



Full wwPDB NMR Structure Validation Report ⓘ

Apr 21, 2024 – 01:06 PM EDT

PDB ID : 2L1G
Title : RDC refined solution structure of the THAP zinc finger of THAP1 in complex with its 16bp RRM1 DNA target
Authors : Campagne, S.; Gervais, V.; Saurel, O.; Milon, A.
Deposited on : 2010-07-28

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	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
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)	:	2.36.2

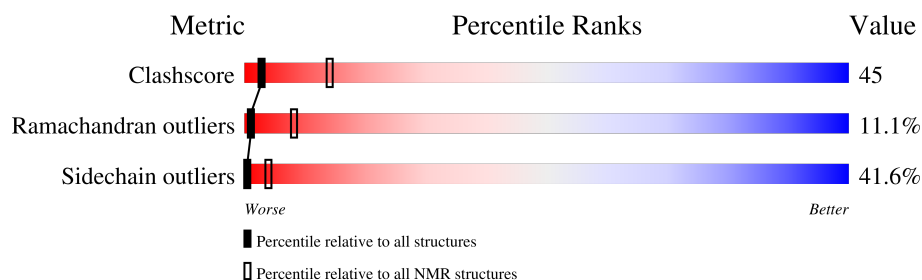
1 Overall quality at a glance

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	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 ≥ 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	87	<div> <div>23%</div> <div>52%</div> <div>16%</div> <div>8%</div> </div>
2	B	16	<div> <div>88%</div> <div>6%</div> <div>6%</div> </div>
3	C	16	<div> <div>12%</div> <div>81%</div> <div>6%</div> </div>

2 Ensemble composition and analysis

This entry contains 17 models. Model 11 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:3-A:82 (80)	0.23	11

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 1 clusters and 3 single-model clusters were found.

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

3 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2447 atoms, of which 1078 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called THAP domain-containing protein 1.

Mol	Chain	Residues	Atoms						Trace
1	A	87	Total	C	H	N	O	S	0
			1434	455	719	127	128	5	

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	62	SER	CYS	engineered mutation	UNP Q9NVV9
A	67	SER	CYS	engineered mutation	UNP Q9NVV9
A	83	GLU	-	expression tag	UNP Q9NVV9
A	84	LEU	-	expression tag	UNP Q9NVV9
A	85	VAL	-	expression tag	UNP Q9NVV9
A	86	PRO	-	expression tag	UNP Q9NVV9
A	87	ARG	-	expression tag	UNP Q9NVV9

- Molecule 2 is a DNA chain called DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3').

Mol	Chain	Residues	Atoms						Trace
2	B	16	Total	C	H	N	O	P	0
			512	157	181	62	97	15	

- Molecule 3 is a DNA chain called DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3').

Mol	Chain	Residues	Atoms						Trace
3	C	16	Total	C	H	N	O	P	0
			500	152	178	61	93	16	

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
4	A	1	Total	Zn
			1	1

4 Residue-property plots [i](#)

4.1 Average score per residue in the NMR ensemble


These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

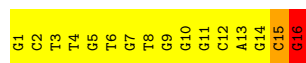
- Molecule 1: THAP domain-containing protein 1

Chain A: 



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B: 



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

Chain C: 

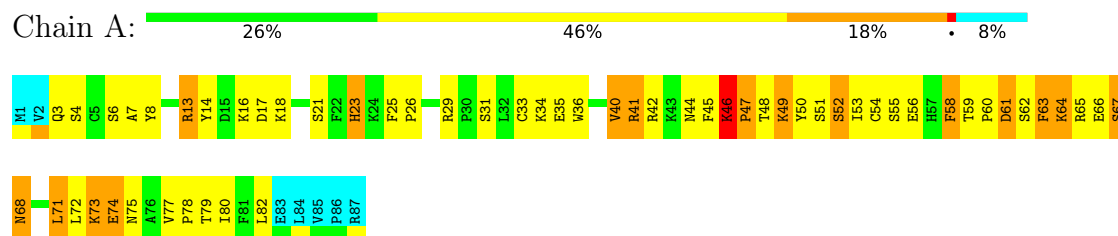


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

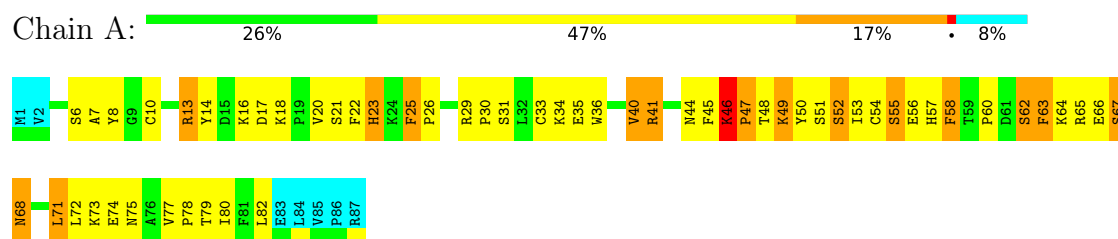


- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

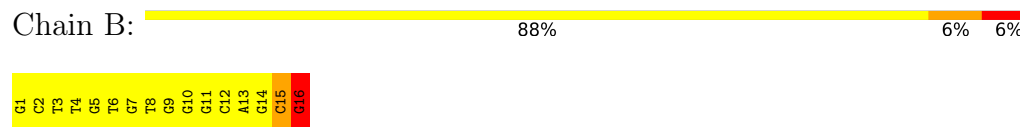


4.2.2 Score per residue for model 2

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')



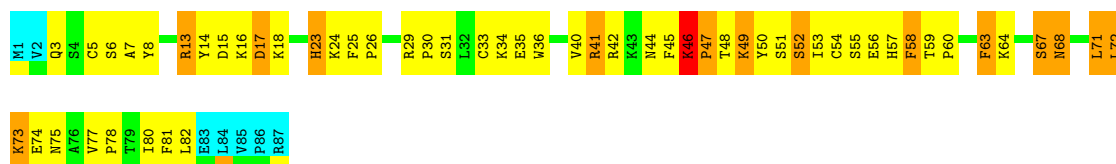
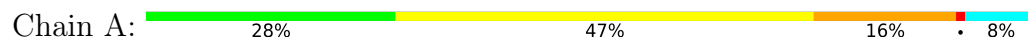
- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')



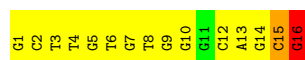


4.2.3 Score per residue for model 3

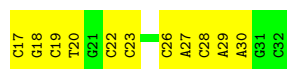
- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

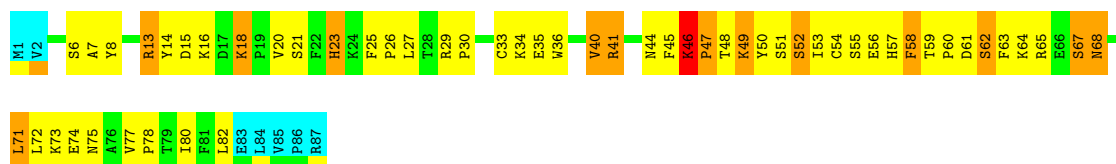


- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

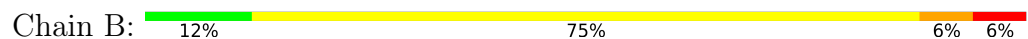


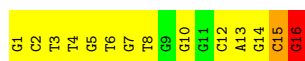
4.2.4 Score per residue for model 4

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')



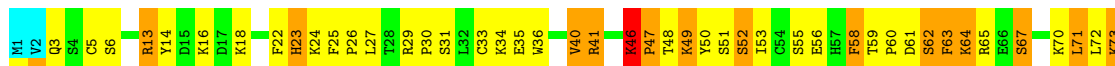
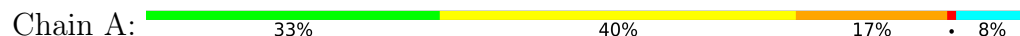


- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

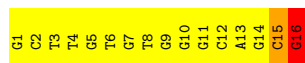
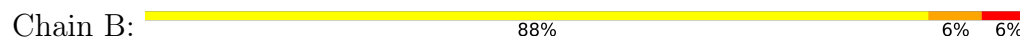


4.2.5 Score per residue for model 5

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')



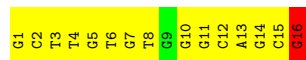
4.2.6 Score per residue for model 6

- Molecule 1: THAP domain-containing protein 1





- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

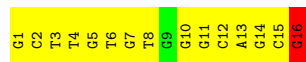


4.2.7 Score per residue for model 7

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

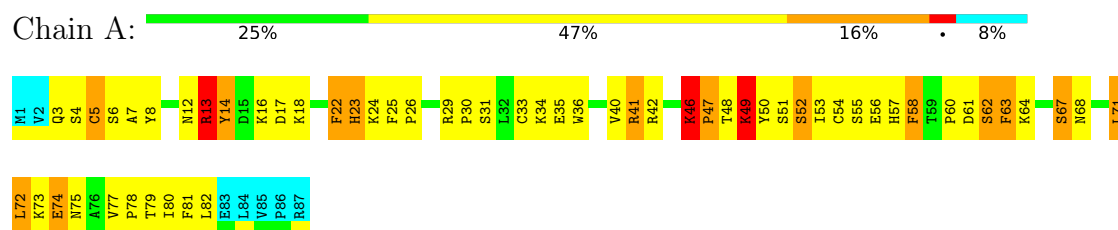


- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

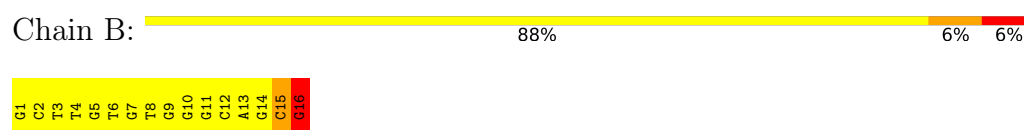


4.2.8 Score per residue for model 8

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

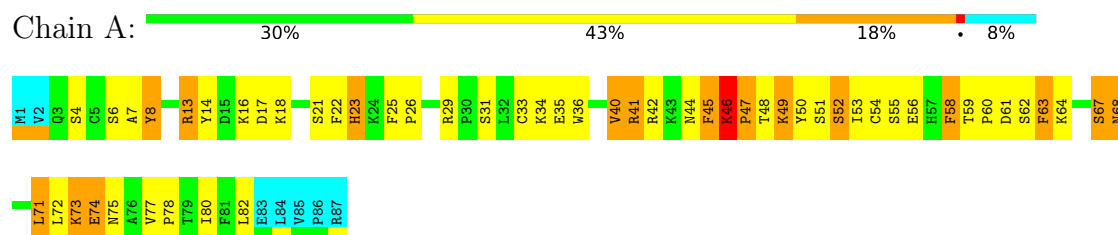


- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

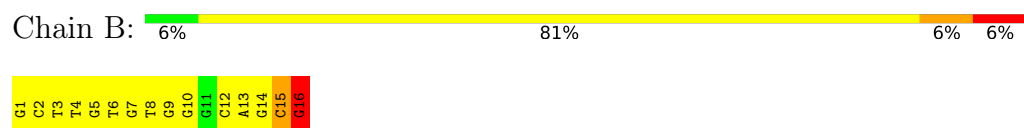


4.2.9 Score per residue for model 9

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

Chain C:  12% 81% 6%

C17
G18
C19
T20
C21
C22
C23
C24
A25
C26
A27
C28
A29
A30
G31
C32

4.2.10 Score per residue for model 10

- Molecule 1: THAP domain-containing protein 1

Chain A:  26% 45% 18% 8%

M1
V2
Q3
S4
C5
S6
A7
Y8
R13
Y14
D15
K16
D17
K18
F22
H23
K24
F25
P26
L27
T28
R29
P30
S31
L32
C33
K34
E35
W36
V40
R41
R42
K43
N44
F45
K46
P47
T48
K49
Y50
S51
S52
I53
C54
S55
E56
H57
F58
T59
P60
D61
S62
F63
K64
R65
E66
S67

N68
R69
K70
L71
L72
K73
E74
N75
A76
V77
P78
T79
I80
F81
L82
E83
L84
V85
P86
R87

- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B:  94% 6%

G1
C2
T3
T4
G5
T6
G7
T8
G9
G10
G11
C12
A13
G14
C15
G16

- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

Chain C:  6% 88% 6%

C17
G18
C19
T20
C21
C22
C23
C24
A25
C26
A27
C28
A29
A30
G31
C32

4.2.11 Score per residue for model 11 (medoid)

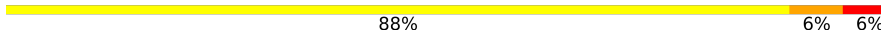
- Molecule 1: THAP domain-containing protein 1

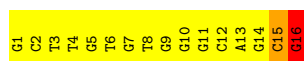
Chain A:  29% 44% 18% 8%

M1
V2
Q3
S4
C5
S6
A7
N12
R13
I80
Y14
D15
K16
D17
K18
S21
F22
H23
K24
F25
R29
P30
S31
L32
C33
K34
E35
W36
V40
R41
R42
K43
N44
F45
K46
P47
T48
K49
Y50
S51
S52
I53
C54
S55
E56
H57
F58
T59
P60
D61
S62
F63
K64
S67

L71
L72
K73
E74
N75
A76
V77
P78
T79
I80
F81
L82
E83
L84
V85
P86
R87

- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B:  88% 6% 6%

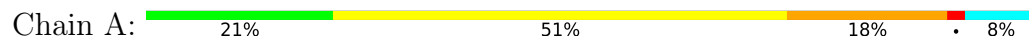


- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')



4.2.12 Score per residue for model 12

- Molecule 1: THAP domain-containing protein 1



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

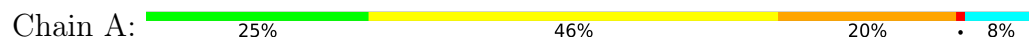


- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')



4.2.13 Score per residue for model 13

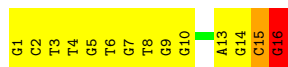
- Molecule 1: THAP domain-containing protein 1





- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B: 12% 75% 6% 6%



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

Chain C: 25% 69% 6%



4.2.14 Score per residue for model 14

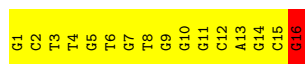
- Molecule 1: THAP domain-containing protein 1

Chain A: 26% 44% 20% 8%



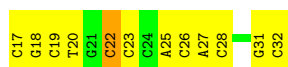
- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B: 94% 6%




- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

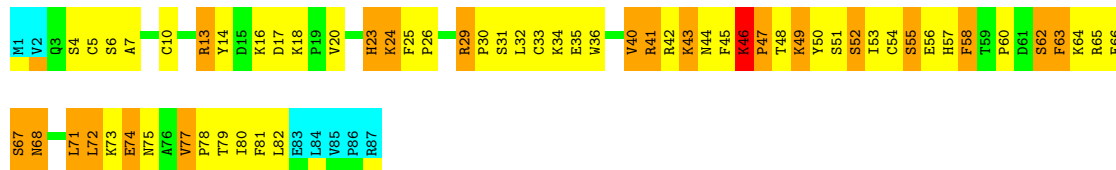
Chain C: 25% 69% 6%



4.2.15 Score per residue for model 15

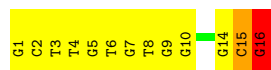
- Molecule 1: THAP domain-containing protein 1

Chain A: 



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B: 



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

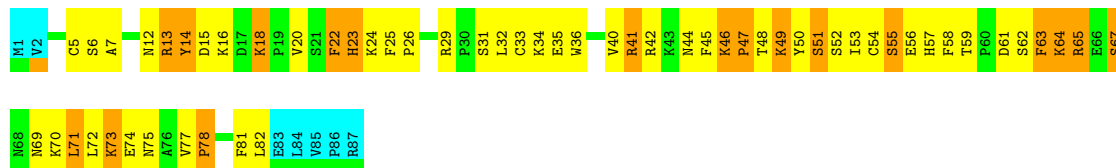
Chain C: 



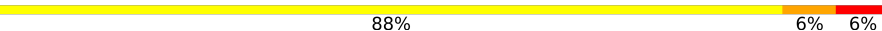
4.2.16 Score per residue for model 16

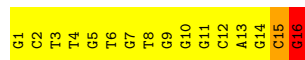
- Molecule 1: THAP domain-containing protein 1

Chain A: 



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B: 



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

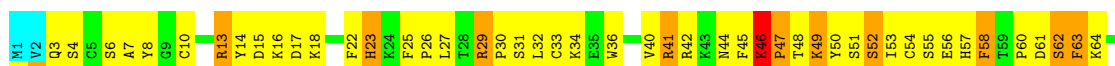
Chain C:  19% 75% 6%



4.2.17 Score per residue for model 17

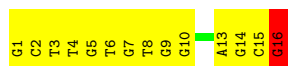
- Molecule 1: THAP domain-containing protein 1

Chain A:  22% 49% 18% 8%



- Molecule 2: DNA (5'-D(*GP*CP*TP*TP*GP*TP*GP*TP*GP*GP*GP*CP*AP*GP*CP*G)-3')

Chain B:  12% 81% 6%



- Molecule 3: DNA (5'-D(P*CP*GP*CP*TP*GP*CP*CP*CP*AP*CP*AP*CP*AP*AP*GP*C)-3')

Chain C:  19% 81%



5 Refinement protocol and experimental data overview

The models were refined using the following method: *Rigid body docking, Semi flexible simulated annealing, Water refinement.*

Of the 200 calculated structures, 17 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
CNS	refinement	1.21

No chemical shift data was provided.

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: ZN

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	0.50±0.02	0±0/676 (0.0± 0.0%)	0.78±0.02	0±0/910 (0.0± 0.0%)
2	B	0.42±0.02	0±0/371 (0.0± 0.0%)	0.84±0.01	2±0/573 (0.3± 0.0%)
3	C	0.31±0.02	0±0/360 (0.0± 0.0%)	0.69±0.01	0±0/551 (0.0± 0.0%)
All	All	0.44	0/23919 (0.0%)	0.77	34/34578 (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
2	B	0.0±0.0	1.6±0.5
3	C	0.0±0.0	0.8±0.5
All	All	0	42

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
2	B	16	DG	O4'-C4'-C3'	-6.70	101.82	104.50	1	17
2	B	16	DG	O4'-C1'-N9	6.46	112.52	108.00	7	17

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)
2	B	16	DG	Sidechain	17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
2	B	15	DC	Sidechain	11
3	C	22	DC	Sidechain	11
3	C	23	DC	Sidechain	3

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	657	653	652	40±5
2	B	331	181	182	47±2
3	C	322	178	178	25±3
All	All	22287	17204	17204	1781

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 45.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:15:DC:H1'	2:B:16:DG:N7	1.12	1.60	7	17
2:B:7:DG:C2'	2:B:8:DT:H71	0.94	1.91	5	10
2:B:14:DG:H2''	2:B:15:DC:C5	0.91	1.99	16	17
2:B:7:DG:H2'	2:B:8:DT:H72	0.91	1.41	6	7
1:A:25:PHE:HD2	1:A:33:CYS:HG	0.89	1.11	7	17
2:B:15:DC:C1'	2:B:16:DG:N7	0.88	2.35	12	17
2:B:1:DG:H2''	2:B:2:DC:C5	0.86	2.04	5	17
3:C:19:DC:C2'	3:C:20:DT:H72	0.84	2.03	11	7
3:C:19:DC:H2''	3:C:20:DT:C5	0.84	2.06	1	17
1:A:71:LEU:HD13	1:A:72:LEU:HD12	0.82	1.52	4	17
2:B:7:DG:H2'	2:B:8:DT:H71	0.82	1.50	15	10
3:C:18:DG:H2''	3:C:19:DC:C5	0.82	2.10	9	17
2:B:3:DT:H2''	2:B:4:DT:H71	0.81	1.50	9	17
2:B:7:DG:H2''	2:B:8:DT:H71	0.76	1.57	12	10
2:B:7:DG:C2'	2:B:8:DT:H72	0.76	2.09	9	7
1:A:67:SER:HB3	1:A:71:LEU:O	0.75	1.81	7	17
2:B:1:DG:H2''	2:B:2:DC:C6	0.75	2.15	5	17
2:B:11:DG:H2''	2:B:12:DC:C5	0.73	2.18	14	9

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:1:DG:C2'	2:B:2:DC:C5	0.73	2.72	5	17
3:C:27:DA:H2''	3:C:28:DC:C5	0.73	2.19	17	17
1:A:55:SER:HB3	1:A:63:PHE:CE1	0.72	2.19	14	1
3:C:19:DC:H2'	3:C:20:DT:H72	0.72	1.60	11	7
2:B:14:DG:C2'	2:B:15:DC:C5	0.72	2.71	16	17
1:A:62:SER:HA	1:A:73:LYS:CD	0.72	2.15	16	1
3:C:18:DG:H2''	3:C:19:DC:C6	0.71	2.21	9	17
2:B:6:DT:H1'	2:B:7:DG:C8	0.70	2.21	11	17
2:B:3:DT:H2''	2:B:4:DT:C7	0.70	2.17	9	17
2:B:2:DC:H2''	2:B:3:DT:H71	0.70	1.63	9	17
2:B:5:DG:H2''	2:B:6:DT:C5	0.69	2.22	17	16
2:B:7:DG:C2'	2:B:8:DT:C7	0.69	2.69	13	17
1:A:67:SER:OG	1:A:73:LYS:HD2	0.69	1.88	14	16
2:B:5:DG:H2''	2:B:6:DT:C7	0.69	2.17	17	17
2:B:15:DC:C2'	2:B:16:DG:N7	0.69	2.55	1	17
3:C:19:DC:H2''	3:C:20:DT:C7	0.69	2.18	5	17
2:B:5:DG:H2''	2:B:6:DT:H71	0.68	1.64	17	17
1:A:47:PRO:HB2	1:A:51:SER:OG	0.68	1.88	10	16
2:B:15:DC:H2''	2:B:16:DG:N7	0.68	2.02	1	17
2:B:3:DT:C2'	2:B:4:DT:H71	0.68	2.19	9	17
2:B:14:DG:H2''	2:B:15:DC:C6	0.68	2.24	7	17
2:B:16:DG:O5'	2:B:16:DG:C8	0.67	2.47	1	16
1:A:23:HIS:HE1	1:A:55:SER:N	0.67	1.88	15	17
1:A:25:PHE:CD1	1:A:53:ILE:HD11	0.66	2.26	16	15
2:B:7:DG:H2'	2:B:8:DT:C7	0.65	2.22	13	16
1:A:25:PHE:HD2	1:A:33:CYS:SG	0.65	2.14	12	17
1:A:24:LYS:HD3	2:B:8:DT:H72	0.65	1.67	5	4
2:B:15:DC:H1'	2:B:16:DG:C8	0.65	2.26	12	17
3:C:17:DC:H2''	3:C:18:DG:C8	0.64	2.26	2	17
3:C:18:DG:C2'	3:C:19:DC:C5	0.64	2.80	9	15
2:B:16:DG:C8	2:B:16:DG:O5'	0.64	2.50	12	1
3:C:19:DC:H2''	3:C:20:DT:H72	0.64	1.68	15	7
3:C:19:DC:C2'	3:C:20:DT:C7	0.63	2.75	5	16
3:C:26:DC:H2''	3:C:27:DA:C8	0.63	2.28	15	9
2:B:15:DC:C2	2:B:16:DG:O6	0.63	2.51	13	17
3:C:19:DC:H2''	3:C:20:DT:C6	0.62	2.29	5	15
1:A:50:TYR:CD2	3:C:22:DC:H2'	0.62	2.29	8	16
2:B:2:DC:C2'	2:B:3:DT:H71	0.62	2.24	8	17
1:A:48:THR:HG21	3:C:22:DC:OP2	0.62	1.95	1	14
1:A:50:TYR:HB2	3:C:22:DC:C5	0.62	2.29	12	17
1:A:36:TRP:O	1:A:40:VAL:HG23	0.61	1.95	14	17

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:5:DG:C2'	2:B:6:DT:H71	0.61	2.25	14	17
2:B:2:DC:H2''	2:B:3:DT:C7	0.61	2.26	6	17
2:B:7:DG:H2''	2:B:8:DT:C6	0.61	2.31	13	17
2:B:16:DG:O5'	2:B:16:DG:H8	0.60	1.79	3	17
1:A:40:VAL:HG22	1:A:81:PHE:HE2	0.60	1.56	16	2
1:A:60:PRO:HA	1:A:63:PHE:CD2	0.60	2.30	7	12
3:C:27:DA:H2''	3:C:28:DC:C6	0.60	2.31	15	17
1:A:62:SER:O	1:A:72:LEU:HD23	0.60	1.97	13	12
1:A:55:SER:HB2	1:A:63:PHE:CE1	0.60	2.31	5	10
1:A:6:SER:HA	1:A:46:LYS:CB	0.60	2.27	9	14
1:A:23:HIS:CE1	1:A:58:PHE:HE2	0.59	2.15	14	4
3:C:19:DC:H2''	3:C:20:DT:H71	0.59	1.74	6	10
1:A:50:TYR:CD1	3:C:23:DC:H5	0.59	2.15	13	16
2:B:6:DT:H2''	2:B:7:DG:C8	0.59	2.31	11	17
2:B:12:DC:H2''	2:B:13:DA:C8	0.59	2.31	6	11
2:B:3:DT:C2'	2:B:4:DT:C7	0.59	2.80	9	15
1:A:42:ARG:HG3	1:A:43:LYS:O	0.59	1.97	14	2
2:B:6:DT:C1'	2:B:7:DG:C8	0.58	2.86	11	17
3:C:28:DC:H2''	3:C:29:DA:C8	0.58	2.33	13	13
1:A:5:CYS:O	1:A:46:LYS:HD3	0.58	1.98	15	4
1:A:64:LYS:O	1:A:67:SER:HB3	0.58	1.99	14	2
1:A:52:SER:C	1:A:53:ILE:HD12	0.57	2.20	6	16
2:B:6:DT:C2	2:B:7:DG:C5	0.57	2.92	11	17
1:A:24:LYS:HG3	2:B:8:DT:H72	0.57	1.74	16	1
2:B:5:DG:C2'	2:B:6:DT:C7	0.57	2.82	17	16
1:A:26:PRO:HB3	1:A:72:LEU:CD1	0.57	2.29	8	12
3:C:19:DC:C2'	3:C:20:DT:C5	0.57	2.86	1	10
1:A:43:LYS:HG2	1:A:44:ASN:N	0.57	2.14	7	2
3:C:29:DA:H2''	3:C:30:DA:C8	0.56	2.36	5	10
2:B:13:DA:H2''	2:B:14:DG:C8	0.56	2.36	6	14
2:B:7:DG:H2''	2:B:8:DT:C7	0.55	2.29	7	12
3:C:27:DA:C2'	3:C:28:DC:C5	0.55	2.89	16	15
1:A:61:ASP:O	1:A:73:LYS:HD3	0.55	2.01	7	6
2:B:2:DC:C2'	2:B:3:DT:C7	0.55	2.84	8	17
1:A:49:LYS:HD2	1:A:49:LYS:H	0.55	1.62	14	1
1:A:47:PRO:HB2	1:A:51:SER:CB	0.55	2.31	8	2
1:A:3:GLN:NE2	3:C:20:DT:H71	0.55	2.17	12	3
3:C:19:DC:C2'	3:C:20:DT:H71	0.54	2.32	5	4
1:A:36:TRP:CE2	1:A:78:PRO:HD2	0.54	2.38	17	17
1:A:23:HIS:HE1	1:A:54:CYS:C	0.54	2.05	17	13
1:A:50:TYR:CE2	3:C:22:DC:H2'	0.54	2.38	8	11

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
3:C:31:DG:H2''	3:C:32:DC:C5	0.54	2.38	2	14
1:A:64:LYS:HD3	2:B:8:DT:OP1	0.54	2.02	7	4
1:A:23:HIS:CE1	1:A:55:SER:N	0.54	2.75	15	17
1:A:27:LEU:HD21	1:A:49:LYS:HG3	0.53	1.79	5	6
1:A:50:TYR:CD2	3:C:23:DC:H5	0.53	2.21	14	1
1:A:45:PHE:HD1	1:A:46:LYS:O	0.53	1.87	14	15
2:B:2:DC:H2''	2:B:3:DT:C5	0.53	2.39	2	16
1:A:25:PHE:CD2	1:A:49:LYS:HB3	0.53	2.39	9	9
1:A:26:PRO:HB3	1:A:72:LEU:HD13	0.53	1.81	2	12
2:B:15:DC:C2	2:B:16:DG:C6	0.52	2.97	1	17
2:B:2:DC:H2''	2:B:3:DT:C6	0.52	2.39	16	12
1:A:15:ASP:OD2	1:A:17:ASP:HB2	0.52	2.05	17	1
3:C:31:DG:H2''	3:C:32:DC:C6	0.52	2.40	17	9
1:A:67:SER:HB2	1:A:73:LYS:CE	0.52	2.34	16	1
3:C:17:DC:OP2	3:C:17:DC:C5	0.52	2.63	16	1
2:B:3:DT:H2''	2:B:4:DT:C5	0.51	2.40	5	8
2:B:9:DG:H2''	2:B:10:DG:OP2	0.51	2.05	8	13
1:A:60:PRO:O	1:A:63:PHE:HB2	0.51	2.04	10	2
1:A:15:ASP:O	1:A:18:LYS:HG2	0.51	2.04	12	3
1:A:57:HIS:O	1:A:78:PRO:HA	0.51	2.06	15	11
3:C:20:DT:H71	3:C:20:DT:OP2	0.51	2.05	16	4
1:A:61:ASP:O	1:A:67:SER:HA	0.50	2.07	14	1
1:A:30:PRO:O	1:A:33:CYS:HB3	0.50	2.07	2	12
3:C:17:DC:OP2	3:C:17:DC:C6	0.49	2.65	16	1
2:B:11:DG:H2''	2:B:12:DC:C6	0.49	2.42	14	1
2:B:5:DG:H2''	2:B:6:DT:C6	0.49	2.41	17	8
2:B:12:DC:H2''	2:B:13:DA:N7	0.49	2.22	6	2
2:B:11:DG:C2'	2:B:12:DC:C5	0.49	2.93	14	2
3:C:28:DC:C4	3:C:29:DA:N6	0.49	2.80	1	9
1:A:15:ASP:OD1	1:A:17:ASP:HB2	0.49	2.07	3	2
3:C:17:DC:H2''	3:C:18:DG:N7	0.49	2.22	6	5
1:A:33:CYS:SG	1:A:49:LYS:HD3	0.49	2.47	15	2
1:A:24:LYS:HB2	2:B:8:DT:H72	0.49	1.84	15	3
3:C:20:DT:H2''	3:C:21:DG:C8	0.49	2.43	12	1
1:A:32:LEU:HD21	1:A:76:ALA:O	0.49	2.07	17	1
1:A:8:TYR:CE1	1:A:42:ARG:HD3	0.49	2.43	6	2
2:B:14:DG:C4	2:B:15:DC:C4	0.48	3.01	13	6
1:A:67:SER:OG	1:A:68:ASN:N	0.48	2.47	13	14
1:A:8:TYR:CE1	1:A:82:LEU:HB2	0.48	2.44	10	1
1:A:65:ARG:HA	1:A:70:LYS:C	0.48	2.27	10	2
1:A:78:PRO:HB2	1:A:81:PHE:CE1	0.48	2.44	14	5

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:6:DT:C2'	2:B:7:DG:C8	0.48	2.96	11	7
1:A:23:HIS:CE1	1:A:58:PHE:CE2	0.48	3.02	1	3
1:A:65:ARG:HA	1:A:70:LYS:O	0.48	2.09	14	1
2:B:5:DG:C2'	2:B:6:DT:C5	0.48	2.96	17	2
1:A:4:SER:HB2	1:A:11:LYS:HA	0.47	1.86	7	3
1:A:23:HIS:HB3	2:B:8:DT:OP1	0.47	2.09	7	4
1:A:64:LYS:HG2	1:A:65:ARG:N	0.47	2.24	14	2
1:A:62:SER:OG	1:A:73:LYS:HG2	0.47	2.09	17	4
1:A:25:PHE:HD1	1:A:53:ILE:HD11	0.47	1.68	16	1
1:A:64:LYS:HG3	1:A:66:GLU:OE1	0.47	2.08	1	1
1:A:23:HIS:CE1	1:A:58:PHE:HE1	0.47	2.27	9	9
1:A:46:LYS:HG3	3:C:20:DT:OP1	0.47	2.10	7	2
1:A:47:PRO:HA	3:C:21:DG:OP2	0.47	2.09	16	2
1:A:62:SER:HA	1:A:73:LYS:HD3	0.47	1.87	16	1
1:A:61:ASP:C	1:A:73:LYS:HD3	0.47	2.30	12	5
2:B:10:DG:H2''	2:B:11:DG:OP2	0.47	2.10	6	2
1:A:29:ARG:HG2	1:A:32:LEU:CB	0.47	2.40	10	3
1:A:6:SER:HA	1:A:46:LYS:HB3	0.47	1.85	16	1
1:A:25:PHE:CE1	1:A:53:ILE:HD11	0.46	2.45	17	6
1:A:20:VAL:HB	1:A:55:SER:OG	0.46	2.10	14	5
1:A:62:SER:CB	1:A:73:LYS:HG2	0.46	2.40	1	3
1:A:23:HIS:CE1	1:A:58:PHE:CE1	0.46	3.03	13	10
1:A:25:PHE:CD1	1:A:49:LYS:HA	0.46	2.45	10	7
3:C:27:DA:C4	3:C:28:DC:C4	0.46	3.04	6	2
1:A:18:LYS:HZ2	1:A:20:VAL:HG22	0.46	1.71	14	1
1:A:49:LYS:HD2	1:A:49:LYS:N	0.46	2.25	14	1
1:A:10:CYS:SG	1:A:10:CYS:O	0.46	2.73	15	1
3:C:25:DA:H2''	3:C:26:DC:OP2	0.46	2.11	6	9
1:A:40:VAL:O	1:A:40:VAL:HG12	0.46	2.10	3	2
1:A:25:PHE:CD2	1:A:33:CYS:SG	0.46	3.08	13	5
1:A:8:TYR:CE2	1:A:42:ARG:HD3	0.45	2.45	12	1
1:A:49:LYS:H	1:A:49:LYS:CD	0.45	2.22	14	1
1:A:6:SER:O	1:A:40:VAL:HG13	0.45	2.11	2	3
2:B:1:DG:HO5'	2:B:1:DG:H8	0.45	1.53	15	1
1:A:72:LEU:HD12	1:A:72:LEU:H	0.45	1.71	12	7
1:A:62:SER:HA	1:A:73:LYS:HD2	0.45	1.88	16	1
1:A:23:HIS:O	1:A:53:ILE:N	0.45	2.46	8	2
2:B:6:DT:H2''	2:B:7:DG:OP2	0.45	2.12	9	5
1:A:10:CYS:O	1:A:10:CYS:SG	0.45	2.75	2	2
1:A:14:TYR:CD2	1:A:22:PHE:HD2	0.45	2.30	16	2
1:A:25:PHE:CG	1:A:49:LYS:HB3	0.44	2.47	3	5

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:6:DT:C1'	2:B:7:DG:N7	0.44	2.81	2	7
1:A:29:ARG:HG2	1:A:32:LEU:HB3	0.44	1.89	10	1
2:B:7:DG:H2''	2:B:8:DT:C5	0.44	2.47	13	3
1:A:64:LYS:HG2	1:A:65:ARG:H	0.44	1.72	14	1
1:A:50:TYR:CD2	3:C:22:DC:C6	0.43	3.06	6	4
1:A:60:PRO:HA	1:A:63:PHE:CE2	0.43	2.47	7	1
1:A:62:SER:HA	1:A:73:LYS:NZ	0.43	2.28	16	1
1:A:59:THR:O	1:A:62:SER:N	0.43	2.51	16	2
3:C:30:DA:H2''	3:C:31:DG:C8	0.43	2.49	16	4
1:A:3:GLN:HE22	3:C:20:DT:H73	0.43	1.73	5	1
1:A:60:PRO:HA	1:A:63:PHE:CG	0.43	2.48	11	4
1:A:71:LEU:CD1	1:A:72:LEU:HD12	0.43	2.40	13	1
1:A:46:LYS:HG3	3:C:20:DT:P	0.43	2.53	2	2
1:A:65:ARG:HD2	2:B:7:DG:H4'	0.43	1.88	16	2
1:A:46:LYS:CB	1:A:47:PRO:CD	0.43	2.97	3	17
1:A:3:GLN:NE2	3:C:20:DT:H73	0.42	2.29	7	3
3:C:21:DG:H2''	3:C:22:DC:C6	0.42	2.48	10	2
3:C:26:DC:C4	3:C:27:DA:N6	0.42	2.87	15	1
1:A:25:PHE:CD1	1:A:53:ILE:CD1	0.42	3.02	13	2
1:A:69:ASN:CG	1:A:70:LYS:N	0.42	2.72	16	1
1:A:48:THR:O	1:A:50:TYR:N	0.42	2.52	10	2
1:A:60:PRO:HA	1:A:63:PHE:CD1	0.42	2.50	17	1
2:B:15:DC:N3	2:B:16:DG:O6	0.42	2.51	9	1
1:A:23:HIS:HB3	2:B:8:DT:P	0.42	2.54	15	1
1:A:48:THR:HG23	3:C:21:DG:H2'	0.42	1.92	9	1
1:A:81:PHE:O	1:A:82:LEU:HG	0.42	2.15	10	1
1:A:50:TYR:HB3	3:C:23:DC:H41	0.41	1.75	16	1
1:A:77:VAL:O	1:A:79:THR:HG23	0.41	2.15	17	1
1:A:50:TYR:O	3:C:23:DC:N4	0.41	2.52	12	1
3:C:19:DC:H2'	3:C:20:DT:C7	0.41	2.45	5	1
1:A:45:PHE:O	3:C:20:DT:H3'	0.41	2.15	10	1
1:A:29:ARG:HD3	1:A:32:LEU:CB	0.41	2.45	15	1
1:A:47:PRO:HB2	1:A:51:SER:HB2	0.41	1.91	16	1
1:A:47:PRO:HB3	3:C:22:DC:N4	0.41	2.30	8	1
1:A:26:PRO:HB3	1:A:72:LEU:HD11	0.41	1.92	8	1
1:A:8:TYR:CZ	1:A:42:ARG:HD3	0.41	2.49	9	1
1:A:29:ARG:HD2	1:A:74:GLU:HB2	0.41	1.92	17	1
2:B:6:DT:C2	2:B:7:DG:N7	0.41	2.89	16	1
1:A:71:LEU:HD23	1:A:71:LEU:HA	0.41	1.77	6	2
1:A:64:LYS:HD2	2:B:8:DT:OP1	0.41	2.16	5	1
2:B:3:DT:H2''	2:B:4:DT:C6	0.41	2.50	11	3

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:36:TRP:CE2	1:A:78:PRO:CD	0.41	3.03	17	1
1:A:29:ARG:N	1:A:30:PRO:CD	0.41	2.83	15	2
1:A:67:SER:OG	1:A:73:LYS:HE3	0.41	2.14	16	1
3:C:21:DG:H2''	3:C:22:DC:C5	0.41	2.51	10	1
1:A:46:LYS:HG3	3:C:20:DT:OP2	0.41	2.15	13	1
1:A:77:VAL:HG23	1:A:78:PRO:HD2	0.41	1.92	15	1
1:A:67:SER:CB	1:A:71:LEU:O	0.41	2.63	16	1
3:C:17:DC:C6	3:C:17:DC:O5'	0.41	2.74	16	1
1:A:57:HIS:HD1	1:A:81:PHE:HA	0.40	1.75	8	1
1:A:50:TYR:CZ	3:C:23:DC:OP2	0.40	2.73	11	1
3:C:17:DC:C5	3:C:17:DC:OP2	0.40	2.75	14	1
1:A:8:TYR:CD1	1:A:42:ARG:HG2	0.40	2.51	12	1
2:B:6:DT:C4	2:B:7:DG:O6	0.40	2.75	2	1
1:A:45:PHE:CD1	1:A:46:LYS:O	0.40	2.72	11	3
1:A:50:TYR:CD1	3:C:23:DC:C5	0.40	3.04	13	1
1:A:67:SER:HB2	1:A:73:LYS:HZ1	0.40	1.76	16	1
1:A:50:TYR:CD2	3:C:22:DC:C2'	0.40	3.04	8	1
1:A:55:SER:HA	1:A:58:PHE:CE1	0.40	2.52	17	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	A	80/87 (92%)	57±2 (71±2%)	14±2 (18±2%)	9±1 (11±1%)	1 8
All	All	1360/1479 (92%)	970 (71%)	239 (18%)	151 (11%)	1 8

All 12 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	41	ARG	17
1	A	46	LYS	17
1	A	47	PRO	17
1	A	64	LYS	17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	67	SER	17
1	A	13	ARG	16
1	A	7	ALA	15
1	A	74	GLU	14
1	A	40	VAL	11
1	A	72	LEU	6
1	A	49	LYS	3
1	A	63	PHE	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	75/82 (91%)	44±2 (58±3%)	31±2 (42±3%)	0	4
All	All	1275/1394 (91%)	744 (58%)	531 (42%)	0	4

All 53 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	13	ARG	17
1	A	14	TYR	17
1	A	16	LYS	17
1	A	23	HIS	17
1	A	34	LYS	17
1	A	41	ARG	17
1	A	49	LYS	17
1	A	52	SER	17
1	A	56	GLU	17
1	A	58	PHE	17
1	A	71	LEU	17
1	A	75	ASN	17
1	A	77	VAL	17
1	A	18	LYS	16
1	A	35	GLU	16
1	A	63	PHE	16
1	A	46	LYS	15

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	80	ILE	15
1	A	82	LEU	15
1	A	29	ARG	14
1	A	31	SER	13
1	A	44	ASN	13
1	A	22	PHE	12
1	A	62	SER	12
1	A	42	ARG	11
1	A	68	ASN	10
1	A	74	GLU	10
1	A	4	SER	9
1	A	8	TYR	9
1	A	17	ASP	9
1	A	73	LYS	9
1	A	21	SER	8
1	A	65	ARG	8
1	A	79	THR	8
1	A	5	CYS	8
1	A	59	THR	7
1	A	61	ASP	7
1	A	12	ASN	7
1	A	55	SER	4
1	A	70	LYS	4
1	A	25	PHE	3
1	A	45	PHE	3
1	A	64	LYS	3
1	A	66	GLU	2
1	A	43	LYS	2
1	A	6	SER	2
1	A	81	PHE	2
1	A	20	VAL	2
1	A	24	LYS	2
1	A	15	ASP	1
1	A	72	LEU	1
1	A	51	SER	1
1	A	78	PRO	1

6.3.3 RNA ⓘ

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 monosaccharides in this entry.

6.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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