



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2023 – 05:54 PM EDT

PDB ID : 2Q4J  
Title : Ensemble refinement of the protein crystal structure of gene product from Arabidopsis thaliana At3g03250, a putative UDP-glucose pyrophosphorylase  
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 : 1.86 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

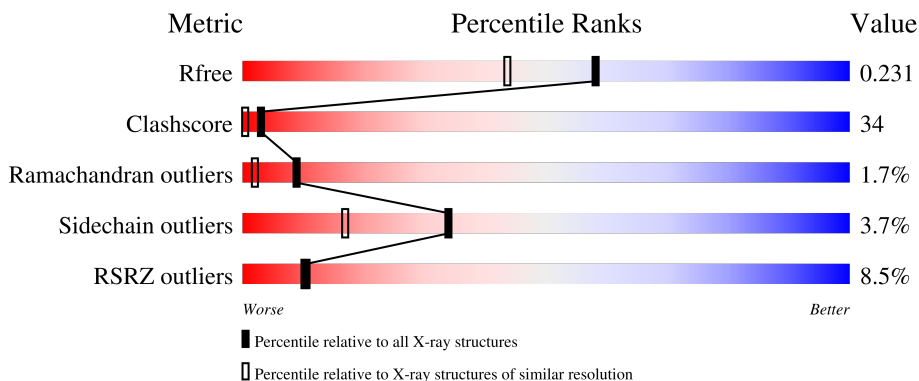
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.86 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 of 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	469	
1	1-B	469	
1	2-A	469	
1	2-B	469	
1	3-A	469	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	3-B	469	
1	4-A	469	
1	4-B	469	
1	5-A	469	
1	5-B	469	
1	6-A	469	
1	6-B	469	
1	7-A	469	
1	7-B	469	
1	8-A	469	
1	8-B	469	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 60344 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 UTP-glucose-1-phosphate uridylyltransferase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	1-A	460	3579	2291	591	687	10	0	0	0
1	2-A	460	3579	2291	591	687	10	0	0	0
1	3-A	460	3579	2291	591	687	10	0	0	0
1	4-A	460	3579	2291	591	687	10	0	0	0
1	5-A	460	3579	2291	591	687	10	0	0	0
1	6-A	460	3579	2291	591	687	10	0	0	0
1	7-A	460	3579	2291	591	687	10	0	0	0
1	8-A	460	3579	2291	591	687	10	0	0	0
1	1-B	455	3543	2271	585	677	10	0	0	0
1	2-B	455	3543	2271	585	677	10	0	0	0
1	3-B	455	3543	2271	585	677	10	0	0	0
1	4-B	455	3543	2271	585	677	10	0	0	0
1	5-B	455	3543	2271	585	677	10	0	0	0
1	6-B	455	3543	2271	585	677	10	0	0	0
1	7-B	455	3543	2271	585	677	10	0	0	0
1	8-B	455	3543	2271	585	677	10	0	0	0

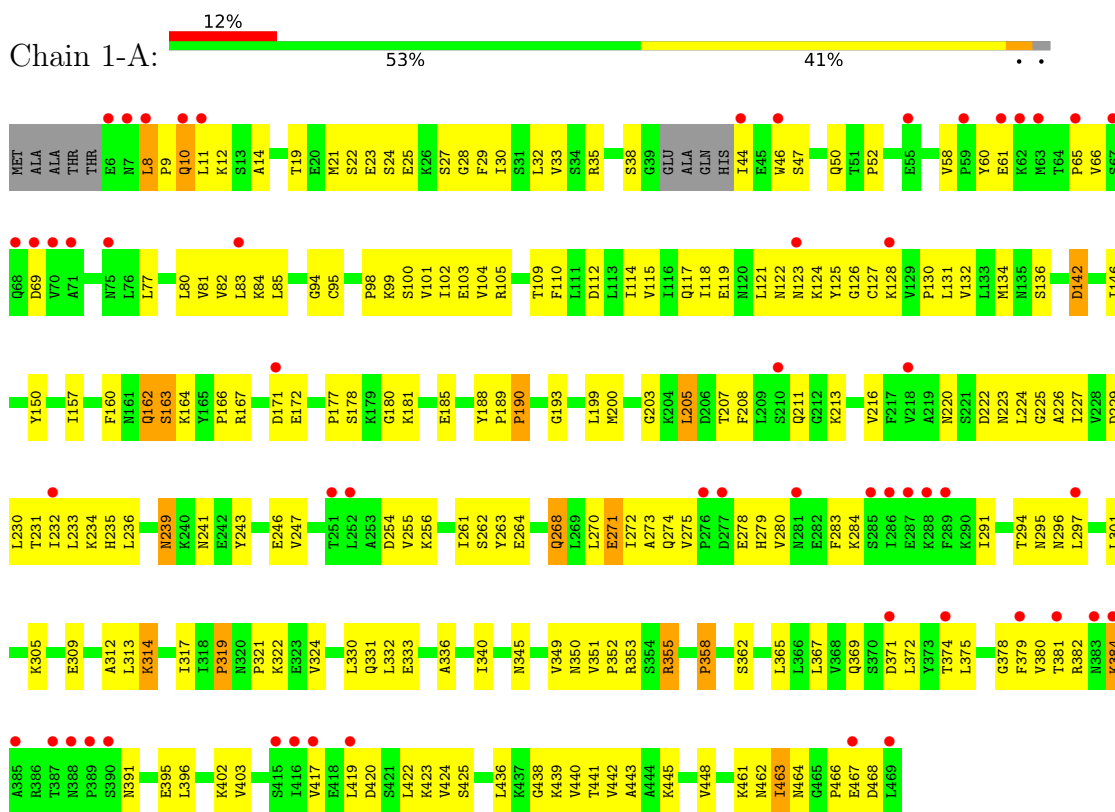
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	1-A	180	Total 180	O 180	0	0
2	2-A	182	Total 182	O 182	0	0
2	3-A	182	Total 182	O 182	0	0
2	4-A	183	Total 183	O 183	0	0
2	5-A	182	Total 182	O 182	0	0
2	6-A	182	Total 182	O 182	0	0
2	7-A	183	Total 183	O 183	0	0
2	8-A	183	Total 183	O 183	0	0
2	1-B	241	Total 241	O 241	0	0
2	2-B	239	Total 239	O 239	0	0
2	3-B	239	Total 239	O 239	0	0
2	4-B	238	Total 238	O 238	0	0
2	5-B	239	Total 239	O 239	0	0
2	6-B	239	Total 239	O 239	0	0
2	7-B	238	Total 238	O 238	0	0
2	8-B	238	Total 238	O 238	0	0

### 3 Residue-property plots

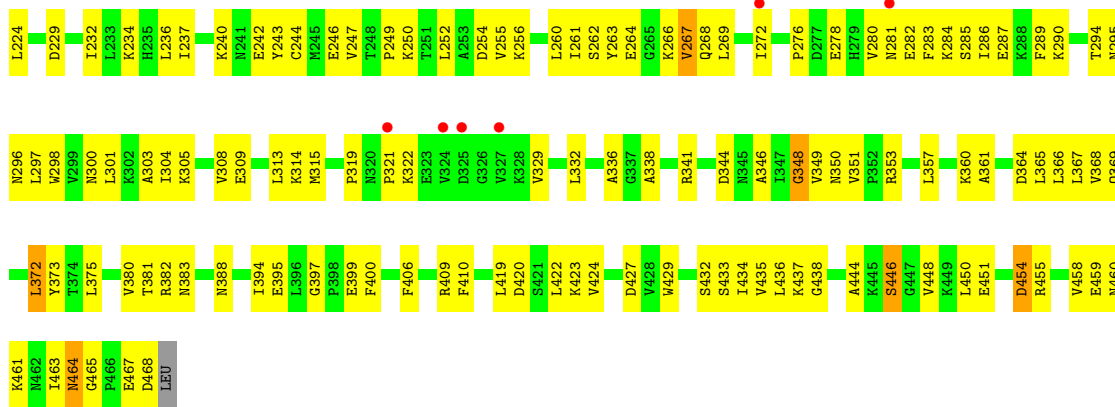
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2

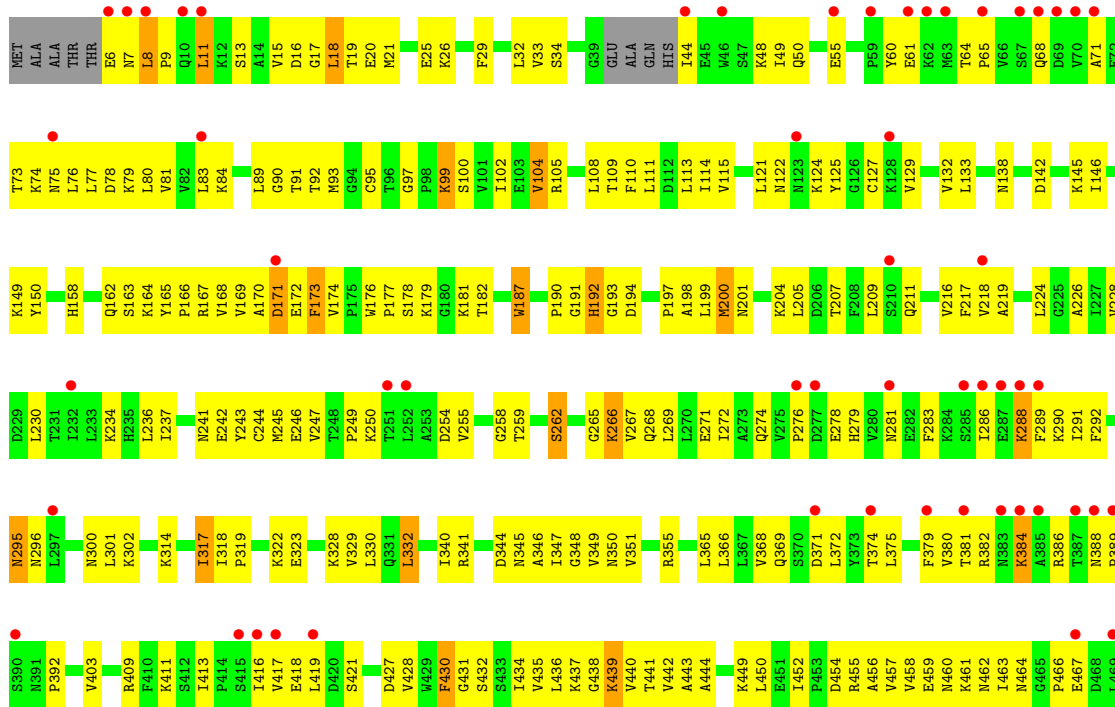


- Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2

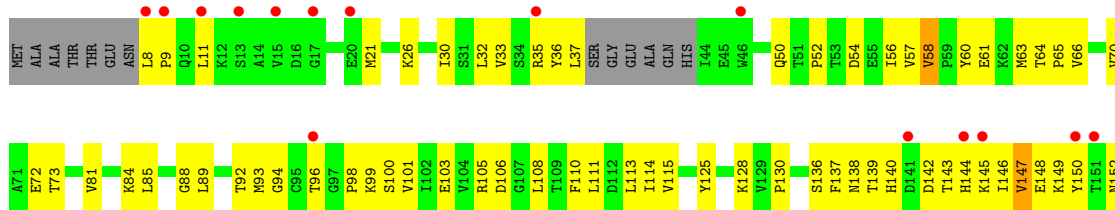


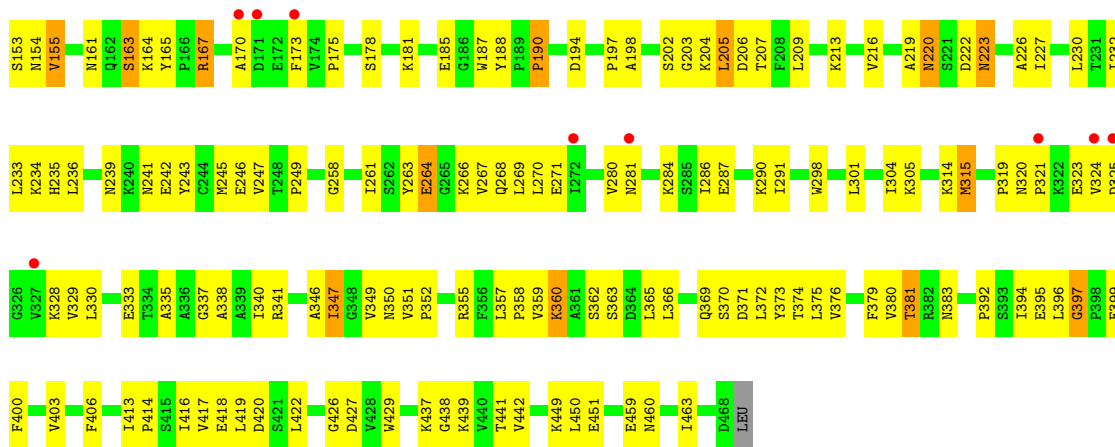


Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2

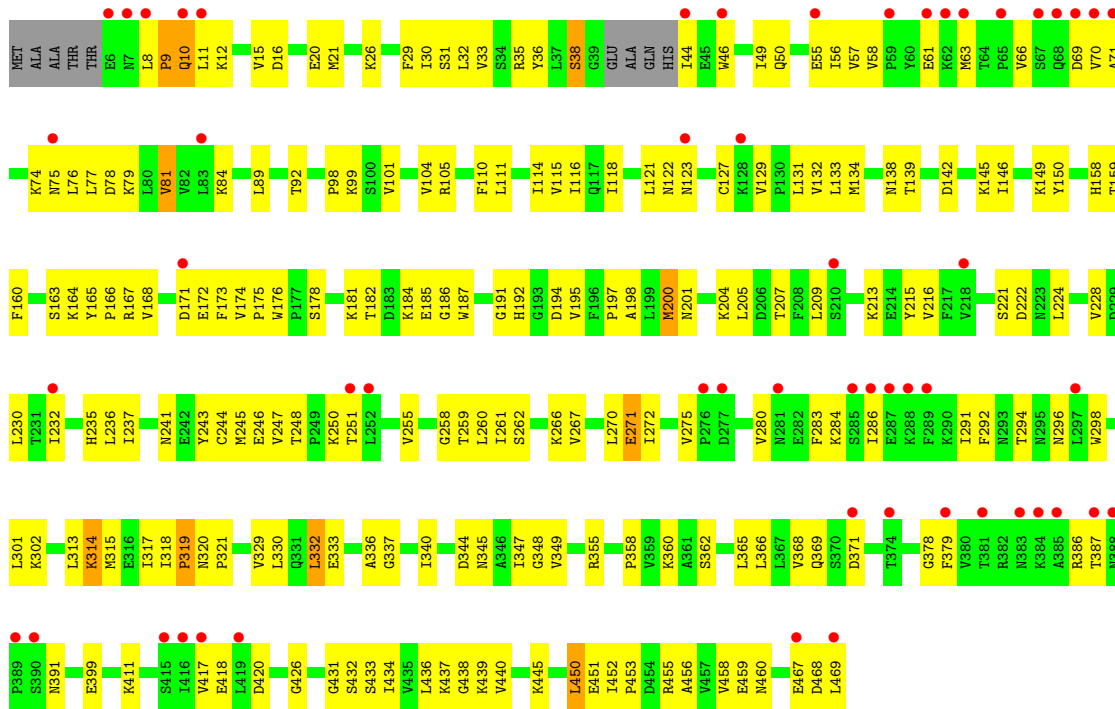


Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2

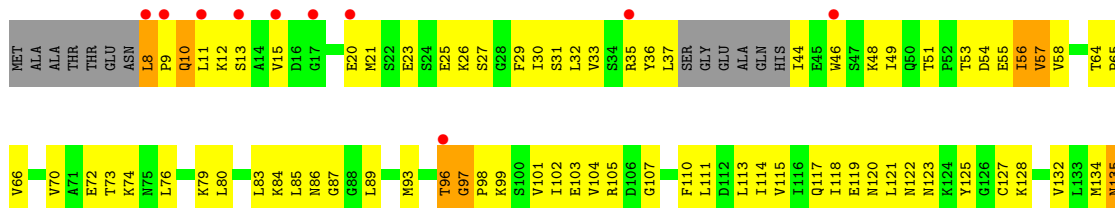




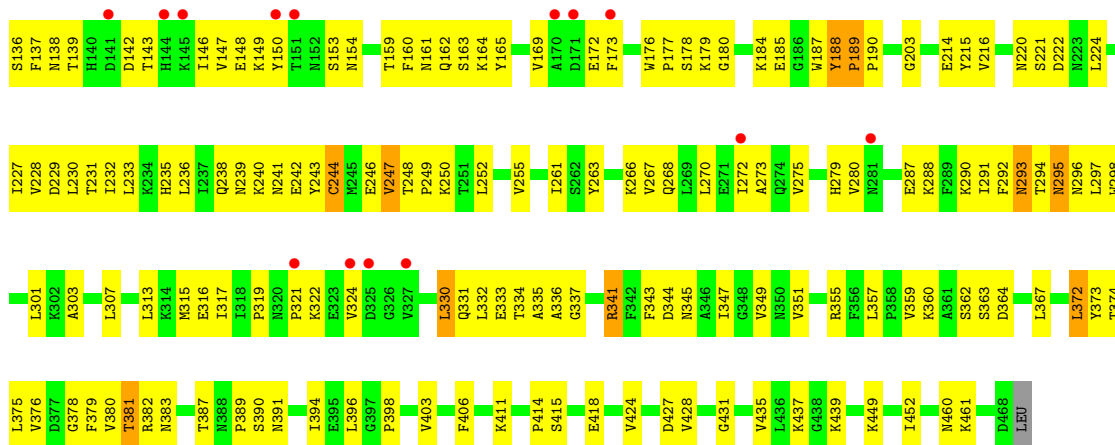
• Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2



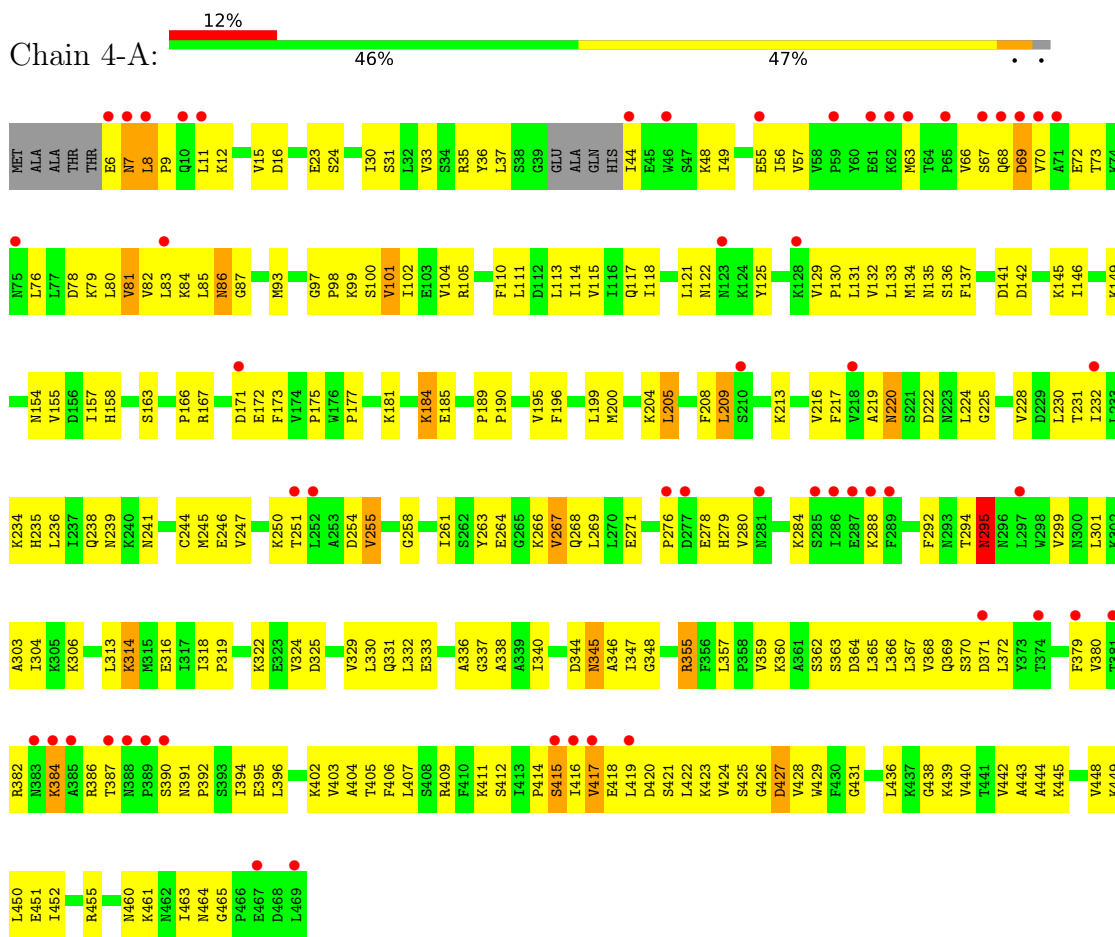
• Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2



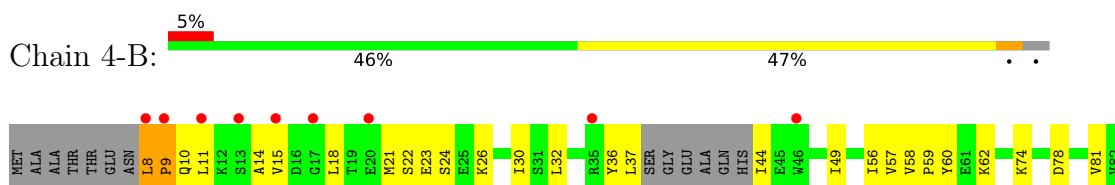


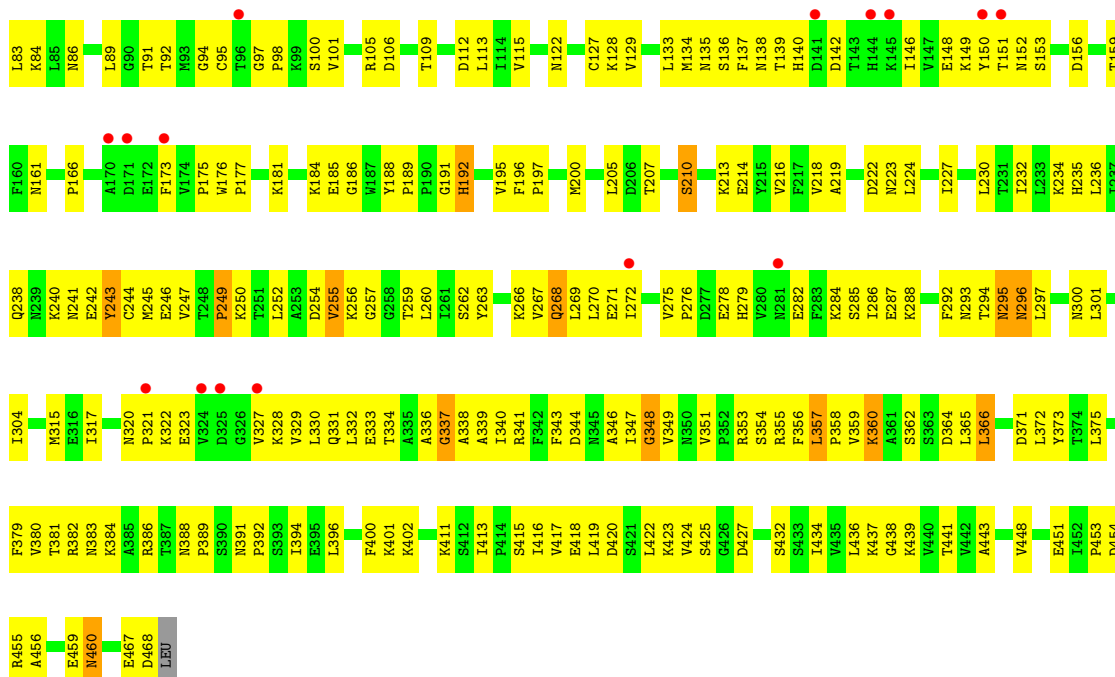


• Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2

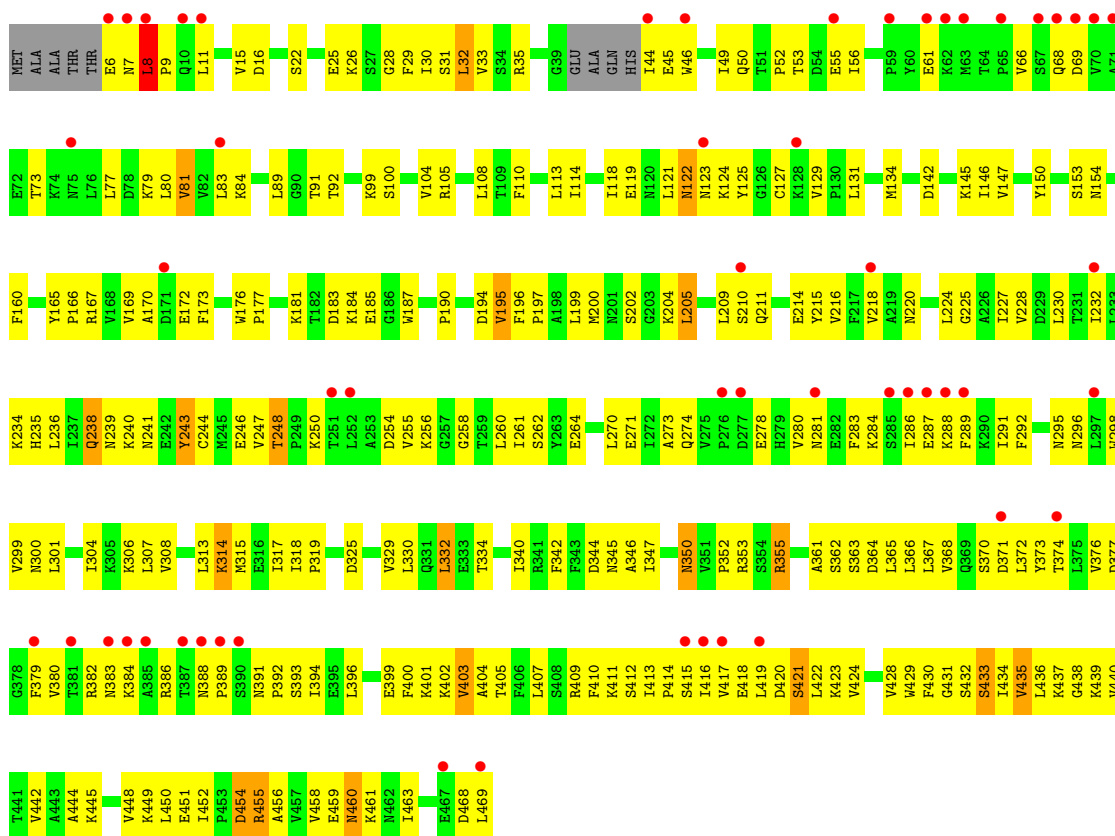


• Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2

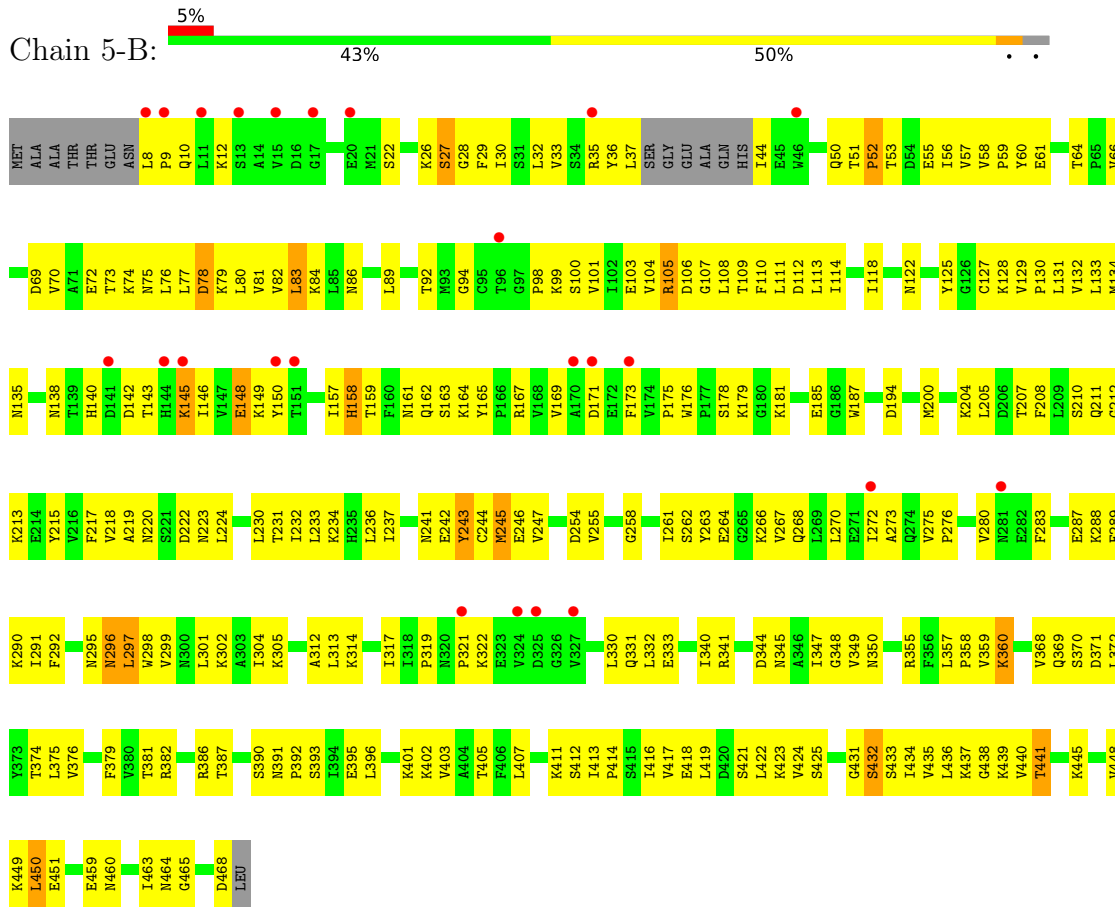




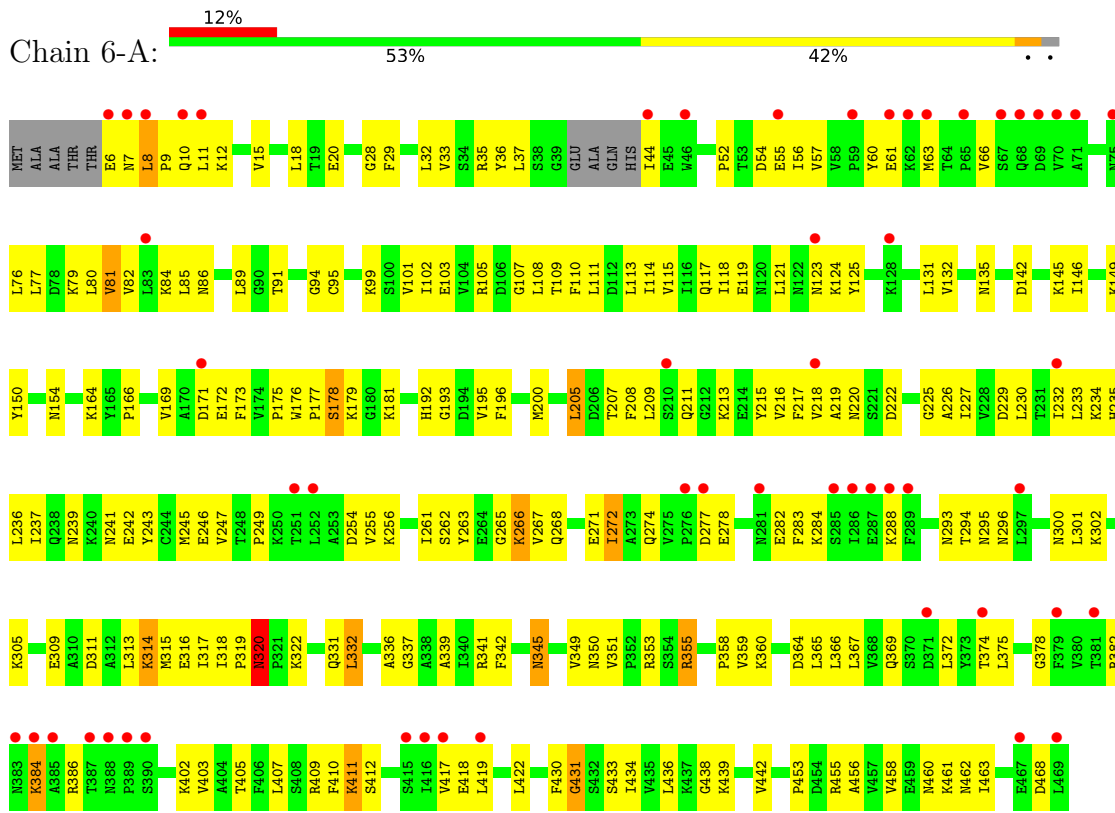
• Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2



• Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2

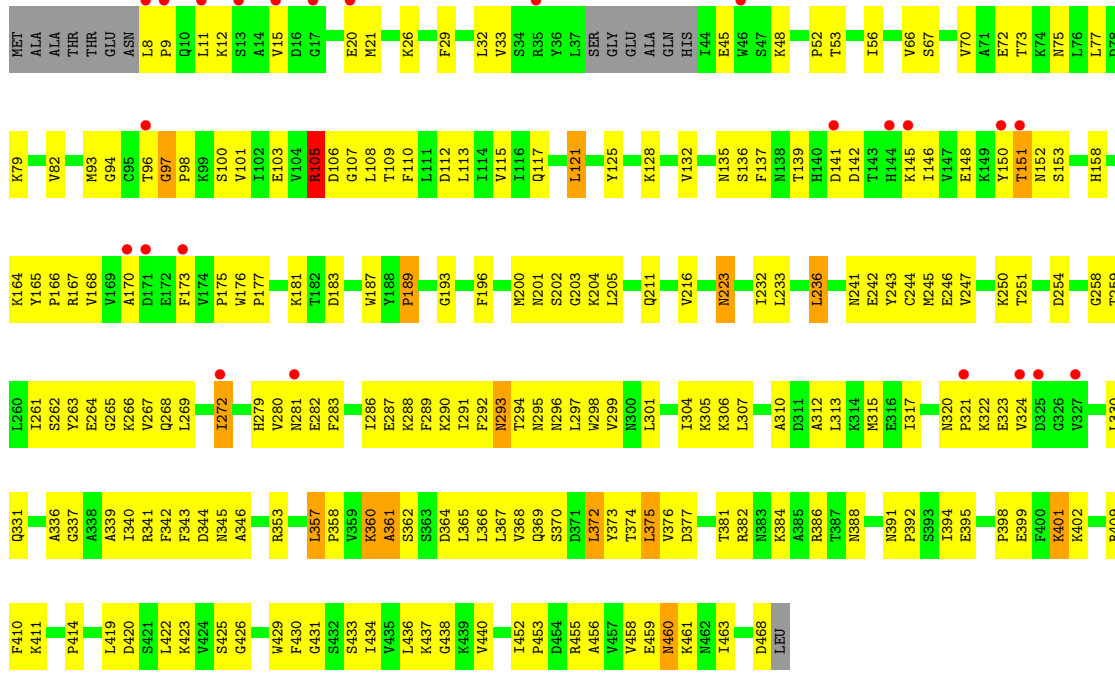


• Molecule 1: Probable UTP-glucose-1-phosphate uridylyltransferase 2









## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	188.68Å 58.86Å 89.86Å 90.00° 100.40° 90.00°	Depositor
Resolution (Å)	37.37 – 1.86 37.37 – 1.86	Depositor EDS
% Data completeness (in resolution range)	97.6 (37.37-1.86) 97.7 (37.37-1.86)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.11 (at 1.87Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.171 , 0.231 0.172 , 0.231	Depositor DCC
$R_{free}$ test set	3986 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.4	Xtrriage
Anisotropy	0.041	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 66.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	60344	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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

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	0.48	0/3646	0.67	1/4942 (0.0%)
1	1-B	0.56	0/3610	0.73	2/4895 (0.0%)
1	2-A	0.49	0/3646	0.66	0/4942
1	2-B	0.55	1/3610 (0.0%)	0.71	1/4895 (0.0%)
1	3-A	0.47	0/3646	0.66	0/4942
1	3-B	0.55	0/3610	0.71	2/4895 (0.0%)
1	4-A	0.49	0/3646	0.66	0/4942
1	4-B	0.56	0/3610	0.73	2/4895 (0.0%)
1	5-A	0.50	0/3646	0.72	0/4942
1	5-B	0.60	0/3610	0.75	1/4895 (0.0%)
1	6-A	0.50	0/3646	0.68	0/4942
1	6-B	0.60	0/3610	0.76	2/4895 (0.0%)
1	7-A	0.51	0/3646	0.70	0/4942
1	7-B	0.57	0/3610	0.75	3/4895 (0.1%)
1	8-A	0.51	0/3646	0.69	0/4942
1	8-B	0.59	0/3610	0.77	4/4895 (0.1%)
All	All	0.53	1/58048 (0.0%)	0.71	18/78696 (0.0%)

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 outliers	#Planarity outliers
1	3-B	0	1
1	4-B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	2-B	242	GLU	CG-CD	5.45	1.60	1.51



The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	1-B	348	GLY	N-CA-C	-6.06	97.95	113.10
1	4-B	348	GLY	N-CA-C	-5.87	98.43	113.10
1	4-B	296	ASN	N-CA-C	-5.74	95.52	111.00
1	6-B	105	ARG	NE-CZ-NH2	-5.73	117.43	120.30
1	7-B	290	LYS	N-CA-C	5.70	126.39	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	3-B	125	TYR	Sidechain
1	4-B	243	TYR	Sidechain

## 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	3579	0	3652	229	0
1	1-B	3543	0	3621	255	0
1	2-A	3579	0	3652	254	0
1	2-B	3543	0	3621	210	0
1	3-A	3579	0	3652	185	0
1	3-B	3543	0	3621	255	0
1	4-A	3579	0	3652	293	0
1	4-B	3543	0	3621	256	0
1	5-A	3579	0	3652	320	0
1	5-B	3543	0	3621	273	0
1	6-A	3579	0	3652	229	0
1	6-B	3543	0	3621	253	0
1	7-A	3579	0	3652	233	0
1	7-B	3543	0	3621	239	0
1	8-A	3579	0	3652	222	0
1	8-B	3543	0	3621	230	0
2	1-A	180	0	0	41	0
2	1-B	241	0	0	32	0
2	2-A	182	0	0	23	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	2-B	239	0	0	38	0
2	3-A	182	0	0	24	0
2	3-B	239	0	0	41	0
2	4-A	183	0	0	32	0
2	4-B	238	0	0	37	0
2	5-A	182	0	0	43	0
2	5-B	239	0	0	37	0
2	6-A	182	0	0	32	0
2	6-B	239	0	0	33	0
2	7-A	183	0	0	27	0
2	7-B	238	0	0	34	0
2	8-A	183	0	0	24	0
2	8-B	238	0	0	41	0
All	All	60344	0	58184	3934	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 34.

The worst 5 of 3934 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:A:415:SER:HB3	2:A:486:HOH:O	1.18	1.29
1:A:363:SER:HB3	2:A:605:HOH:O	1.42	1.19
1:B:341:ARG:HA	2:B:568:HOH:O	1.47	1.15
1:A:355:ARG:HG3	1:A:355:ARG:HH11	1.12	1.13
1:A:344:ASP:HB3	2:A:570:HOH:O	1.44	1.13

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 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 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	456/469 (97%)	399 (88%)	48 (10%)	9 (2%)	7	1
1	1-B	451/469 (96%)	408 (90%)	38 (8%)	5 (1%)	14	4
1	2-A	456/469 (97%)	407 (89%)	44 (10%)	5 (1%)	14	4
1	2-B	451/469 (96%)	405 (90%)	38 (8%)	8 (2%)	8	2
1	3-A	456/469 (97%)	415 (91%)	33 (7%)	8 (2%)	8	2
1	3-B	451/469 (96%)	410 (91%)	34 (8%)	7 (2%)	9	2
1	4-A	456/469 (97%)	403 (88%)	42 (9%)	11 (2%)	6	1
1	4-B	451/469 (96%)	411 (91%)	33 (7%)	7 (2%)	9	2
1	5-A	456/469 (97%)	406 (89%)	37 (8%)	13 (3%)	4	0
1	5-B	451/469 (96%)	406 (90%)	41 (9%)	4 (1%)	17	6
1	6-A	456/469 (97%)	402 (88%)	48 (10%)	6 (1%)	12	3
1	6-B	451/469 (96%)	399 (88%)	40 (9%)	12 (3%)	5	0
1	7-A	456/469 (97%)	415 (91%)	34 (8%)	7 (2%)	10	2
1	7-B	451/469 (96%)	407 (90%)	38 (8%)	6 (1%)	12	3
1	8-A	456/469 (97%)	403 (88%)	46 (10%)	7 (2%)	10	2
1	8-B	451/469 (96%)	409 (91%)	34 (8%)	8 (2%)	8	2
All	All	7256/7504 (97%)	6505 (90%)	628 (9%)	123 (2%)	9	2

5 of 123 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1-A	8	LEU
1	1-B	170	ALA
1	3-A	38	SER
1	3-A	319	PRO
1	3-B	96	THR

### 5.3.2 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 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	408/414 (99%)	394 (97%)	14 (3%)	37	19
1	1-B	404/414 (98%)	390 (96%)	14 (4%)	36	18
1	2-A	408/414 (99%)	389 (95%)	19 (5%)	26	10
1	2-B	404/414 (98%)	392 (97%)	12 (3%)	41	24
1	3-A	408/414 (99%)	396 (97%)	12 (3%)	42	26
1	3-B	404/414 (98%)	389 (96%)	15 (4%)	34	17
1	4-A	408/414 (99%)	388 (95%)	20 (5%)	25	9
1	4-B	404/414 (98%)	392 (97%)	12 (3%)	41	24
1	5-A	408/414 (99%)	392 (96%)	16 (4%)	32	15
1	5-B	404/414 (98%)	385 (95%)	19 (5%)	26	10
1	6-A	408/414 (99%)	395 (97%)	13 (3%)	39	22
1	6-B	404/414 (98%)	389 (96%)	15 (4%)	34	17
1	7-A	408/414 (99%)	395 (97%)	13 (3%)	39	22
1	7-B	404/414 (98%)	385 (95%)	19 (5%)	26	10
1	8-A	408/414 (99%)	392 (96%)	16 (4%)	32	15
1	8-B	404/414 (98%)	391 (97%)	13 (3%)	39	22
All	All	6496/6624 (98%)	6254 (96%)	242 (4%)	34	17

5 of 242 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	4-B	360	LYS
1	8-A	244	CYS
1	5-B	220	ASN
1	8-A	209	LEU
1	8-B	357	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 126 such sidechains are listed below:

Mol	Chain	Res	Type
1	4-A	117	GLN
1	7-B	152	ASN
1	4-B	140	HIS
1	7-B	75	ASN
1	8-A	460	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	1-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
1	2-A	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	2-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
1	3-A	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	3-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
1	4-A	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	4-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
1	5-A	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	5-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
1	6-A	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	6-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
1	7-A	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	7-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
1	8-A	460/469 (98%)	0.61	54 (11%) 4 4	9, 24, 48, 69	460 (100%)
1	8-B	455/469 (97%)	0.10	24 (5%) 26 25	3, 17, 39, 55	455 (100%)
All	All	7320/7504 (97%)	0.36	624 (8%) 10 10	3, 21, 46, 69	7320 (100%)

The worst 5 of 624 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	1-A	387	THR	7.6
1	2-A	387	THR	7.6
1	3-A	387	THR	7.6
1	4-A	387	THR	7.6
1	5-A	387	THR	7.6

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.