



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 01:02 am BST

PDB ID : 2Q4L  
Title : Ensemble refinement of the crystal structure of GALT-like protein from *Arabidopsis thaliana* At5g18200  
Authors : Levin, E.J.; Kondrashov, D.A.; Wesenberg, G.E.; Phillips Jr., G.N.; Center for Eukaryotic Structural Genomics (CESG)  
Deposited on : 2007-05-31  
Resolution : 2.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

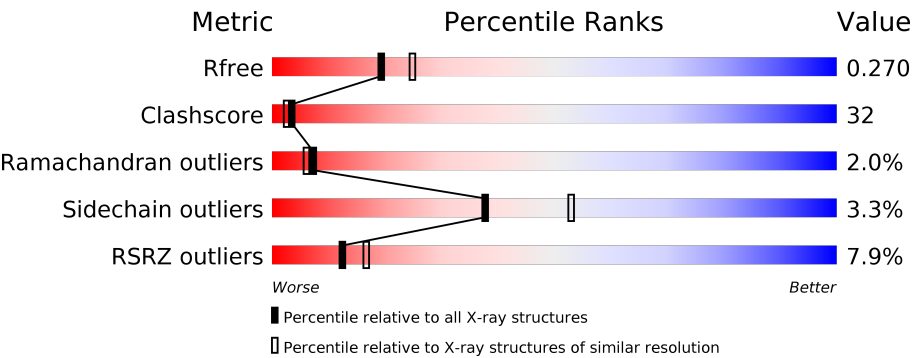
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 Å.

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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	1-A	351	<div><div>7%</div><div><div></div><div>37%</div><div>46%</div><div>•</div><div>14%</div></div></div>
1	1-B	351	<div><div>7%</div><div><div></div><div>41%</div><div>42%</div><div>•</div><div>15%</div></div></div>
1	2-A	351	<div><div>7%</div><div><div></div><div>42%</div><div>40%</div><div>•</div><div>14%</div></div></div>
1	2-B	351	<div><div>7%</div><div><div></div><div>35%</div><div>48%</div><div>•</div><div>15%</div></div></div>
1	3-A	351	<div><div>7%</div><div><div></div><div>39%</div><div>42%</div><div>•</div><div>14%</div></div></div>
1	3-B	351	<div><div>7%</div><div><div></div><div>37%</div><div>45%</div><div>•</div><div>15%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	4-A	351	
1	4-B	351	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 19940 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Probable galactose-1-phosphate uridyl transferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1-A	301	Total	C	N	O	S	0	0	0
			2388	1536	403	436	13			
1	2-A	301	Total	C	N	O	S	0	0	0
			2388	1536	403	436	13			
1	3-A	301	Total	C	N	O	S	0	0	0
			2388	1536	403	436	13			
1	4-A	301	Total	C	N	O	S	0	0	0
			2388	1536	403	436	13			
1	1-B	299	Total	C	N	O	S	0	0	0
			2368	1524	397	434	13			
1	2-B	299	Total	C	N	O	S	0	0	0
			2368	1524	397	434	13			
1	3-B	299	Total	C	N	O	S	0	0	0
			2368	1524	397	434	13			
1	4-B	299	Total	C	N	O	S	0	0	0
			2368	1524	397	434	13			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	3-A	2	Total	Zn	0	0
			2	2		
2	3-B	2	Total	Zn	0	0
			2	2		
2	4-A	2	Total	Zn	0	0
			2	2		
2	1-B	2	Total	Zn	0	0
			2	2		
2	4-B	2	Total	Zn	0	0
			2	2		
2	1-A	2	Total	Zn	0	0
			2	2		

*Continued on next page...*

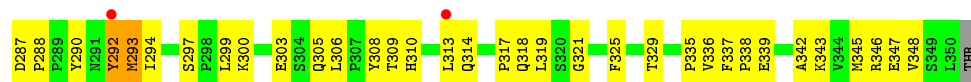
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	2-B	2	Total 2	Zn 2	0	0
2	2-A	2	Total 2	Zn 2	0	0

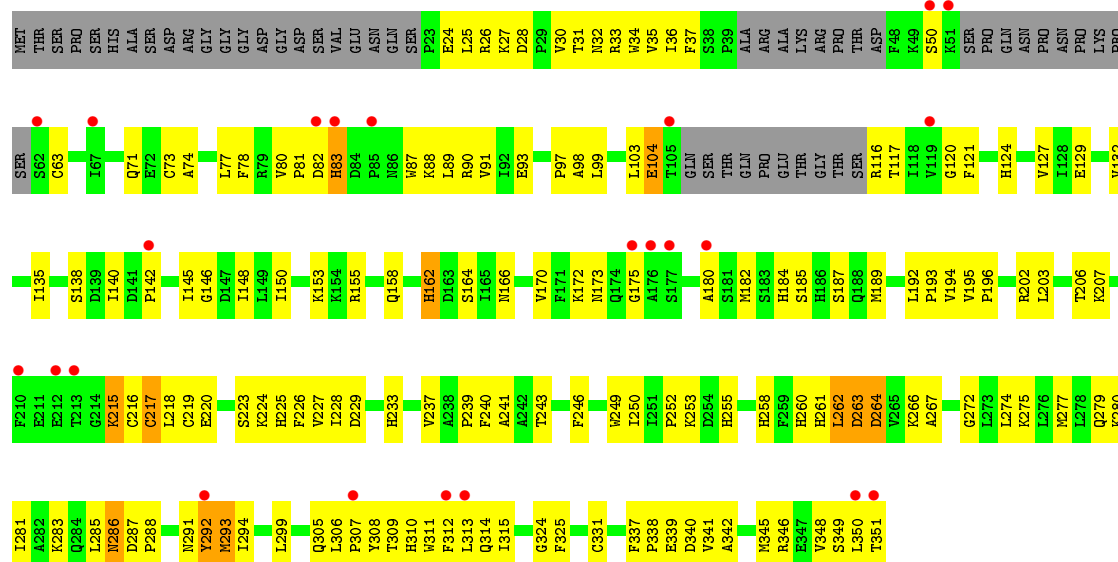
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	1-A	109	Total 109	O 109	0	0
3	2-A	114	Total 114	O 114	0	0
3	3-A	112	Total 112	O 112	0	0
3	4-A	114	Total 114	O 114	0	0
3	1-B	116	Total 116	O 116	0	0
3	2-B	111	Total 111	O 111	0	0
3	3-B	113	Total 113	O 113	0	0
3	4-B	111	Total 111	O 111	0	0

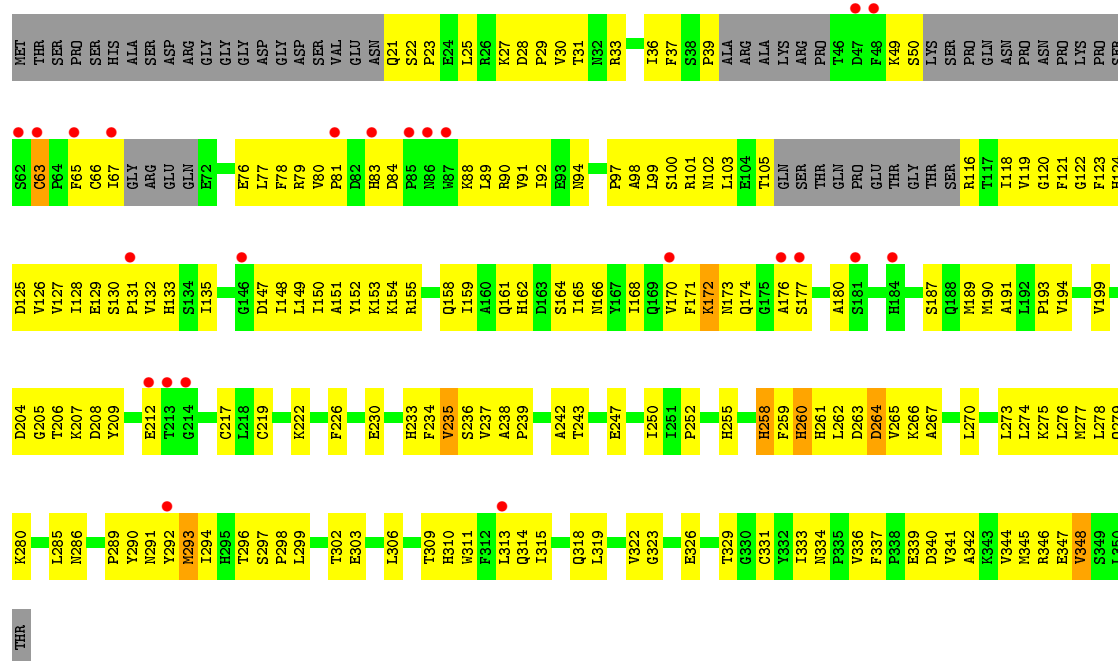




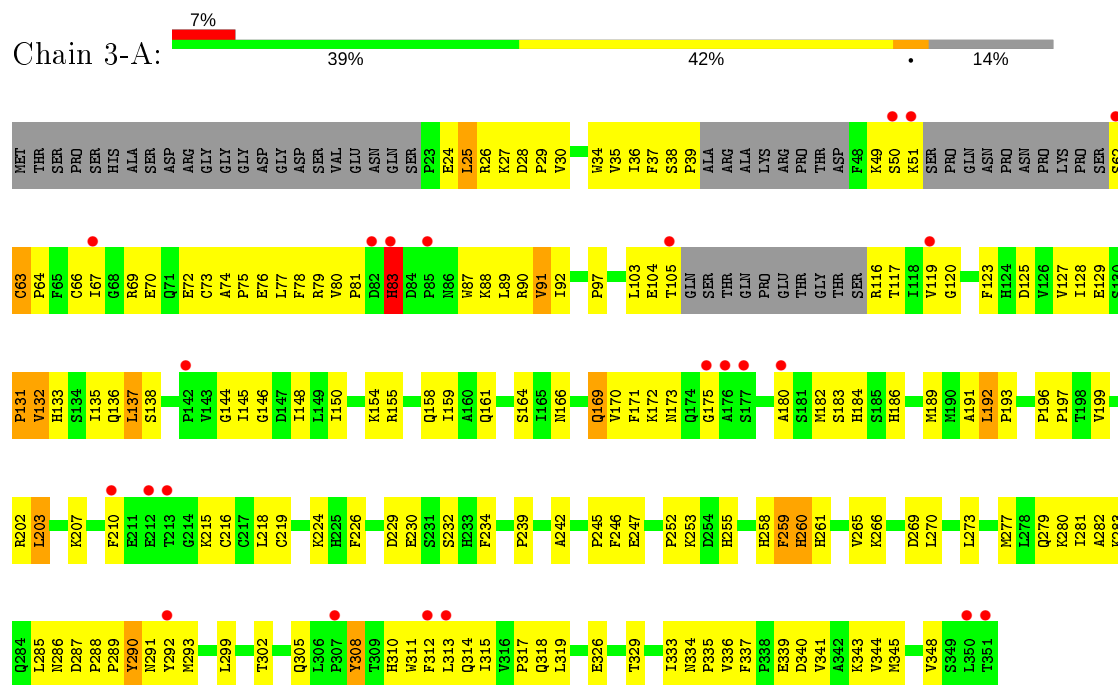
- Molecule 1: Probable galactose-1-phosphate uridyl transferase



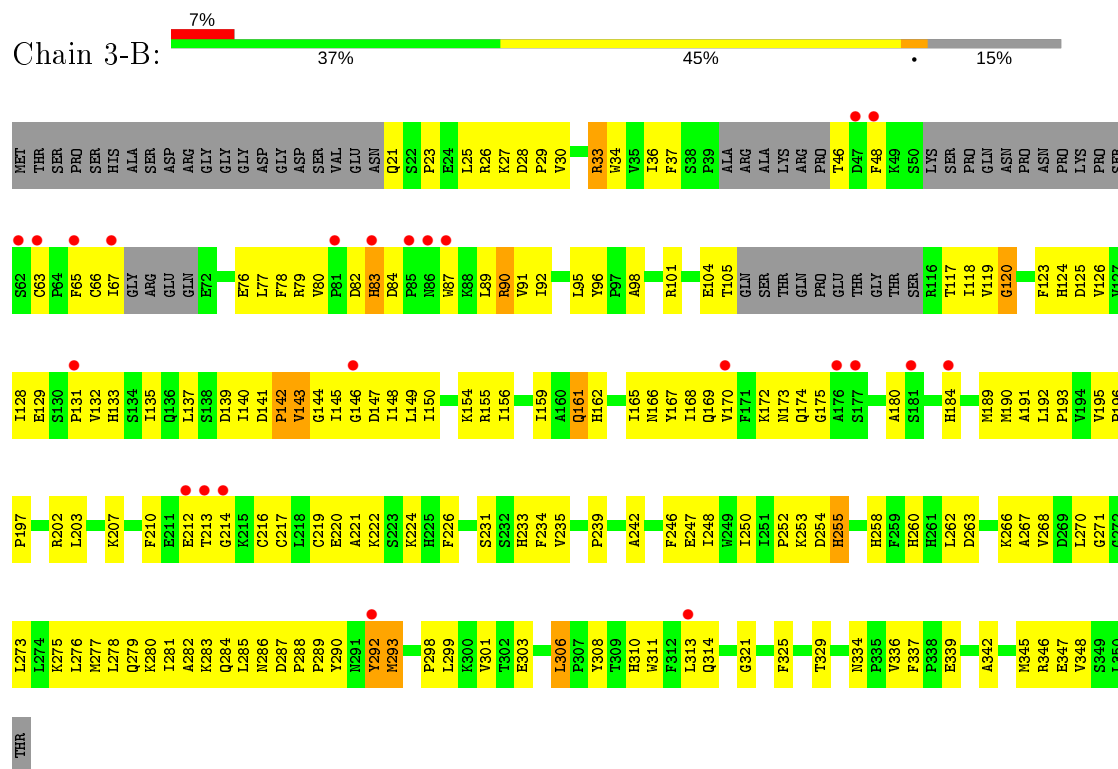
- Molecule 1: Probable galactose-1-phosphate uridyl transferase



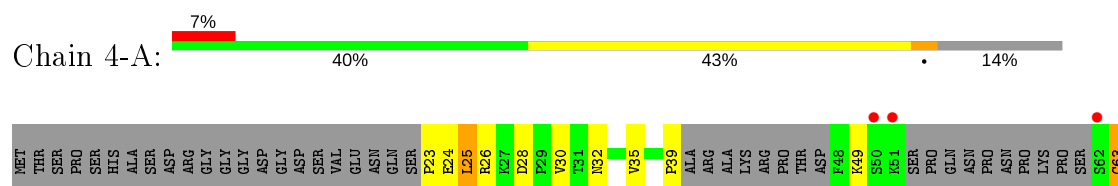
- Molecule 1: Probable galactose-1-phosphate uridyl transferase



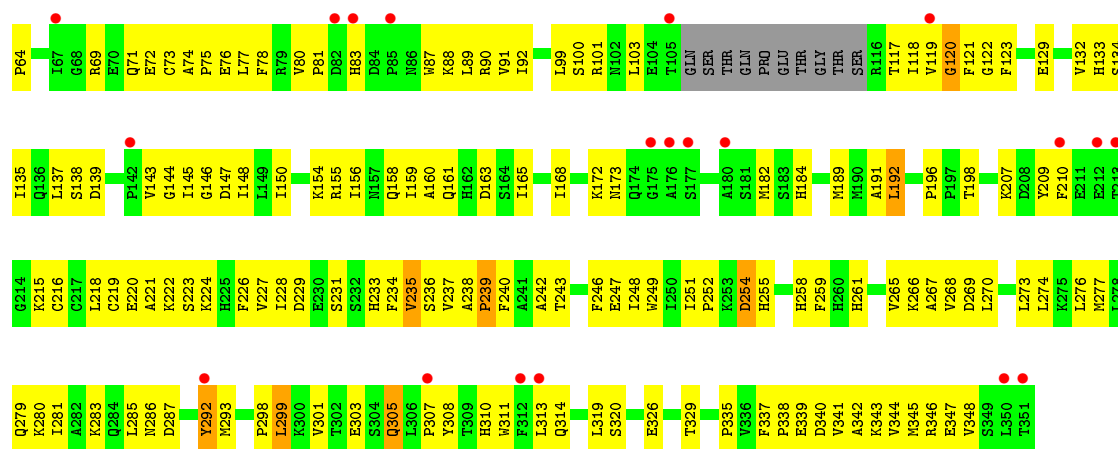
- Molecule 1: Probable galactose-1-phosphate uridyl transferase



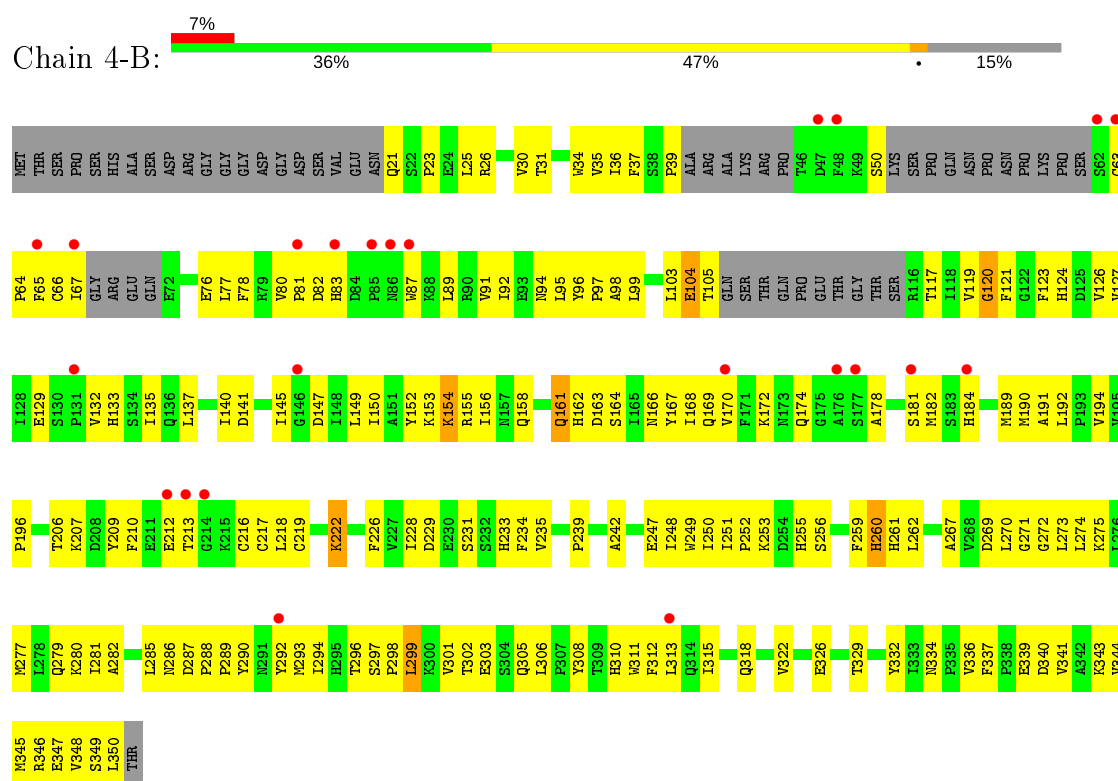
- Molecule 1: Probable galactose-1-phosphate uridyl transferase







- Molecule 1: Probable galactose-1-phosphate uridyl transferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.98Å 95.53Å 110.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.26 – 2.30 27.59 – 2.30	Depositor EDS
% Data completeness (in resolution range)	89.8 (27.26-2.30) 89.9 (27.59-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.79 (at 2.31Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.190 , 0.270 0.199 , 0.270	Depositor DCC
$R_{free}$ test set	2581 reflections (9.37%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.9	Xtriage
Anisotropy	0.447	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 55.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	19940	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 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 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	1-A	0.35	0/2454	0.59	0/3331
1	1-B	0.35	0/2433	0.61	0/3306
1	2-A	0.35	0/2454	0.59	0/3331
1	2-B	0.35	0/2433	0.59	1/3306 (0.0%)
1	3-A	0.35	0/2454	0.60	0/3331
1	3-B	0.35	0/2433	0.59	0/3306
1	4-A	0.35	0/2454	0.61	0/3331
1	4-B	0.35	0/2433	0.61	0/3306
All	All	0.35	0/19548	0.60	1/26548 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	2-B	238	ALA	N-CA-C	-5.02	97.46	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	1-A	2388	0	2353	159	0
1	1-B	2368	0	2325	157	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	2-A	2388	0	2355	136	0
1	2-B	2368	0	2326	173	0
1	3-A	2388	0	2353	150	0
1	3-B	2368	0	2325	174	0
1	4-A	2388	0	2354	140	0
1	4-B	2368	0	2325	170	0
2	1-A	2	0	0	0	0
2	1-B	2	0	0	0	0
2	2-A	2	0	0	0	0
2	2-B	2	0	0	0	0
2	3-A	2	0	0	0	0
2	3-B	2	0	0	0	0
2	4-A	2	0	0	0	0
2	4-B	2	0	0	0	0
3	1-A	109	0	0	6	0
3	1-B	116	0	0	7	0
3	2-A	114	0	0	9	0
3	2-B	111	0	0	10	0
3	3-A	112	0	0	6	0
3	3-B	113	0	0	10	0
3	4-A	114	0	0	7	0
3	4-B	111	0	0	11	0
All	All	19940	0	18716	1215	0

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 32.

The worst 5 of 1215 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:105:THR:HG23	1:B:119:VAL:HG21	1.34	1.07
1:B:239:PRO:HG2	1:B:242:ALA:HB2	1.34	1.06
1:B:239:PRO:HG2	1:B:242:ALA:HB2	1.36	1.04
1:A:172:LYS:HB3	1:A:292:TYR:HB3	1.42	1.01
1:A:81:PRO:HD3	1:A:88:LYS:HD3	1.47	0.96

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 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 X-ray entries followed by that with respect to entries of similar resolution.

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	1-A	293/351 (84%)	253 (86%)	33 (11%)	7 (2%)	6	4
1	1-B	289/351 (82%)	255 (88%)	28 (10%)	6 (2%)	7	5
1	2-A	293/351 (84%)	262 (89%)	26 (9%)	5 (2%)	9	8
1	2-B	289/351 (82%)	249 (86%)	36 (12%)	4 (1%)	11	11
1	3-A	293/351 (84%)	262 (89%)	26 (9%)	5 (2%)	9	8
1	3-B	289/351 (82%)	256 (89%)	25 (9%)	8 (3%)	5	3
1	4-A	293/351 (84%)	254 (87%)	32 (11%)	7 (2%)	6	4
1	4-B	289/351 (82%)	255 (88%)	30 (10%)	4 (1%)	11	11
All	All	2328/2808 (83%)	2046 (88%)	236 (10%)	46 (2%)	7	6

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1-B	83	HIS
1	1-B	104	GLU
1	3-A	137	LEU
1	3-B	142	PRO
1	1-A	50	SER

### 5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	1-A	270/312 (86%)	265 (98%)	5 (2%)	57	73

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	1-B	269/312 (86%)	262 (97%)	7 (3%)	46	63
1	2-A	270/312 (86%)	256 (95%)	14 (5%)	23	32
1	2-B	269/312 (86%)	262 (97%)	7 (3%)	46	63
1	3-A	270/312 (86%)	256 (95%)	14 (5%)	23	32
1	3-B	269/312 (86%)	261 (97%)	8 (3%)	41	57
1	4-A	270/312 (86%)	260 (96%)	10 (4%)	34	48
1	4-B	269/312 (86%)	263 (98%)	6 (2%)	52	69
All	All	2156/2496 (86%)	2085 (97%)	71 (3%)	38	53

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	2-B	348	VAL
1	3-A	203	LEU
1	4-B	141	ASP
1	3-A	25	LEU
1	3-A	91	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	3-A	169	GLN
1	3-B	21	GLN
1	4-B	260	HIS
1	3-A	260	HIS
1	1-B	174	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 16 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	1-A	301/351 (85%)	0.31	23 (7%) 13 18	21, 40, 74, 93	301 (100%)
1	1-B	299/351 (85%)	0.27	23 (7%) 13 17	21, 38, 74, 84	299 (100%)
1	2-A	301/351 (85%)	0.31	23 (7%) 13 18	21, 40, 74, 93	301 (100%)
1	2-B	299/351 (85%)	0.27	23 (7%) 13 17	21, 38, 74, 84	299 (100%)
1	3-A	301/351 (85%)	0.31	23 (7%) 13 18	21, 40, 74, 93	301 (100%)
1	3-B	299/351 (85%)	0.27	23 (7%) 13 17	21, 38, 74, 84	299 (100%)
1	4-A	301/351 (85%)	0.31	23 (7%) 13 18	21, 40, 74, 93	301 (100%)
1	4-B	299/351 (85%)	0.27	23 (7%) 13 17	21, 38, 74, 84	299 (100%)
All	All	2400/2808 (85%)	0.29	184 (7%) 12 17	21, 39, 74, 93	2400 (100%)

The worst 5 of 184 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	1-A	351	THR	7.1
1	2-A	351	THR	7.1
1	3-A	351	THR	7.1
1	4-A	351	THR	7.1
1	1-A	350	LEU	5.3

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.



## 6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	ZN	4-B	402	1/1	0.98	0.06	60,60,60,60	1
2	ZN	1-B	402	1/1	0.98	0.06	60,60,60,60	1
2	ZN	2-B	402	1/1	0.98	0.06	55,55,55,55	1
2	ZN	3-B	402	1/1	0.98	0.06	53,53,53,53	1
2	ZN	4-A	400	1/1	0.99	0.04	49,49,49,49	1
2	ZN	2-B	403	1/1	0.99	0.02	26,26,26,26	1
2	ZN	4-B	403	1/1	0.99	0.02	35,35,35,35	1
2	ZN	1-B	403	1/1	0.99	0.02	36,36,36,36	1
2	ZN	3-A	400	1/1	0.99	0.04	55,55,55,55	1
2	ZN	1-A	400	1/1	0.99	0.04	53,53,53,53	1
2	ZN	2-A	400	1/1	0.99	0.04	52,52,52,52	1
2	ZN	1-A	401	1/1	0.99	0.08	42,42,42,42	1
2	ZN	4-A	401	1/1	0.99	0.08	32,32,32,32	1
2	ZN	3-A	401	1/1	0.99	0.08	40,40,40,40	1
2	ZN	2-A	401	1/1	0.99	0.08	25,25,25,25	1
2	ZN	3-B	403	1/1	0.99	0.02	32,32,32,32	1

## 6.5 Other polymers

There are no such residues in this entry.