



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 15, 2020 – 06:00 am BST

PDB ID : 4Q3T  
Title : Crystal structure of Schistosoma mansoni arginase in complex with inhibitor NOHA  
Authors : Hai, Y.; Edwards, J.E.; Van Zandt, M.C.; Hoffmann, K.F.; Christianson, D.W.  
Deposited on : 2014-04-12  
Resolution : 2.14 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

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

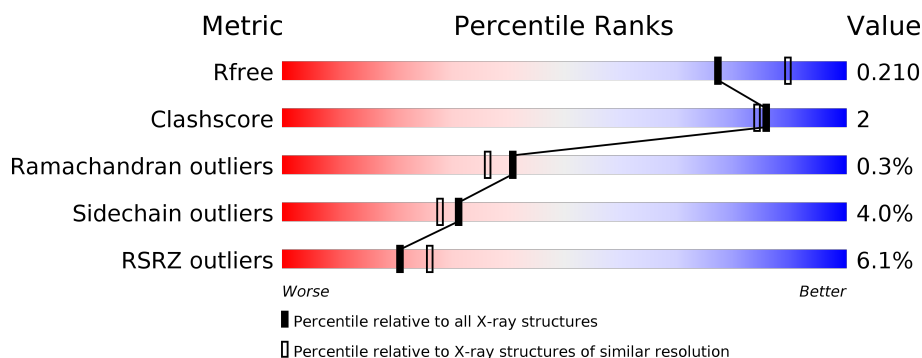
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.14 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	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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	385	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 82%, yellow 5%, grey 13%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>82%</span> <span>5%</span> <span>13%</span> </div> </div>
1	B	385	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, green 81%, yellow 6%, grey 13%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>81%</span> <span>6%</span> <span>13%</span> </div> </div>
1	C	385	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 82%, yellow 5%, grey 13%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>82%</span> <span>5%</span> <span>13%</span> </div> </div>
1	D	385	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 17%, green 76%, yellow 9%, grey 15%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>17%</span> <span>76%</span> <span>9%</span> <span>15%</span> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10985 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
			2583	1627	451	490	15			
1	B	335	Total	C	N	O	S	0	1	0
			2573	1621	448	489	15			
1	C	334	Total	C	N	O	S	0	1	0
			2565	1617	447	485	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).

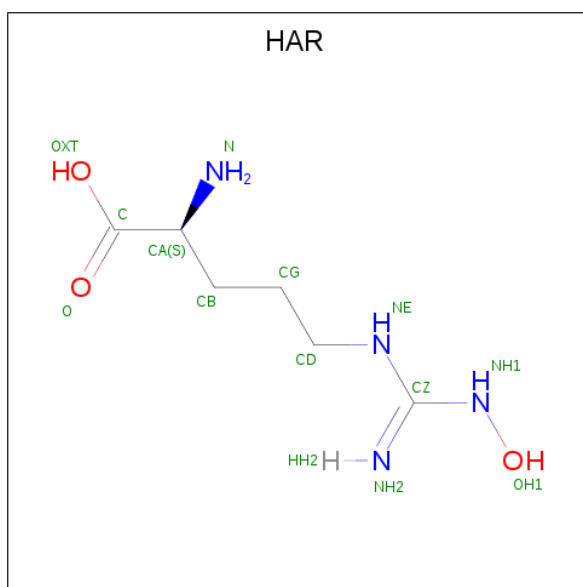
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Mn 2 2	0	0
2	A	2	Total Mn 2 2	0	0
2	D	2	Total Mn 2 2	0	0
2	C	2	Total Mn 2 2	0	0

- Molecule 3 is 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 N-OMEGA-HYDROXY-L-ARGININE (three-letter code: HAR) (formula:  $C_6H_{14}N_4O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	1
			26	12	8	6		
4	B	1	Total	C	N	O	0	1
			26	12	8	6		
4	C	1	Total	C	N	O	0	1
			26	12	8	6		
4	D	1	Total	C	N	O	0	0
			13	6	4	3		

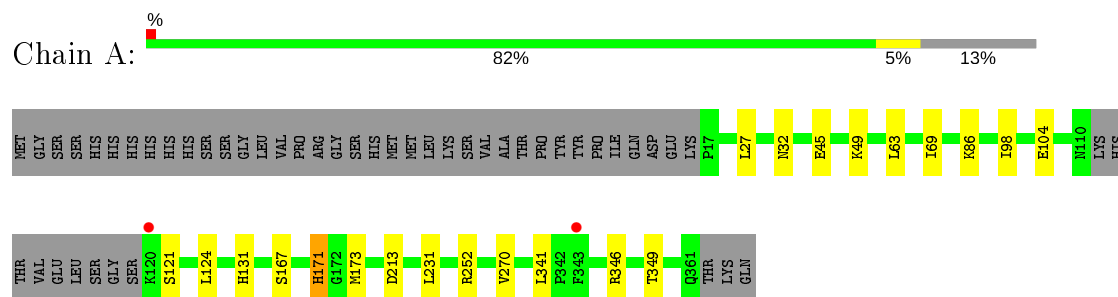
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	206	Total	O	0	0
			206	206		
5	B	218	Total	O	0	0
			218	218		
5	C	143	Total	O	0	0
			143	143		
5	D	40	Total	O	0	0
			40	40		

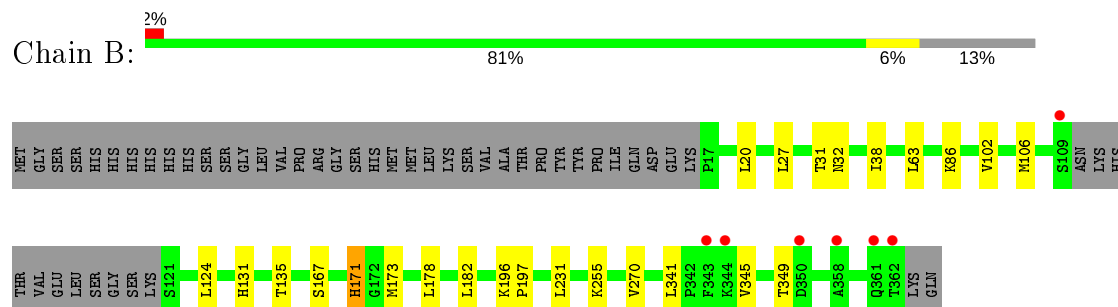
### 3 Residue-property plots [i](#)

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

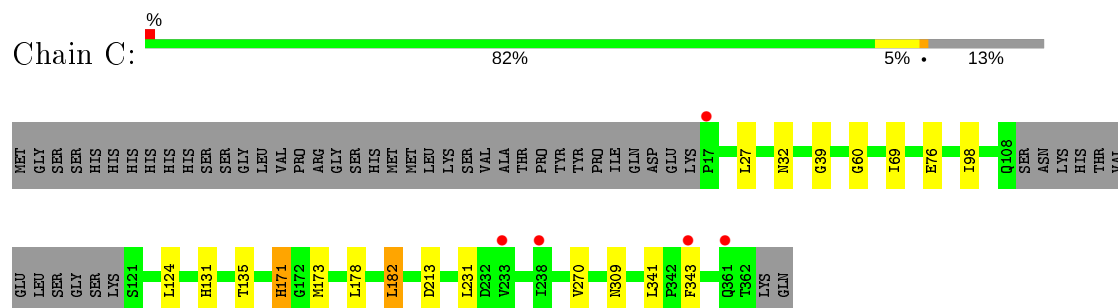
- Molecule 1: Arginase



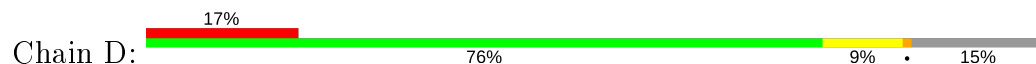
- Molecule 1: Arginase



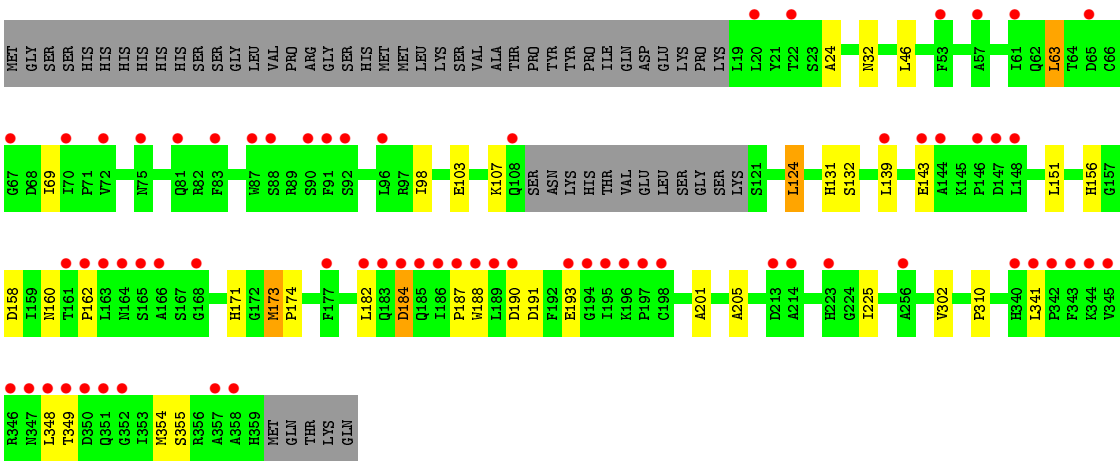
- Molecule 1: Arginase



- Molecule 1: Arginase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	177.76Å 177.76Å 177.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.30 – 2.14 49.30 – 2.14	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.30-2.14) 99.7 (49.30-2.14)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 2.14Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.176 , 0.209 0.177 , 0.210	Depositor DCC
$R_{free}$ test set	5117 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.7	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.024 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10985	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.70% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, HAR, 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	0.28	0/2634	0.48	0/3569
1	B	0.27	0/2624	0.48	0/3557
1	C	0.24	0/2616	0.46	0/3546
1	D	0.23	0/2569	0.43	0/3484
All	All	0.26	0/10443	0.46	0/14156

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	2583	0	2597	6	0
1	B	2573	0	2585	8	0
1	C	2565	0	2581	10	0
1	D	2522	0	2531	20	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	0	0
4	A	26	0	22	3	0
4	B	26	0	24	2	0
4	C	26	0	22	4	0
4	D	13	0	11	2	0
5	A	206	0	0	0	0
5	B	218	0	0	2	0
5	C	143	0	0	1	0
5	D	40	0	0	0	0
All	All	10985	0	10421	46	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 (46) 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:184:ASP:OD1	1:D:184:ASP:N	2.29	0.65
1:A:213[A]:ASP:OD2	4:A:405[A]:HAR:N	2.33	0.61
1:D:158:ASP:HB3	1:D:174:PRO:HD2	1.83	0.61
1:D:188:TRP:CH2	1:D:193:GLU:HG3	2.37	0.60
1:D:348:LEU:HD22	1:D:355:SER:HB3	1.84	0.59
1:C:213:ASP:OD2	4:C:405[B]:HAR:N	2.36	0.59
1:B:255:LYS:NZ	5:B:679:HOH:O	2.36	0.57
1:A:171:HIS:CE1	4:A:405[A]:HAR:HE	2.24	0.56
1:B:196:LYS:HD2	1:B:197:PRO:HD2	1.87	0.56
1:D:173:MET:N	1:D:174:PRO:HD3	2.21	0.55
1:D:190:ASP:HA	1:D:193:GLU:OE2	2.07	0.55
1:D:188:TRP:HH2	1:D:193:GLU:HG3	1.74	0.53
1:D:160:ASN:ND2	4:D:401:HAR:OXT	2.41	0.52
1:A:69:ILE:HD13	1:A:98:ILE:HA	1.91	0.52
1:C:69:ILE:HD13	1:C:98:ILE:HA	1.91	0.52
1:D:162:PRO:HG2	1:D:187:PRO:HD2	1.93	0.51
1:C:213:ASP:OD2	4:C:405[A]:HAR:N	2.43	0.51
1:C:171:HIS:CE1	4:C:405[B]:HAR:HE	2.30	0.50
1:D:103:GLU:OE2	1:D:107:LYS:NZ	2.44	0.49
1:D:69:ILE:HD13	1:D:98:ILE:HA	1.94	0.49
1:C:60:GLY:HA3	1:C:343:PHE:CZ	2.48	0.49
1:D:124:LEU:HB3	1:D:302:VAL:HG22	1.95	0.48
1:D:46:LEU:HD22	1:D:310:PRO:HG3	1.95	0.48

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:139:LEU:O	1:D:143:GLU:HG2	2.13	0.48
1:D:24:ALA:HB3	1:D:63:LEU:HD23	1.95	0.48
1:B:171:HIS:CE1	4:B:405[A]:HAR:HE	2.31	0.47
1:D:191:ASP:OD1	1:D:191:ASP:N	2.47	0.47
1:D:156:HIS:HB2	4:D:401:HAR:HH1	1.80	0.47
1:B:86:LYS:NZ	1:B:167:SER:O	2.48	0.47
1:B:102:VAL:O	1:B:106:MET:HG2	2.16	0.46
1:A:86:LYS:NZ	1:A:167:SER:O	2.47	0.45
4:A:405[A]:HAR:HD2	4:A:405[A]:HAR:HH1	1.61	0.45
1:C:135:THR:HA	1:C:178:LEU:HD11	1.99	0.44
1:B:135:THR:HA	1:B:178:LEU:HD11	1.99	0.44
4:B:405[A]:HAR:HH1	4:B:405[A]:HAR:HD2	1.44	0.43
1:A:45:GLU:HG3	1:A:49:LYS:HD3	2.00	0.43
4:C:405[B]:HAR:HD2	4:C:405[B]:HAR:HH1	1.53	0.43
1:D:151:LEU:HD23	1:D:205:ALA:HB3	2.01	0.42
1:C:182:LEU:HA	1:C:182:LEU:HD12	1.88	0.42
1:C:39:GLY:HA3	1:C:309:ASN:OD1	2.19	0.42
1:C:213:ASP:HB3	5:C:618:HOH:O	2.20	0.42
1:D:201:ALA:HB1	1:D:225:ILE:HG12	2.01	0.42
1:C:76:GLU:N	1:C:76:GLU:OE1	2.41	0.41
1:D:132:SER:HA	1:D:174:PRO:HG3	2.02	0.41
1:B:38:ILE:HG22	5:B:600:HOH:O	2.20	0.41
1:A:252:ARG:NE	1:B:20:LEU:HD21	2.37	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%)	327 (98%)	5 (2%)	1 (0%)	41 36
1	B	332/385 (86%)	326 (98%)	5 (2%)	1 (0%)	41 36

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	331/385 (86%)	323 (98%)	7 (2%)	1 (0%)	41	36
1	D	325/385 (84%)	315 (97%)	9 (3%)	1 (0%)	41	36
All	All	1321/1540 (86%)	1291 (98%)	26 (2%)	4 (0%)	41	36

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	173	MET
1	C	173	MET
1	D	173	MET
1	B	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%)	271 (95%)	13 (5%)	27	22
1	B	283/327 (86%)	270 (95%)	13 (5%)	27	22
1	C	282/327 (86%)	273 (97%)	9 (3%)	39	37
1	D	276/327 (84%)	266 (96%)	10 (4%)	35	32
All	All	1125/1308 (86%)	1080 (96%)	45 (4%)	31	28

All (45) 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	104	GLU
1	A	121	SER
1	A	124	LEU
1	A	131	HIS
1	A	171	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	231	LEU
1	A	270	VAL
1	A	341	LEU
1	A	346	ARG
1	A	349	THR
1	B	27	LEU
1	B	31	THR
1	B	32	ASN
1	B	63	LEU
1	B	124	LEU
1	B	131	HIS
1	B	171	HIS
1	B	182	LEU
1	B	231	LEU
1	B	270	VAL
1	B	341	LEU
1	B	345	VAL
1	B	349	THR
1	C	27	LEU
1	C	32	ASN
1	C	124	LEU
1	C	131	HIS
1	C	171	HIS
1	C	182	LEU
1	C	231	LEU
1	C	270	VAL
1	C	341	LEU
1	D	32	ASN
1	D	63	LEU
1	D	124	LEU
1	D	131	HIS
1	D	171	HIS
1	D	182	LEU
1	D	184	ASP
1	D	341	LEU
1	D	349	THR
1	D	354	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	200	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	160	ASN
1	D	185	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 21 ligands modelled in this entry, 8 are monoatomic - leaving 13 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	C	404	-	5,5,5	0.36	0	5,5,5	0.36	0
3	GOL	A	404	-	5,5,5	0.39	0	5,5,5	0.28	0
4	HAR	A	405[A]	2	7,12,12	2.72	2 (28%)	6,14,14	1.35	1 (16%)
4	HAR	D	401	2	7,12,12	2.50	2 (28%)	6,14,14	0.91	0
4	HAR	A	405[B]	2	7,12,12	2.66	2 (28%)	6,14,14	1.20	1 (16%)
3	GOL	B	404	-	5,5,5	0.39	0	5,5,5	0.23	0
4	HAR	C	405[B]	2	7,12,12	2.60	2 (28%)	6,14,14	1.53	1 (16%)
4	HAR	B	405[B]	2	7,12,12	2.66	2 (28%)	6,14,14	1.12	0
3	GOL	B	403	-	5,5,5	0.30	0	5,5,5	0.37	0
3	GOL	A	403	-	5,5,5	0.32	0	5,5,5	0.49	0
3	GOL	C	403	-	5,5,5	0.36	0	5,5,5	0.25	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	HAR	B	405[A]	2	7,12,12	2.73	2 (28%)	6,14,14	1.91	3 (50%)
4	HAR	C	405[A]	2	7,12,12	2.61	2 (28%)	6,14,14	1.23	1 (16%)

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	C	404	-	-	2/4/4/4	-
3	GOL	A	404	-	-	2/4/4/4	-
4	HAR	A	405[A]	2	-	1/7/13/13	-
4	HAR	D	401	2	-	0/7/13/13	-
4	HAR	A	405[B]	2	-	3/7/13/13	-
3	GOL	B	404	-	-	2/4/4/4	-
4	HAR	C	405[B]	2	-	1/7/13/13	-
4	HAR	B	405[B]	2	-	3/7/13/13	-
3	GOL	B	403	-	-	2/4/4/4	-
3	GOL	A	403	-	-	0/4/4/4	-
3	GOL	C	403	-	-	2/4/4/4	-
4	HAR	B	405[A]	2	-	2/7/13/13	-
4	HAR	C	405[A]	2	-	4/7/13/13	-

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	405[A]	HAR	CZ-NH1	6.68	1.50	1.34
4	B	405[A]	HAR	CZ-NH1	6.63	1.50	1.34
4	B	405[B]	HAR	CZ-NH1	6.61	1.50	1.34
4	A	405[B]	HAR	CZ-NH1	6.61	1.49	1.34
4	C	405[A]	HAR	CZ-NH1	6.34	1.49	1.34
4	C	405[B]	HAR	CZ-NH1	6.24	1.49	1.34
4	D	401	HAR	CZ-NH1	6.05	1.48	1.34
4	B	405[A]	HAR	OH1-NH1	-2.70	1.33	1.40
4	C	405[B]	HAR	OH1-NH1	-2.65	1.33	1.40
4	C	405[A]	HAR	OH1-NH1	-2.51	1.33	1.40
4	D	401	HAR	OH1-NH1	-2.45	1.33	1.40
4	A	405[A]	HAR	OH1-NH1	-2.45	1.33	1.40
4	B	405[B]	HAR	OH1-NH1	-2.29	1.34	1.40
4	A	405[B]	HAR	OH1-NH1	-2.25	1.34	1.40

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	405[A]	HAR	NH1-CZ-NE	-2.82	110.25	116.82
4	C	405[B]	HAR	NH1-CZ-NE	-2.74	110.46	116.82
4	B	405[A]	HAR	CD-NE-CZ	-2.37	119.18	123.50
4	A	405[A]	HAR	NH1-CZ-NE	-2.26	111.56	116.82
4	B	405[A]	HAR	CB-CG-CD	-2.23	105.37	112.05
4	C	405[A]	HAR	NH1-CZ-NE	-2.19	111.73	116.82
4	A	405[B]	HAR	NH1-CZ-NE	-2.07	112.00	116.82

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	404	GOL	O1-C1-C2-C3
4	A	405[A]	HAR	C-CA-CB-CG
3	A	404	GOL	O1-C1-C2-C3
3	B	403	GOL	O1-C1-C2-O2
3	B	403	GOL	O1-C1-C2-C3
3	A	404	GOL	O1-C1-C2-O2
4	B	405[A]	HAR	NE-CD-CG-CB
3	B	404	GOL	O1-C1-C2-C3
3	C	403	GOL	O1-C1-C2-C3
3	C	404	GOL	O1-C1-C2-O2
3	B	404	GOL	O1-C1-C2-O2
3	C	403	GOL	O1-C1-C2-O2
4	C	405[A]	HAR	NE-CD-CG-CB
4	A	405[B]	HAR	C-CA-CB-CG
4	C	405[A]	HAR	C-CA-CB-CG
4	C	405[B]	HAR	C-CA-CB-CG
4	B	405[B]	HAR	C-CA-CB-CG
4	B	405[B]	HAR	NE-CD-CG-CB
4	A	405[B]	HAR	NE-CD-CG-CB
4	A	405[B]	HAR	CG-CD-NE-CZ
4	C	405[A]	HAR	CG-CD-NE-CZ
4	B	405[B]	HAR	CG-CD-NE-CZ
4	B	405[A]	HAR	CG-CD-NE-CZ
4	C	405[A]	HAR	CA-CB-CG-CD

There are no ring outliers.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	405[A]	HAR	3	0
4	D	401	HAR	2	0
4	C	405[B]	HAR	3	0
4	B	405[A]	HAR	2	0
4	C	405[A]	HAR	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%)	-0.43	2 (0%) 89 91	10, 18, 41, 57	7 (2%)
1	B	335/385 (87%)	-0.16	7 (2%) 63 69	11, 18, 43, 66	6 (1%)
1	C	334/385 (86%)	-0.29	5 (1%) 73 78	18, 27, 47, 79	2 (0%)
1	D	329/385 (85%)	1.02	67 (20%) 1 1	31, 47, 62, 71	32 (9%)
All	All	1334/1540 (86%)	0.03	81 (6%) 21 26	10, 26, 58, 79	47 (3%)

All (81) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	350	ASP	6.9
1	D	346	ARG	6.7
1	D	357	ALA	5.9
1	D	348	LEU	5.7
1	D	347	ASN	5.5
1	D	343	PHE	5.4
1	D	351	GLN	5.2
1	D	341	LEU	4.9
1	D	108	GLN	4.9
1	D	195	ILE	4.8
1	D	182	LEU	4.7
1	D	163	LEU	4.3
1	D	349	THR	4.3
1	D	190	ASP	4.2
1	D	345	VAL	4.1
1	D	344	LYS	4.0
1	B	361	GLN	3.9
1	D	148	LEU	3.9
1	D	358	ALA	3.7
1	C	361	GLN	3.6
1	D	96	LEU	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	362	THR	3.6
1	D	194	GLY	3.6
1	B	343	PHE	3.5
1	B	109	SER	3.4
1	D	187	PRO	3.4
1	D	184	ASP	3.4
1	D	166	ALA	3.3
1	D	340	HIS	3.3
1	D	189	LEU	3.2
1	D	70	ILE	3.2
1	D	183	GLN	3.2
1	D	161	THR	3.2
1	D	91	PHE	3.1
1	D	342	PRO	3.0
1	D	186	ILE	3.0
1	D	165	SER	3.0
1	D	146	PRO	3.0
1	D	147	ASP	3.0
1	D	185	GLN	3.0
1	D	352	GLY	2.9
1	D	139	LEU	2.9
1	D	87	TRP	2.8
1	D	193	GLU	2.8
1	D	72	VAL	2.8
1	D	213	ASP	2.8
1	D	177	PHE	2.8
1	D	61	ILE	2.7
1	D	162	PRO	2.7
1	C	17	PRO	2.6
1	D	83	PHE	2.6
1	D	88	SER	2.6
1	D	90	SER	2.5
1	D	53	PHE	2.5
1	B	358	ALA	2.5
1	D	75	ASN	2.5
1	D	223	HIS	2.5
1	D	57	ALA	2.5
1	D	144	ALA	2.5
1	D	214	ALA	2.5
1	A	120	LYS	2.5
1	D	92	SER	2.4
1	B	350	ASP	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	343	PHE	2.4
1	D	198	CYS	2.4
1	D	188	TRP	2.3
1	A	343	PHE	2.3
1	D	168	GLY	2.3
1	D	256	ALA	2.3
1	D	65	ASP	2.3
1	C	233	VAL	2.3
1	D	143	GLU	2.2
1	D	81	GLN	2.2
1	D	22	THR	2.2
1	D	197	PRO	2.1
1	B	344	LYS	2.1
1	C	238	ILE	2.1
1	D	164	ASN	2.1
1	D	67	GLY	2.1
1	D	20	LEU	2.0
1	D	196	LYS	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. 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
4	HAR	D	401	13/13	0.86	0.21	41,49,58,59	13
3	GOL	B	404	6/6	0.92	0.11	21,32,38,45	0
4	HAR	B	405[B]	13/13	0.92	0.24	37,49,56,65	13
4	HAR	B	405[A]	13/13	0.92	0.24	31,48,55,66	13
3	GOL	B	403	6/6	0.93	0.11	24,32,36,37	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GOL	A	404	6/6	0.93	0.10	23,26,34,50	0
3	GOL	C	404	6/6	0.94	0.13	33,41,46,46	0
4	HAR	A	405[B]	13/13	0.94	0.21	36,53,64,64	13
3	GOL	A	403	6/6	0.94	0.12	20,28,41,43	0
3	GOL	C	403	6/6	0.94	0.10	28,35,40,43	0
4	HAR	A	405[A]	13/13	0.94	0.21	26,52,63,64	13
4	HAR	C	405[B]	13/13	0.96	0.11	22,32,35,35	13
4	HAR	C	405[A]	13/13	0.96	0.11	25,32,35,36	13
2	MN	D	403	1/1	0.98	0.04	43,43,43,43	0
2	MN	D	402	1/1	0.99	0.04	38,38,38,38	0
2	MN	B	402	1/1	1.00	0.12	15,15,15,15	0
2	MN	B	401	1/1	1.00	0.13	14,14,14,14	0
2	MN	C	402	1/1	1.00	0.12	25,25,25,25	0
2	MN	A	402	1/1	1.00	0.11	15,15,15,15	0
2	MN	A	401	1/1	1.00	0.10	14,14,14,14	0
2	MN	C	401	1/1	1.00	0.09	21,21,21,21	0

## 6.5 Other polymers

There are no such residues in this entry.