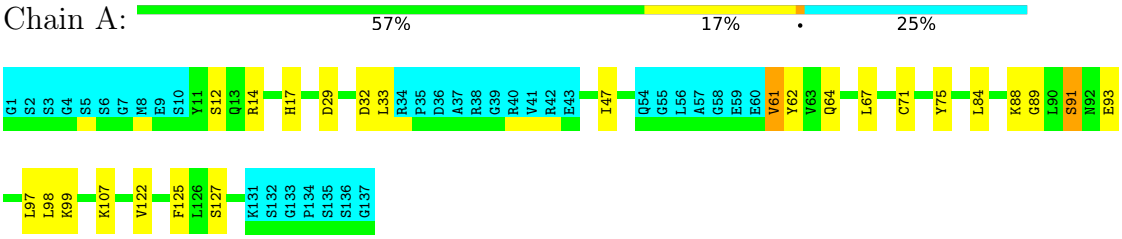
4.2.12 Score per residue for model 12

- Molecule 1: GCN2 eIF2alpha kinase



4.2.13 Score per residue for model 13

- Molecule 1: GCN2 eIF2alpha kinase



4.2.14 Score per residue for model 14

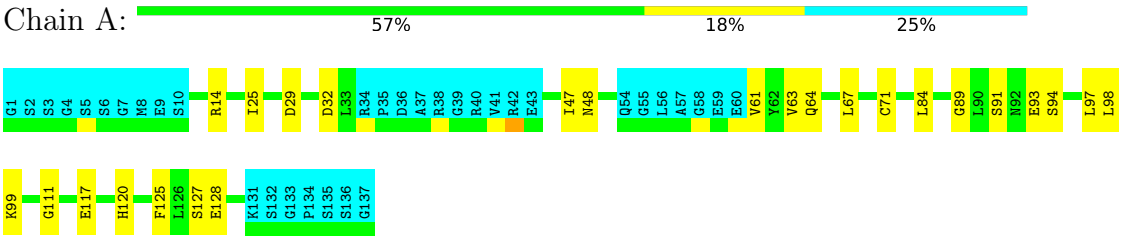
• Molecule 1: GCN2 eIF2alpha kinase

Chain A:

4.2.15 Score per residue for model 15

• Molecule 1: GCN2 eIF2alpha kinase

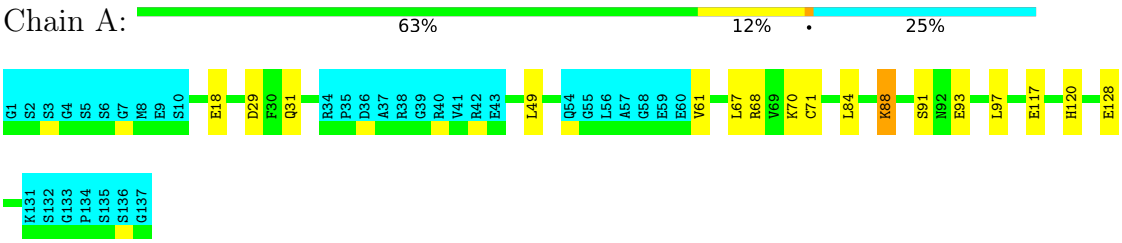
Chain A:



4.2.16 Score per residue for model 16 (medoid)

• Molecule 1: GCN2 eIF2alpha kinase

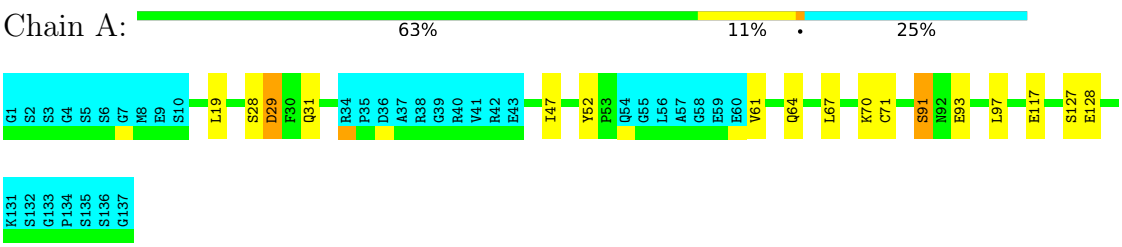
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: GCN2 eIF2alpha kinase

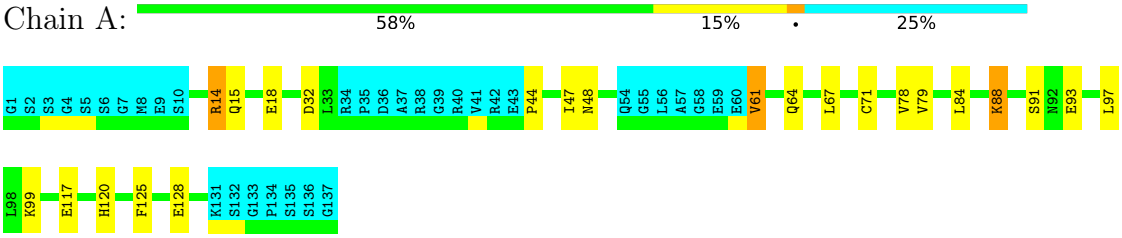
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: GCN2 eIF2alpha kinase

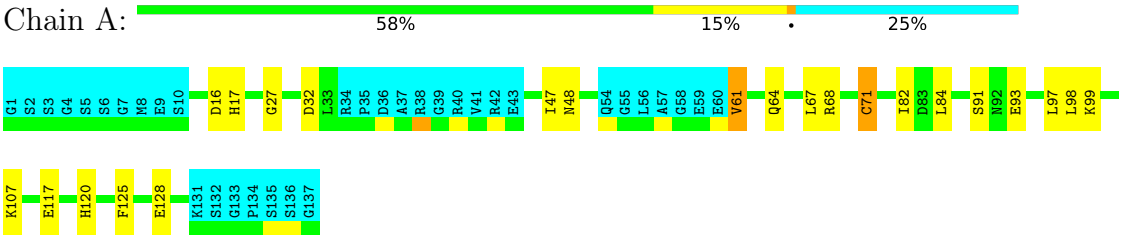
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: GCN2 eIF2alpha kinase

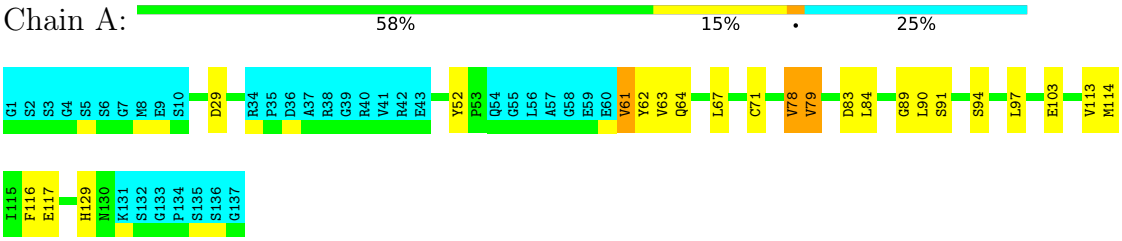
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: GCN2 eIF2alpha kinase

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics, restrained molecular dynamics.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations, structures with the lowest energy, target function.*

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

Software name	Classification	Version
CYANA	structure solution	1.0.7
OPALp	refinement	

No chemical shift data was provided.

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	0.70±0.01	0±0/850 (0.0± 0.0%)	1.41±0.04	2±2/1156 (0.2± 0.2%)
All	All	0.70	1/17000 (0.0%)	1.42	40/23120 (0.2%)

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	0.8±0.8
All	All	0	16

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	44	PRO	CA-C	5.27	1.54	1.51	18	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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	16	ASP	CA-CB-CG	10.14	122.74	112.60	14	2
1	A	29	ASP	CA-CB-CG	9.58	122.18	112.60	14	6
1	A	113	VAL	N-CA-C	6.41	113.07	106.21	11	2
1	A	86	ASN	CA-CB-CG	6.33	118.93	112.60	7	3
1	A	116	PHE	CA-CB-CG	6.32	120.12	113.80	3	2
1	A	104	GLU	CA-C-N	5.70	128.23	120.54	11	2
1	A	104	GLU	C-N-CA	5.70	128.23	120.54	11	2
1	A	130	ASN	CA-CB-CG	5.49	118.09	112.60	5	2
1	A	15	GLN	CA-C-N	5.44	127.57	120.28	3	2
1	A	15	GLN	C-N-CA	5.44	127.57	120.28	3	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	127	SER	CA-C-N	5.38	127.49	120.28	11	2
1	A	127	SER	C-N-CA	5.38	127.49	120.28	11	2
1	A	53	PRO	CA-C-N	5.36	131.78	121.54	7	1
1	A	53	PRO	C-N-CA	5.36	131.78	121.54	7	1
1	A	93	GLU	CA-C-N	5.31	131.68	121.54	10	1
1	A	93	GLU	C-N-CA	5.31	131.68	121.54	10	1
1	A	89	GLY	CA-C-N	5.30	127.82	120.29	13	1
1	A	89	GLY	C-N-CA	5.30	127.82	120.29	13	1
1	A	29	ASP	CA-C-N	5.13	128.76	121.42	17	1
1	A	29	ASP	C-N-CA	5.13	128.76	121.42	17	1
1	A	62	TYR	CA-C-N	5.13	129.23	122.51	20	1
1	A	62	TYR	C-N-CA	5.13	129.23	122.51	20	1
1	A	61	VAL	CA-CB-CG2	5.11	119.08	110.40	2	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	52	TYR	Sidechain	5
1	A	14	ARG	Sidechain	4
1	A	68	ARG	Sidechain	2
1	A	30	PHE	Peptide	1
1	A	26	TYR	Sidechain	1
1	A	125	PHE	Sidechain	1
1	A	70	LYS	Peptide	1
1	A	75	TYR	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	830	815	815	3±1
All	All	16600	16300	16300	65

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:98:LEU:HD21	1:A:122:VAL:HG22	0.66	1.65	13	5
1:A:98:LEU:HD22	1:A:125:PHE:CD1	0.62	2.30	4	1
1:A:78:VAL:HG22	1:A:79:VAL:H	0.61	1.56	20	2
1:A:98:LEU:HD22	1:A:125:PHE:CD2	0.59	2.32	13	6
1:A:90:LEU:HD22	1:A:94:SER:OG	0.59	1.96	20	3
1:A:97:LEU:HD22	1:A:125:PHE:CE2	0.57	2.33	12	3
1:A:16:ASP:HA	1:A:19:LEU:HD22	0.55	1.77	1	2
1:A:63:VAL:HG12	1:A:89:GLY:HA3	0.52	1.81	6	3
1:A:97:LEU:HD22	1:A:125:PHE:CE1	0.51	2.41	1	5
1:A:50:VAL:HG12	1:A:52:TYR:HE1	0.51	1.66	7	1
1:A:97:LEU:C	1:A:97:LEU:HD23	0.49	2.33	1	20
1:A:50:VAL:HG12	1:A:52:TYR:CE1	0.48	2.42	7	1
1:A:82:ILE:HG22	1:A:99:LYS:HG3	0.46	1.88	19	1
1:A:90:LEU:HD22	1:A:94:SER:HB2	0.46	1.86	4	1
1:A:33:LEU:CD1	1:A:47:ILE:HA	0.43	2.43	13	1
1:A:75:TYR:CD1	1:A:75:TYR:C	0.43	2.96	3	3
1:A:61:VAL:HG12	1:A:62:TYR:H	0.43	1.74	13	1
1:A:19:LEU:HD22	1:A:47:ILE:HG21	0.42	1.91	17	1
1:A:75:TYR:CE2	1:A:114:MET:SD	0.42	3.12	10	2
1:A:125:PHE:C	1:A:125:PHE:CD1	0.41	2.98	5	1
1:A:13:GLN:HA	1:A:16:ASP:CG	0.41	2.40	14	1
1:A:65:VAL:HG23	1:A:87:ALA:HB2	0.40	1.94	7	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	103/137 (75%)	91±2 (88±2%)	8±2 (8±2%)	4±1 (4±1%)	4	29
All	All	2060/2740 (75%)	1812 (88%)	163 (8%)	85 (4%)	4	29

All 13 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	61	VAL	20
1	A	71	CYS	20
1	A	93	GLU	15
1	A	91	SER	12
1	A	88	LYS	4
1	A	28	SER	3
1	A	27	GLY	3
1	A	92	ASN	2
1	A	111	GLY	2
1	A	94	SER	1
1	A	11	TYR	1
1	A	78	VAL	1
1	A	113	VAL	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	95/119 (80%)	84±2 (88±3%)	11±2 (12±3%)	7 49
All	All	1900/2380 (80%)	1676 (88%)	224 (12%)	7 49

All 46 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	67	LEU	16
1	A	117	GLU	16
1	A	84	LEU	15
1	A	32	ASP	13
1	A	128	GLU	12
1	A	64	GLN	10
1	A	120	HIS	9
1	A	29	ASP	8
1	A	47	ILE	8
1	A	68	ARG	8
1	A	18	GLU	8
1	A	17	HIS	7
1	A	107	LYS	7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	61	VAL	6
1	A	129	HIS	6
1	A	31	GLN	5
1	A	127	SER	5
1	A	114	MET	4
1	A	83	ASP	4
1	A	94	SER	4
1	A	49	LEU	4
1	A	91	SER	4
1	A	48	ASN	4
1	A	99	LYS	4
1	A	12	SER	3
1	A	118	LEU	3
1	A	70	LYS	3
1	A	88	LYS	3
1	A	19	LEU	2
1	A	75	TYR	2
1	A	14	ARG	2
1	A	23	GLU	2
1	A	16	ASP	2
1	A	93	GLU	2
1	A	25	ILE	2
1	A	46	GLU	1
1	A	28	SER	1
1	A	76	PRO	1
1	A	86	ASN	1
1	A	13	GLN	1
1	A	92	ASN	1
1	A	81	GLU	1
1	A	71	CYS	1
1	A	52	TYR	1
1	A	79	VAL	1
1	A	103	GLU	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

There are no oligosaccharides 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

No chemical shift data were provided