



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:51 PM GMT

PDB ID : 4Q3P  
Title : Crystal structure of Schistosoma mansoni arginase  
Authors : Hai, Y.; Edwards, J.E.; Van Zandt, M.C.; Hoffmann, K.F.; Christianson, D.W.  
Deposited on : 2014-04-12  
Resolution : 2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

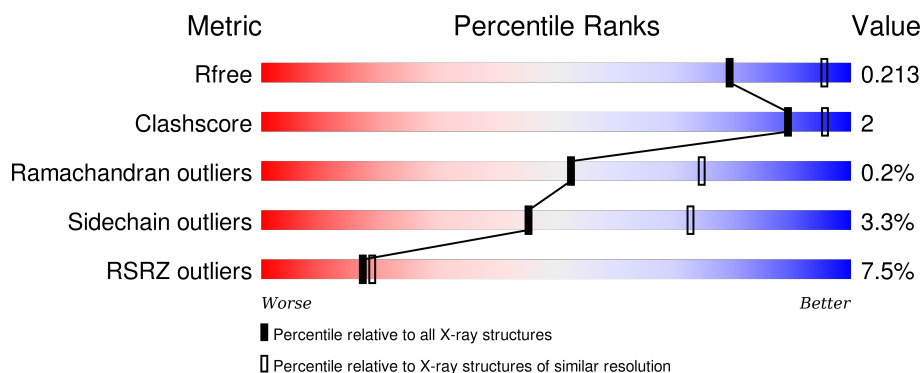
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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. 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	385	<div> <div>81%</div> <div>6%</div> <div>13%</div> </div>
1	B	385	<div> <div>%</div> <div>79%</div> <div>7%</div> <div>•</div> <div>13%</div> </div>
1	C	385	<div> <div>80%</div> <div>6%</div> <div>•</div> <div>13%</div> </div>
1	D	385	<div> <div>24%</div> <div>77%</div> <div>9%</div> <div>•</div> <div>14%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10810 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 Arginase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	336	Total	C	N	O	S	0	1	0
			2581	1625	451	489	16			
1	B	335	Total	C	N	O	S	0	1	0
			2573	1621	449	487	16			
1	C	335	Total	C	N	O	S	0	1	0
			2571	1620	448	487	16			
1	D	332	Total	C	N	O	S	0	0	0
			2546	1605	444	482	15			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	EXPRESSION TAG	UNP Q6WVP6
A	-19	GLY	-	EXPRESSION TAG	UNP Q6WVP6
A	-18	SER	-	EXPRESSION TAG	UNP Q6WVP6
A	-17	SER	-	EXPRESSION TAG	UNP Q6WVP6
A	-16	HIS	-	EXPRESSION TAG	UNP Q6WVP6
A	-15	HIS	-	EXPRESSION TAG	UNP Q6WVP6
A	-14	HIS	-	EXPRESSION TAG	UNP Q6WVP6
A	-13	HIS	-	EXPRESSION TAG	UNP Q6WVP6
A	-12	HIS	-	EXPRESSION TAG	UNP Q6WVP6
A	-11	HIS	-	EXPRESSION TAG	UNP Q6WVP6
A	-10	SER	-	EXPRESSION TAG	UNP Q6WVP6
A	-9	SER	-	EXPRESSION TAG	UNP Q6WVP6
A	-8	GLY	-	EXPRESSION TAG	UNP Q6WVP6
A	-7	LEU	-	EXPRESSION TAG	UNP Q6WVP6
A	-6	VAL	-	EXPRESSION TAG	UNP Q6WVP6
A	-5	PRO	-	EXPRESSION TAG	UNP Q6WVP6
A	-4	ARG	-	EXPRESSION TAG	UNP Q6WVP6
A	-3	GLY	-	EXPRESSION TAG	UNP Q6WVP6
A	-2	SER	-	EXPRESSION TAG	UNP Q6WVP6
A	-1	HIS	-	EXPRESSION TAG	UNP Q6WVP6
A	0	MET	-	EXPRESSION TAG	UNP Q6WVP6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	MET	-	EXPRESSION TAG	UNP Q6WVP6
B	-19	GLY	-	EXPRESSION TAG	UNP Q6WVP6
B	-18	SER	-	EXPRESSION TAG	UNP Q6WVP6
B	-17	SER	-	EXPRESSION TAG	UNP Q6WVP6
B	-16	HIS	-	EXPRESSION TAG	UNP Q6WVP6
B	-15	HIS	-	EXPRESSION TAG	UNP Q6WVP6
B	-14	HIS	-	EXPRESSION TAG	UNP Q6WVP6
B	-13	HIS	-	EXPRESSION TAG	UNP Q6WVP6
B	-12	HIS	-	EXPRESSION TAG	UNP Q6WVP6
B	-11	HIS	-	EXPRESSION TAG	UNP Q6WVP6
B	-10	SER	-	EXPRESSION TAG	UNP Q6WVP6
B	-9	SER	-	EXPRESSION TAG	UNP Q6WVP6
B	-8	GLY	-	EXPRESSION TAG	UNP Q6WVP6
B	-7	LEU	-	EXPRESSION TAG	UNP Q6WVP6
B	-6	VAL	-	EXPRESSION TAG	UNP Q6WVP6
B	-5	PRO	-	EXPRESSION TAG	UNP Q6WVP6
B	-4	ARG	-	EXPRESSION TAG	UNP Q6WVP6
B	-3	GLY	-	EXPRESSION TAG	UNP Q6WVP6
B	-2	SER	-	EXPRESSION TAG	UNP Q6WVP6
B	-1	HIS	-	EXPRESSION TAG	UNP Q6WVP6
B	0	MET	-	EXPRESSION TAG	UNP Q6WVP6
C	-20	MET	-	EXPRESSION TAG	UNP Q6WVP6
C	-19	GLY	-	EXPRESSION TAG	UNP Q6WVP6
C	-18	SER	-	EXPRESSION TAG	UNP Q6WVP6
C	-17	SER	-	EXPRESSION TAG	UNP Q6WVP6
C	-16	HIS	-	EXPRESSION TAG	UNP Q6WVP6
C	-15	HIS	-	EXPRESSION TAG	UNP Q6WVP6
C	-14	HIS	-	EXPRESSION TAG	UNP Q6WVP6
C	-13	HIS	-	EXPRESSION TAG	UNP Q6WVP6
C	-12	HIS	-	EXPRESSION TAG	UNP Q6WVP6
C	-11	HIS	-	EXPRESSION TAG	UNP Q6WVP6
C	-10	SER	-	EXPRESSION TAG	UNP Q6WVP6
C	-9	SER	-	EXPRESSION TAG	UNP Q6WVP6
C	-8	GLY	-	EXPRESSION TAG	UNP Q6WVP6
C	-7	LEU	-	EXPRESSION TAG	UNP Q6WVP6
C	-6	VAL	-	EXPRESSION TAG	UNP Q6WVP6
C	-5	PRO	-	EXPRESSION TAG	UNP Q6WVP6
C	-4	ARG	-	EXPRESSION TAG	UNP Q6WVP6
C	-3	GLY	-	EXPRESSION TAG	UNP Q6WVP6
C	-2	SER	-	EXPRESSION TAG	UNP Q6WVP6
C	-1	HIS	-	EXPRESSION TAG	UNP Q6WVP6
C	0	MET	-	EXPRESSION TAG	UNP Q6WVP6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-20	MET	-	EXPRESSION TAG	UNP Q6WVP6
D	-19	GLY	-	EXPRESSION TAG	UNP Q6WVP6
D	-18	SER	-	EXPRESSION TAG	UNP Q6WVP6
D	-17	SER	-	EXPRESSION TAG	UNP Q6WVP6
D	-16	HIS	-	EXPRESSION TAG	UNP Q6WVP6
D	-15	HIS	-	EXPRESSION TAG	UNP Q6WVP6
D	-14	HIS	-	EXPRESSION TAG	UNP Q6WVP6
D	-13	HIS	-	EXPRESSION TAG	UNP Q6WVP6
D	-12	HIS	-	EXPRESSION TAG	UNP Q6WVP6
D	-11	HIS	-	EXPRESSION TAG	UNP Q6WVP6
D	-10	SER	-	EXPRESSION TAG	UNP Q6WVP6
D	-9	SER	-	EXPRESSION TAG	UNP Q6WVP6
D	-8	GLY	-	EXPRESSION TAG	UNP Q6WVP6
D	-7	LEU	-	EXPRESSION TAG	UNP Q6WVP6
D	-6	VAL	-	EXPRESSION TAG	UNP Q6WVP6
D	-5	PRO	-	EXPRESSION TAG	UNP Q6WVP6
D	-4	ARG	-	EXPRESSION TAG	UNP Q6WVP6
D	-3	GLY	-	EXPRESSION TAG	UNP Q6WVP6
D	-2	SER	-	EXPRESSION TAG	UNP Q6WVP6
D	-1	HIS	-	EXPRESSION TAG	UNP Q6WVP6
D	0	MET	-	EXPRESSION TAG	UNP Q6WVP6

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Mn 2 2	0	0
2	A	2	Total Mn 2 2	0	0
2	D	2	Total Mn 2 2	0	0
2	C	2	Total Mn 2 2	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		

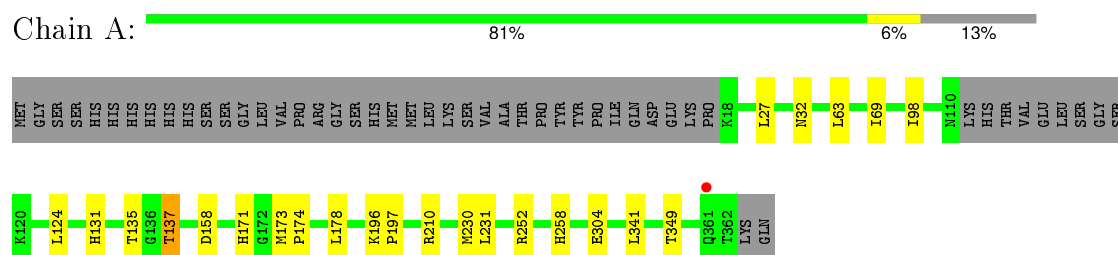
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	195	Total	O	0	0
			195	195		
4	B	172	Total	O	0	0
			172	172		
4	C	118	Total	O	0	0
			118	118		
4	D	10	Total	O	0	0
			10	10		

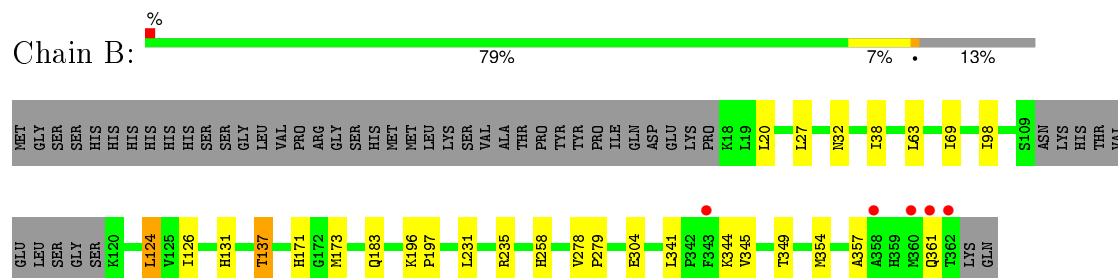
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

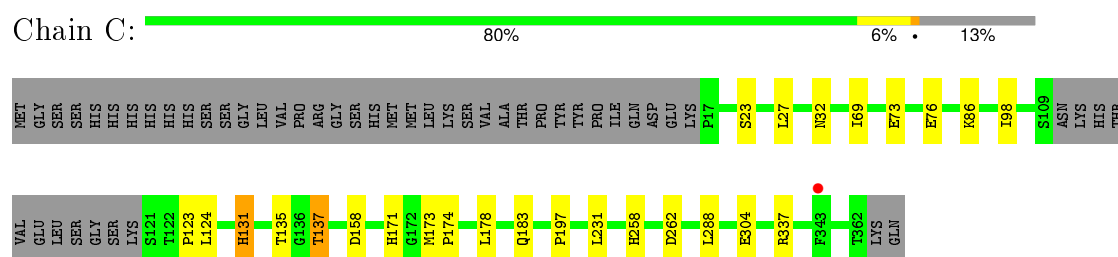
#### • Molecule 1: Arginase



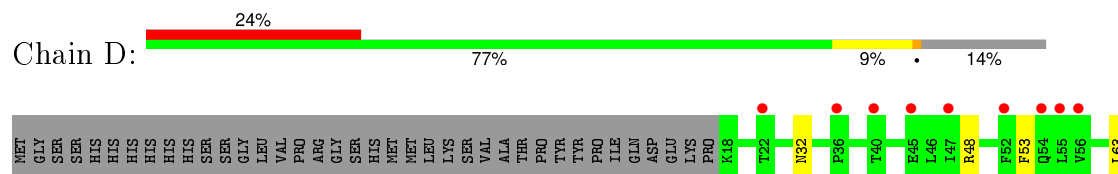
#### • Molecule 1: Arginase

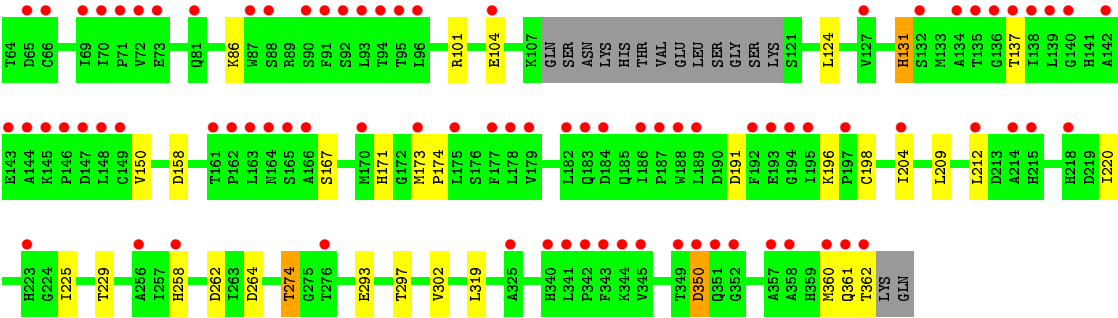


#### • Molecule 1: Arginase



#### • Molecule 1: Arginase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	178.38Å 178.38Å 178.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.48 – 2.50 49.47 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.48-2.50) 99.7 (49.47-2.50)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.35 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.177 , 0.213 0.173 , 0.213	Depositor DCC
$R_{free}$ test set	3301 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 43.5	EDS
Estimated twinning fraction	0.023 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 65285 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10810	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.83% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, MN

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	0.23	0/2631	0.43	0/3565
1	B	0.23	0/2623	0.43	0/3554
1	C	0.22	0/2622	0.43	0/3554
1	D	0.21	0/2593	0.40	0/3515
All	All	0.22	0/10469	0.42	0/14188

There are no bond length outliers.

There are no bond angle outliers.

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	A	2581	0	2597	8	0
1	B	2573	0	2591	12	0
1	C	2571	0	2586	11	0
1	D	2546	0	2560	19	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	12	0	16	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	12	0	16	0	0
3	C	12	0	16	2	0
4	A	195	0	0	1	0
4	B	172	0	0	1	0
4	C	118	0	0	0	0
4	D	10	0	0	0	0
All	All	10810	0	10382	50	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.

All (50) 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:344:LYS:NZ	1:B:345:VAL:O	2.32	0.62
1:D:264:ASP:HB3	1:D:274:THR:HG21	1.83	0.61
1:D:158:ASP:HB3	1:D:174:PRO:HD2	1.83	0.60
1:B:137:THR:HG23	1:B:258:HIS:HE1	1.70	0.57
1:D:137:THR:HG22	1:D:258:HIS:HE1	1.71	0.56
1:C:86:LYS:HG2	3:C:403:GOL:H31	1.91	0.53
1:A:137:THR:HG23	1:A:258:HIS:HE1	1.75	0.52
1:D:209:LEU:HB2	1:D:229:THR:HG22	1.92	0.52
1:B:137:THR:HG21	1:B:304:GLU:OE2	2.09	0.52
1:C:69:ILE:HD13	1:C:98:ILE:HA	1.91	0.51
1:A:69:ILE:HD13	1:A:98:ILE:HA	1.93	0.51
1:D:101:ARG:NH1	1:D:104:GLU:OE2	2.42	0.50
1:D:361:GLN:HG3	1:D:362:THR:H	1.77	0.50
1:C:137:THR:HG21	1:C:304:GLU:OE2	2.12	0.49
1:A:137:THR:HG21	1:A:304:GLU:OE2	2.13	0.49
1:D:220:ILE:HG23	1:D:225:ILE:HG23	1.95	0.48
1:B:235:ARG:NH1	4:B:664:HOH:O	2.46	0.48
1:D:124:LEU:HB3	1:D:302:VAL:HG22	1.96	0.48
1:D:173:MET:N	1:D:174:PRO:HD3	2.28	0.48
1:B:137:THR:HG23	1:B:258:HIS:CE1	2.49	0.47
1:D:131:HIS:CE1	1:D:262:ASP:HB2	2.49	0.47
1:C:183:GLN:HG3	1:C:197:PRO:HG3	1.96	0.47
1:B:69:ILE:HD13	1:B:98:ILE:HA	1.96	0.47
1:C:137:THR:HG23	1:C:258:HIS:HE1	1.81	0.45
1:D:293:GLU:O	1:D:297:THR:HG23	2.17	0.45
1:C:135:THR:HA	1:C:178:LEU:HD11	1.98	0.45
1:A:135:THR:HA	1:A:178:LEU:HD11	1.99	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:124:LEU:HD13	1:B:126:ILE:HD11	2.00	0.44
1:C:131:HIS:CE1	1:C:262:ASP:HB2	2.53	0.44
1:A:252:ARG:NH2	4:A:651:HOH:O	2.51	0.44
1:B:196:LYS:HA	1:B:197:PRO:HD3	1.90	0.44
1:A:158:ASP:HB3	1:A:174:PRO:HD2	1.99	0.43
1:B:20:LEU:HA	1:B:20:LEU:HD23	1.84	0.43
1:C:288:LEU:HD13	1:C:337:ARG:HD3	2.00	0.43
1:C:23:SER:O	1:C:123:PRO:HD2	2.19	0.43
1:D:191:ASP:N	1:D:191:ASP:OD2	2.51	0.42
1:A:196:LYS:HA	1:A:197:PRO:HD3	1.91	0.42
1:D:150:VAL:HB	1:D:204:ILE:HG23	2.00	0.42
1:D:319:LEU:HD23	1:D:319:LEU:HA	1.91	0.42
1:D:48:ARG:HA	1:D:53:PHE:HE2	1.85	0.41
1:D:209:LEU:HD13	1:D:212:LEU:HD21	2.02	0.41
1:A:210:ARG:HD3	1:A:230:MET:HG3	2.02	0.41
1:B:278:VAL:HA	1:B:279:PRO:HD3	1.93	0.41
1:B:183:GLN:HG3	1:B:197:PRO:HG2	2.02	0.41
1:D:350:ASP:N	1:D:350:ASP:OD2	2.52	0.41
1:C:73:GLU:OE1	3:C:404:GOL:O2	2.36	0.41
1:B:357:ALA:O	1:B:361:GLN:HG2	2.21	0.41
1:D:86:LYS:HE2	1:D:167:SER:O	2.21	0.41
1:D:196:LYS:HE2	1:D:198:CYS:HB3	2.03	0.40
1:C:158:ASP:HB3	1:C:174:PRO:HD2	2.03	0.40

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	333/385 (86%)	324 (97%)	8 (2%)	1 (0%)	46 68
1	B	332/385 (86%)	324 (98%)	7 (2%)	1 (0%)	46 68

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	332/385 (86%)	324 (98%)	7 (2%)	1 (0%)	46	68
1	D	328/385 (85%)	317 (97%)	11 (3%)	0	100	100
All	All	1325/1540 (86%)	1289 (97%)	33 (2%)	3 (0%)	52	75

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	173	MET
1	B	173	MET
1	C	173	MET

### 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	A	284/327 (87%)	274 (96%)	10 (4%)	43	70
1	B	283/327 (86%)	271 (96%)	12 (4%)	36	62
1	C	283/327 (86%)	275 (97%)	8 (3%)	51	78
1	D	279/327 (85%)	272 (98%)	7 (2%)	55	82
All	All	1129/1308 (86%)	1092 (97%)	37 (3%)	45	73

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	LEU
1	A	32	ASN
1	A	63	LEU
1	A	124	LEU
1	A	131	HIS
1	A	137	THR
1	A	171	HIS
1	A	231	LEU
1	A	341	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	349	THR
1	B	27	LEU
1	B	32	ASN
1	B	38	ILE
1	B	63	LEU
1	B	124	LEU
1	B	131	HIS
1	B	137	THR
1	B	171	HIS
1	B	231	LEU
1	B	341	LEU
1	B	349	THR
1	B	354	MET
1	C	27	LEU
1	C	32	ASN
1	C	76	GLU
1	C	124	LEU
1	C	131	HIS
1	C	137	THR
1	C	171	HIS
1	C	231	LEU
1	D	32	ASN
1	D	63	LEU
1	D	131	HIS
1	D	171	HIS
1	D	274	THR
1	D	350	ASP
1	D	360	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	359	HIS

### 5.3.3 RNA ⓘ

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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 8 are monoatomic - leaving 6 for Mogul analysis.

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	GOL	A	403	-	5,5,5	0.38	0	5,5,5	0.23	0
3	GOL	A	404	-	5,5,5	0.36	0	5,5,5	0.17	0
3	GOL	B	403	-	5,5,5	0.36	0	5,5,5	0.26	0
3	GOL	B	404	-	5,5,5	0.32	0	5,5,5	0.33	0
3	GOL	C	403	-	5,5,5	0.35	0	5,5,5	0.24	0
3	GOL	C	404	-	5,5,5	0.36	0	5,5,5	0.22	0

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	GOL	A	403	-	-	0/4/4/4	0/0/0/0
3	GOL	A	404	-	-	0/4/4/4	0/0/0/0
3	GOL	B	403	-	-	0/4/4/4	0/0/0/0
3	GOL	B	404	-	-	0/4/4/4	0/0/0/0
3	GOL	C	403	-	-	0/4/4/4	0/0/0/0
3	GOL	C	404	-	-	0/4/4/4	0/0/0/0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	403	GOL	1	0
3	C	404	GOL	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 ⓘ

### 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	A	336/385 (87%)	-0.36	1 (0%)	94 95	12, 20, 47, 75	0
1	B	335/385 (87%)	-0.18	5 (1%)	76 79	13, 21, 50, 73	5 (1%)
1	C	335/385 (87%)	-0.25	1 (0%)	94 95	19, 28, 48, 66	4 (1%)
1	D	332/385 (86%)	1.45	93 (28%)	1 0	50, 65, 81, 98	109 (32%)
All	All	1338/1540 (86%)	0.16	100 (7%)	17 19	12, 28, 73, 98	118 (8%)

All (100) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	360	MET	6.8
1	D	357	ALA	6.5
1	D	350	ASP	6.2
1	D	144	ALA	6.0
1	D	96	LEU	5.7
1	D	194	GLY	5.1
1	D	139	LEU	5.1
1	D	166	ALA	4.8
1	B	361	GLN	4.5
1	D	177	PHE	4.5
1	D	92	SER	4.5
1	D	146	PRO	4.4
1	D	351	GLN	4.3
1	D	137	THR	4.3
1	D	70	ILE	4.3
1	D	195	ILE	4.3
1	D	358	ALA	4.3
1	D	182	LEU	4.2
1	D	361	GLN	4.2
1	D	192	PHE	4.1
1	D	72	VAL	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	132	SER	4.0
1	D	22	THR	4.0
1	D	136	GLY	4.0
1	D	161	THR	3.9
1	D	344	LYS	3.9
1	D	162	PRO	3.8
1	D	91	PHE	3.8
1	D	184	ASP	3.7
1	D	164	ASN	3.6
1	D	352	GLY	3.5
1	D	140	GLY	3.5
1	D	165	SER	3.5
1	D	193	GLU	3.4
1	D	276	THR	3.4
1	D	349	THR	3.4
1	D	143	GLU	3.4
1	D	340	HIS	3.4
1	D	175	LEU	3.3
1	D	95	THR	3.3
1	D	142	ALA	3.3
1	D	189	LEU	3.2
1	D	212	LEU	3.2
1	D	362	THR	3.2
1	D	183	GLN	3.2
1	D	88	SER	3.2
1	D	345	VAL	3.1
1	D	135	THR	3.1
1	B	358	ALA	3.1
1	D	188	TRP	3.1
1	A	361	GLN	3.1
1	D	170	MET	3.0
1	D	163	LEU	3.0
1	D	341	LEU	3.0
1	D	47	ILE	3.0
1	D	93	LEU	2.9
1	D	138	ILE	2.9
1	D	186	ILE	2.9
1	D	66	CYS	2.9
1	D	71	PRO	2.8
1	B	362	THR	2.8
1	D	69	ILE	2.8
1	D	218	HIS	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	36	PRO	2.7
1	D	343	PHE	2.7
1	D	55	LEU	2.7
1	D	134	ALA	2.7
1	D	204	ILE	2.7
1	D	40	THR	2.6
1	B	360	MET	2.6
1	D	214	ALA	2.6
1	D	342	PRO	2.6
1	D	65	ASP	2.5
1	D	179	VAL	2.4
1	D	87	TRP	2.4
1	D	258	HIS	2.4
1	D	256	ALA	2.4
1	D	56	VAL	2.3
1	D	90	SER	2.3
1	D	149	CYS	2.3
1	C	343	PHE	2.3
1	D	187	PRO	2.3
1	D	148	LEU	2.3
1	D	215	HIS	2.3
1	D	223	HIS	2.2
1	D	94	THR	2.2
1	D	127	VAL	2.2
1	D	73	GLU	2.2
1	D	104	GLU	2.2
1	D	178	LEU	2.2
1	D	147	ASP	2.1
1	D	81	GLN	2.1
1	D	145	LYS	2.1
1	D	325	ALA	2.1
1	D	173	MET	2.1
1	B	343	PHE	2.1
1	D	52	PHE	2.1
1	D	197	PRO	2.0
1	D	54	GLN	2.0
1	D	45	GLU	2.0

## 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 [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	GOL	A	403	6/6	0.93	0.17	1.92	29,33,35,39	1
2	MN	B	401	1/1	0.99	0.15	1.37	17,17,17,17	0
3	GOL	C	403	6/6	0.94	0.18	1.33	32,41,46,49	1
3	GOL	B	404	6/6	0.96	0.14	0.48	25,29,40,51	0
2	MN	C	401	1/1	1.00	0.10	-1.04	23,23,23,23	0
2	MN	C	402	1/1	1.00	0.11	-1.09	25,25,25,25	0
2	MN	A	401	1/1	1.00	0.11	-1.37	17,17,17,17	0
2	MN	B	402	1/1	0.99	0.12	-1.61	16,16,16,16	0
2	MN	A	402	1/1	0.99	0.10	-2.28	15,15,15,15	0
2	MN	D	401	1/1	0.95	0.06	-4.13	59,59,59,59	0
2	MN	D	400	1/1	0.97	0.04	-9.08	57,57,57,57	0
3	GOL	C	404	6/6	0.93	0.20	-	35,38,48,48	0
3	GOL	B	403	6/6	0.90	0.20	-	29,35,41,48	0
3	GOL	A	404	6/6	0.93	0.15	-	28,31,36,41	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.