Full wwPDB NMR Structure Validation Report (i)

Apr 21, 2024 – 05:57 PM EDT

PDB ID	:	2LEX
BMRB ID	:	17732
Title	:	Complex of the C-terminal WRKY domain of AtWRKY4 and a W-box DNA
Authors	:	Yamasaki, K.; Kigawa, T.; Watanabe, S.; Inoue, M.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2011-06-24

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:

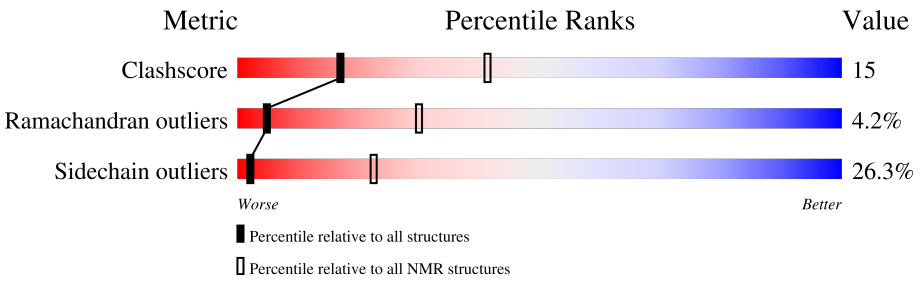
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 55%.

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	78	36% 31% • 12% 19%
2	B	16	38% 25% 38%
3	C	16	44% 19% 38%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:412-A:465 (54)	0.56	1

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 4 clusters. No single-model clusters were found.

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

3 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1636 atoms, of which 723 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Probable WRKY transcription factor 4.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	63	1006	320	498	95	91	2	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	392	GLY	-	expression tag	UNP Q9XI90
A	393	SER	-	expression tag	UNP Q9XI90
A	394	SER	-	expression tag	UNP Q9XI90
A	395	GLY	-	expression tag	UNP Q9XI90
A	396	SER	-	expression tag	UNP Q9XI90
A	397	SER	-	expression tag	UNP Q9XI90
A	398	GLY	-	expression tag	UNP Q9XI90

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

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	10	310	96	113	33	59	9	0

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

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
3	C	10	319	99	112	42	57	9	0

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

Mol	Chain	Residues	Atoms	
			Total	Zn
4	A	1	1	1

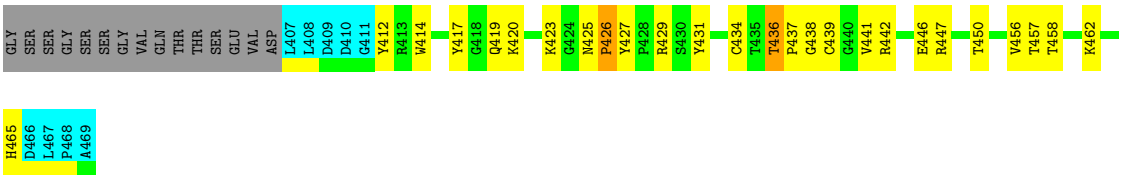
4 Residue-property plots (i)

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.

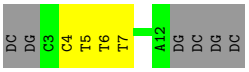
• Molecule 1: Probable WRKY transcription factor 4

Chain A: 36% 31% • 12% 19%



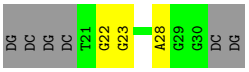
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 38% 25% 38%



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

Chain C: 44% 19% 38%



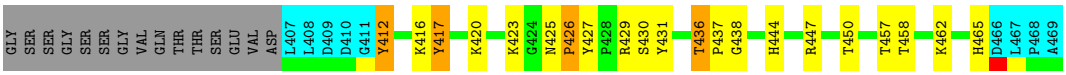
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

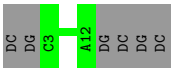
• Molecule 1: Probable WRKY transcription factor 4

Chain A: 42% 22% 5% 12% 19%



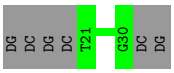
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 62% 38%



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

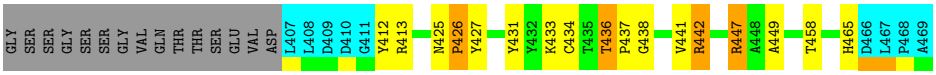
Chain C: 62% 38%



4.2.2 Score per residue for model 2

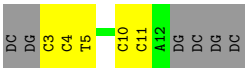
• Molecule 1: Probable WRKY transcription factor 4

Chain A: 47% 17% 5% 12% 19%



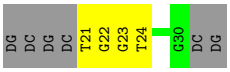
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 31% 31% 38%



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

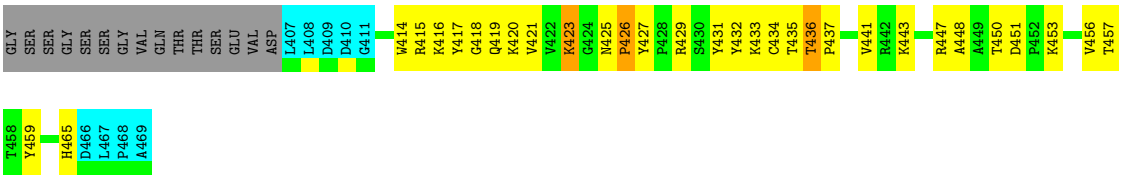
Chain C: 38% 25% 38%



4.2.3 Score per residue for model 3

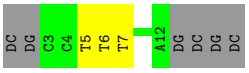
• Molecule 1: Probable WRKY transcription factor 4

Chain A: 29% 36% • 12% 19%



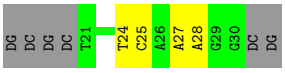
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 44% 19% 38%



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

Chain C: 38% 25% 38%



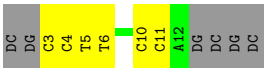
4.2.4 Score per residue for model 4

• Molecule 1: Probable WRKY transcription factor 4

Chain A: 36% 29% • 12% 19%

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

Chain B: 25% 38% 38%



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

Chain C: 38% 25% 38%

DG DC DG DC T21 G22 G23 A27 A28 G29 G30 DC DG

4.2.5 Score per residue for model 5

• Molecule 1: Probable WRKY transcription factor 4

Chain A: 29% 33% 6% 12% 19%

GLY SER SER GLY SER SER GLY VAL GLN THR THR SER GLU VAL ASP L407 L408 D409 D410 G411 Y412 R413 W414 R415 K416 Y417 G418 Q419 K420 V421 V422 K423 G424 N425 P426 Y427 P428 R429 S430 Y431 Y432 K433 C434 T435 T436 P437 G438 V441 R442 K443 H444 V445 E446 R447 A448 A449 T450 D451

T457 E460 G461 K462 H463 D466 L467 P468 A469

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

Chain B: 38% 25% 38%

DC DG C3 C4 T5 T6 T7 A12 DG DC DG DC

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

Chain C: 38% 25% 38%

DG DC DG DC T21 G22 G23 A28 G29 G30 DC DG

4.2.6 Score per residue for model 6

• Molecule 1: Probable WRKY transcription factor 4

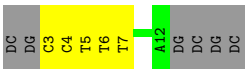
Chain A: 35% 27% 8% 12% 19%

GLY SER SER GLY SER SER GLY VAL GLN THR THR SER GLU VAL ASP L407 L408 D409 D410 G411 Y412 R413 W414 R415 K416 Y417 G418 Q419 K420 V421 V422 K423 G424 N425 P426 Y427 S430 Y431 Y432 K433 C434 T435 T436 P437 G438 C439 G440 V441 E446 A449 K453 V456 T457 T458

H463 N464 H465 D466 L467 P468 A469

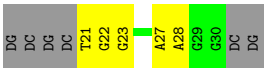
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 31% 31% 38%



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

Chain C: 31% 31% 38%



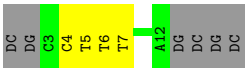
4.2.7 Score per residue for model 7

• Molecule 1: Probable WRKY transcription factor 4

Chain A: 31% 31% 8% 12% 19%

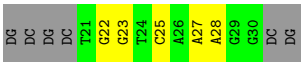
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 38% 25% 38%



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

Chain C: 31% 31% 38%



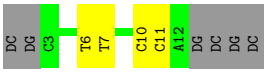
4.2.8 Score per residue for model 8

• Molecule 1: Probable WRKY transcription factor 4

Chain A: 38% 26% 5% 12% 19%

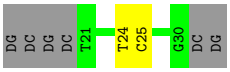
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 38% 25% 38%



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

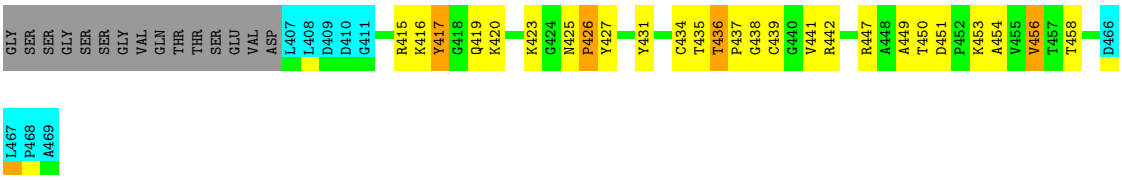
Chain C: 50% 12% 38%



4.2.9 Score per residue for model 9

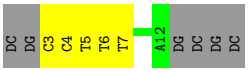
• Molecule 1: Probable WRKY transcription factor 4

Chain A: 36% 28% 5% 12% 19%



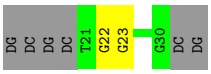
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 31% 31% 38%



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

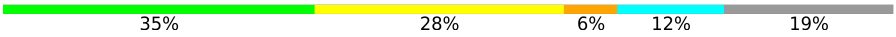
Chain C: 50% 12% 38%

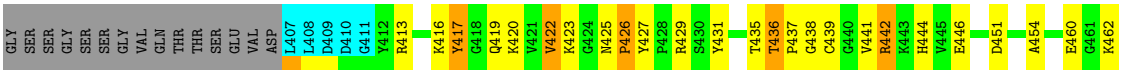


4.2.10 Score per residue for model 10

• Molecule 1: Probable WRKY transcription factor 4

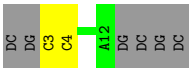
Chain A:





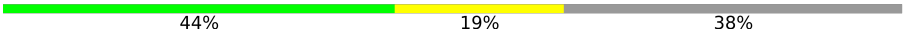
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

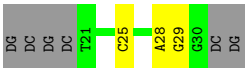
Chain B:



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

Chain C:

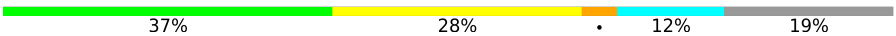


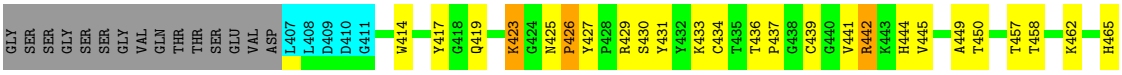


4.2.11 Score per residue for model 11

• Molecule 1: Probable WRKY transcription factor 4

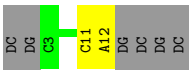
Chain A:





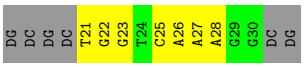
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B:



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

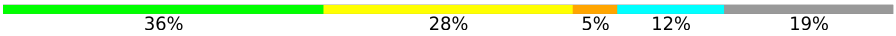
Chain C:

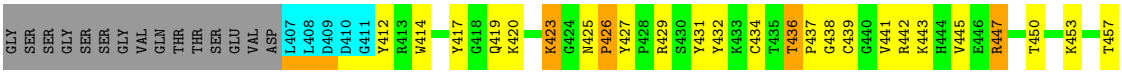


4.2.12 Score per residue for model 12

• Molecule 1: Probable WRKY transcription factor 4

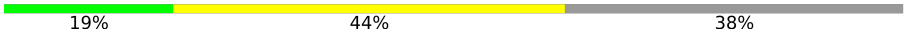
Chain A:

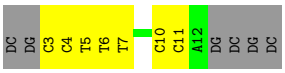




• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

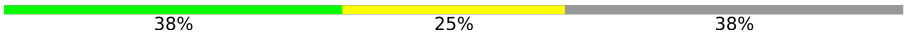
Chain B:

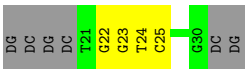




• Molecule 3: DNA (5'-D(*GP*CP*GP*C*TP*GP*GP*TP*CP*AP*AP*AP*GP*GP*CP*G)-3')

Chain C:

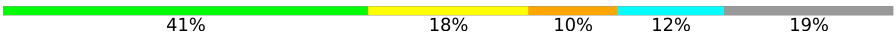


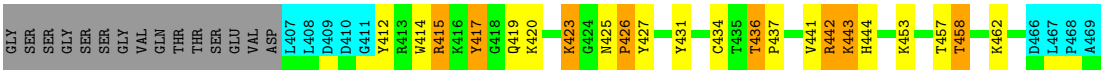


4.2.13 Score per residue for model 13

• Molecule 1: Probable WRKY transcription factor 4

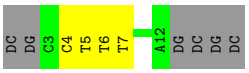
Chain A:





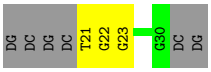
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 38% 25% 38%



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

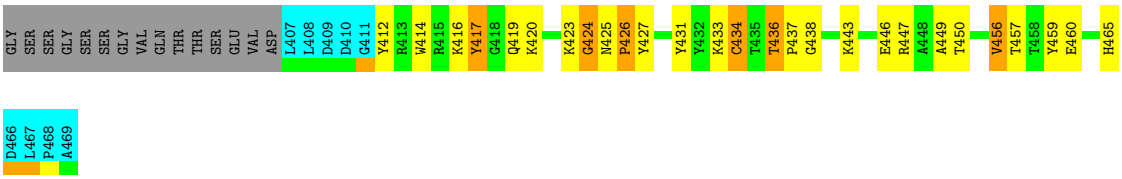
Chain C: 44% 19% 38%



4.2.14 Score per residue for model 14

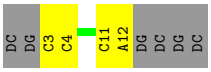
• Molecule 1: Probable WRKY transcription factor 4

Chain A: 35% 27% 8% 12% 19%



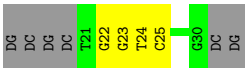
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 38% 25% 38%



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

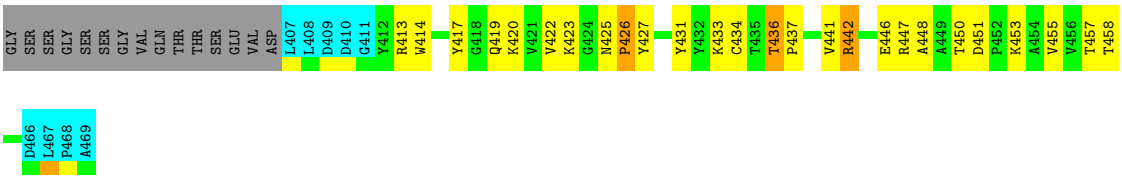
Chain C: 38% 25% 38%



4.2.15 Score per residue for model 15

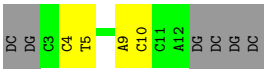
• Molecule 1: Probable WRKY transcription factor 4

Chain A: 36% 29% • 12% 19%



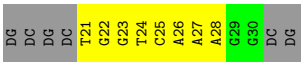
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 38% 25% 38%



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

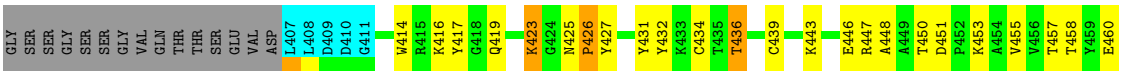
Chain C: 12% 50% 38%

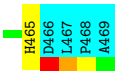


4.2.16 Score per residue for model 16

• Molecule 1: Probable WRKY transcription factor 4

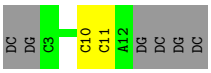
Chain A: 37% 28% • 12% 19%





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

Chain B: 50% 12% 38%



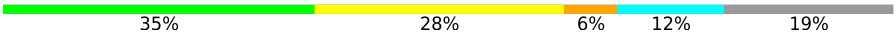
• Molecule 3: DNA (5'-D(*GP*CP*GP*C*TP*GP*GP*TP*CP*AP*AP*AP*GP*GP*CP*G)-3')

Chain C: 19% 44% 38%

4.2.17 Score per residue for model 17

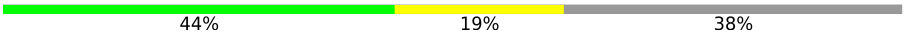
• Molecule 1: Probable WRKY transcription factor 4

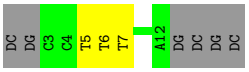
Chain A:



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

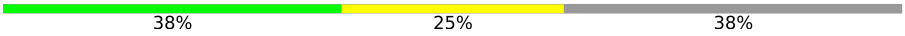
Chain B:

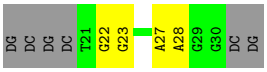




• Molecule 3: DNA (5'-D(*GP*CP*GP*C*TP*GP*GP*TP*CP*AP*AP*AP*GP*GP*CP*G)-3')

Chain C:

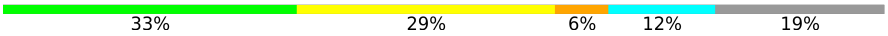


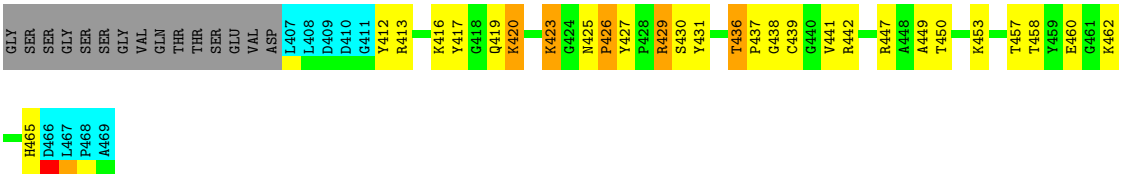


4.2.18 Score per residue for model 18

• Molecule 1: Probable WRKY transcription factor 4

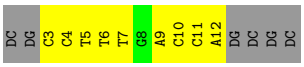
Chain A:





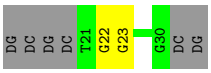
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B:



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

Chain C: 50% 12% 38%



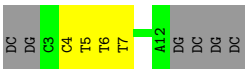
4.2.19 Score per residue for model 19

• Molecule 1: Probable WRKY transcription factor 4

Chain A: 37% 27% 5% 12% 19%

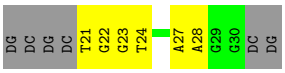
• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

Chain B: 38% 25% 38%



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

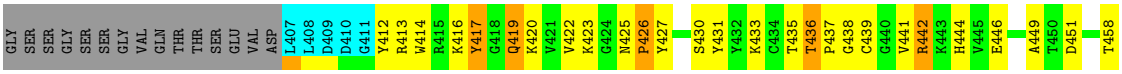
Chain C: 25% 38% 38%

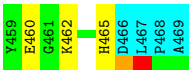


4.2.20 Score per residue for model 20

• Molecule 1: Probable WRKY transcription factor 4

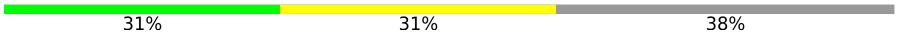
Chain A: 31% 32% 6% 12% 19%

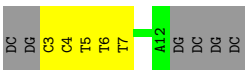




• Molecule 2: DNA (5'-D(*CP*G*CP*CP*TP*TP*TP*GP*AP*CP*CP*AP*GP*CP*GP*C)-3')

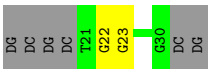
Chain B:





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

Chain C:



5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 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	structure solution	1.1
CNS	refinement	1.1

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	857
Number of shifts mapped to atoms	712
Number of unparsed shifts	0
Number of shifts with mapping errors	145
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	55%

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

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	443	438	437	16±3
2	B	197	113	115	3±2
3	C	207	112	114	5±3
All	All	16960	13260	13320	466

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 15.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:448:ALA:HB3	1:A:451:ASP:O	0.76	1.80	16	4
1:A:415:ARG:NH1	2:B:5:DT:H71	0.74	1.97	13	1
3:C:25:DC:H2"	3:C:26:DA:O5'	0.73	1.84	16	3
1:A:436:THR:HG21	1:A:465:HIS:CE1	0.71	2.20	1	3
3:C:22:DG:H2'	3:C:23:DG:C8	0.71	2.21	11	11
3:C:22:DG:H2"	3:C:23:DG:O4'	0.71	1.86	18	4
1:A:425:ASN:O	1:A:427:TYR:N	0.70	2.24	5	20
3:C:21:DT:H2"	3:C:22:DG:O5'	0.69	1.87	15	3
1:A:426:PRO:O	1:A:427:TYR:CD1	0.68	2.47	2	20
1:A:421:VAL:HG12	1:A:428:PRO:HB3	0.67	1.67	4	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:425:ASN:OD1	1:A:449:ALA:HB2	0.67	1.88	18	6
2:B:5:DT:C6	2:B:6:DT:H72	0.65	2.26	3	4
1:A:436:THR:HG21	1:A:465:HIS:NE2	0.64	2.07	7	3
2:B:4:DC:H2"	2:B:5:DT:O5'	0.64	1.91	18	4
2:B:3:DC:H2"	2:B:4:DC:O5'	0.62	1.94	14	7
1:A:436:THR:HG21	1:A:465:HIS:CG	0.62	2.30	10	5
3:C:21:DT:H2"	3:C:22:DG:N7	0.61	2.10	2	3
3:C:22:DG:H2"	3:C:23:DG:O5'	0.60	1.96	12	7
1:A:436:THR:HG21	1:A:465:HIS:CD2	0.60	2.32	7	10
3:C:24:DT:H2"	3:C:25:DC:O5'	0.60	1.96	8	4
1:A:436:THR:HG22	1:A:437:PRO:HD2	0.59	1.73	13	4
2:B:6:DT:H2"	2:B:7:DT:O5'	0.57	1.99	13	10
1:A:425:ASN:CG	1:A:449:ALA:HB2	0.57	2.18	14	1
1:A:436:THR:HG22	1:A:437:PRO:CD	0.57	2.29	3	2
1:A:415:ARG:HH12	1:A:435:THR:HG21	0.56	1.60	3	1
2:B:4:DC:C2'	2:B:5:DT:H71	0.56	2.30	9	1
1:A:425:ASN:N	1:A:426:PRO:CD	0.56	2.68	15	17
1:A:423:LYS:C	1:A:425:ASN:H	0.56	2.04	14	17
1:A:415:ARG:NH1	1:A:435:THR:HG21	0.55	2.16	3	1
3:C:27:DA:C5	3:C:28:DA:N6	0.55	2.75	4	7
1:A:425:ASN:HA	1:A:449:ALA:HB2	0.55	1.79	20	3
1:A:422:VAL:HG13	1:A:429:ARG:HD2	0.55	1.77	10	1
3:C:27:DA:C6	3:C:28:DA:N6	0.54	2.75	7	1
2:B:3:DC:H2"	2:B:4:DC:O4'	0.54	2.03	18	2
3:C:27:DA:C5	3:C:28:DA:C6	0.54	2.96	15	6
3:C:27:DA:N7	3:C:28:DA:N6	0.54	2.56	16	4
1:A:412:TYR:O	1:A:414:TRP:CD1	0.53	2.61	5	6
1:A:451:ASP:HB3	1:A:454:ALA:HB3	0.53	1.80	9	2
1:A:417:TYR:CD2	1:A:431:TYR:O	0.53	2.62	6	15
3:C:22:DG:C2'	3:C:23:DG:C8	0.53	2.92	16	2
1:A:417:TYR:CD1	1:A:431:TYR:O	0.52	2.63	15	2
2:B:11:DC:C5	2:B:12:DA:N6	0.52	2.78	14	2
1:A:412:TYR:CZ	1:A:465:HIS:CE1	0.51	2.99	6	1
1:A:431:TYR:CD1	3:C:24:DT:C7	0.51	2.93	2	1
3:C:25:DC:O5'	3:C:25:DC:H6	0.51	1.89	11	5
1:A:436:THR:O	1:A:438:GLY:N	0.50	2.44	6	14
3:C:22:DG:C2'	3:C:23:DG:O5'	0.50	2.60	4	9
2:B:4:DC:H2"	2:B:5:DT:H71	0.49	1.83	2	3
1:A:412:TYR:CE2	1:A:463:HIS:CD2	0.49	3.00	5	1
2:B:3:DC:C2'	2:B:4:DC:C6	0.49	2.96	9	2
1:A:420:LYS:CE	3:C:23:DG:N7	0.48	2.75	6	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
3:C:25:DC:O5'	3:C:25:DC:C6	0.48	2.65	11	4
1:A:412:TYR:CE2	1:A:463:HIS:CG	0.48	3.01	7	1
1:A:426:PRO:C	1:A:427:TYR:CD1	0.48	2.87	2	19
1:A:463:HIS:CD2	1:A:463:HIS:N	0.48	2.82	6	1
1:A:427:TYR:CD2	1:A:447:ARG:HB2	0.48	2.44	2	4
1:A:414:TRP:CE3	1:A:434:CYS:N	0.48	2.82	15	6
1:A:427:TYR:CD2	1:A:447:ARG:CB	0.47	2.98	15	3
1:A:432:TYR:O	1:A:443:LYS:N	0.47	2.47	16	5
1:A:451:ASP:OD2	1:A:454:ALA:HB2	0.47	2.09	10	1
1:A:465:HIS:O	1:A:465:HIS:ND1	0.47	2.48	18	2
1:A:414:TRP:CE3	1:A:434:CYS:HB2	0.47	2.44	6	1
2:B:11:DC:C4	2:B:12:DA:N6	0.47	2.82	14	3
3:C:22:DG:O5'	3:C:22:DG:C8	0.47	2.68	2	2
1:A:441:VAL:CG1	1:A:442:ARG:N	0.47	2.78	20	11
1:A:417:TYR:CE1	1:A:431:TYR:CB	0.47	2.98	12	1
1:A:434:CYS:N	1:A:441:VAL:O	0.47	2.48	2	10
3:C:25:DC:H2'	3:C:26:DA:C8	0.47	2.45	15	2
1:A:417:TYR:CZ	1:A:431:TYR:CB	0.46	2.98	3	6
2:B:6:DT:C2'	2:B:7:DT:C6	0.46	2.98	20	3
1:A:421:VAL:HG12	1:A:428:PRO:CB	0.46	2.40	4	1
1:A:431:TYR:CD1	1:A:431:TYR:N	0.46	2.84	6	4
2:B:10:DC:H2"	2:B:11:DC:C6	0.46	2.45	16	5
2:B:6:DT:H2'	2:B:7:DT:H72	0.46	1.86	12	1
1:A:424:GLY:O	1:A:425:ASN:OD1	0.46	2.34	14	1
1:A:417:TYR:CD1	1:A:417:TYR:C	0.46	2.89	13	6
1:A:420:LYS:O	1:A:429:ARG:N	0.46	2.48	18	1
2:B:9:DA:H2"	2:B:10:DC:OP2	0.46	2.10	18	1
1:A:436:THR:HG23	1:A:437:PRO:HD2	0.45	1.89	11	1
1:A:414:TRP:CZ3	1:A:434:CYS:HB2	0.45	2.47	19	8
1:A:417:TYR:C	1:A:417:TYR:CD1	0.45	2.89	9	2
1:A:423:LYS:C	1:A:425:ASN:N	0.45	2.69	14	5
1:A:448:ALA:HB3	1:A:451:ASP:HB3	0.45	1.89	3	1
1:A:443:LYS:CD	1:A:459:TYR:CE1	0.45	3.00	3	1
3:C:22:DG:H2"	3:C:23:DG:C8	0.45	2.46	16	3
2:B:4:DC:C2'	2:B:5:DT:H72	0.45	2.42	4	4
1:A:423:LYS:O	1:A:425:ASN:N	0.44	2.50	14	1
1:A:463:HIS:N	1:A:463:HIS:CD2	0.44	2.85	19	1
1:A:417:TYR:CZ	1:A:431:TYR:HB3	0.44	2.48	9	4
1:A:443:LYS:HA	1:A:458:THR:O	0.44	2.12	13	2
2:B:4:DC:C2'	2:B:5:DT:O5'	0.44	2.65	18	1
1:A:446:GLU:HB2	1:A:456:VAL:HG22	0.43	1.89	4	4

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:417:TYR:CE2	1:A:431:TYR:HB3	0.43	2.48	1	1
1:A:418:GLY:N	2:B:7:DT:H71	0.43	2.29	8	1
1:A:456:VAL:O	1:A:456:VAL:CG2	0.43	2.66	9	1
1:A:415:ARG:NH1	2:B:5:DT:C7	0.43	2.78	13	1
1:A:425:ASN:O	1:A:425:ASN:ND2	0.43	2.51	2	4
1:A:412:TYR:CE2	1:A:463:HIS:CE1	0.43	3.07	4	1
1:A:427:TYR:CE1	1:A:449:ALA:HA	0.43	2.49	5	2
1:A:431:TYR:CE2	3:C:24:DT:C7	0.43	3.02	14	2
1:A:443:LYS:HD2	1:A:459:TYR:CZ	0.43	2.49	14	1
2:B:5:DT:C5	2:B:6:DT:H73	0.43	2.48	17	1
1:A:426:PRO:O	1:A:427:TYR:CG	0.43	2.72	2	1
1:A:415:ARG:HG3	2:B:5:DT:H71	0.42	1.90	6	1
1:A:417:TYR:CE2	1:A:431:TYR:CB	0.42	3.02	14	2
2:B:6:DT:H2'	2:B:7:DT:C6	0.42	2.50	8	1
3:C:22:DG:C2'	3:C:23:DG:O4'	0.42	2.67	15	2
1:A:447:ARG:CD	1:A:447:ARG:N	0.42	2.83	8	1
1:A:417:TYR:CE1	1:A:433:LYS:CD	0.42	3.03	15	1
2:B:6:DT:H2"	2:B:7:DT:C6	0.42	2.50	20	3
1:A:425:ASN:N	1:A:426:PRO:HD3	0.41	2.30	2	2
2:B:9:DA:H2"	2:B:10:DC:C5	0.41	2.50	15	1
1:A:417:TYR:O	1:A:431:TYR:HB2	0.41	2.15	7	1
2:B:3:DC:H2"	2:B:4:DC:C6	0.41	2.51	6	1
1:A:417:TYR:O	1:A:417:TYR:CD1	0.41	2.73	17	1
3:C:21:DT:H2"	3:C:22:DG:C8	0.41	2.50	13	1
1:A:436:THR:HG21	1:A:465:HIS:ND1	0.41	2.31	17	1
1:A:436:THR:OG1	1:A:439:CYS:HB3	0.41	2.16	16	1
1:A:420:LYS:HE3	3:C:23:DG:N7	0.41	2.31	6	1
3:C:24:DT:H2'	3:C:25:DC:C6	0.41	2.50	3	1
3:C:28:DA:H2"	3:C:29:DG:OP2	0.41	2.16	10	2
2:B:6:DT:C2'	2:B:7:DT:O5'	0.41	2.69	7	1
1:A:417:TYR:CE1	1:A:431:TYR:HB2	0.41	2.51	12	1
2:B:3:DC:H2'	2:B:4:DC:C6	0.41	2.51	18	2
1:A:418:GLY:O	2:B:7:DT:H73	0.41	2.16	6	1
1:A:434:CYS:SG	1:A:463:HIS:CE1	0.41	3.14	19	1
1:A:441:VAL:HG22	1:A:461:GLY:C	0.40	2.37	17	1
1:A:425:ASN:OD1	1:A:449:ALA:CB	0.40	2.68	4	1
2:B:4:DC:H2"	2:B:5:DT:C6	0.40	2.51	15	1
1:A:441:VAL:HB	1:A:463:HIS:CE1	0.40	2.52	6	1
1:A:429:ARG:HG2	1:A:430:SER:N	0.40	2.31	7	1
1:A:441:VAL:HG12	1:A:442:ARG:N	0.40	2.31	13	1
1:A:412:TYR:OH	1:A:463:HIS:CG	0.40	2.75	19	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
3:C:23:DG:H2"	3:C:24:DT:C6	0.40	2.50	19	1
1:A:444:HIS:CE1	1:A:446:GLU:HB2	0.40	2.52	5	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	54/78 (69%)	45±2 (83±4%)	7±2 (13±3%)	2±1 (4±2%)	5 30
All	All	1080/1560 (69%)	892 (83%)	143 (13%)	45 (4%)	5 30

All 9 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	426	PRO	20
1	A	437	PRO	14
1	A	455	VAL	3
1	A	412	TYR	2
1	A	419	GLN	2
1	A	418	GLY	1
1	A	424	GLY	1
1	A	465	HIS	1
1	A	451	ASP	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	46/65 (71%)	34±2 (74±5%)	12±2 (26±5%)	2 22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	920/1300 (71%)	678 (74%)	242 (26%)	2 22

All 32 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	436	THR	17
1	A	457	THR	15
1	A	419	GLN	15
1	A	420	LYS	13
1	A	447	ARG	12
1	A	450	THR	12
1	A	462	LYS	12
1	A	423	LYS	12
1	A	458	THR	11
1	A	439	CYS	11
1	A	453	LYS	10
1	A	416	LYS	9
1	A	429	ARG	9
1	A	442	ARG	9
1	A	417	TYR	8
1	A	430	SER	7
1	A	433	LYS	7
1	A	446	GLU	7
1	A	444	HIS	6
1	A	413	ARG	6
1	A	456	VAL	6
1	A	460	GLU	6
1	A	422	VAL	5
1	A	435	THR	5
1	A	415	ARG	3
1	A	421	VAL	2
1	A	434	CYS	2
1	A	412	TYR	1
1	A	431	TYR	1
1	A	464	ASN	1
1	A	443	LYS	1
1	A	463	HIS	1

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates (i)

There are no 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 (i)

The completeness of assignment taking into account all chemical shift lists is 55% for the well-defined parts and 56% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *shiftFile_1*

7.1.1 Bookkeeping (i)

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	857
Number of shifts mapped to atoms	712
Number of unparsed shifts	0
Number of shifts with mapping errors	145
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. All 145 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	1	DC	H1'	5.79	.01	1
1	B	1	DC	H2'	1.97	.01	1
1	B	1	DC	H2"	2.43	.01	1
1	B	1	DC	H5	5.9	.01	1
1	B	1	DC	H6	7.64	.01	1
1	B	2	DG	H1	13.08	.01	1
1	B	2	DG	H1'	5.94	.01	1
1	B	2	DG	H2'	2.71	.01	1
1	B	2	DG	H2"	2.76	.01	1
1	B	2	DG	H8	7.97	.01	1
1	B	13	DG	H1	12.79	.01	1
1	B	13	DG	H1'	5.69	.01	1
1	B	13	DG	H2'	2.51	.01	1
1	B	13	DG	H2"	2.61	.01	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	13	DG	H8	7.69	.01	1
1	B	14	DC	H1'	5.68	.01	1
1	B	14	DC	H2'	1.88	.01	1
1	B	14	DC	H2"	2.33	.01	1
1	B	14	DC	H5	5.3	.01	1
1	B	14	DC	H6	7.23	.01	1
1	B	14	DC	H41	8.18	.01	1
1	B	14	DC	H42	6.3	.01	1
1	B	15	DG	H1	13.08	.01	1
1	B	15	DG	H1'	5.93	.01	1
1	B	15	DG	H2'	2.59	.01	1
1	B	15	DG	H2"	2.71	.01	1
1	B	15	DG	H8	7.87	.01	1
1	B	16	DC	H1'	6.17	.01	1
1	B	16	DC	H2'	2.15	.01	1
1	B	16	DC	H2"	2.19	.01	1
1	B	16	DC	H5	5.48	.01	1
1	B	16	DC	H6	7.43	.01	1
1	C	17	DG	H1'	5.96	.01	1
1	C	17	DG	H2'	2.59	.01	1
1	C	17	DG	H2"	2.76	.01	1
1	C	17	DG	H8	7.94	.01	1
1	C	18	DC	H1'	5.74	.01	1
1	C	18	DC	H2'	2.11	.01	1
1	C	18	DC	H2"	2.45	.01	1
1	C	18	DC	H5	5.39	.01	1
1	C	18	DC	H6	7.41	.01	1
1	C	19	DG	H1	12.95	.01	1
1	C	19	DG	H1'	5.92	.01	1
1	C	19	DG	H2'	2.64	.01	1
1	C	19	DG	H2"	2.73	.01	1
1	C	19	DG	H8	7.9	.01	1
1	C	20	DC	H1'	5.97	.01	1
1	C	20	DC	H2'	1.9	.01	1
1	C	20	DC	H2"	2.44	.01	1
1	C	20	DC	H3'	4.83	.01	1
1	C	20	DC	H5	5.31	.01	1
1	C	20	DC	H6	7.37	.01	1
1	C	31	DC	H1'	5.77	.01	1
1	C	31	DC	H2'	1.93	.01	1
1	C	31	DC	H2"	2.34	.01	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	C	31	DC	H3'	4.82	.01	1
1	C	31	DC	H5	5.28	.01	1
1	C	31	DC	H6	7.28	.01	1
1	C	32	DG	H1'	6.18	.01	1
1	C	32	DG	H2'	2.38	.01	1
1	C	32	DG	H2"	2.61	.01	1
1	C	32	DG	H3'	4.68	.01	1
1	C	32	DG	H4'	4.19	.01	1
1	C	32	DG	H5'	4.08	.01	2
1	C	32	DG	H8	7.92	.01	1
1	A	399	VAL	H	8.0	.01	1
1	A	399	VAL	HA	4.19	.01	1
1	A	399	VAL	HB	2.08	.01	1
1	A	399	VAL	HG11	0.93	.01	2
1	A	399	VAL	HG12	0.93	.01	2
1	A	399	VAL	HG13	0.93	.01	2
1	A	399	VAL	HG21	0.93	.01	2
1	A	399	VAL	HG22	0.93	.01	2
1	A	399	VAL	HG23	0.93	.01	2
1	A	399	VAL	CA	62.3	0.1	1
1	A	399	VAL	N	119.6	0.1	1
1	A	400	GLN	H	8.56	.01	1
1	A	400	GLN	HA	4.57	.01	1
1	A	400	GLN	HB2	2.01	.01	2
1	A	400	GLN	HB3	2.13	.01	2
1	A	400	GLN	HG2	2.38	.01	2
1	A	400	GLN	HG3	2.38	.01	2
1	A	400	GLN	HE21	6.83	.01	2
1	A	400	GLN	HE22	7.48	.01	2
1	A	400	GLN	CA	55.6	0.1	1
1	A	400	GLN	N	124.8	0.1	1
1	A	400	GLN	NE2	112.5	0.1	1
1	A	401	THR	H	8.38	.01	1
1	A	401	THR	HA	4.53	.01	1
1	A	401	THR	HB	4.24	.01	1
1	A	401	THR	HG21	1.21	.01	1
1	A	401	THR	HG22	1.21	.01	1
1	A	401	THR	HG23	1.21	.01	1
1	A	401	THR	CA	61.6	0.1	1
1	A	401	THR	N	116.7	0.1	1
1	A	402	THR	H	8.28	.01	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	402	THR	HA	4.51	.01	1
1	A	402	THR	HB	4.25	.01	1
1	A	402	THR	HG21	1.21	.01	1
1	A	402	THR	HG22	1.21	.01	1
1	A	402	THR	HG23	1.21	.01	1
1	A	402	THR	CA	61.7	0.1	1
1	A	402	THR	N	116.6	0.1	1
1	A	403	SER	H	8.43	.01	1
1	A	403	SER	HA	4.56	.01	1
1	A	403	SER	HB2	3.89	.01	2
1	A	403	SER	HB3	3.89	.01	2
1	A	403	SER	CA	58.3	0.1	1
1	A	403	SER	N	118.4	0.1	1
1	A	404	GLU	H	8.55	.01	1
1	A	404	GLU	HA	4.34	.01	1
1	A	404	GLU	HB2	1.96	.01	2
1	A	404	GLU	HB3	2.08	.01	2
1	A	404	GLU	HG2	2.27	.01	2
1	A	404	GLU	HG3	2.3	.01	2
1	A	404	GLU	CA	57.1	0.1	1
1	A	404	GLU	N	123.3	0.1	1
1	A	405	VAL	H	8.0	.01	1
1	A	405	VAL	HA	4.1	.01	1
1	A	405	VAL	HB	2.05	.01	1
1	A	405	VAL	HG11	0.9	.01	2
1	A	405	VAL	HG12	0.9	.01	2
1	A	405	VAL	HG13	0.9	.01	2
1	A	405	VAL	HG21	0.9	.01	2
1	A	405	VAL	HG22	0.9	.01	2
1	A	405	VAL	HG23	0.9	.01	2
1	A	405	VAL	CA	62.3	0.1	1
1	A	405	VAL	N	119.1	0.1	1
1	A	406	ASP	H	8.25	.01	1
1	A	406	ASP	HA	4.58	.01	1
1	A	406	ASP	HB2	2.58	.01	1
1	A	406	ASP	HB3	2.7	.01	1
1	A	406	ASP	CA	54.5	0.1	1
1	A	406	ASP	N	123.4	0.1	1
1	B	2	DG	H3'	4.99	.01	1
1	B	13	DG	H3'	4.95	.01	1
1	B	14	DC	H3'	4.78	.01	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	15	DG	H3'	4.95	.01	1
1	B	16	DC	H3'	4.48	.01	1
1	B	16	DC	H4'	4.03	.01	1
1	C	17	DG	H3'	4.84	.01	1
1	C	18	DC	H3'	4.87	.01	1
1	C	18	DC	H4'	4.19	.01	1
1	C	19	DG	H3'	4.99	.01	1
1	C	31	DC	H4'	4.15	.01	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	70	-0.05 ± 0.25	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	65	-0.90 ± 0.67	None needed (imprecise)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 55%, i.e. 633 atoms were assigned a chemical shift out of a possible 1150. 0 out of 6 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	212/267 (79%)	109/109 (100%)	53/108 (49%)	50/50 (100%)
Sidechain	243/400 (61%)	235/256 (92%)	0/119 (0%)	8/25 (32%)
Aromatic	36/88 (41%)	35/42 (83%)	0/41 (0%)	1/5 (20%)
Sugar	92/240 (38%)	92/140 (66%)	0/100 (0%)	0/0 (—%)
Base	50/155 (32%)	45/95 (47%)	0/35 (0%)	5/25 (20%)
Overall	633/1150 (55%)	516/642 (80%)	53/403 (13%)	64/105 (61%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 56%, i.e. 707 atoms were assigned a chemical shift out of a possible 1258. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	246/311 (79%)	127/127 (100%)	62/126 (49%)	57/58 (98%)

Continued on next page...

Continued from previous page...

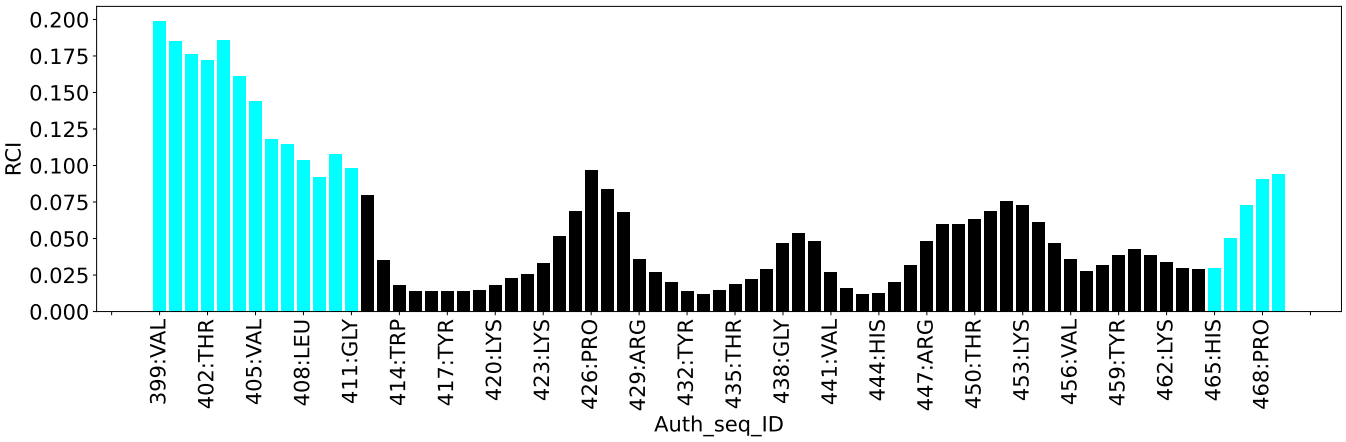
	Total	^{1}H	^{13}C	^{15}N
Sidechain	283/464 (61%)	275/298 (92%)	0/141 (0%)	8/25 (32%)
Aromatic	36/88 (41%)	35/42 (83%)	0/41 (0%)	1/5 (20%)
Sugar	92/240 (38%)	92/140 (66%)	0/100 (0%)	0/0 (—%)
Base	50/155 (32%)	45/95 (47%)	0/35 (0%)	5/25 (20%)
Overall	707/1258 (56%)	574/702 (82%)	62/443 (14%)	71/113 (63%)

7.1.4 Statistically unusual chemical shifts (i)

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.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	443	LYS	HE3	0.57	1.92 – 3.89	-11.9
1	A	443	LYS	HD3	-0.18	0.54 – 2.65	-8.4
1	A	443	LYS	HD2	-0.07	0.58 – 2.64	-8.1
1	A	443	LYS	HE2	1.51	1.95 – 3.88	-7.3
1	A	416	LYS	H	11.12	5.24 – 11.12	5.0

7.1.5 Random Coil Index (RCI) plots (i)

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: