



# Full wwPDB NMR Structure Validation Report ⓘ

Jan 5, 2021 – 08:44 AM GMT

PDB ID : 6TK0  
Title : Cytochrome C from Thioalkalivibrio paradoxus  
Authors : Bocharov, E.V.; Altukhov, D.A.; Timofeev, V.I.; Dergousova, N.I.; Rakitina, T.V.; Tikhonova, T.V.; Popov, V.O.  
Deposited on : 2019-11-27

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.16
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

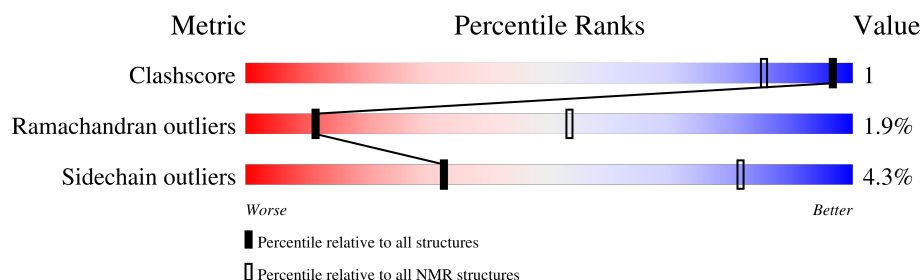
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 37%.

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  $\geq 3$ , 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	X	153	

## 2 Ensemble composition and analysis

This entry contains 15 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: *closest to the average*.

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	X:1-X:18, X:38-X:153 (134)	0.55	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 3 clusters and 4 single-model clusters were found.

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

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2290 atoms, of which 1067 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cytochrome c, mono-and diheme variants family.

Mol	Chain	Residues	Atoms						Trace
1	X	153	Total	C	H	N	O	S	0
			2217	715	1037	235	225	5	

There are 43 discrepancies between the modelled and reference sequences:

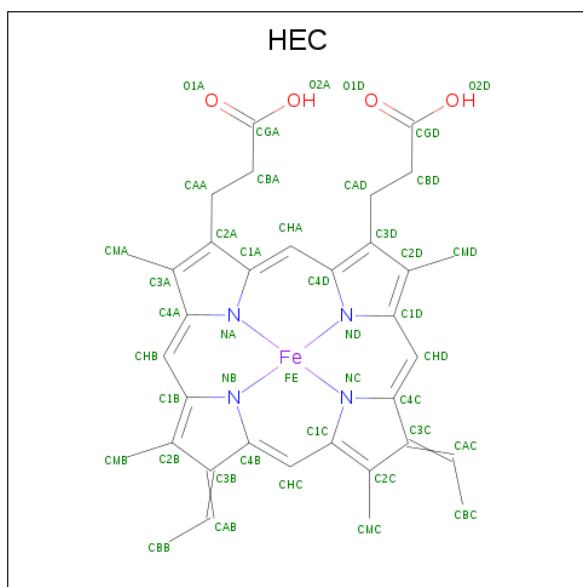
Chain	Residue	Modelled	Actual	Comment	Reference
X	1	MET	-	initiating methionine	UNP L0DXU7
X	2	ASP	-	expression tag	UNP L0DXU7
X	3	ILE	-	expression tag	UNP L0DXU7
X	4	GLY	-	expression tag	UNP L0DXU7
X	5	ILE	-	expression tag	UNP L0DXU7
X	6	ASN	-	expression tag	UNP L0DXU7
X	7	SER	-	expression tag	UNP L0DXU7
X	8	ASP	-	expression tag	UNP L0DXU7
X	9	PRO	-	expression tag	UNP L0DXU7
X	10	HIS	-	expression tag	UNP L0DXU7
X	11	PRO	-	expression tag	UNP L0DXU7
X	12	PRO	-	expression tag	UNP L0DXU7
X	13	HIS	-	expression tag	UNP L0DXU7
X	14	HIS	-	expression tag	UNP L0DXU7
X	15	HIS	-	expression tag	UNP L0DXU7
X	16	ASP	-	expression tag	UNP L0DXU7
X	17	HIS	-	expression tag	UNP L0DXU7
X	18	HIS	-	expression tag	UNP L0DXU7
X	19	GLY	-	expression tag	UNP L0DXU7
X	20	HIS	-	expression tag	UNP L0DXU7
X	21	GLY	-	expression tag	UNP L0DXU7
X	25	GLU	LYS	conflict	UNP L0DXU7
X	32	HIS	GLN	conflict	UNP L0DXU7
X	88	SER	THR	conflict	UNP L0DXU7
X	133	GLY	GLU	conflict	UNP L0DXU7
X	135	ASN	ASP	conflict	UNP L0DXU7
X	137	HIS	-	expression tag	UNP L0DXU7
X	138	SER	-	expression tag	UNP L0DXU7
X	139	PRO	-	expression tag	UNP L0DXU7
X	140	ASP	-	expression tag	UNP L0DXU7
X	141	HIS	-	expression tag	UNP L0DXU7

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
X	142	ALA	-	expression tag	UNP L0DXU7
X	143	ALA	-	expression tag	UNP L0DXU7
X	144	GLY	-	expression tag	UNP L0DXU7
X	145	ASP	-	expression tag	UNP L0DXU7
X	146	HIS	-	expression tag	UNP L0DXU7
X	147	HIS	-	expression tag	UNP L0DXU7
X	148	HIS	-	expression tag	UNP L0DXU7
X	149	GLY	-	expression tag	UNP L0DXU7
X	150	ASP	-	expression tag	UNP L0DXU7
X	151	HIS	-	expression tag	UNP L0DXU7
X	152	HIS	-	expression tag	UNP L0DXU7
X	153	HIS	-	expression tag	UNP L0DXU7

- Molecule 2 is HEME C (three-letter code: HEC) (formula:  $\text{C}_{34}\text{H}_{34}\text{FeN}_4\text{O}_4$ ).



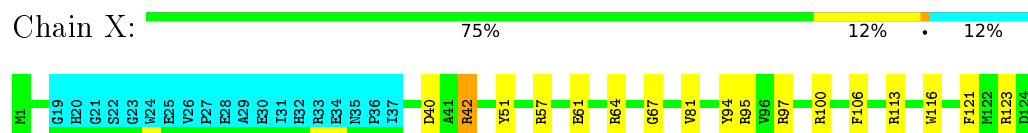
Mol	Chain	Residues	Atoms					
2	X	1	Total	C	Fe	H	N	O
			73	34	1	30	4	4

## 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: Cytochrome c, mono-and diheme variants family

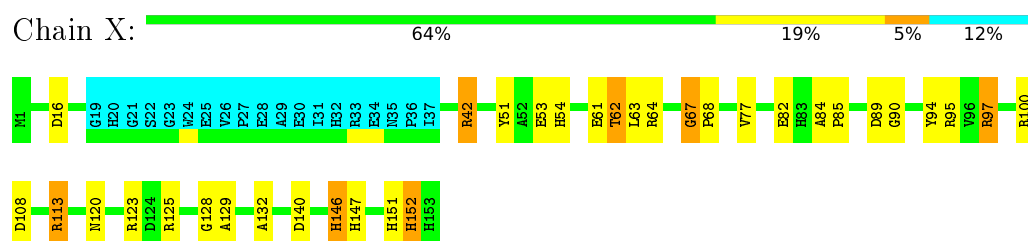


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

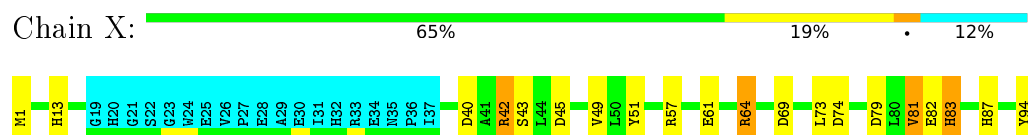
#### 4.2.1 Score per residue for model 1

- Molecule 1: Cytochrome c, mono-and diheme variants family



#### 4.2.2 Score per residue for model 2

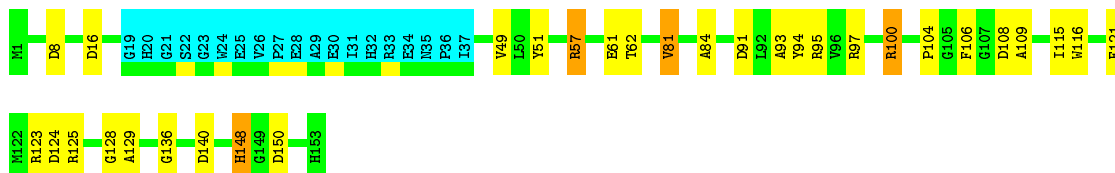
- Molecule 1: Cytochrome c, mono-and diheme variants family





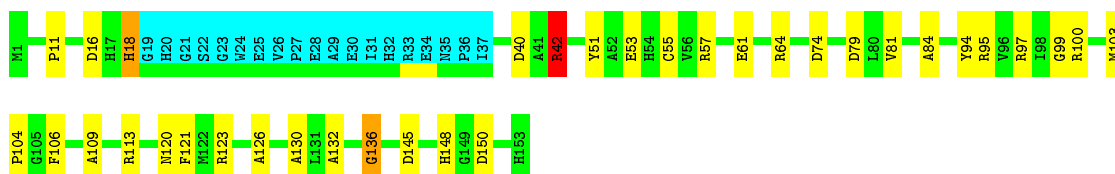
### 4.2.3 Score per residue for model 3

- Molecule 1: Cytochrome c, mono-and diheme variants family



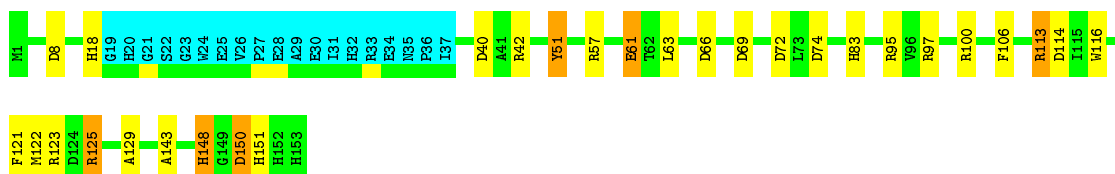
### 4.2.4 Score per residue for model 4

- Molecule 1: Cytochrome c, mono-and diheme variants family



### 4.2.5 Score per residue for model 5

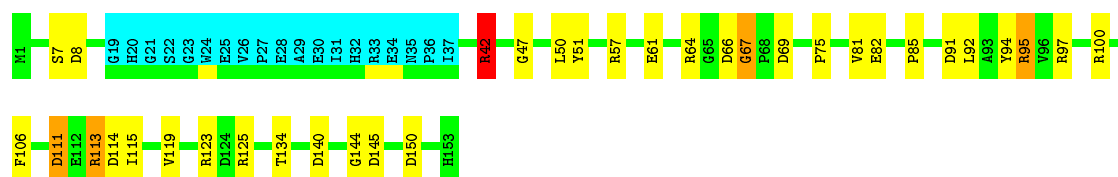
- Molecule 1: Cytochrome c, mono-and diheme variants family



### 4.2.6 Score per residue for model 6

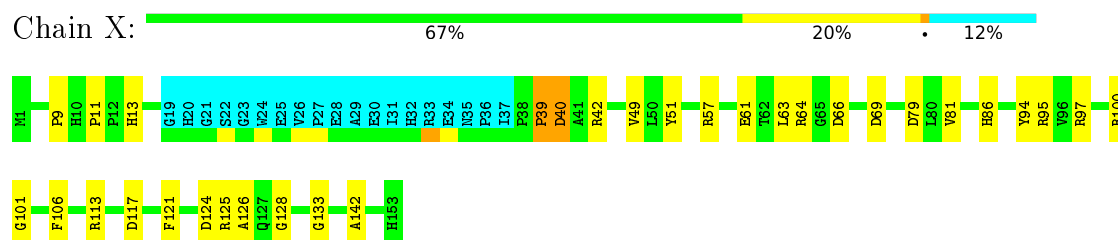
- Molecule 1: Cytochrome c, mono-and diheme variants family





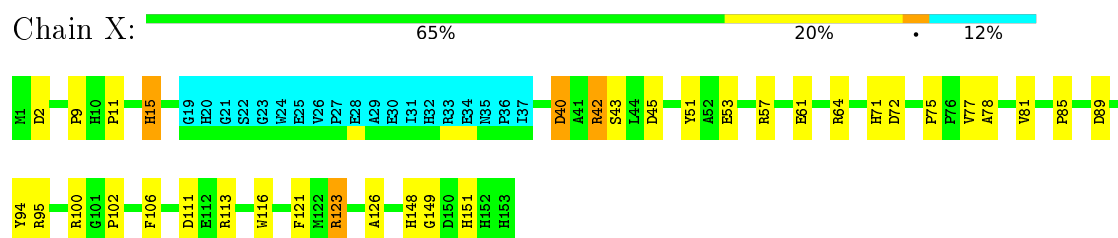
#### 4.2.7 Score per residue for model 7

- Molecule 1: Cytochrome c, mono-and diheme variants family



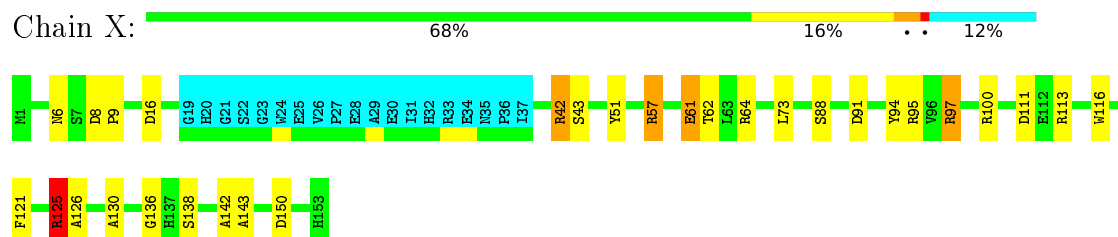
#### 4.2.8 Score per residue for model 8

- Molecule 1: Cytochrome c, mono-and diheme variants family



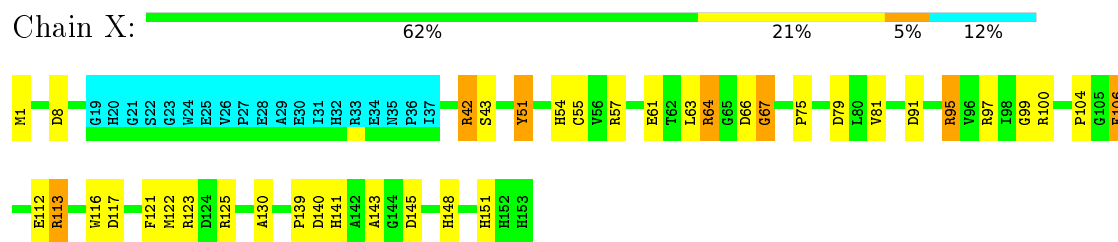
#### 4.2.9 Score per residue for model 9

- Molecule 1: Cytochrome c, mono-and diheme variants family



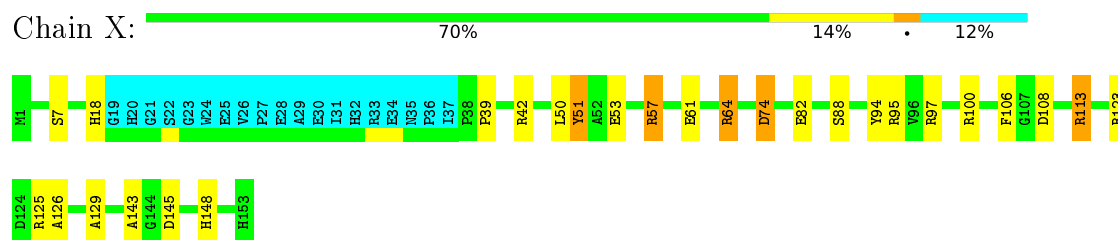
#### 4.2.10 Score per residue for model 10

- Molecule 1: Cytochrome c, mono-and diheme variants family



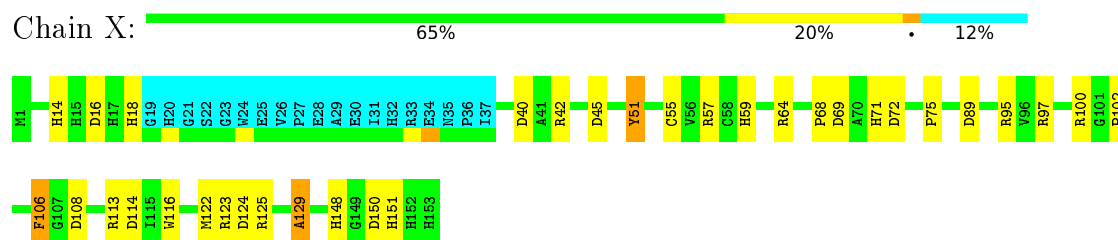
#### 4.2.11 Score per residue for model 11

- Molecule 1: Cytochrome c, mono-and diheme variants family



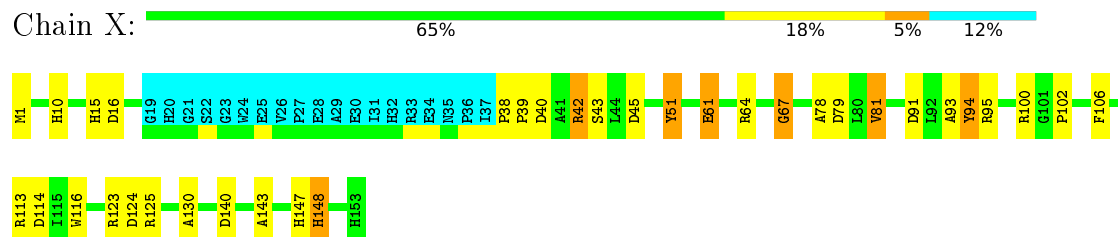
#### 4.2.12 Score per residue for model 12

- Molecule 1: Cytochrome c, mono-and diheme variants family



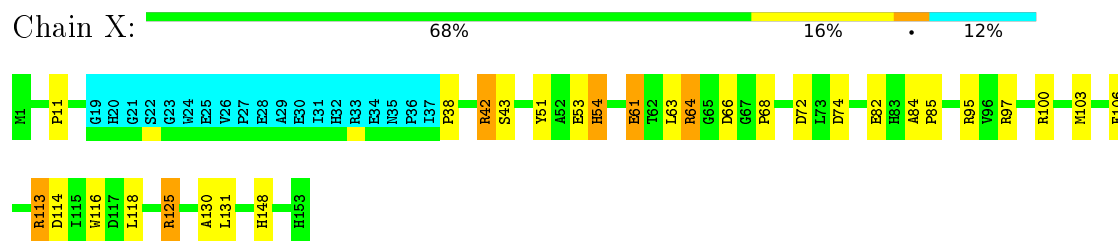
#### 4.2.13 Score per residue for model 13

- Molecule 1: Cytochrome c, mono-and diheme variants family



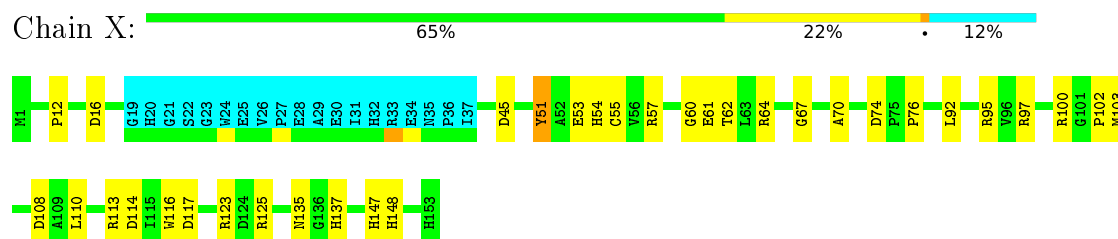
#### 4.2.14 Score per residue for model 14

- Molecule 1: Cytochrome c, mono-and diheme variants family



#### 4.2.15 Score per residue for model 15 (medoid)

- Molecule 1: Cytochrome c, mono-and diheme variants family



## 5 Refinement protocol and experimental data overview

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

Of the 5000 calculated structures, 15 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
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	699
Number of shifts mapped to atoms	699
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	37%

## 6 Model quality i

### 6.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: HEC

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	X	1.51±0.04	5±2/1068 ( 0.5± 0.2%)	2.02±0.06	31±4/1454 ( 2.1± 0.2%)
All	All	1.51	80/16020 ( 0.5%)	2.02	464/21810 ( 2.1%)

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	X	0.0±0.0	3.7±2.1
All	All	0	55

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	X	97	ARG	CZ-NH1	-7.80	1.23	1.33	6	1
1	X	94	TYR	CB-CG	7.80	1.63	1.51	11	1
1	X	57	ARG	CZ-NH1	-7.80	1.23	1.33	11	3
1	X	39	PRO	N-CD	-7.61	1.37	1.47	11	1
1	X	123	ARG	CZ-NH2	-7.32	1.23	1.33	12	1
1	X	88	SER	CB-OG	-7.21	1.32	1.42	11	2
1	X	57	ARG	CZ-NH2	-6.93	1.24	1.33	10	2
1	X	100	ARG	CZ-NH2	-6.86	1.24	1.33	7	4
1	X	43	SER	CB-OG	-6.67	1.33	1.42	14	2
1	X	95	ARG	CZ-NH1	-6.36	1.24	1.33	14	1
1	X	134	THR	CB-OG1	-6.30	1.30	1.43	6	1
1	X	136	GLY	CA-C	6.29	1.61	1.51	3	1
1	X	68	PRO	N-CD	-6.19	1.39	1.47	14	2
1	X	106	PHE	CG-CD1	6.10	1.47	1.38	8	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	X	99	GLY	CA-C	6.09	1.61	1.51	4	2
1	X	133	GLY	CA-C	5.95	1.61	1.51	7	1
1	X	113	ARG	CZ-NH2	-5.92	1.25	1.33	11	1
1	X	138	SER	CB-OG	5.91	1.50	1.42	9	2
1	X	82	GLU	CG-CD	-5.90	1.43	1.51	2	1
1	X	38	PRO	N-CD	-5.89	1.39	1.47	14	1
1	X	18	HIS	CB-CG	5.89	1.60	1.50	12	2
1	X	97	ARG	CZ-NH2	-5.87	1.25	1.33	1	1
1	X	128	GLY	N-CA	5.80	1.54	1.46	7	2
1	X	51	TYR	CG-CD2	5.80	1.46	1.39	3	1
1	X	116	TRP	CD2-CE2	5.73	1.48	1.41	10	1
1	X	13	HIS	CB-CG	5.73	1.60	1.50	7	1
1	X	125	ARG	CZ-NH1	-5.73	1.25	1.33	3	2
1	X	95	ARG	CZ-NH2	-5.72	1.25	1.33	12	3
1	X	136	GLY	N-CA	5.67	1.54	1.46	4	1
1	X	43	SER	CA-CB	5.66	1.61	1.52	10	1
1	X	90	GLY	CA-C	5.60	1.60	1.51	1	1
1	X	50	LEU	N-CA	5.58	1.57	1.46	11	1
1	X	116	TRP	NE1-CE2	5.54	1.44	1.37	9	1
1	X	95	ARG	N-CA	5.54	1.57	1.46	10	1
1	X	123	ARG	CZ-NH1	-5.45	1.25	1.33	6	1
1	X	121	PHE	CE1-CZ	5.43	1.47	1.37	7	1
1	X	53	GLU	CD-OE2	5.41	1.31	1.25	14	1
1	X	106	PHE	CG-CD2	5.40	1.46	1.38	3	1
1	X	113	ARG	CZ-NH1	-5.37	1.26	1.33	8	4
1	X	82	GLU	CA-CB	5.37	1.65	1.53	6	1
1	X	117	ASP	CB-CG	5.31	1.62	1.51	2	1
1	X	12	PRO	N-CD	-5.28	1.40	1.47	15	1
1	X	67	GLY	CA-C	5.24	1.60	1.51	6	1
1	X	94	TYR	CG-CD1	5.22	1.46	1.39	3	1
1	X	102	PRO	N-CD	-5.21	1.40	1.47	12	2
1	X	106	PHE	CE2-CZ	5.17	1.47	1.37	7	1
1	X	95	ARG	NE-CZ	-5.16	1.26	1.33	12	1
1	X	47	GLY	N-CA	5.16	1.53	1.46	6	1
1	X	101	GLY	N-CA	5.15	1.53	1.46	7	1
1	X	61	GLU	N-CA	5.11	1.56	1.46	5	1
1	X	144	GLY	N-CA	5.11	1.53	1.46	6	1
1	X	42	ARG	CZ-NH1	-5.11	1.26	1.33	5	1
1	X	82	GLU	CD-OE2	-5.10	1.20	1.25	11	1
1	X	66	ASP	CB-CG	5.09	1.62	1.51	6	1
1	X	94	TYR	CG-CD2	5.07	1.45	1.39	13	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	X	7	SER	CB-OG	-5.06	1.35	1.42	11	1
1	X	135	ASN	N-CA	5.05	1.56	1.46	15	1
1	X	51	TYR	CG-CD1	5.04	1.45	1.39	8	1
1	X	55	CYS	CB-SG	5.00	1.90	1.82	4	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	X	113	ARG	NE-CZ-NH2	28.09	134.34	120.30	6	11
1	X	125	ARG	NE-CZ-NH2	-19.41	110.60	120.30	3	7
1	X	113	ARG	NE-CZ-NH1	19.03	129.82	120.30	14	9
1	X	42	ARG	NE-CZ-NH1	17.68	129.14	120.30	7	7
1	X	95	ARG	NE-CZ-NH1	16.56	128.58	120.30	7	8
1	X	42	ARG	NE-CZ-NH2	-15.85	112.37	120.30	13	6
1	X	125	ARG	NE-CZ-NH1	15.43	128.01	120.30	1	10
1	X	123	ARG	NE-CZ-NH2	-15.06	112.77	120.30	3	7
1	X	123	ARG	NE-CZ-NH1	15.04	127.82	120.30	10	7
1	X	100	ARG	NE-CZ-NH1	14.64	127.62	120.30	3	7
1	X	97	ARG	NE-CZ-NH2	-14.03	113.28	120.30	15	9
1	X	57	ARG	NE-CZ-NH1	13.18	126.89	120.30	12	9
1	X	114	ASP	CB-CG-OD1	12.09	129.18	118.30	5	4
1	X	97	ARG	NE-CZ-NH1	12.01	126.31	120.30	4	10
1	X	94	TYR	CB-CG-CD1	-11.97	113.82	121.00	1	5
1	X	64	ARG	NE-CZ-NH1	11.76	126.18	120.30	8	6
1	X	57	ARG	NE-CZ-NH2	11.68	126.14	120.30	5	6
1	X	106	PHE	CB-CG-CD2	-11.36	112.85	120.80	11	7
1	X	82	GLU	OE1-CD-OE2	-11.19	109.87	123.30	1	3
1	X	113	ARG	NH1-CZ-NH2	-11.18	107.10	119.40	6	6
1	X	95	ARG	NE-CZ-NH2	11.18	125.89	120.30	8	7
1	X	94	TYR	CB-CG-CD2	-11.04	114.38	121.00	6	3
1	X	145	ASP	CB-CG-OD2	10.48	127.73	118.30	4	2
1	X	100	ARG	NE-CZ-NH2	10.07	125.34	120.30	2	5
1	X	64	ARG	NE-CZ-NH2	9.77	125.18	120.30	12	6
1	X	106	PHE	CB-CG-CD1	9.75	127.63	120.80	13	3
1	X	114	ASP	CB-CG-OD2	9.66	127.00	118.30	14	3
1	X	79	ASP	CB-CG-OD2	9.57	126.92	118.30	4	2
1	X	145	ASP	CB-CG-OD1	9.55	126.89	118.30	6	3
1	X	91	ASP	CB-CG-OD2	-9.31	109.92	118.30	10	3
1	X	79	ASP	CB-CG-OD1	8.93	126.33	118.30	10	4
1	X	51	TYR	CD1-CE1-CZ	-8.69	111.98	119.80	8	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	X	95	ARG	NH1-CZ-NH2	-8.60	109.94	119.40	3	3
1	X	40	ASP	CB-CG-OD2	8.54	125.98	118.30	13	3
1	X	111	ASP	CB-CG-OD2	-8.49	110.66	118.30	8	2
1	X	95	ARG	CD-NE-CZ	8.34	135.28	123.60	11	6
1	X	150	ASP	CB-CG-OD1	8.32	125.79	118.30	4	5
1	X	51	TYR	CB-CG-CD2	-8.14	116.12	121.00	12	3
1	X	125	ARG	CD-NE-CZ	8.12	134.97	123.60	2	3
1	X	64	ARG	CD-NE-CZ	8.11	134.96	123.60	9	5
1	X	8	ASP	CB-CG-OD2	8.02	125.52	118.30	3	3
1	X	125	ARG	NH1-CZ-NH2	-7.86	110.76	119.40	15	4
1	X	116	TRP	CD1-CG-CD2	-7.81	100.05	106.30	10	2
1	X	57	ARG	NH1-CZ-NH2	-7.71	110.92	119.40	11	2
1	X	51	TYR	CG-CD1-CE1	-7.68	115.16	121.30	11	4
1	X	74	ASP	CB-CG-OD1	7.66	125.19	118.30	2	2
1	X	121	PHE	CB-CG-CD1	7.66	126.16	120.80	9	4
1	X	109	ALA	N-CA-CB	-7.61	99.44	110.10	3	1
1	X	140	ASP	CB-CG-OD1	7.54	125.08	118.30	1	4
1	X	74	ASP	CB-CG-OD2	7.53	125.07	118.30	14	3
1	X	11	PRO	N-CA-CB	7.49	112.29	103.30	14	3
1	X	69	ASP	CB-CG-OD1	7.39	124.95	118.30	5	4
1	X	123	ARG	NH1-CZ-NH2	-7.38	111.28	119.40	8	1
1	X	113	ARG	CD-NE-CZ	7.35	133.89	123.60	5	2
1	X	130	ALA	CB-CA-C	7.30	121.05	110.10	9	3
1	X	51	TYR	CB-CG-CD1	-7.30	116.62	121.00	9	6
1	X	42	ARG	N-CA-CB	7.24	123.63	110.60	7	1
1	X	64	ARG	NH1-CZ-NH2	-7.23	111.44	119.40	12	1
1	X	68	PRO	N-CD-CG	7.19	113.99	103.20	1	1
1	X	116	TRP	CD1-NE1-CE2	7.18	115.46	109.00	12	4
1	X	62	THR	CA-CB-CG2	7.16	122.43	112.40	15	2
1	X	16	ASP	CB-CG-OD2	7.15	124.74	118.30	9	3
1	X	16	ASP	CB-CG-OD1	7.12	124.70	118.30	15	4
1	X	12	PRO	N-CD-CG	7.06	113.80	103.20	15	1
1	X	57	ARG	CD-NE-CZ	7.04	133.46	123.60	11	2
1	X	72	ASP	CB-CG-OD1	7.04	124.64	118.30	14	2
1	X	116	TRP	CE2-CD2-CG	6.99	112.89	107.30	10	3
1	X	124	ASP	CB-CG-OD1	6.96	124.56	118.30	2	2
1	X	78	ALA	N-CA-CB	-6.94	100.38	110.10	8	2
1	X	51	TYR	CG-CD2-CE2	-6.94	115.75	121.30	15	1
1	X	116	TRP	CG-CD2-CE3	6.92	140.13	133.90	5	1
1	X	108	ASP	CB-CG-OD1	6.91	124.52	118.30	12	1
1	X	9	PRO	N-CA-CB	6.87	111.55	103.30	7	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	X	104	PRO	N-CD-CG	6.84	113.45	103.20	10	1
1	X	124	ASP	CB-CG-OD2	6.79	124.41	118.30	12	3
1	X	66	ASP	CB-CG-OD2	6.77	124.40	118.30	14	1
1	X	104	PRO	N-CA-CB	6.73	111.37	103.30	1	1
1	X	116	TRP	CB-CG-CD2	6.70	135.31	126.60	3	2
1	X	75	PRO	CA-N-CD	-6.67	102.16	111.50	6	3
1	X	7	SER	N-CA-CB	6.60	120.40	110.50	6	1
1	X	81	VAL	CG1-CB-CG2	-6.59	100.36	110.90	13	3
1	X	40	ASP	CB-CG-OD1	6.53	124.17	118.30	7	2
1	X	116	TRP	NE1-CE2-CD2	-6.53	100.77	107.30	12	2
1	X	112	GLU	OE1-CD-OE2	-6.53	115.47	123.30	10	1
1	X	84	ALA	CB-CA-C	6.52	119.88	110.10	14	2
1	X	106	PHE	CD1-CE1-CZ	-6.51	112.29	120.10	13	1
1	X	61	GLU	OE1-CD-OE2	-6.47	115.53	123.30	9	2
1	X	97	ARG	NH1-CZ-NH2	-6.46	112.29	119.40	5	3
1	X	100	ARG	NH1-CZ-NH2	-6.46	112.29	119.40	9	2
1	X	94	TYR	CG-CD2-CE2	-6.44	116.14	121.30	1	1
1	X	69	ASP	CB-CG-OD2	6.43	124.09	118.30	7	1
1	X	116	TRP	CD2-CE3-CZ3	6.41	127.13	118.80	5	1
1	X	117	ASP	CB-CG-OD2	6.36	124.02	118.30	7	1
1	X	142	ALA	N-CA-CB	-6.34	101.23	110.10	7	1
1	X	51	TYR	CD1-CG-CD2	6.30	124.83	117.90	12	1
1	X	94	TYR	CA-CB-CG	6.29	125.35	113.40	1	1
1	X	100	ARG	CD-NE-CZ	6.29	132.41	123.60	12	2
1	X	42	ARG	CD-NE-CZ	6.26	132.36	123.60	12	2
1	X	53	GLU	OE1-CD-OE2	-6.22	115.84	123.30	11	3
1	X	120	ASN	CB-CA-C	6.22	122.84	110.40	1	2
1	X	83	HIS	CA-CB-CG	6.16	124.07	113.60	2	1
1	X	126	ALA	N-CA-CB	-6.15	101.49	110.10	8	2
1	X	45	ASP	CB-CG-OD1	6.11	123.80	118.30	15	2
1	X	49	VAL	CG1-CB-CG2	-6.09	101.15	110.90	3	2
1	X	121	PHE	CB-CG-CD2	6.09	125.06	120.80	4	2
1	X	39	PRO	N-CA-CB	6.08	110.59	103.30	13	1
1	X	66	ASP	C-N-CA	6.07	135.05	122.30	6	1
1	X	9	PRO	CA-N-CD	-6.06	103.02	111.50	7	1
1	X	143	ALA	CB-CA-C	6.00	119.10	110.10	5	3
1	X	81	VAL	CA-CB-CG2	5.97	119.86	110.90	13	3
1	X	147	HIS	O-C-N	-5.95	113.19	122.70	1	1
1	X	119	VAL	CA-C-O	-5.93	107.64	120.10	6	1
1	X	76	PRO	N-CA-CB	5.93	110.42	103.30	15	1
1	X	6	ASN	O-C-N	-5.90	113.26	122.70	9	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	X	116	TRP	CE3-CZ3-CH2	-5.87	114.74	121.20	5	2
1	X	63	LEU	CB-CA-C	5.83	121.28	110.20	14	4
1	X	89	ASP	CB-CG-OD2	5.82	123.54	118.30	1	2
1	X	70	ALA	CB-CA-C	5.81	118.82	110.10	15	1
1	X	140	ASP	CB-CG-OD2	5.80	123.52	118.30	3	2
1	X	42	ARG	NH1-CZ-NH2	-5.80	113.02	119.40	4	3
1	X	81	VAL	CA-CB-CG1	5.79	119.59	110.90	3	1
1	X	51	TYR	CZ-CE2-CD2	-5.79	114.59	119.80	7	1
1	X	100	ARG	C-N-CA	5.77	134.42	122.30	2	1
1	X	1	MET	CA-CB-CG	5.76	123.09	113.30	13	2
1	X	116	TRP	CZ3-CH2-CZ2	-5.75	114.69	121.60	14	2
1	X	84	ALA	N-CA-C	5.74	126.50	111.00	3	1
1	X	9	PRO	N-CD-CG	5.74	111.81	103.20	8	1
1	X	66	ASP	CB-CG-OD1	5.70	123.43	118.30	5	2
1	X	117	ASP	CB-CG-OD1	5.70	123.43	118.30	15	2
1	X	94	TYR	CD1-CE1-CZ	5.70	124.93	119.80	9	1
1	X	45	ASP	CB-CG-OD2	5.69	123.42	118.30	12	3
1	X	146	HIS	CA-CB-CG	5.69	123.28	113.60	1	1
1	X	139	PRO	CA-N-CD	-5.65	103.59	111.50	10	1
1	X	38	PRO	N-CD-CG	5.63	111.64	103.20	13	1
1	X	63	LEU	CB-CG-CD1	5.61	120.54	111.00	1	3
1	X	8	ASP	CB-CG-OD1	-5.60	113.26	118.30	9	4
1	X	106	PHE	CG-CD1-CE1	-5.57	114.67	120.80	4	1
1	X	126	ALA	CB-CA-C	5.56	118.45	110.10	9	2
1	X	116	TRP	CB-CG-CD1	-5.55	119.78	127.00	3	1
1	X	119	VAL	CA-C-N	5.54	129.39	117.20	6	1
1	X	91	ASP	CB-CG-OD1	5.54	123.28	118.30	13	1
1	X	141	HIS	O-C-N	-5.53	113.85	122.70	10	1
1	X	89	ASP	CB-CG-OD1	5.53	123.27	118.30	12	1
1	X	18	HIS	CA-CB-CG	5.51	122.96	113.60	11	1
1	X	102	PRO	O-C-N	-5.50	113.91	122.70	15	1
1	X	130	ALA	N-CA-CB	-5.49	102.41	110.10	14	1
1	X	129	ALA	N-CA-CB	5.46	117.75	110.10	12	1
1	X	62	THR	O-C-N	-5.45	113.99	122.70	3	1
1	X	100	ARG	N-CA-C	5.44	125.69	111.00	4	1
1	X	142	ALA	CB-CA-C	5.43	118.24	110.10	9	1
1	X	11	PRO	N-CD-CG	5.38	111.27	103.20	4	1
1	X	85	PRO	N-CD-CG	5.38	111.26	103.20	6	2
1	X	10	HIS	CA-CB-CG	-5.38	104.46	113.60	13	1
1	X	93	ALA	CB-CA-C	5.36	118.14	110.10	13	1
1	X	106	PHE	CG-CD2-CE2	-5.34	114.93	120.80	6	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	X	73	LEU	CB-CG-CD2	5.33	120.06	111.00	2	1
1	X	102	PRO	N-CA-CB	5.33	109.69	103.30	13	1
1	X	83	HIS	O-C-N	-5.31	114.21	122.70	5	1
1	X	75	PRO	N-CD-CG	5.28	111.12	103.20	8	1
1	X	122	MET	CA-CB-CG	-5.28	104.33	113.30	5	1
1	X	123	ARG	CD-NE-CZ	5.26	130.96	123.60	4	1
1	X	51	TYR	O-C-N	-5.24	114.32	122.70	10	1
1	X	116	TRP	CH2-CZ2-CE2	5.23	122.63	117.40	14	1
1	X	9	PRO	O-C-N	-5.23	114.34	122.70	9	1
1	X	50	LEU	CA-C-O	-5.22	109.14	120.10	6	1
1	X	92	LEU	CB-CG-CD2	5.22	119.87	111.00	15	1
1	X	121	PHE	N-CA-CB	-5.21	101.21	110.60	3	1
1	X	111	ASP	CB-CG-OD1	5.20	122.98	118.30	2	2
1	X	72	ASP	C-N-CA	5.20	134.69	121.70	5	1
1	X	82	GLU	C-N-CA	5.20	134.69	121.70	14	1
1	X	75	PRO	N-CA-CB	5.19	109.53	103.30	10	1
1	X	54	HIS	C-N-CA	5.18	134.65	121.70	15	1
1	X	122	MET	O-C-N	-5.18	114.42	122.70	10	1
1	X	2	ASP	CB-CG-OD2	5.16	122.95	118.30	8	1
1	X	1	MET	CG-SD-CE	5.16	108.45	100.20	2	1
1	X	97	ARG	CD-NE-CZ	5.13	130.79	123.60	6	1
1	X	116	TRP	CG-CD1-NE1	5.13	115.23	110.10	10	1
1	X	77	VAL	CG1-CB-CG2	-5.12	102.71	110.90	8	1
1	X	103	MET	CB-CA-C	5.11	120.62	110.40	14	1
1	X	13	HIS	O-C-N	-5.11	114.53	122.70	2	1
1	X	151	HIS	N-CA-CB	-5.10	101.42	110.60	8	1
1	X	77	VAL	CA-CB-CG2	5.08	118.52	110.90	1	1
1	X	148	HIS	CA-CB-CG	5.08	122.23	113.60	13	1
1	X	39	PRO	CA-N-CD	-5.07	104.40	111.50	7	1
1	X	82	GLU	O-C-N	5.05	130.78	122.70	2	1
1	X	94	TYR	N-CA-CB	-5.04	101.52	110.60	4	1
1	X	53	GLU	CA-C-N	5.04	128.28	117.20	15	1
1	X	122	MET	CG-SD-CE	5.03	108.25	100.20	12	1
1	X	124	ASP	OD1-CG-OD2	-5.03	113.75	123.30	12	1
1	X	16	ASP	OD1-CG-OD2	-5.02	113.77	123.30	13	1
1	X	115	ILE	CG1-CB-CG2	-5.00	100.39	111.40	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	X	51	TYR	Sidechain	6
1	X	151	HIS	Sidechain	3
1	X	57	ARG	Sidechain	3
1	X	113	ARG	Sidechain	3
1	X	67	GLY	Peptide,Mainchain	3
1	X	106	PHE	Sidechain	2
1	X	123	ARG	Sidechain	2
1	X	94	TYR	Sidechain	2
1	X	43	SER	Mainchain	2
1	X	18	HIS	Sidechain	2
1	X	97	ARG	Sidechain	2
1	X	136	GLY	Peptide,Mainchain	2
1	X	42	ARG	Sidechain	2
1	X	121	PHE	Sidechain	1
1	X	137	HIS	Sidechain	1
1	X	14	HIS	Sidechain	1
1	X	59	HIS	Sidechain	1
1	X	95	ARG	Sidechain	1
1	X	60	GLY	Peptide	1
1	X	110	LEU	Peptide	1
1	X	146	HIS	Sidechain	1
1	X	132	ALA	Mainchain	1
1	X	152	HIS	Sidechain	1
1	X	91	ASP	Sidechain	1
1	X	87	HIS	Sidechain	1
1	X	100	ARG	Sidechain	1
1	X	125	ARG	Sidechain	1
1	X	53	GLU	Sidechain	1
1	X	148	HIS	Sidechain	1
1	X	64	ARG	Sidechain	1
1	X	86	HIS	Sidechain	1
1	X	109	ALA	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	X	1031	906	906	1±1
2	X	43	30	30	1±1

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes
All	All	16110	14040	14040	21

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:X:103:MET:HB3	2:X:201:HEC:HMD2	0.58	1.75	4	1
1:X:100:ARG:HE	2:X:201:HEC:CGA	0.54	2.16	14	1
1:X:51:TYR:CD1	1:X:55:CYS:HB2	0.54	2.37	15	2
1:X:100:ARG:HE	2:X:201:HEC:CGD	0.51	2.19	3	1
2:X:201:HEC:HMB1	2:X:201:HEC:HBB3	0.49	1.84	7	1
1:X:103:MET:O	2:X:201:HEC:HBC2	0.49	2.07	15	1
1:X:71:HIS:ND1	2:X:201:HEC:HMA3	0.48	2.23	12	1
1:X:51:TYR:CD1	1:X:55:CYS:CB	0.47	2.98	15	1
1:X:55:CYS:SG	2:X:201:HEC:HMB1	0.47	2.50	10	1
1:X:93:ALA:O	1:X:115:ILE:HD11	0.46	2.10	3	1
1:X:61:GLU:N	1:X:61:GLU:CD	0.45	2.69	14	1
1:X:54:HIS:CE1	2:X:201:HEC:HMC2	0.45	2.46	1	2
1:X:151:HIS:CE1	1:X:152:HIS:CD2	0.44	3.05	1	1
1:X:84:ALA:HB3	1:X:85:PRO:HD3	0.44	1.88	1	1
1:X:103:MET:H	1:X:104:PRO:HD2	0.43	1.73	4	1
1:X:51:TYR:CG	1:X:121:PHE:CE1	0.41	3.08	5	1
1:X:129:ALA:O	1:X:132:ALA:HB3	0.41	2.16	1	1
1:X:95:ARG:HE	2:X:201:HEC:CGA	0.40	2.29	10	1
1:X:125:ARG:HA	1:X:125:ARG:CZ	0.40	2.46	14	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	X	132/153 (86%)	116±4 (88±3%)	14±4 (10±3%)	2±2 (2±1%)	11	53
All	All	1980/2295 (86%)	1739 (88%)	204 (10%)	37 (2%)	11	53

All 16 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	X	81	VAL	5
1	X	67	GLY	5
1	X	129	ALA	4
1	X	108	ASP	4
1	X	148	HIS	3
1	X	143	ALA	3
1	X	147	HIS	2
1	X	15	HIS	2
1	X	54	HIS	2
1	X	149	GLY	1
1	X	72	ASP	1
1	X	131	LEU	1
1	X	71	HIS	1
1	X	130	ALA	1
1	X	128	GLY	1
1	X	126	ALA	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	X	105/120 (88%)	101±2 (96±2%)	4±2 (4±2%)	33	81
All	All	1575/1800 (88%)	1508 (96%)	67 (4%)	33	81

All 27 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	X	61	GLU	14
1	X	148	HIS	9
1	X	42	ARG	9
1	X	40	ASP	4
1	X	113	ARG	3
1	X	64	ARG	3
1	X	74	ASP	2
1	X	150	ASP	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	X	62	THR	2
1	X	125	ARG	2
1	X	1	MET	1
1	X	49	VAL	1
1	X	85	PRO	1
1	X	39	PRO	1
1	X	43	SER	1
1	X	104	PRO	1
1	X	15	HIS	1
1	X	66	ASP	1
1	X	111	ASP	1
1	X	72	ASP	1
1	X	92	LEU	1
1	X	102	PRO	1
1	X	131	LEU	1
1	X	73	LEU	1
1	X	118	LEU	1
1	X	83	HIS	1
1	X	81	VAL	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 monosaccharides in this entry.

## 6.6 Ligand geometry ⓘ

1 ligand is modelled in this entry.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 37% for the well-defined parts and 37% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *Cyt\_BMRB.tab*

#### 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	699
Number of shifts mapped to atoms	699
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

#### 7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	101	$1.31 \pm 0.12$	Should be applied
$^{13}\text{C}_\beta$	87	$1.94 \pm 0.17$	Should be applied
$^{13}\text{C}'$	99	$2.34 \pm 0.13$	Should be applied
$^{15}\text{N}$	106	$0.92 \pm 0.23$	Should be applied

#### 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 37%, i.e. 583 atoms were assigned a chemical shift out of a possible 1555. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	458/646 (71%)	184/256 (72%)	179/268 (67%)	95/122 (78%)
Sidechain	125/723 (17%)	49/428 (11%)	76/263 (29%)	0/32 (0%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	0/186 (0%)	0/104 (0%)	0/61 (0%)	0/21 (0%)
Overall	583/1555 (37%)	233/788 (30%)	255/592 (43%)	95/175 (54%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 37%, i.e. 652 atoms were assigned a chemical shift out of a possible 1781. 0 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	511/737 (69%)	205/292 (70%)	200/306 (65%)	106/139 (76%)
Sidechain	141/832 (17%)	54/493 (11%)	87/303 (29%)	0/36 (0%)
Aromatic	0/212 (0%)	0/118 (0%)	0/70 (0%)	0/24 (0%)
Overall	652/1781 (37%)	259/903 (29%)	287/679 (42%)	106/199 (53%)

#### 7.1.4 Statistically unusual chemical shifts ⓘ

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
???	X	79	ASP	HB2	4.66	4.07 – 1.37	7.2
???	X	79	ASP	HB3	4.66	4.07 – 1.27	7.1
???	X	35	ASN	HA	2.39	6.52 – 2.82	-6.2
???	X	64	ARG	HD2	4.29	4.27 – 1.97	5.1

#### 7.1.5 Random Coil Index (RCI) plots ⓘ

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.

Random coil index (RCI) for chain X:

