Full wwPDB NMR Structure Validation Report (i)

Nov 7, 2023 – 10:19 AM EST

PDB ID	:	2LTO
BMRB ID	:	18490
Title	:	TDRD3 complex
Authors	:	Sikorsky, T.
Deposited on	:	2012-05-30

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:

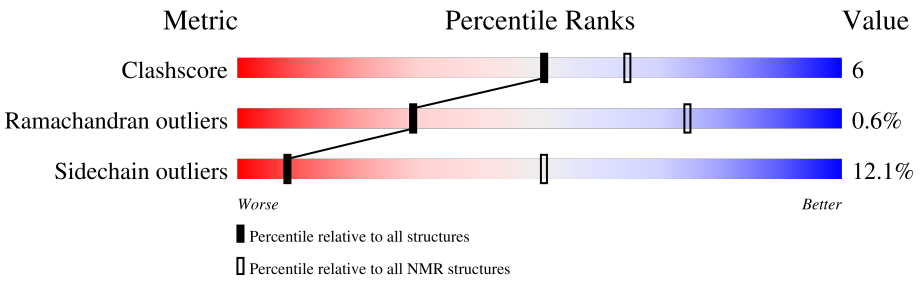
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 78%.

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	58	67% 16% 17%
2	B	13	100%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 5 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:556-A:580, A:586-A:608 (48)	0.33	5

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 3 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1149 atoms, of which 563 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Tudor domain-containing protein 3.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	58	943	313	463	73	90	4	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	553	MET	-	expression tag	UNP Q9H7E2
A	609	LEU	-	expression tag	UNP Q9H7E2
A	610	GLU	-	expression tag	UNP Q9H7E2

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	13	206	67	100	17	22	0

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

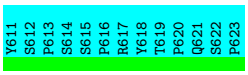
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for 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: Tudor domain-containing protein 3

Chain A: 67% 16% 17%

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%



4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

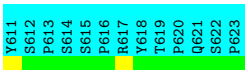
4.2.1 Score per residue for model 1

• Molecule 1: Tudor domain-containing protein 3

Chain A: 62% 17% • 17%

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

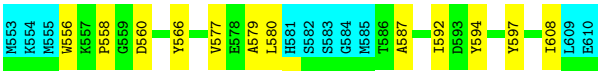
Chain B: 100%



4.2.2 Score per residue for model 2

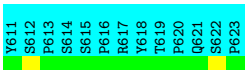
• Molecule 1: Tudor domain-containing protein 3

Chain A: 62% 21% 17%



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

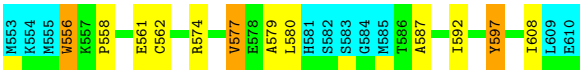
Chain B: 100%



4.2.3 Score per residue for model 3

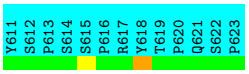
• Molecule 1: Tudor domain-containing protein 3

Chain A: 62% 16% 5% 17%



• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%



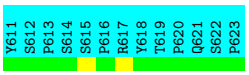
4.2.4 Score per residue for model 4

• Molecule 1: Tudor domain-containing protein 3

Chain A: 62% 17% • 17%

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%



4.2.5 Score per residue for model 5 (medoid)

• Molecule 1: Tudor domain-containing protein 3

Chain A: 62% 17% • 17%

M553 K554 M555 W556 D560 F572 Y573 R574 A579 L580 H581 S582 S583 G584 M585 T586 A587 I592 Y597 L602 K606 P607 I608 L609 E610

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.6 Score per residue for model 6

• Molecule 1: Tudor domain-containing protein 3

Chain A: 57% 26% 17%

M553 K554 M555 W556 K557 D560 E561 K571 R574 V577 E578 A579 L580 H581 S582 S583 G584 M585 T586 A587 I592 D593 Y594 Y597 I605 I608 L609 E610

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.7 Score per residue for model 7

• Molecule 1: Tudor domain-containing protein 3

Chain A: 69% 10% • • 17%

M553 K554 M555 W556 R574 L580 H581 S582 S583 G584 M585 T586 A587 N596 Y597 L601 I608 L609 E610

• Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.8 Score per residue for model 8

- Molecule 1: Tudor domain-containing protein 3

Chain A: 69% 10% • 17%

M553 K554 M555 P558 Y573 V577 E578 A579 L580 H581 S582 S583 G584 M585 T586 A587 Y597 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.9 Score per residue for model 9

- Molecule 1: Tudor domain-containing protein 3

Chain A: 67% 14% • 17%

M553 K554 M555 W556 R574 A579 L580 H581 S582 S583 G584 M585 T586 A587 I592 Y597 L602 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.10 Score per residue for model 10

- Molecule 1: Tudor domain-containing protein 3

Chain A: 62% 19% • 17%

M553 K554 M555 W556 C562 Y566 R574 A579 L580 H581 S582 S583 G584 M585 T586 A587 Y597 E598 E599 S603 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.11 Score per residue for model 11

- Molecule 1: Tudor domain-containing protein 3

Chain A: 64% 17% • 17%

M553 K554 M555 W556 K557 P558 R574 V577 L580 H581 S582 S583 G584 M585 T586 A587 I592 D593 Y597 L601 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.12 Score per residue for model 12

- Molecule 1: Tudor domain-containing protein 3

Chain A: 60% 22% 17%

M553 K554 M555 W556 K557 Y566 F572 Y573 R574 A579 L580 H581 S582 S583 G584 M585 T586 A587 Y594 G595 L601 K606 P607 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.13 Score per residue for model 13

- Molecule 1: Tudor domain-containing protein 3

Chain A: 60% 22% 17%

M553 K554 M555 W556 D560 Y566 W567 E568 R574 A579 L580 H581 S582 S583 G584 M585 T586 A587 Y597 E598 E599 V600 L601 L602 S603 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.14 Score per residue for model 14

- Molecule 1: Tudor domain-containing protein 3

Chain A: 55% 26% • 17%

M553 K554 M555 W556 K557 P558 G559 D560 P563 Y566 Y573 R574 V577 E578 A579 L580 H581 S582 S583 G584 M585 T586 A587 Y597 E598 E599 V600 L601 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.15 Score per residue for model 15

- Molecule 1: Tudor domain-containing protein 3

Chain A: 67% 14% • 17%

M553 K554 M555 W556 K557 R574 A579 L580 H581 S582 S583 G584 M585 T586 A587 Y597 L601 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.16 Score per residue for model 16

- Molecule 1: Tudor domain-containing protein 3

Chain A: 66% 17% 17%

M553 K554 M555 W556 R574 V577 E578 A579 L580 H581 S582 S583 G584 M585 T586 A587 I592 L601 L602 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.17 Score per residue for model 17

- Molecule 1: Tudor domain-containing protein 3

Chain A: 69% 12% • 17%

M553 K554 M555 W556 V577 L580 H581 S582 S583 G584 M585 T586 A587 I592 Y597 L601 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.18 Score per residue for model 18

- Molecule 1: Tudor domain-containing protein 3

Chain A: 64% 17% • 17%

M553 K554 M555 W556 F563 K571 R574 V577 L580 H581 S582 S583 G584 M585 T586 A587 I592 D593 Y594 Y597 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.19 Score per residue for model 19

- Molecule 1: Tudor domain-containing protein 3

Chain A: 60% 19% • 17%

M553 K554 M555 W556 D560 R574 V577 E578 A579 L580 H581 S582 S583 G584 M585 T586 A587 I592 N596 Y597 L601 L602 I608 L609 E610

- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

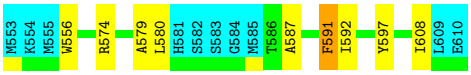
Chain B: 100%

Y611 S612 P613 S614 S615 P616 R617 Y618 T619 P620 Q621 S622 P623

4.2.20 Score per residue for model 20

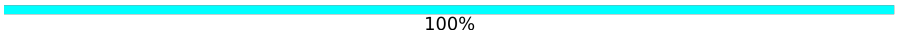
- Molecule 1: Tudor domain-containing protein 3

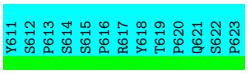
Chain A:



- Molecule 2: DNA-directed RNA polymerase II subunit RPB1

Chain B:





5 Refinement protocol and experimental data overview (i)

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

Of the 20 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
Amber	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	575
Number of shifts mapped to atoms	575
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	78%

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: DA2

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	0.83±0.02	0±0/416 (0.0± 0.0%)	1.35±0.06	2±1/567 (0.4± 0.2%)
All	All	0.83	0/8320 (0.0%)	1.35	41/11340 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.1±0.2
All	All	0	1

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	574	ARG	NE-CZ-NH1	14.41	127.51	120.30	7	16
1	A	597	TYR	CB-CG-CD1	-9.57	115.26	121.00	17	4
1	A	566	TYR	CB-CG-CD2	-7.22	116.67	121.00	14	2
1	A	574	ARG	NH1-CZ-NH2	-7.21	111.47	119.40	6	4
1	A	573	TYR	CB-CG-CD2	-6.92	116.85	121.00	8	2
1	A	597	TYR	CB-CG-CD2	6.85	125.11	121.00	17	2
1	A	608	ILE	CA-CB-CG2	6.12	123.15	110.90	1	1
1	A	574	ARG	NE-CZ-NH2	-6.10	117.25	120.30	7	2
1	A	566	TYR	CB-CG-CD1	-5.72	117.57	121.00	13	1
1	A	591	PHE	CB-CG-CD1	-5.67	116.83	120.80	20	1
1	A	597	TYR	CA-CB-CG	5.62	124.08	113.40	1	1
1	A	587	ALA	N-CA-CB	-5.44	102.48	110.10	19	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	594	TYR	CB-CG-CD2	-5.24	117.86	121.00	2	1
1	A	600	VAL	CG1-CB-CG2	-5.20	102.59	110.90	14	1
1	A	580	LEU	CB-CG-CD2	5.10	119.67	111.00	18	2

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	574	ARG	Sidechain	1

6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	403	384	384	5±2
2	B	0	0	0	0±0
All	All	8060	7680	7680	93

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 6.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:580:LEU:HD22	1:A:587:ALA:HB1	1.09	1.23	11	20
1:A:556:TRP:HB3	1:A:580:LEU:HD12	0.69	1.62	4	12
1:A:579:ALA:C	1:A:580:LEU:HD23	0.68	2.09	20	16
1:A:556:TRP:CB	1:A:580:LEU:HD12	0.62	2.24	4	2
1:A:572:PHE:CG	1:A:606:LYS:HE2	0.60	2.31	12	2
1:A:587:ALA:HB2	1:A:602:LEU:HD23	0.59	1.73	16	4
1:A:580:LEU:HD22	1:A:587:ALA:CB	0.56	2.27	6	2
1:A:556:TRP:CG	1:A:580:LEU:HD12	0.51	2.40	7	2
1:A:556:TRP:CH2	1:A:562:CYS:SG	0.51	3.03	10	2
1:A:580:LEU:CD2	1:A:587:ALA:HB1	0.50	2.26	5	3
1:A:587:ALA:HB2	1:A:602:LEU:CD2	0.46	2.40	1	3
1:A:577:VAL:HG13	1:A:580:LEU:HG	0.45	1.88	2	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:558:PRO:HA	1:A:577:VAL:HG12	0.45	1.87	4	6
1:A:580:LEU:HD13	1:A:587:ALA:HB1	0.44	1.89	17	2
1:A:587:ALA:HB3	1:A:605:ILE:CD1	0.44	2.43	6	1
1:A:577:VAL:HG13	1:A:580:LEU:CD2	0.44	2.43	14	4
1:A:580:LEU:HD23	1:A:580:LEU:N	0.43	2.29	13	3
1:A:557:LYS:O	1:A:577:VAL:HG11	0.42	2.14	14	2
1:A:597:TYR:CD1	1:A:597:TYR:N	0.41	2.88	10	3
1:A:577:VAL:HG22	1:A:580:LEU:HD11	0.40	1.91	2	1
1:A:577:VAL:HG13	1:A:580:LEU:HD21	0.40	1.92	14	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	48/58 (83%)	44±1 (91±3%)	4±1 (8±3%)	0±0 (1±1%)	29	74
2	B	0	-	-	-	-	
All	All	960/1420 (68%)	877 (91%)	77 (8%)	6 (1%)	29	74

All 5 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	577	VAL	2
1	A	603	SER	1
1	A	556	TRP	1
1	A	595	GLY	1
1	A	591	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	42/51 (82%)	37±1 (88±3%)	5±1 (12±3%)	8 51
2	B	0	-	-	-
All	All	840/1260 (67%)	738 (88%)	102 (12%)	8 51

All 20 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	608	ILE	19
1	A	597	TYR	15
1	A	592	ILE	13
1	A	601	LEU	9
1	A	580	LEU	7
1	A	556	TRP	6
1	A	560	ASP	6
1	A	561	GLU	3
1	A	568	GLU	3
1	A	566	TYR	3
1	A	563	PHE	3
1	A	599	GLU	3
1	A	574	ARG	2
1	A	571	LYS	2
1	A	594	TYR	2
1	A	596	ASN	2
1	A	606	LYS	1
1	A	593	ASP	1
1	A	603	SER	1
1	A	557	LYS	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)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is

considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	DA2	B	617	2	10,12,13	0.73±0.17	0±0 (1±3%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	DA2	B	617	2	7,14,16	1.02±0.18	0±0 (3±6%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	DA2	B	617	2	-	0±0,12,13,15	-

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	B	617	DA2	CB-CA	2.57	1.57	1.53	12	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	617	DA2	CB-CG-CD	2.75	103.80	112.05	10	1
2	B	617	DA2	C1-NH1-CZ	2.37	125.51	120.55	16	1
2	B	617	DA2	C2-NH1-CZ	2.16	125.06	120.55	8	3

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

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 78% for the well-defined parts and 61% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_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	575
Number of shifts mapped to atoms	575
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	51	3.13 ± 0.25	Should be applied
$^{13}C_{\beta}$	47	2.43 ± 0.27	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	51	1.04 ± 1.01	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 78%, i.e. 528 atoms were assigned a chemical shift out of a possible 677. 0 out of 8 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	183/238 (77%)	93/96 (97%)	45/96 (47%)	45/46 (98%)
Sidechain	270/349 (77%)	187/226 (83%)	83/113 (73%)	0/10 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	75/90 (83%)	38/43 (88%)	35/45 (78%)	2/2 (100%)
Overall	528/677 (78%)	318/365 (87%)	163/254 (64%)	47/58 (81%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 61%, i.e. 575 atoms were assigned a chemical shift out of a possible 947. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	208/341 (61%)	106/137 (77%)	51/140 (36%)	51/64 (80%)
Sidechain	292/491 (59%)	202/320 (63%)	90/159 (57%)	0/12 (0%)
Aromatic	75/115 (65%)	38/55 (69%)	35/57 (61%)	2/3 (67%)
Overall	575/947 (61%)	346/512 (68%)	176/356 (49%)	53/79 (67%)

7.1.4 Statistically unusual chemical shifts (i)

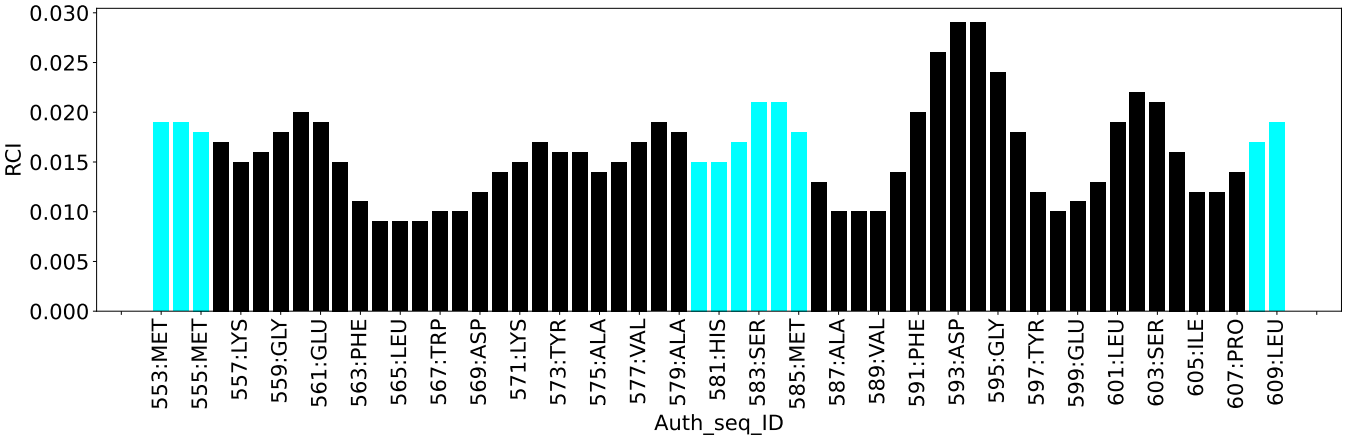
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	568	GLU	HB3	-0.30	0.95 – 3.05	-11.0
1	A	578	GLU	CG	26.18	30.20 – 42.01	-8.4
1	A	568	GLU	H	3.79	5.45 – 11.20	-7.9
1	A	568	GLU	HB2	0.85	1.00 – 3.05	-5.7
1	A	574	ARG	CG	33.52	21.24 – 33.19	5.3
1	A	590	LYS	HG2	0.13	0.13 – 2.61	-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:


0.030
0.025
0.020
0.015
0.010
0.005
0.000
RCI
553:MET
555:MET
557:LYS
559:GLY
561:GLU
563:PHE
565:LEU
567:TRP
569:ASP
571:LYS
573:TYR
575:ALA
577:VAL
579:ALA
581:HIS
583:SER
585:MET
587:ALA
589:VAL
591:PHE
593:ASP
595:GLY
597:TYR
599:GLU
601:LEU
603:SER
605:ILE
607:PRO
609:LEU
Auth_seq_ID