



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2020 – 01:58 PM BST

PDB ID : 3UG4  
Title : Crystal structure of alpha-L-arabinofuranosidase from *Thermotoga maritima* arabinose complex  
Authors : Im, D.-H.; Miyazaki, K.; Wakagi, T.; Fushinobu, S.  
Deposited on : 2011-11-02  
Resolution : 2.15 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

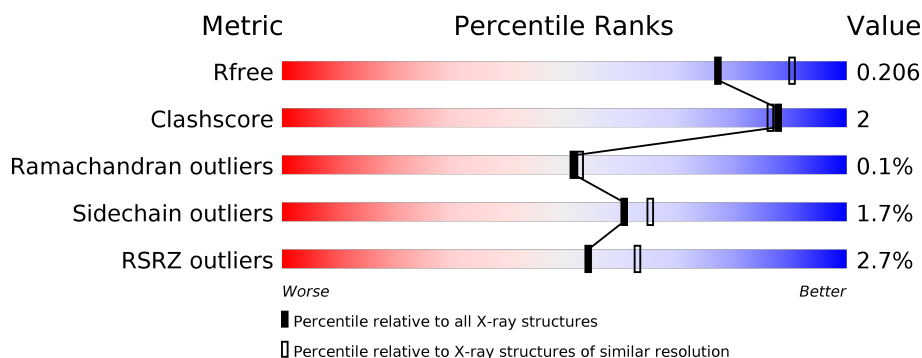
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	A	504	<div> <div>2%</div> <div> <div></div> <div>90%</div> <div>6%</div> <div>.</div> </div> </div>
1	B	504	<div> <div>6%</div> <div> <div></div> <div>88%</div> <div>8%</div> <div>.</div> </div> </div>
1	C	504	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>6%</div> <div>.</div> </div> </div>
1	D	504	<div> <div>3%</div> <div> <div></div> <div>88%</div> <div>6%</div> <div>.</div> <div>5%</div> </div> </div>
1	E	504	<div> <div>3%</div> <div> <div></div> <div>89%</div> <div>6%</div> <div>.</div> <div>.</div> </div> </div>
1	F	504	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>5%</div> <div>.</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	TRS	B	485	-	X	-	-
2	TRS	F	485	-	X	-	-
2	TRS	F	486	-	X	-	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 25486 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 Alpha-L-arabinofuranosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	482	Total	C	N	O	S	0	0	0
			3876	2488	644	729	15			
1	B	482	Total	C	N	O	S	0	0	0
			3876	2488	644	729	15			
1	C	484	Total	C	N	O	S	0	0	0
			3893	2498	646	733	16			
1	D	481	Total	C	N	O	S	0	0	0
			3868	2482	643	728	15			
1	E	482	Total	C	N	O	S	0	0	0
			3876	2488	644	729	15			
1	F	482	Total	C	N	O	S	0	0	0
			3876	2488	644	729	15			

There are 126 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP Q9WYB7
A	-18	GLY	-	expression tag	UNP Q9WYB7
A	-17	SER	-	expression tag	UNP Q9WYB7
A	-16	SER	-	expression tag	UNP Q9WYB7
A	-15	HIS	-	expression tag	UNP Q9WYB7
A	-14	HIS	-	expression tag	UNP Q9WYB7
A	-13	HIS	-	expression tag	UNP Q9WYB7
A	-12	HIS	-	expression tag	UNP Q9WYB7
A	-11	HIS	-	expression tag	UNP Q9WYB7
A	-10	HIS	-	expression tag	UNP Q9WYB7
A	-9	SER	-	expression tag	UNP Q9WYB7
A	-8	SER	-	expression tag	UNP Q9WYB7
A	-7	GLY	-	expression tag	UNP Q9WYB7
A	-6	LEU	-	expression tag	UNP Q9WYB7
A	-5	VAL	-	expression tag	UNP Q9WYB7
A	-4	PRO	-	expression tag	UNP Q9WYB7
A	-3	ARG	-	expression tag	UNP Q9WYB7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q9WYB7
A	-1	SER	-	expression tag	UNP Q9WYB7
A	0	HIS	-	expression tag	UNP Q9WYB7
A	4	GLY	ARG	engineered mutation	UNP Q9WYB7
B	-19	MET	-	expression tag	UNP Q9WYB7
B	-18	GLY	-	expression tag	UNP Q9WYB7
B	-17	SER	-	expression tag	UNP Q9WYB7
B	-16	SER	-	expression tag	UNP Q9WYB7
B	-15	HIS	-	expression tag	UNP Q9WYB7
B	-14	HIS	-	expression tag	UNP Q9WYB7
B	-13	HIS	-	expression tag	UNP Q9WYB7
B	-12	HIS	-	expression tag	UNP Q9WYB7
B	-11	HIS	-	expression tag	UNP Q9WYB7
B	-10	HIS	-	expression tag	UNP Q9WYB7
B	-9	SER	-	expression tag	UNP Q9WYB7
B	-8	SER	-	expression tag	UNP Q9WYB7
B	-7	GLY	-	expression tag	UNP Q9WYB7
B	-6	LEU	-	expression tag	UNP Q9WYB7
B	-5	VAL	-	expression tag	UNP Q9WYB7
B	-4	PRO	-	expression tag	UNP Q9WYB7
B	-3	ARG	-	expression tag	UNP Q9WYB7
B	-2	GLY	-	expression tag	UNP Q9WYB7
B	-1	SER	-	expression tag	UNP Q9WYB7
B	0	HIS	-	expression tag	UNP Q9WYB7
B	4	GLY	ARG	engineered mutation	UNP Q9WYB7
C	-19	MET	-	expression tag	UNP Q9WYB7
C	-18	GLY	-	expression tag	UNP Q9WYB7
C	-17	SER	-	expression tag	UNP Q9WYB7
C	-16	SER	-	expression tag	UNP Q9WYB7
C	-15	HIS	-	expression tag	UNP Q9WYB7
C	-14	HIS	-	expression tag	UNP Q9WYB7
C	-13	HIS	-	expression tag	UNP Q9WYB7
C	-12	HIS	-	expression tag	UNP Q9WYB7
C	-11	HIS	-	expression tag	UNP Q9WYB7
C	-10	HIS	-	expression tag	UNP Q9WYB7
C	-9	SER	-	expression tag	UNP Q9WYB7
C	-8	SER	-	expression tag	UNP Q9WYB7
C	-7	GLY	-	expression tag	UNP Q9WYB7
C	-6	LEU	-	expression tag	UNP Q9WYB7
C	-5	VAL	-	expression tag	UNP Q9WYB7
C	-4	PRO	-	expression tag	UNP Q9WYB7
C	-3	ARG	-	expression tag	UNP Q9WYB7

*Continued on next page...*

*Continued from previous page...*

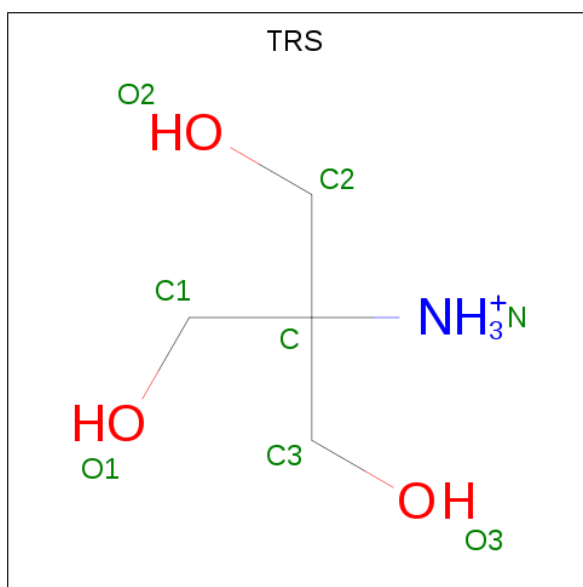
Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	GLY	-	expression tag	UNP Q9WYB7
C	-1	SER	-	expression tag	UNP Q9WYB7
C	0	HIS	-	expression tag	UNP Q9WYB7
C	4	GLY	ARG	engineered mutation	UNP Q9WYB7
D	-19	MET	-	expression tag	UNP Q9WYB7
D	-18	GLY	-	expression tag	UNP Q9WYB7
D	-17	SER	-	expression tag	UNP Q9WYB7
D	-16	SER	-	expression tag	UNP Q9WYB7
D	-15	HIS	-	expression tag	UNP Q9WYB7
D	-14	HIS	-	expression tag	UNP Q9WYB7
D	-13	HIS	-	expression tag	UNP Q9WYB7
D	-12	HIS	-	expression tag	UNP Q9WYB7
D	-11	HIS	-	expression tag	UNP Q9WYB7
D	-10	HIS	-	expression tag	UNP Q9WYB7
D	-9	SER	-	expression tag	UNP Q9WYB7
D	-8	SER	-	expression tag	UNP Q9WYB7
D	-7	GLY	-	expression tag	UNP Q9WYB7
D	-6	LEU	-	expression tag	UNP Q9WYB7
D	-5	VAL	-	expression tag	UNP Q9WYB7
D	-4	PRO	-	expression tag	UNP Q9WYB7
D	-3	ARG	-	expression tag	UNP Q9WYB7
D	-2	GLY	-	expression tag	UNP Q9WYB7
D	-1	SER	-	expression tag	UNP Q9WYB7
D	0	HIS	-	expression tag	UNP Q9WYB7
D	4	GLY	ARG	engineered mutation	UNP Q9WYB7
E	-19	MET	-	expression tag	UNP Q9WYB7
E	-18	GLY	-	expression tag	UNP Q9WYB7
E	-17	SER	-	expression tag	UNP Q9WYB7
E	-16	SER	-	expression tag	UNP Q9WYB7
E	-15	HIS	-	expression tag	UNP Q9WYB7
E	-14	HIS	-	expression tag	UNP Q9WYB7
E	-13	HIS	-	expression tag	UNP Q9WYB7
E	-12	HIS	-	expression tag	UNP Q9WYB7
E	-11	HIS	-	expression tag	UNP Q9WYB7
E	-10	HIS	-	expression tag	UNP Q9WYB7
E	-9	SER	-	expression tag	UNP Q9WYB7
E	-8	SER	-	expression tag	UNP Q9WYB7
E	-7	GLY	-	expression tag	UNP Q9WYB7
E	-6	LEU	-	expression tag	UNP Q9WYB7
E	-5	VAL	-	expression tag	UNP Q9WYB7
E	-4	PRO	-	expression tag	UNP Q9WYB7
E	-3	ARG	-	expression tag	UNP Q9WYB7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	GLY	-	expression tag	UNP Q9WYB7
E	-1	SER	-	expression tag	UNP Q9WYB7
E	0	HIS	-	expression tag	UNP Q9WYB7
E	4	GLY	ARG	engineered mutation	UNP Q9WYB7
F	-19	MET	-	expression tag	UNP Q9WYB7
F	-18	GLY	-	expression tag	UNP Q9WYB7
F	-17	SER	-	expression tag	UNP Q9WYB7
F	-16	SER	-	expression tag	UNP Q9WYB7
F	-15	HIS	-	expression tag	UNP Q9WYB7
F	-14	HIS	-	expression tag	UNP Q9WYB7
F	-13	HIS	-	expression tag	UNP Q9WYB7
F	-12	HIS	-	expression tag	UNP Q9WYB7
F	-11	HIS	-	expression tag	UNP Q9WYB7
F	-10	HIS	-	expression tag	UNP Q9WYB7
F	-9	SER	-	expression tag	UNP Q9WYB7
F	-8	SER	-	expression tag	UNP Q9WYB7
F	-7	GLY	-	expression tag	UNP Q9WYB7
F	-6	LEU	-	expression tag	UNP Q9WYB7
F	-5	VAL	-	expression tag	UNP Q9WYB7
F	-4	PRO	-	expression tag	UNP Q9WYB7
F	-3	ARG	-	expression tag	UNP Q9WYB7
F	-2	GLY	-	expression tag	UNP Q9WYB7
F	-1	SER	-	expression tag	UNP Q9WYB7
F	0	HIS	-	expression tag	UNP Q9WYB7
F	4	GLY	ARG	engineered mutation	UNP Q9WYB7

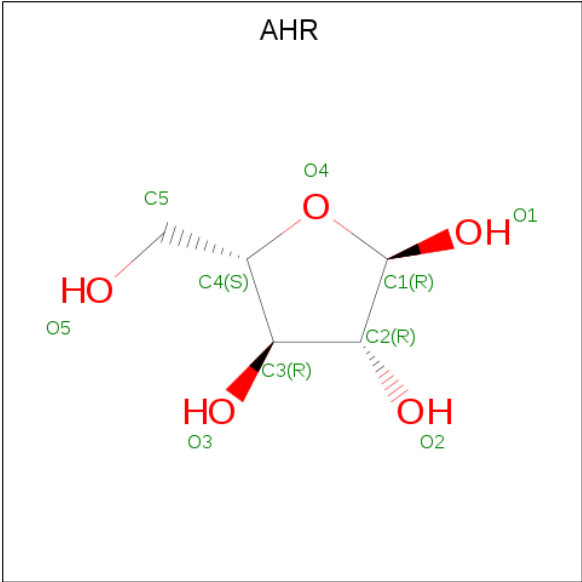
- Molecule 2 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			8	4	1	3		
2	A	1	Total	C	N	O	0	0
			8	4	1	3		
2	B	1	Total	C	N	O	0	0
			8	4	1	3		
2	B	1	Total	C	N	O	0	0
			8	4	1	3		
2	C	1	Total	C	N	O	0	0
			8	4	1	3		
2	D	1	Total	C	N	O	0	0
			8	4	1	3		
2	D	1	Total	C	N	O	0	0
			8	4	1	3		
2	E	1	Total	C	N	O	0	0
			8	4	1	3		
2	E	1	Total	C	N	O	0	0
			8	4	1	3		
2	F	1	Total	C	N	O	0	0
			8	4	1	3		
2	F	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 3 is alpha-L-arabinofuranose (three-letter code: AHR) (formula: C<sub>5</sub>H<sub>10</sub>O<sub>5</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	5	5		
3	A	1	Total	C	O	0	0
			10	5	5		
3	A	1	Total	C	O	0	0
			10	5	5		
3	A	1	Total	C	O	0	0
			10	5	5		
3	A	1	Total	C	O	0	0
			10	5	5		
3	A	1	Total	C	O	0	0
			10	5	5		
3	B	1	Total	C	O	0	0
			10	5	5		
3	B	1	Total	C	O	0	0
			10	5	5		
3	B	1	Total	C	O	0	0
			10	5	5		
3	B	1	Total	C	O	0	0
			10	5	5		
3	C	1	Total	C	O	0	0
			10	5	5		
3	C	1	Total	C	O	0	0
			10	5	5		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			10	5	5		
3	C	1	Total	C	O	0	0
			10	5	5		
3	C	1	Total	C	O	0	0
			10	5	5		
3	D	1	Total	C	O	0	0
			10	5	5		
3	D	1	Total	C	O	0	0
			10	5	5		
3	D	1	Total	C	O	0	0
			10	5	5		
3	D	1	Total	C	O	0	0
			10	5	5		
3	D	1	Total	C	O	0	0
			10	5	5		
3	E	1	Total	C	O	0	0
			10	5	5		
3	E	1	Total	C	O	0	0
			10	5	5		
3	E	1	Total	C	O	0	0
			10	5	5		
3	F	1	Total	C	O	0	0
			10	5	5		
3	F	1	Total	C	O	0	0
			10	5	5		
3	F	1	Total	C	O	0	0
			10	5	5		
3	F	1	Total	C	O	0	0
			10	5	5		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	342	Total	O	0	0
			342	342		
4	B	252	Total	O	0	0
			252	252		

*Continued on next page...*

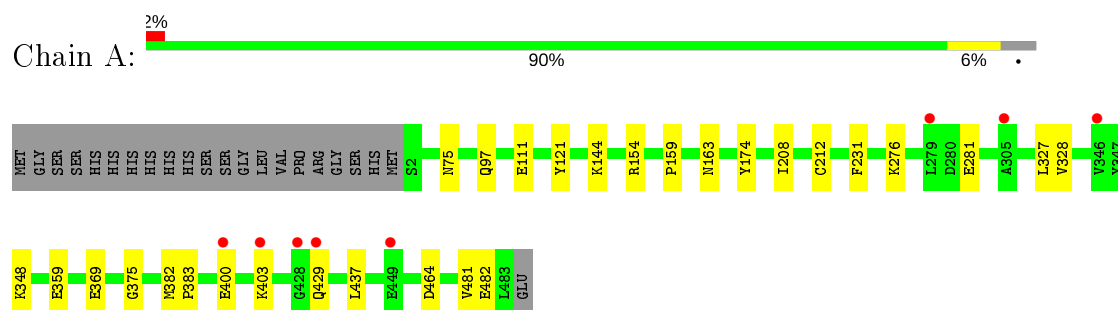
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	361	Total 361	O 361	0	0
4	D	293	Total 293	O 293	0	0
4	E	260	Total 260	O 260	0	0
4	F	315	Total 315	O 315	0	0

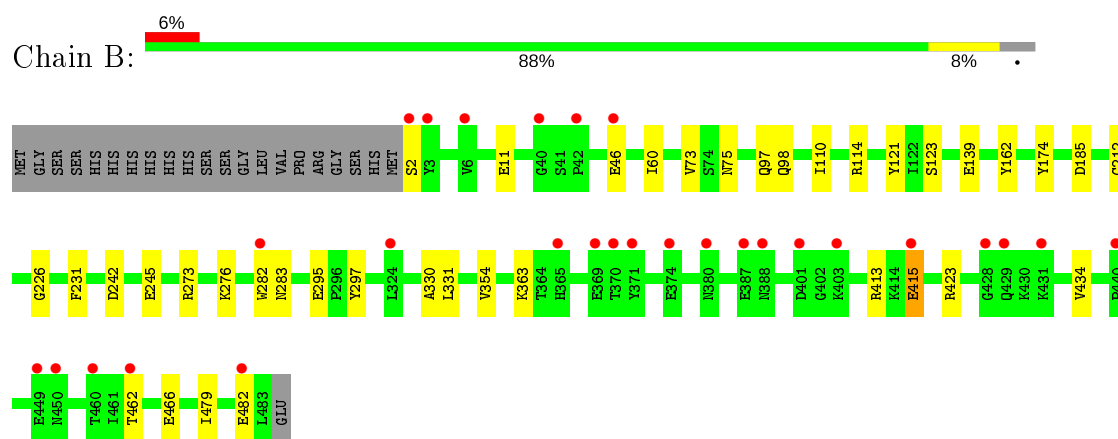
### 3 Residue-property plots [i](#)

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.

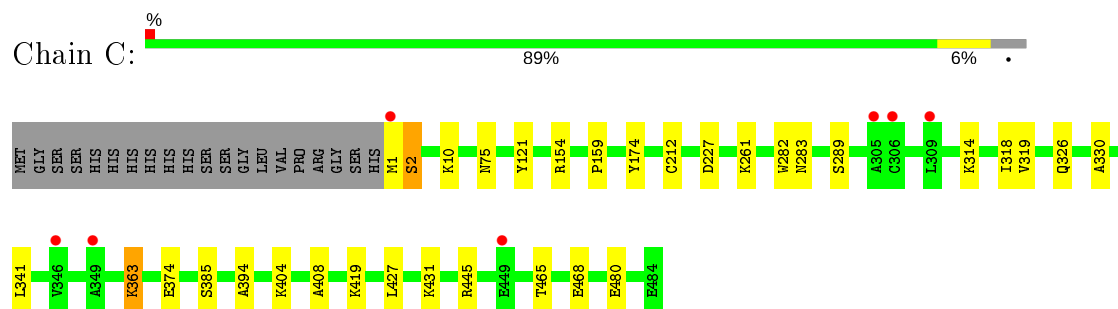
- Molecule 1: Alpha-L-arabinofuranosidase



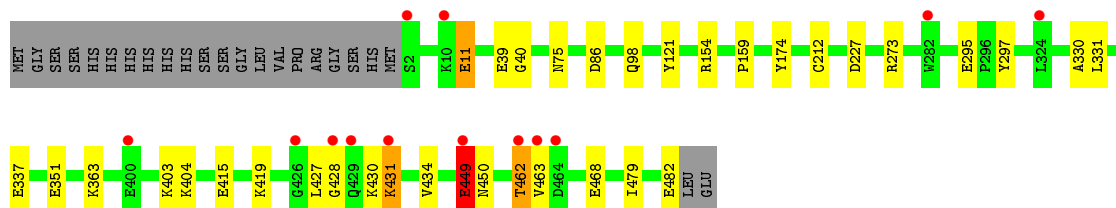
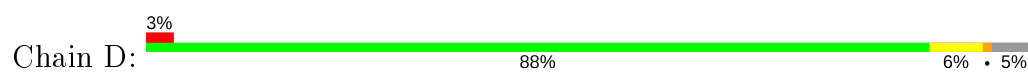
- Molecule 1: Alpha-L-arabinofuranosidase



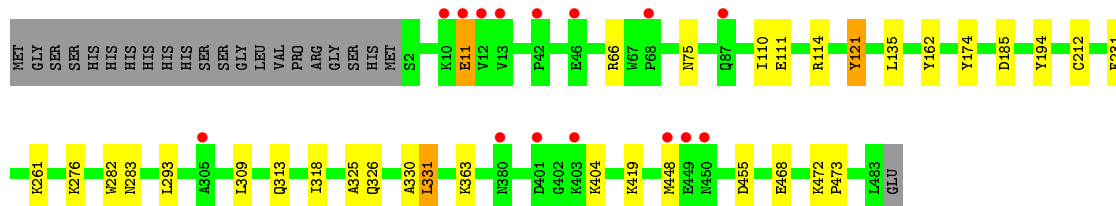
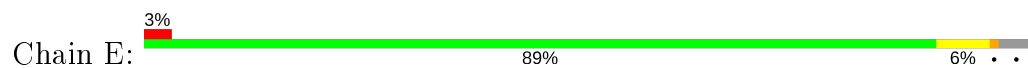
- Molecule 1: Alpha-L-arabinofuranosidase



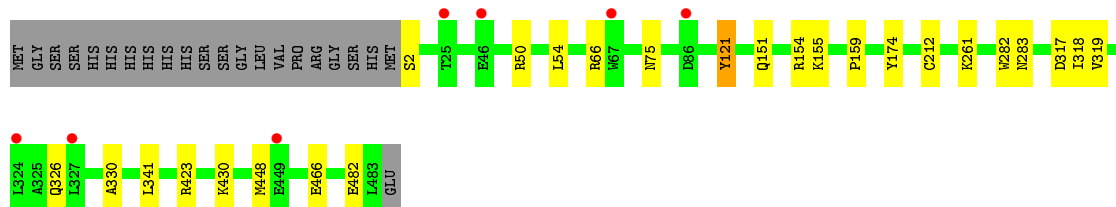
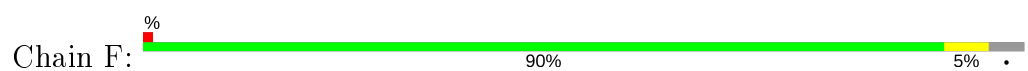
- Molecule 1: Alpha-L-arabinofuranosidase



- Molecule 1: Alpha-L-arabinofuranosidase



- Molecule 1: Alpha-L-arabinofuranosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.48Å 160.64Å 154.84Å 90.00° 92.44° 90.00°	Depositor
Resolution (Å)	29.99 – 2.15 29.99 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.6 (29.99-2.15) 99.6 (29.99-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.22 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, $R_{free}$	0.167 , 0.207 0.167 , 0.206	Depositor DCC
$R_{free}$ test set	13365 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.9	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 52.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for -h,-l,-k 0.000 for -h,l,k 0.008 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	25486	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: AHR, TRS

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	A	0.95	3/3973 (0.1%)	0.82	2/5392 (0.0%)
1	B	0.94	1/3973 (0.0%)	0.81	1/5392 (0.0%)
1	C	0.91	0/3990	0.80	1/5414 (0.0%)
1	D	0.94	3/3965 (0.1%)	0.81	2/5381 (0.0%)
1	E	0.92	2/3973 (0.1%)	0.81	2/5392 (0.0%)
1	F	0.88	0/3973	0.80	2/5392 (0.0%)
All	All	0.93	9/23847 (0.0%)	0.81	10/32363 (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	D	0	1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	359	GLU	CG-CD	6.42	1.61	1.51
1	A	400	GLU	CG-CD	6.01	1.60	1.51
1	D	337	GLU	CG-CD	5.88	1.60	1.51
1	E	111	GLU	CG-CD	5.80	1.60	1.51
1	E	194	TYR	CD1-CE1	5.36	1.47	1.39

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	382	MET	CG-SD-CE	6.54	110.67	100.20
1	D	39	GLU	C-N-CA	-6.44	108.78	122.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	317	ASP	CB-CG-OD1	5.84	123.56	118.30
1	A	464	ASP	CB-CG-OD1	5.29	123.06	118.30
1	D	86	ASP	CB-CG-OD1	5.25	123.03	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	40	GLY	Peptide

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in 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	A	3876	0	3793	11	0
1	B	3876	0	3793	16	0
1	C	3893	0	3811	18	0
1	D	3868	0	3782	16	0
1	E	3876	0	3793	15	0
1	F	3876	0	3793	11	0
2	A	16	0	24	0	0
2	B	16	0	23	0	0
2	C	8	0	12	0	0
2	D	16	0	24	0	0
2	E	16	0	24	0	0
2	F	16	0	23	1	0
3	A	70	0	0	1	0
3	B	50	0	0	0	0
3	C	50	0	0	1	0
3	D	60	0	0	2	0
3	E	30	0	0	0	0
3	F	50	0	0	1	0
4	A	342	0	0	3	0
4	B	252	0	0	2	0
4	C	361	0	0	3	0
4	D	293	0	0	1	0
4	E	260	0	0	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	315	0	0	1	0
All	All	25486	0	22895	91	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 2.

The worst 5 of 91 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:163:ASN:HB2	4:A:1472:HOH:O	1.65	0.96
1:E:185:ASP:HB2	4:E:545:HOH:O	1.72	0.88
1:C:419:LYS:HE2	1:C:468:GLU:OE1	1.78	0.83
1:A:369:GLU:HG3	4:A:567:HOH:O	1.86	0.75
1:D:154:ARG:HG2	1:D:159:PRO:HA	1.71	0.71

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	A	480/504 (95%)	460 (96%)	20 (4%)	0	100	100
1	B	480/504 (95%)	457 (95%)	23 (5%)	0	100	100
1	C	482/504 (96%)	463 (96%)	18 (4%)	1 (0%)	47	46
1	D	479/504 (95%)	450 (94%)	28 (6%)	1 (0%)	47	46
1	E	480/504 (95%)	458 (95%)	22 (5%)	0	100	100
1	F	480/504 (95%)	457 (95%)	23 (5%)	0	100	100
All	All	2881/3024 (95%)	2745 (95%)	134 (5%)	2 (0%)	51	53

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	2	SER
1	D	428	GLY

### 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	A	417/436 (96%)	411 (99%)	6 (1%)	67	72
1	B	417/436 (96%)	408 (98%)	9 (2%)	52	55
1	C	419/436 (96%)	414 (99%)	5 (1%)	71	76
1	D	416/436 (95%)	406 (98%)	10 (2%)	49	51
1	E	417/436 (96%)	411 (99%)	6 (1%)	67	72
1	F	417/436 (96%)	411 (99%)	6 (1%)	67	72
All	All	2503/2616 (96%)	2461 (98%)	42 (2%)	60	65

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	289	SER
1	D	404	LYS
1	F	75	ASN
1	C	363	LYS
1	D	75	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

42 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or 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 length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	AHR	B	489	-	10,10,10	3.22	2 (20%)	13,14,14	3.95	5 (38%)
2	TRS	F	486	-	7,7,7	0.55	0	9,9,9	1.77	3 (33%)
2	TRS	A	486	-	7,7,7	0.58	0	9,9,9	0.72	0
3	AHR	F	488	-	10,10,10	3.75	3 (30%)	13,14,14	4.33	7 (53%)
3	AHR	D	488	-	10,10,10	1.87	4 (40%)	13,14,14	1.48	2 (15%)
3	AHR	D	492	-	10,10,10	2.02	1 (10%)	13,14,14	4.31	5 (38%)
3	AHR	C	490	-	10,10,10	2.39	2 (20%)	13,14,14	3.73	9 (69%)
3	AHR	B	490	-	10,10,10	3.42	2 (20%)	13,14,14	4.72	5 (38%)
2	TRS	C	485	-	7,7,7	0.66	0	9,9,9	0.91	0
2	TRS	E	486	-	7,7,7	0.47	0	9,9,9	1.41	2 (22%)
2	TRS	F	485	-	7,7,7	0.70	0	9,9,9	2.03	4 (44%)
3	AHR	E	489	-	10,10,10	3.06	2 (20%)	13,14,14	5.14	7 (53%)
3	AHR	E	487	-	10,10,10	1.54	2 (20%)	13,14,14	1.74	2 (15%)
3	AHR	A	487	-	10,10,10	1.17	1 (10%)	13,14,14	1.95	1 (7%)
3	AHR	D	485	-	10,10,10	3.26	4 (40%)	13,14,14	2.80	3 (23%)
2	TRS	E	485	-	7,7,7	0.43	0	9,9,9	0.70	0
3	AHR	D	489	-	10,10,10	4.11	3 (30%)	13,14,14	5.01	7 (53%)
3	AHR	B	488	-	10,10,10	2.92	3 (30%)	13,14,14	4.52	5 (38%)
2	TRS	B	486	-	7,7,7	0.54	0	9,9,9	0.81	0
3	AHR	A	491	-	10,10,10	1.92	1 (10%)	13,14,14	4.21	5 (38%)
3	AHR	D	491	-	10,10,10	2.84	2 (20%)	13,14,14	4.96	6 (46%)
3	AHR	C	486	-	10,10,10	1.56	2 (20%)	13,14,14	1.04	1 (7%)
3	AHR	A	490	-	10,10,10	3.47	2 (20%)	13,14,14	5.98	9 (69%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	AHR	D	490	-	10,10,10	4.47	3 (30%)	13,14,14	6.36	5 (38%)
3	AHR	C	487	-	10,10,10	3.51	3 (30%)	13,14,14	4.48	6 (46%)
3	AHR	F	490	-	10,10,10	3.85	3 (30%)	13,14,14	5.81	8 (61%)
3	AHR	F	491	-	10,10,10	2.29	1 (10%)	13,14,14	2.94	5 (38%)
3	AHR	B	491	-	10,10,10	0.96	1 (10%)	13,14,14	2.18	4 (30%)
3	AHR	A	493	-	10,10,10	2.12	3 (30%)	13,14,14	1.93	4 (30%)
2	TRS	A	485	-	7,7,7	0.60	0	9,9,9	0.91	0
3	AHR	F	489	-	10,10,10	3.80	3 (30%)	13,14,14	4.89	6 (46%)
3	AHR	C	488	-	10,10,10	4.01	3 (30%)	13,14,14	5.64	8 (61%)
2	TRS	D	486	-	7,7,7	0.31	0	9,9,9	0.83	0
3	AHR	A	488	-	10,10,10	3.56	3 (30%)	13,14,14	5.80	8 (61%)
3	AHR	E	488	-	10,10,10	3.31	3 (30%)	13,14,14	4.56	4 (30%)
3	AHR	C	489	-	10,10,10	2.95	3 (30%)	13,14,14	6.16	8 (61%)
3	AHR	A	489	-	10,10,10	4.58	3 (30%)	13,14,14	7.26	7 (53%)
2	TRS	B	485	-	7,7,7	1.13	1 (14%)	9,9,9	2.19	3 (33%)
2	TRS	D	487	-	7,7,7	0.41	0	9,9,9	1.10	0
3	AHR	A	492	-	10,10,10	2.62	4 (40%)	13,14,14	2.30	6 (46%)
3	AHR	F	487	-	10,10,10	1.63	2 (20%)	13,14,14	0.94	0
3	AHR	B	487	-	10,10,10	2.54	2 (20%)	13,14,14	1.45	2 (15%)

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
3	AHR	B	489	-	-	2/2/18/18	0/1/1/1
2	TRS	F	486	-	-	9/9/9/9	-
2	TRS	A	486	-	-	0/9/9/9	-
3	AHR	F	488	-	-	2/2/18/18	0/1/1/1
3	AHR	D	488	-	-	0/2/18/18	0/1/1/1
3	AHR	D	492	-	-	0/2/18/18	0/1/1/1
3	AHR	C	490	-	-	0/2/18/18	0/1/1/1
3	AHR	B	490	-	-	0/2/18/18	0/1/1/1
2	TRS	C	485	-	-	0/9/9/9	-
2	TRS	E	486	-	-	0/9/9/9	-
2	TRS	F	485	-	-	9/9/9/9	-
3	AHR	E	489	-	-	2/2/18/18	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	AHR	E	487	-	-	0/2/18/18	0/1/1/1
3	AHR	A	487	-	-	0/2/18/18	0/1/1/1
3	AHR	D	485	-	-	0/2/18/18	0/1/1/1
2	TRS	E	485	-	-	0/9/9/9	-
3	AHR	D	489	-	-	2/2/18/18	0/1/1/1
3	AHR	B	488	-	-	2/2/18/18	0/1/1/1
2	TRS	B	486	-	-	0/9/9/9	-
3	AHR	A	491	-	-	0/2/18/18	0/1/1/1
3	AHR	D	491	-	-	0/2/18/18	0/1/1/1
3	AHR	C	486	-	-	0/2/18/18	0/1/1/1
3	AHR	A	490	-	-	2/2/18/18	0/1/1/1
3	AHR	D	490	-	-	1/2/18/18	0/1/1/1
3	AHR	C	487	-	-	2/2/18/18	0/1/1/1
3	AHR	F	490	-	-	0/2/18/18	0/1/1/1
3	AHR	F	491	-	-	1/2/18/18	0/1/1/1
3	AHR	B	491	-	-	0/2/18/18	0/1/1/1
3	AHR	A	493	-	-	2/2/18/18	0/1/1/1
2	TRS	A	485	-	-	0/9/9/9	-
3	AHR	F	489	-	-	2/2/18/18	0/1/1/1
3	AHR	C	488	-	-	0/2/18/18	0/1/1/1
2	TRS	D	486	-	-	0/9/9/9	-
3	AHR	A	488	-	-	2/2/18/18	0/1/1/1
3	AHR	E	488	-	-	2/2/18/18	0/1/1/1
3	AHR	C	489	-	-	2/2/18/18	0/1/1/1
3	AHR	A	489	-	-	0/2/18/18	0/1/1/1
2	TRS	B	485	-	-	9/9/9/9	-
2	TRS	D	487	-	-	6/9/9/9	-
3	AHR	A	492	-	-	0/2/18/18	0/1/1/1
3	AHR	F	487	-	-	0/2/18/18	0/1/1/1
3	AHR	B	487	-	-	0/2/18/18	0/1/1/1

The worst 5 of 77 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	490	AHR	O4-C1	13.49	1.59	1.43
3	D	489	AHR	O4-C1	11.77	1.57	1.43
3	A	489	AHR	O4-C1	11.59	1.57	1.43
3	C	488	AHR	O4-C1	11.48	1.57	1.43
3	F	489	AHR	O4-C1	11.20	1.56	1.43

The worst 5 of 172 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	D	490	AHR	O1-C1-O4	21.24	138.33	111.13
3	A	489	AHR	O1-C1-O4	19.35	135.91	111.13
3	C	488	AHR	O1-C1-O4	16.09	131.74	111.13
3	F	489	AHR	O1-C1-O4	15.34	130.78	111.13
3	A	490	AHR	O1-C1-O4	14.90	130.22	111.13

There are no chirality outliers.

5 of 59 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	486	TRS	C2-C-C1-O1
2	F	486	TRS	C3-C-C1-O1
2	F	486	TRS	C1-C-C2-O2
2	F	486	TRS	C3-C-C2-O2
2	F	486	TRS	N-C-C2-O2

There are no ring outliers.

10 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	490	AHR	1	0
2	F	485	TRS	1	0
3	D	489	AHR	1	0
3	A	491	AHR	1	0
3	D	491	AHR	1	0
3	A	490	AHR	1	0
3	D	490	AHR	1	0
3	F	490	AHR	1	0
3	F	491	AHR	1	0
3	C	489	AHR	1	0

## 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	A	482/504 (95%)	-0.21	8 (1%) 70 76	18, 27, 45, 58	0
1	B	482/504 (95%)	-0.04	28 (5%) 23 31	20, 31, 48, 61	0
1	C	484/504 (96%)	-0.23	7 (1%) 75 80	19, 27, 44, 62	0
1	D	481/504 (95%)	-0.15	13 (2%) 54 63	20, 29, 47, 60	0
1	E	482/504 (95%)	-0.18	15 (3%) 49 58	21, 30, 46, 61	0
1	F	482/504 (95%)	-0.19	7 (1%) 73 79	21, 29, 46, 60	0
All	All	2893/3024 (95%)	-0.17	78 (2%) 54 63	18, 29, 46, 62	0

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	449	GLU	5.2
1	D	431	LYS	4.7
1	F	46	GLU	4.5
1	D	429	GLN	4.3
1	B	449	GLU	4.2

### 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

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
3	AHR	D	485	10/10	0.65	0.18	59,64,66,67	0
3	AHR	A	492	10/10	0.66	0.23	64,69,71,71	0
2	TRS	D	487	8/8	0.80	0.21	41,46,47,48	0
3	AHR	A	493	10/10	0.80	0.17	62,69,70,70	0
3	AHR	B	489	10/10	0.83	0.13	40,45,47,49	0
3	AHR	F	491	10/10	0.85	0.13	49,52,54,56	0
3	AHR	D	492	10/10	0.86	0.16	60,64,65,66	0
3	AHR	D	490	10/10	0.86	0.13	22,36,39,42	0
3	AHR	B	491	10/10	0.87	0.15	58,61,64,65	0
2	TRS	A	486	8/8	0.87	0.16	31,36,38,39	0
3	AHR	A	488	10/10	0.88	0.13	31,35,37,37	0
3	AHR	E	488	10/10	0.88	0.12	38,41,43,44	0
3	AHR	D	489	10/10	0.89	0.10	34,38,40,42	0
3	AHR	A	491	10/10	0.89	0.14	47,51,55,56	0
3	AHR	E	489	10/10	0.89	0.13	34,41,43,45	0
3	AHR	C	488	10/10	0.89	0.11	37,42,44,44	0
3	AHR	A	489	10/10	0.91	0.12	27,30,34,36	0
3	AHR	C	490	10/10	0.91	0.13	44,49,51,51	0
3	AHR	C	487	10/10	0.91	0.10	38,44,45,45	0
2	TRS	B	486	8/8	0.92	0.13	40,42,43,44	0
2	TRS	E	486	8/8	0.92	0.10	34,36,36,38	0
3	AHR	A	490	10/10	0.92	0.08	30,34,39,40	0
2	TRS	F	486	8/8	0.92	0.13	33,38,40,42	0
3	AHR	F	488	10/10	0.92	0.11	32,38,41,41	0
3	AHR	B	490	10/10	0.92	0.10	43,46,48,49	0
3	AHR	B	488	10/10	0.92	0.10	36,38,41,41	0
2	TRS	B	485	8/8	0.93	0.12	28,30,32,32	0
2	TRS	F	485	8/8	0.93	0.12	32,33,34,36	0
3	AHR	F	489	10/10	0.93	0.11	31,41,44,46	0
3	AHR	D	491	10/10	0.94	0.08	39,42,45,45	0
3	AHR	C	489	10/10	0.94	0.09	29,33,35,35	0
3	AHR	F	490	10/10	0.94	0.08	32,37,40,41	0
3	AHR	D	488	10/10	0.95	0.13	24,27,29,29	0
3	AHR	E	487	10/10	0.95	0.10	27,30,31,32	0
2	TRS	E	485	8/8	0.96	0.10	27,29,32,32	0
3	AHR	B	487	10/10	0.96	0.10	25,28,30,31	0
2	TRS	D	486	8/8	0.97	0.07	26,27,29,30	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	AHR	C	486	10/10	0.97	0.07	25,29,32,33	0
2	TRS	A	485	8/8	0.97	0.08	27,28,30,31	0
3	AHR	A	487	10/10	0.98	0.09	22,24,27,28	0
3	AHR	F	487	10/10	0.98	0.07	23,27,29,31	0
2	TRS	C	485	8/8	0.98	0.06	24,25,25,26	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.