



Full wwPDB NMR Structure Validation Report ⓘ

Feb 18, 2018 – 11:53 pm GMT

PDB ID : 2K3X
Title : Solution structure of EAF3 chromo barrel domain
Authors : Mer, G.; Xu, C.
Deposited on : 2008-05-19

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 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
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk30686
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30686

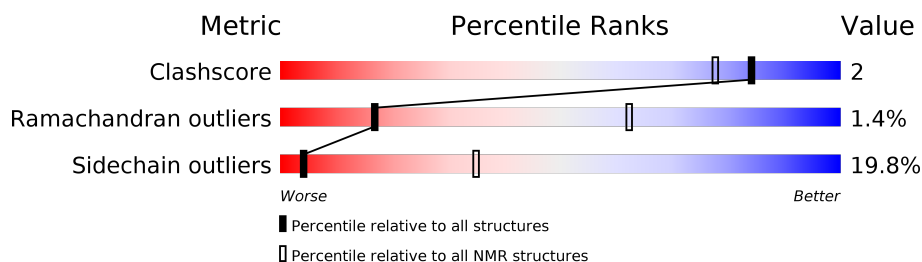
1 Overall quality at a glance

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	136279	12091
Ramachandran outliers	132675	10835
Sidechain outliers	132484	10811

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	121	

2 Ensemble composition and analysis

This entry contains 20 models. Model 15 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:6-A:41, A:53-A:115 (99)	0.30	15

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 2 single-model clusters were found.

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1842 atoms, of which 910 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Chromatin modification-related protein EAF3.

Mol	Chain	Residues	Atoms						Trace
1	A	115	Total	C	H	N	O	S	0
			1842	596	910	155	175	6	

There are 8 discrepancies between the modelled and reference sequences:

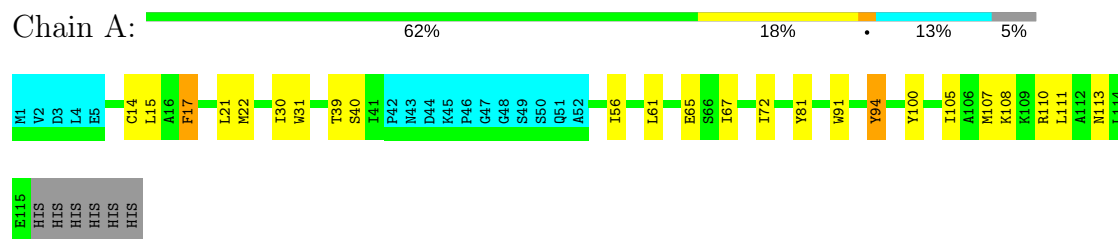
Chain	Residue	Modelled	Actual	Comment	Reference
A	114	LEU	-	EXPRESSION TAG	UNP Q12432
A	115	GLU	-	EXPRESSION TAG	UNP Q12432
A	116	HIS	-	EXPRESSION TAG	UNP Q12432
A	117	HIS	-	EXPRESSION TAG	UNP Q12432
A	118	HIS	-	EXPRESSION TAG	UNP Q12432
A	119	HIS	-	EXPRESSION TAG	UNP Q12432
A	120	HIS	-	EXPRESSION TAG	UNP Q12432
A	121	HIS	-	EXPRESSION TAG	UNP Q12432

4 Residue-property plots [i](#)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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: Chromatin modification-related protein EAF3



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

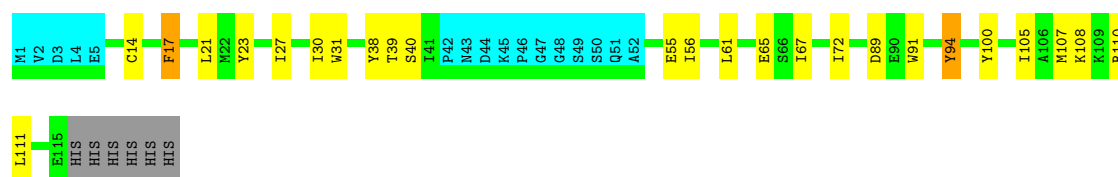
- Molecule 1: Chromatin modification-related protein EAF3



4.2.2 Score per residue for model 2

- Molecule 1: Chromatin modification-related protein EAF3

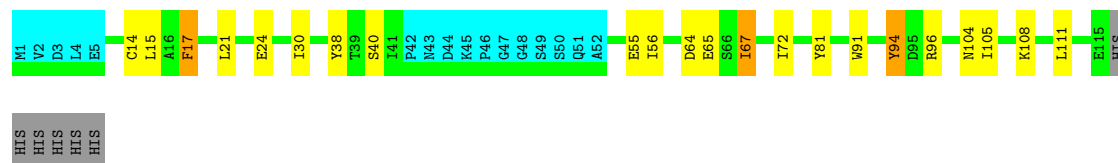




4.2.3 Score per residue for model 3

- Molecule 1: Chromatin modification-related protein EAF3

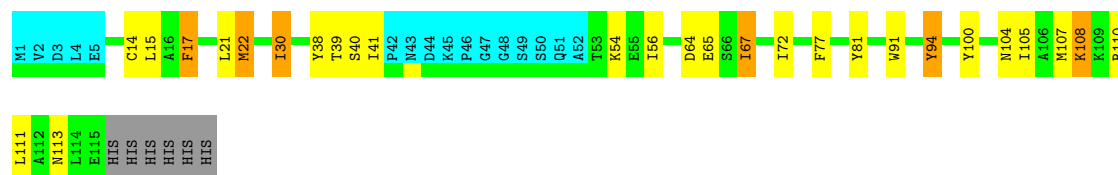
Chain A:



4.2.4 Score per residue for model 4

- Molecule 1: Chromatin modification-related protein EAF3

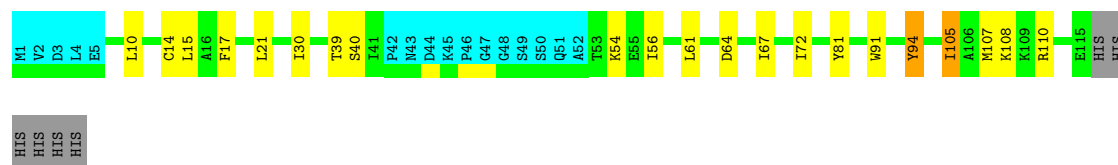
Chain A:



4.2.5 Score per residue for model 5

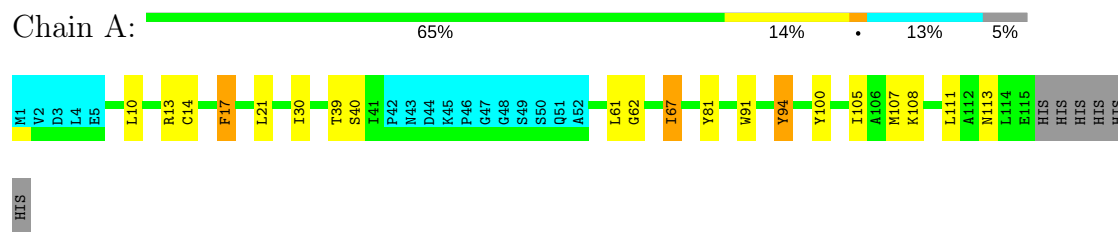
- Molecule 1: Chromatin modification-related protein EAF3

Chain A:



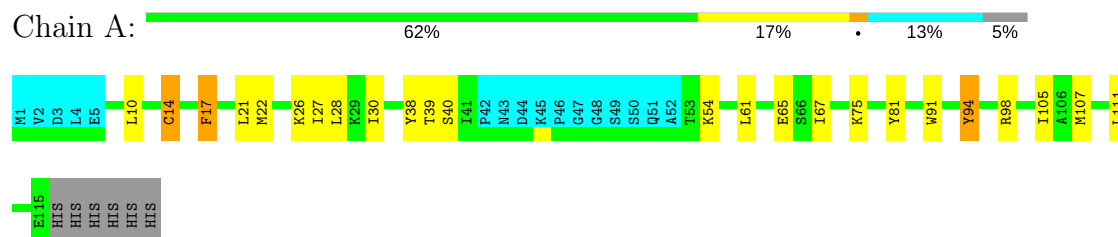
4.2.6 Score per residue for model 6

- Molecule 1: Chromatin modification-related protein EAF3



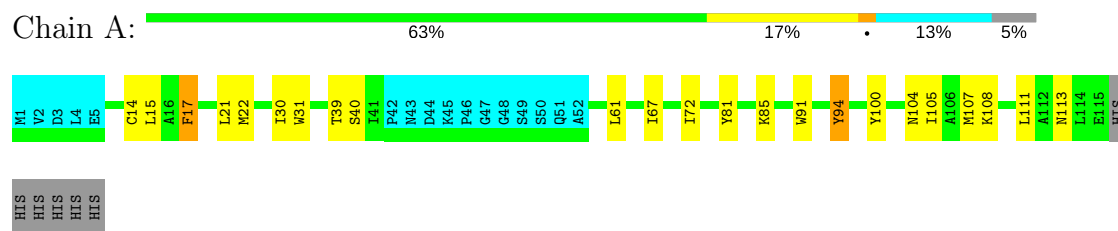
4.2.7 Score per residue for model 7

- Molecule 1: Chromatin modification-related protein EAF3



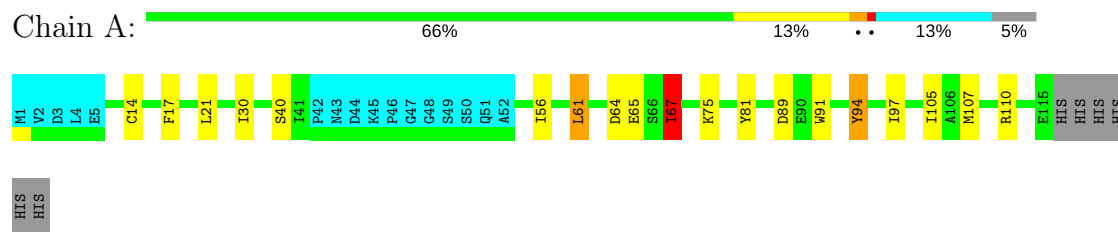
4.2.8 Score per residue for model 8

- Molecule 1: Chromatin modification-related protein EAF3



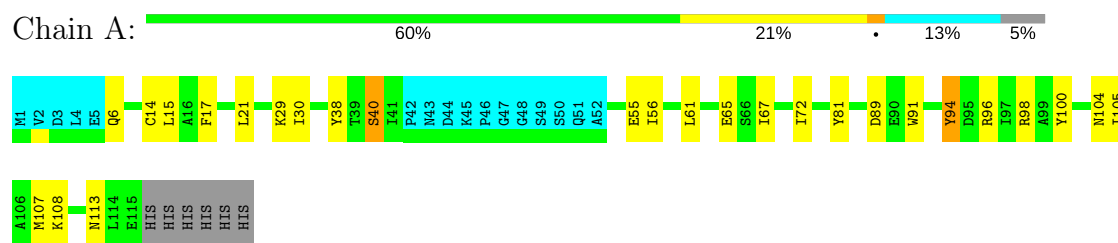
4.2.9 Score per residue for model 9

- Molecule 1: Chromatin modification-related protein EAF3



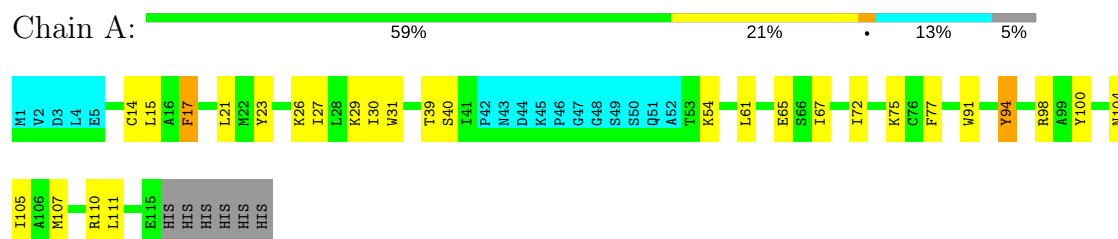
4.2.10 Score per residue for model 10

- Molecule 1: Chromatin modification-related protein EAF3



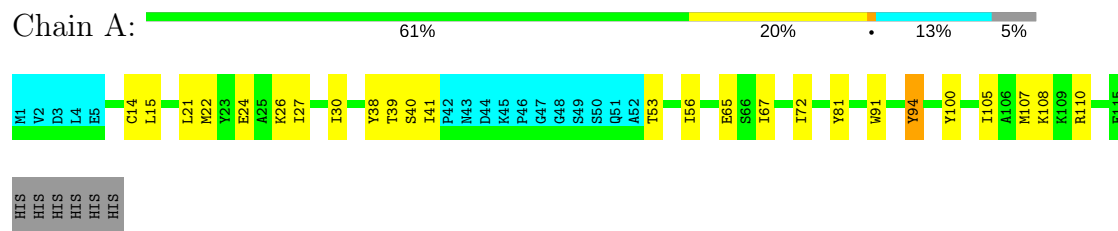
4.2.11 Score per residue for model 11

- Molecule 1: Chromatin modification-related protein EAF3



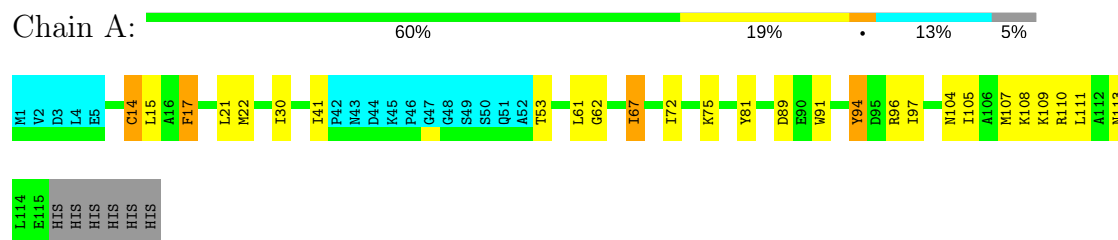
4.2.12 Score per residue for model 12

- Molecule 1: Chromatin modification-related protein EAF3



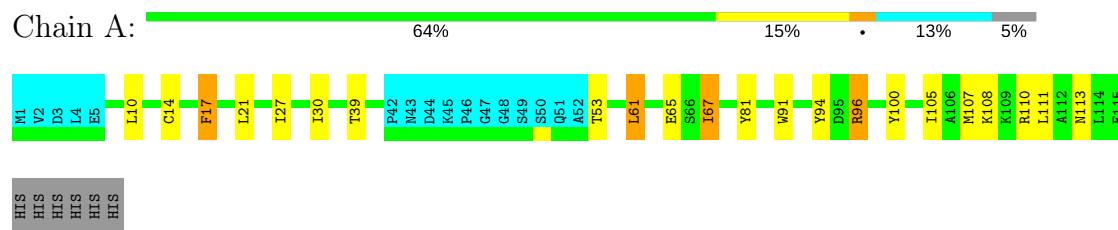
4.2.13 Score per residue for model 13

- Molecule 1: Chromatin modification-related protein EAF3



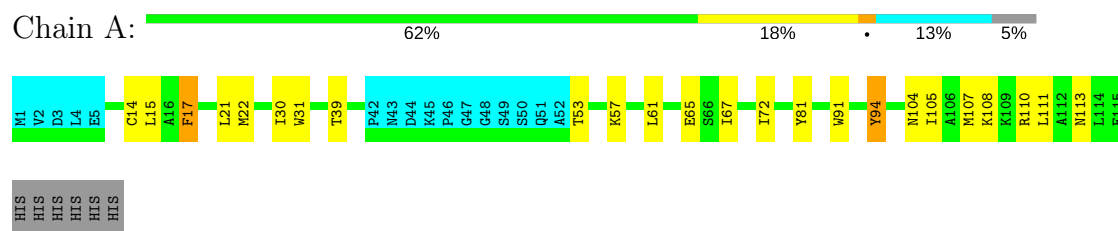
4.2.14 Score per residue for model 14

- Molecule 1: Chromatin modification-related protein EAF3



4.2.15 Score per residue for model 15 (medoid)

- Molecule 1: Chromatin modification-related protein EAF3



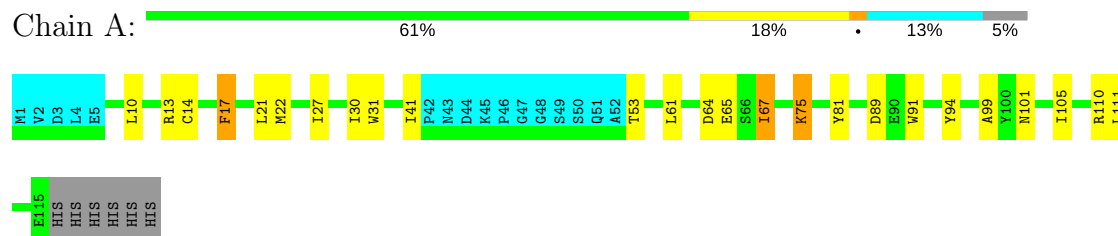
4.2.16 Score per residue for model 16

- Molecule 1: Chromatin modification-related protein EAF3



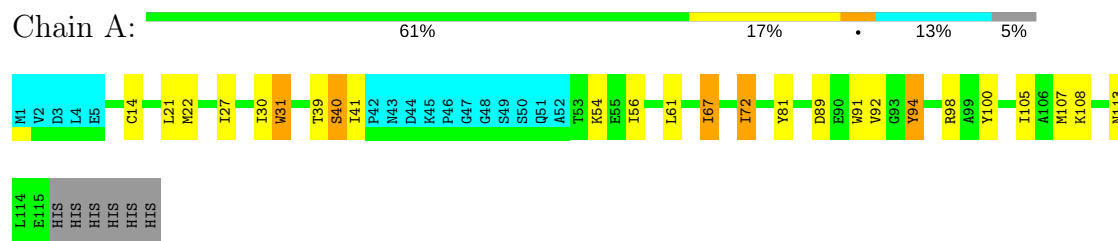
4.2.17 Score per residue for model 17

- Molecule 1: Chromatin modification-related protein EAF3



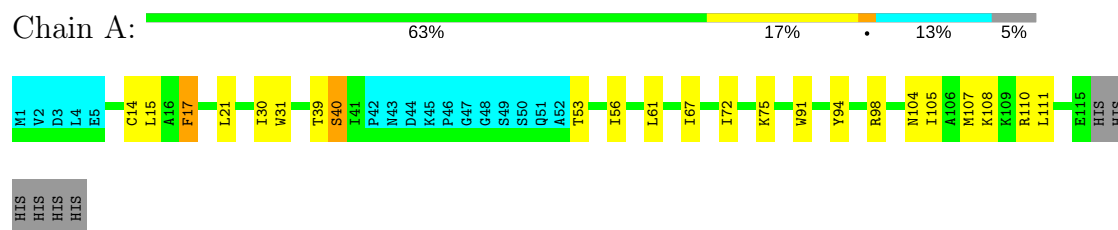
4.2.18 Score per residue for model 18

- Molecule 1: Chromatin modification-related protein EAF3



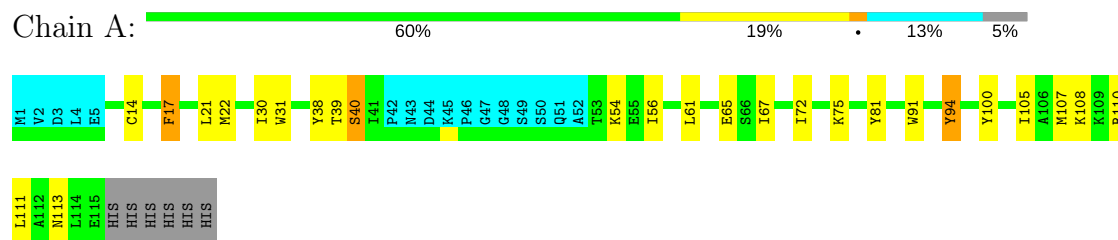
4.2.19 Score per residue for model 19

- Molecule 1: Chromatin modification-related protein EAF3



4.2.20 Score per residue for model 20

- Molecule 1: Chromatin modification-related protein EAF3



5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *simulated annealing*.

Of the 200 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 solution	
AMBER	refinement	

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

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	0.69±0.00	0±0/841 (0.0±0.0%)	1.16±0.02	1±1/1132 (0.1±0.1%)
All	All	0.69	0/16820 (0.0%)	1.16	21/22640 (0.1%)

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	1.9±0.6
All	All	0	38

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	67	ILE	CA-CB-CG1	7.14	124.56	111.00	3	6
1	A	38	TYR	CB-CG-CD2	-6.39	117.17	121.00	16	8
1	A	96	ARG	NE-CZ-NH1	5.64	123.12	120.30	14	1
1	A	38	TYR	CB-CG-CD1	-5.27	117.84	121.00	4	1
1	A	96	ARG	NE-CZ-NH2	-5.25	117.67	120.30	13	2
1	A	72	ILE	CA-CB-CG1	-5.21	101.10	111.00	19	2
1	A	13	ARG	NE-CZ-NH2	-5.06	117.77	120.30	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	81	TYR	Sidechain	17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	94	TYR	Sidechain	16
1	A	23	TYR	Sidechain	3
1	A	96	ARG	Sidechain	2

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	819	803	803	4±1
All	All	16380	16060	16060	80

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:14:CYS:SG	1:A:27:ILE:HD11	0.72	2.24	18	8
1:A:17:PHE:CZ	1:A:111:LEU:HD11	0.61	2.30	20	14
1:A:40:SER:HB3	1:A:56:ILE:HD12	0.60	1.73	3	8
1:A:40:SER:CB	1:A:56:ILE:HD12	0.57	2.30	2	10
1:A:15:LEU:HD23	1:A:104:ASN:CB	0.53	2.32	16	9
1:A:41:ILE:O	1:A:41:ILE:HD12	0.51	2.07	17	4
1:A:31:TRP:CZ2	1:A:72:ILE:HG12	0.50	2.41	18	1
1:A:14:CYS:SG	1:A:27:ILE:CD1	0.46	3.03	2	4
1:A:17:PHE:CE1	1:A:111:LEU:HD11	0.45	2.46	20	1
1:A:14:CYS:SG	1:A:97:ILE:HD11	0.45	2.51	9	2
1:A:22:MET:HE1	1:A:108:LYS:HB2	0.44	1.90	4	1
1:A:14:CYS:SG	1:A:99:ALA:HA	0.44	2.52	16	2
1:A:67:ILE:HD13	1:A:67:ILE:H	0.43	1.73	13	1
1:A:41:ILE:HD12	1:A:41:ILE:O	0.43	2.13	12	3
1:A:61:LEU:HD11	1:A:67:ILE:CD1	0.42	2.45	14	2
1:A:15:LEU:HD23	1:A:104:ASN:HB3	0.42	1.92	3	2
1:A:15:LEU:HD11	1:A:105:ILE:CD1	0.42	2.45	5	1
1:A:30:ILE:HG22	1:A:77:PHE:CD1	0.41	2.50	4	1
1:A:30:ILE:HG22	1:A:77:PHE:HD1	0.41	1.75	4	1
1:A:75:LYS:HB3	1:A:77:PHE:CZ	0.41	2.51	11	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:15:LEU:HD23	1:A:24:GLU:HA	0.41	1.90	12	1
1:A:29:LYS:C	1:A:30:ILE:HD12	0.41	2.36	10	1
1:A:40:SER:HB3	1:A:56:ILE:CD1	0.41	2.46	12	1
1:A:40:SER:HB2	1:A:56:ILE:HD12	0.40	1.93	10	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	98/121 (81%)	88±2 (90±2%)	9±1 (9±1%)	1±1 (1±1%)	17	62
All	All	1960/2420 (81%)	1763 (90%)	170 (9%)	27 (1%)	17	62

All 7 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	53	THR	8
1	A	54	LYS	6
1	A	64	ASP	5
1	A	55	GLU	3
1	A	75	LYS	2
1	A	62	GLY	2
1	A	6	GLN	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	86/105 (82%)	69±2 (80±3%)	17±2 (20±3%)	4	35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1720/2100 (82%)	1380 (80%)	340 (20%)	4 35

All 35 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	94	TYR	20
1	A	105	ILE	20
1	A	21	LEU	20
1	A	91	TRP	20
1	A	67	ILE	20
1	A	30	ILE	19
1	A	107	MET	18
1	A	17	PHE	18
1	A	61	LEU	17
1	A	108	LYS	15
1	A	65	GLU	14
1	A	39	THR	14
1	A	72	ILE	13
1	A	110	ARG	13
1	A	14	CYS	12
1	A	100	TYR	11
1	A	113	ASN	11
1	A	22	MET	10
1	A	40	SER	9
1	A	31	TRP	9
1	A	89	ASP	6
1	A	10	LEU	6
1	A	75	LYS	6
1	A	98	ARG	5
1	A	26	LYS	3
1	A	85	LYS	2
1	A	28	LEU	1
1	A	109	LYS	1
1	A	57	LYS	1
1	A	29	LYS	1
1	A	92	VAL	1
1	A	24	GLU	1
1	A	13	ARG	1
1	A	101	ASN	1
1	A	96	ARG	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 carbohydrates 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