



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:50 PM GMT

PDB ID : 4Q41  
Title : Crystal structure of Schistosoma mansoni arginase in complex with L-lysine  
Authors : Hai, Y.; Edwards, J.E.; Van Zandt, M.C.; Hoffmann, K.F.; Christianson, D.W.  
Deposited on : 2014-04-12  
Resolution : 2.20 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

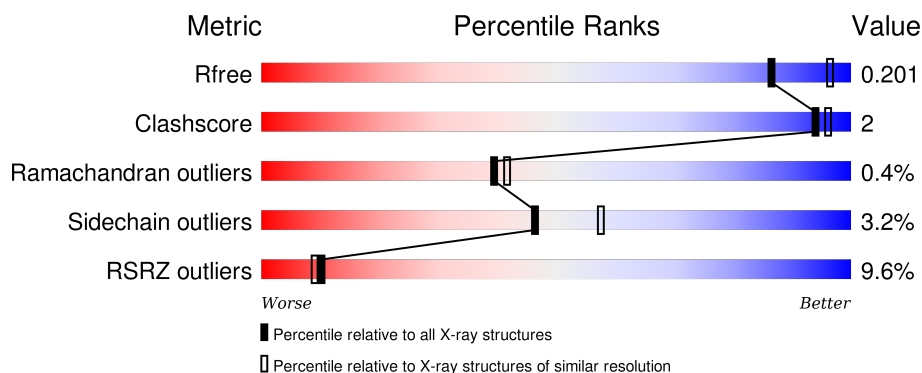
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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>3%</div> <div>81% 5% 14%</div> </div>
1	B	385	<div> <div>%</div> <div>82% 5% • 13%</div> </div>
1	C	385	<div> <div>4%</div> <div>80% 6% 14%</div> </div>
1	D	385	<div> <div>25%</div> <div>78% 6% • 15%</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	MN	A	402	-	-	-	X
2	MN	B	402	-	-	-	X
3	LYS	A	403	-	-	-	X
4	GOL	A	404	-	-	-	X
4	GOL	A	405	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10684 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	333	Total	C	N	O	S	0	0	0
			2557	1612	447	483	15			
1	B	336	Total	C	N	O	S	0	1	0
			2586	1627	452	492	15			
1	C	333	Total	C	N	O	S	0	1	0
			2558	1612	446	484	16			
1	D	329	Total	C	N	O	S	0	0	0
			2522	1590	440	478	14			

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...*

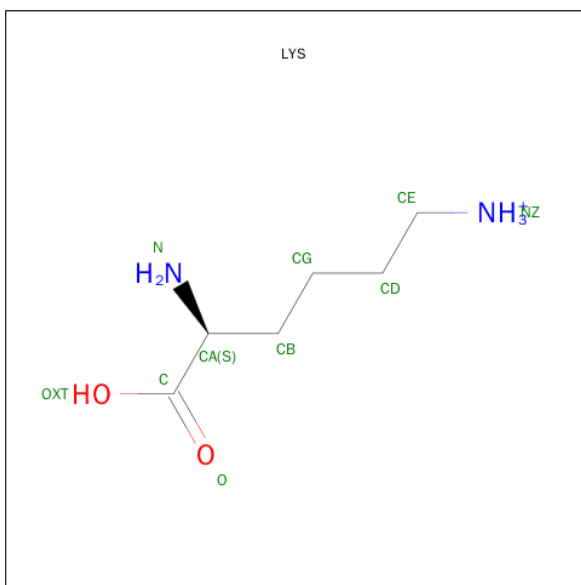
*Continued from previous page...*

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 LYSINE (three-letter code: LYS) (formula: C<sub>6</sub>H<sub>15</sub>N<sub>2</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			10	6	2	2		
3	B	1	Total	C	N	O	0	0
			10	6	2	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 5 is water.

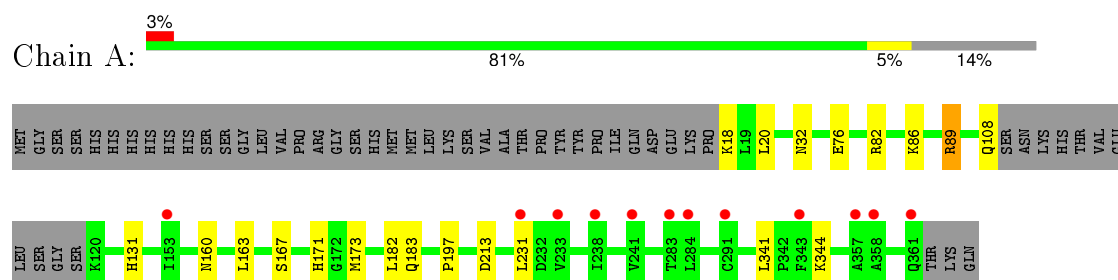
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	155	Total	O	0	0
			155	155		
5	B	138	Total	O	0	0
			138	138		
5	C	82	Total	O	0	0
			82	82		
5	D	22	Total	O	0	0
			22	22		



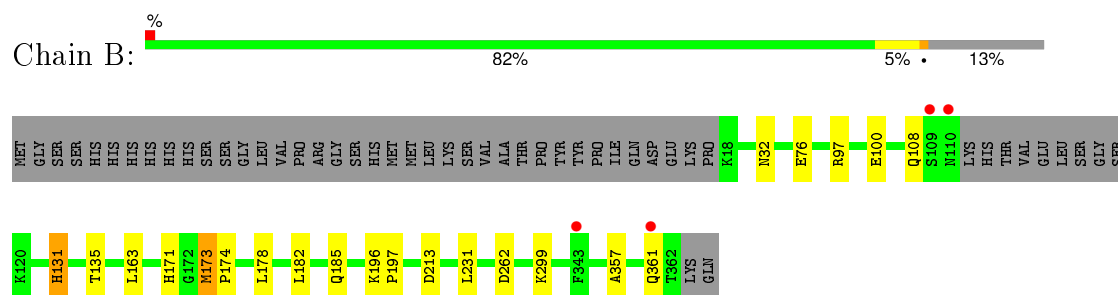
### 3 Residue-property plots [i](#)

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.

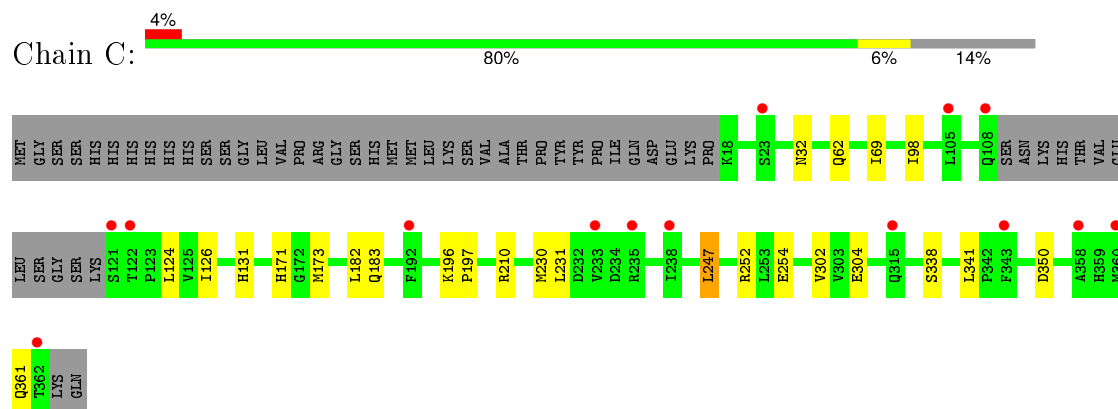
#### • Molecule 1: Arginase



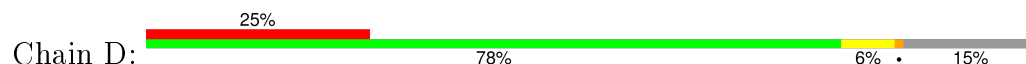
#### • Molecule 1: Arginase

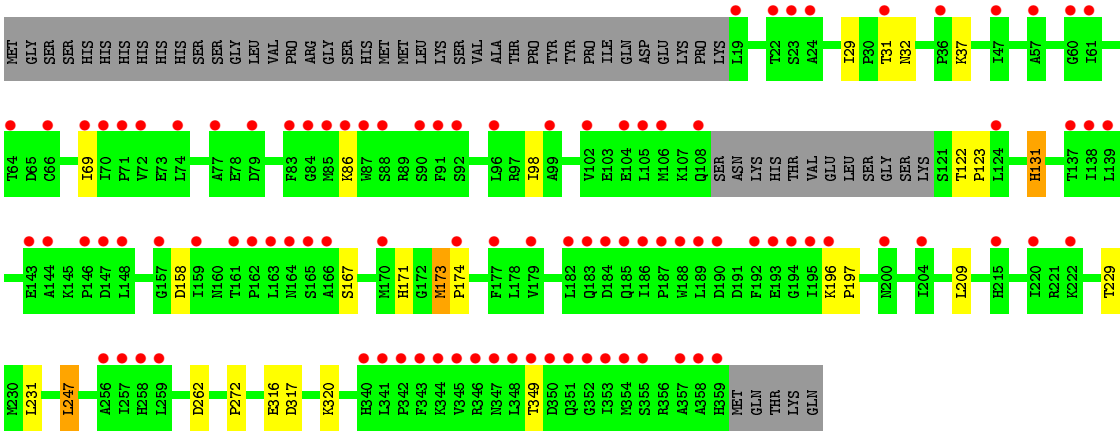


#### • Molecule 1: Arginase



#### • Molecule 1: Arginase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	177.45Å 177.45Å 177.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.22 – 2.20 49.22 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.22-2.20) 100.0 (49.22-2.20)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.21 (at 2.20Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.3_1479)	Depositor
R, $R_{free}$	0.172 , 0.201 0.173 , 0.201	Depositor DCC
$R_{free}$ test set	4732 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.8	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 42.0	EDS
Estimated twinning fraction	0.033 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 94297 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10684	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.68% 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	0.45	0/2604	0.58	0/3528
1	B	0.44	0/2633	0.58	0/3568
1	C	0.36	0/2608	0.50	0/3535
1	D	0.32	0/2569	0.45	0/3484
All	All	0.40	0/10414	0.53	0/14115

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	2557	0	2574	8	0
1	B	2586	0	2595	8	0
1	C	2558	0	2573	9	0
1	D	2522	0	2531	13	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	10	0	12	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	10	0	12	1	0
4	A	12	0	16	0	0
4	B	12	0	16	0	0
4	C	12	0	16	0	0
5	A	155	0	0	2	0
5	B	138	0	0	0	0
5	C	82	0	0	0	0
5	D	22	0	0	0	0
All	All	10684	0	10345	38	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 (38) 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:D:131:HIS:CE1	1:D:262:ASP:HB2	2.37	0.59
1:B:213[A]:ASP:OD2	3:B:403:LYS:N	2.35	0.59
1:D:158:ASP:HB3	1:D:174:PRO:HD2	1.90	0.53
1:C:69:ILE:HD13	1:C:98:ILE:HA	1.91	0.53
1:C:210:ARG:HD3	1:C:230:MET:HG3	1.92	0.50
1:A:160:ASN:ND2	3:A:403:LYS:OXT	2.45	0.49
1:D:173:MET:N	1:D:174:PRO:HD3	2.27	0.49
1:D:69:ILE:HD13	1:D:98:ILE:HA	1.95	0.49
1:D:316:GLU:HG2	1:D:320:LYS:HE3	1.95	0.48
1:D:29:ILE:HG22	1:D:31:THR:HG23	1.95	0.47
1:C:252:ARG:NH1	1:C:254:GLU:OE1	2.48	0.47
1:B:357:ALA:O	1:B:361:GLN:HG2	2.15	0.46
1:A:183:GLN:HG3	1:A:197:PRO:HG2	1.98	0.46
3:A:403:LYS:NZ	5:A:655:HOH:O	2.46	0.45
1:B:76:GLU:CD	1:B:76:GLU:H	2.20	0.45
1:D:86:LYS:HE2	1:D:167:SER:O	2.16	0.45
1:D:209:LEU:HB2	1:D:229:THR:HG22	1.99	0.45
1:A:82:ARG:NH2	1:A:89:ARG:HG3	2.32	0.44
1:A:213:ASP:HB3	5:A:625:HOH:O	2.18	0.44
1:A:18:LYS:HE3	1:A:18:LYS:HB2	1.84	0.44
1:A:86:LYS:NZ	1:A:167:SER:O	2.49	0.42
1:D:196:LYS:HA	1:D:197:PRO:HD3	1.92	0.42
1:B:131:HIS:CE1	1:B:262:ASP:HB2	2.54	0.42
1:C:196:LYS:HA	1:C:197:PRO:HD3	1.79	0.42
1:D:122:THR:HA	1:D:123:PRO:HD3	1.91	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:20:LEU:HD12	1:A:20:LEU:HA	1.87	0.41
1:C:247:LEU:HD12	1:C:247:LEU:HA	1.89	0.41
1:A:163:LEU:HA	1:A:163:LEU:HD13	1.86	0.41
1:B:196:LYS:HA	1:B:197:PRO:HD3	1.92	0.41
1:B:97:ARG:NE	1:C:350:ASP:O	2.53	0.41
1:D:158:ASP:O	1:D:174:PRO:HD2	2.21	0.41
1:D:247:LEU:HD12	1:D:247:LEU:HA	1.84	0.41
1:C:126:ILE:O	1:C:304:GLU:HA	2.21	0.40
1:C:124:LEU:HB3	1:C:302:VAL:HG22	2.02	0.40
1:B:173:MET:N	1:B:174:PRO:CD	2.85	0.40
1:B:135:THR:HA	1:B:178:LEU:HD11	2.03	0.40
1:D:272:PRO:HB2	1:D:317:ASP:HB3	2.04	0.40
1:C:183:GLN:HG3	1:C:197:PRO:HG2	2.04	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	329/385 (86%)	322 (98%)	6 (2%)	1 (0%)	46	50
1	B	333/385 (86%)	327 (98%)	5 (2%)	1 (0%)	46	50
1	C	330/385 (86%)	323 (98%)	5 (2%)	2 (1%)	30	29
1	D	325/385 (84%)	315 (97%)	9 (3%)	1 (0%)	46	50
All	All	1317/1540 (86%)	1287 (98%)	25 (2%)	5 (0%)	39	42

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	361	GLN
1	A	173	MET

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	173	MET
1	C	173	MET
1	D	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	280/327 (86%)	270 (96%)	10 (4%)	42	52
1	B	284/327 (87%)	274 (96%)	10 (4%)	43	53
1	C	281/327 (86%)	272 (97%)	9 (3%)	46	57
1	D	276/327 (84%)	269 (98%)	7 (2%)	55	67
All	All	1121/1308 (86%)	1085 (97%)	36 (3%)	46	57

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	32	ASN
1	A	76	GLU
1	A	89	ARG
1	A	108	GLN
1	A	131	HIS
1	A	171	HIS
1	A	182	LEU
1	A	231	LEU
1	A	341	LEU
1	A	344	LYS
1	B	32	ASN
1	B	100	GLU
1	B	108	GLN
1	B	131	HIS
1	B	163	LEU
1	B	171	HIS
1	B	182	LEU
1	B	185	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	231	LEU
1	B	299	LYS
1	C	32	ASN
1	C	62	GLN
1	C	131	HIS
1	C	171	HIS
1	C	182	LEU
1	C	231	LEU
1	C	247	LEU
1	C	338	SER
1	C	341	LEU
1	D	32	ASN
1	D	37	LYS
1	D	131	HIS
1	D	171	HIS
1	D	231	LEU
1	D	247	LEU
1	D	349	THR

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	200	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 carbohydrates in this entry.

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

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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	LYS	A	403	-	6,9,9	0.28	0	4,10,10	0.70	0
4	GOL	A	404	-	5,5,5	0.45	0	5,5,5	0.46	0
4	GOL	A	405	-	5,5,5	0.22	0	5,5,5	0.98	0
3	LYS	B	403	-	6,9,9	0.27	0	4,10,10	0.37	0
4	GOL	B	404	-	5,5,5	0.32	0	5,5,5	0.66	0
4	GOL	B	405	-	5,5,5	0.30	0	5,5,5	0.48	0
4	GOL	C	403	-	5,5,5	0.36	0	5,5,5	0.31	0
4	GOL	C	404	-	5,5,5	0.32	0	5,5,5	0.20	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	LYS	A	403	-	-	0/5/9/9	0/0/0/0
4	GOL	A	404	-	-	0/4/4/4	0/0/0/0
4	GOL	A	405	-	-	0/4/4/4	0/0/0/0
3	LYS	B	403	-	-	0/5/9/9	0/0/0/0
4	GOL	B	404	-	-	0/4/4/4	0/0/0/0
4	GOL	B	405	-	-	0/4/4/4	0/0/0/0
4	GOL	C	403	-	-	0/4/4/4	0/0/0/0
4	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 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	403	LYS	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	403	LYS	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	333/385 (86%)	0.31	12 (3%) 46 45	20, 31, 59, 88	0
1	B	336/385 (87%)	0.03	4 (1%) 81 80	23, 34, 57, 84	0
1	C	333/385 (86%)	0.17	14 (4%) 40 39	38, 49, 70, 84	1 (0%)
1	D	329/385 (85%)	1.34	98 (29%) 1 0	46, 71, 93, 105	13 (3%)
All	All	1331/1540 (86%)	0.46	128 (9%) 10 9	20, 44, 84, 105	14 (1%)

All (128) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	345	VAL	7.6
1	D	343	PHE	7.5
1	D	195	ILE	5.8
1	D	184	ASP	5.8
1	D	91	PHE	5.7
1	D	166	ALA	5.5
1	D	139	LEU	5.2
1	D	350	ASP	5.2
1	D	351	GLN	4.9
1	D	164	ASN	4.9
1	D	349	THR	4.8
1	C	362	THR	4.8
1	D	352	GLY	4.7
1	D	358	ALA	4.6
1	D	22	THR	4.6
1	D	344	LYS	4.6
1	D	182	LEU	4.5
1	D	341	LEU	4.3
1	D	354	MET	4.3
1	D	192	PHE	4.3
1	D	90	SER	4.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	177	PHE	4.2
1	D	138	ILE	3.9
1	D	346	ARG	3.9
1	D	96	LEU	3.8
1	D	189	LEU	3.8
1	D	146	PRO	3.7
1	D	163	LEU	3.7
1	A	343	PHE	3.6
1	D	161	THR	3.6
1	D	186	ILE	3.6
1	D	188	TRP	3.5
1	D	353	ILE	3.4
1	D	256	ALA	3.3
1	C	233	VAL	3.3
1	C	343	PHE	3.3
1	D	162	PRO	3.3
1	D	108	GLN	3.2
1	D	196	LYS	3.2
1	D	355	SER	3.2
1	D	148	LEU	3.2
1	C	122	THR	3.2
1	D	36	PRO	3.2
1	D	147	ASP	3.2
1	D	187	PRO	3.2
1	C	105	LEU	3.2
1	A	238	ILE	3.2
1	D	347	ASN	3.1
1	D	102	VAL	3.1
1	D	74	LEU	3.1
1	D	342	PRO	3.1
1	D	348	LEU	3.0
1	D	92	SER	3.0
1	D	87	TRP	3.0
1	D	357	ALA	3.0
1	D	183	GLN	2.9
1	D	215	HIS	2.9
1	D	190	ASP	2.9
1	D	144	ALA	2.9
1	D	220	ILE	2.8
1	D	72	VAL	2.8
1	A	357	ALA	2.8
1	D	170	MET	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	358	ALA	2.8
1	D	185	GLN	2.7
1	D	83	PHE	2.7
1	D	24	ALA	2.7
1	D	70	ILE	2.7
1	A	361	GLN	2.7
1	D	165	SER	2.7
1	D	84	GLY	2.6
1	A	241	VAL	2.6
1	D	60	GLY	2.6
1	D	57	ALA	2.6
1	A	283	THR	2.5
1	D	69	ILE	2.5
1	D	85	MET	2.5
1	D	99	ALA	2.5
1	B	110	ASN	2.5
1	D	79	ASP	2.5
1	D	61	ILE	2.5
1	A	358	ALA	2.4
1	D	157	GLY	2.4
1	B	343	PHE	2.4
1	D	222	LYS	2.4
1	D	104	GLU	2.4
1	C	192	PHE	2.4
1	D	143	GLU	2.4
1	A	233	VAL	2.3
1	B	109	SER	2.3
1	D	194	GLY	2.3
1	A	291	CYS	2.3
1	D	19	LEU	2.3
1	D	159	ILE	2.3
1	D	23	SER	2.3
1	C	108	GLN	2.3
1	D	105	LEU	2.3
1	D	193	GLU	2.3
1	C	23	SER	2.3
1	D	66	CYS	2.3
1	D	47	ILE	2.3
1	D	258	HIS	2.3
1	C	360	MET	2.2
1	B	361	GLN	2.2
1	D	86	LYS	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	121	SER	2.2
1	A	231	LEU	2.2
1	D	31	THR	2.2
1	D	340	HIS	2.2
1	C	235	ARG	2.2
1	D	106	MET	2.2
1	D	179	VAL	2.2
1	A	284	LEU	2.2
1	D	71	PRO	2.2
1	D	174	PRO	2.2
1	C	315	GLN	2.1
1	D	257	ILE	2.1
1	D	359	HIS	2.1
1	D	64	THR	2.1
1	D	137	THR	2.1
1	D	77	ALA	2.1
1	D	88	SER	2.1
1	C	238	ILE	2.1
1	D	124	LEU	2.1
1	D	259	LEU	2.1
1	A	153	ILE	2.0
1	D	204	ILE	2.0
1	D	200	ASN	2.0

## 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 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( $\text{\AA}^2$ )	Q<0.9
4	GOL	A	404	6/6	0.95	0.19	5.13	42,45,53,55	0
3	LYS	A	403	10/10	0.81	0.19	3.39	36,44,49,54	0
2	MN	A	402	1/1	1.00	0.19	2.91	25,25,25,25	0
2	MN	B	402	1/1	1.00	0.14	2.28	27,27,27,27	0
4	GOL	A	405	6/6	0.91	0.19	2.15	36,50,57,61	0
3	LYS	B	403	10/10	0.85	0.15	1.97	41,49,58,58	0
4	GOL	C	404	6/6	0.91	0.20	1.74	57,65,69,69	0
2	MN	B	401	1/1	1.00	0.14	1.38	28,28,28,28	0
4	GOL	B	405	6/6	0.95	0.14	0.32	41,51,58,65	0
2	MN	A	401	1/1	1.00	0.14	0.20	26,26,26,26	0
2	MN	C	402	1/1	1.00	0.10	-0.88	40,40,40,40	0
2	MN	C	401	1/1	0.99	0.08	-1.10	41,41,41,41	0
2	MN	D	400	1/1	0.96	0.10	-1.77	48,48,48,48	1
2	MN	D	401	1/1	0.99	0.06	-3.47	53,53,53,53	0
4	GOL	B	404	6/6	0.92	0.11	-	45,51,58,63	0
4	GOL	C	403	6/6	0.88	0.24	-	58,60,64,70	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.