



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 01:14 am GMT

PDB ID : 1HMW  
Title : ACTIVE SITE OF CHONDROITINASE AC LYASE REVEALED BY THE  
STRUCTURE OF ENZYME-OLIGOSACCHARIDE COMPLEXES AND  
MUTAGENESIS  
Authors : Huang, W.; Boju, L.; Tkalec, L.; Su, H.; Yang, H.O.; Gunay, N.S.; Linhardt,  
R.J.; Kim, Y.S.; Matte, A.; Cygler, M.  
Deposited on : 2000-12-05  
Resolution : 2.30 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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

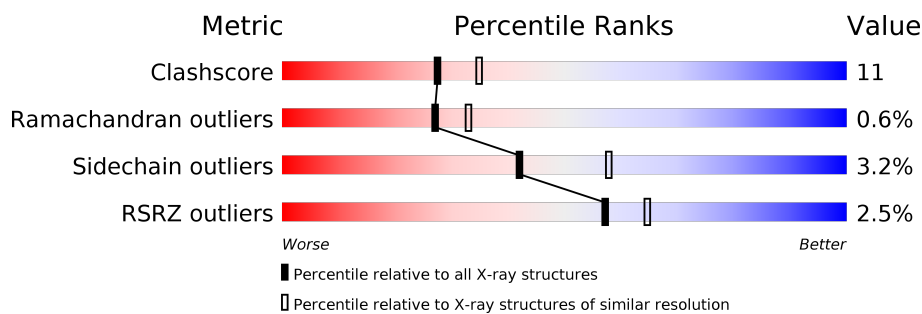
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 2.30 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))
Clashscore	122126	5072 (2.30-2.30)
Ramachandran outliers	120053	5022 (2.30-2.30)
Sidechain outliers	120020	5021 (2.30-2.30)
RSRZ outliers	108989	4374 (2.30-2.30)

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	700	

## 2 Entry composition [i](#)

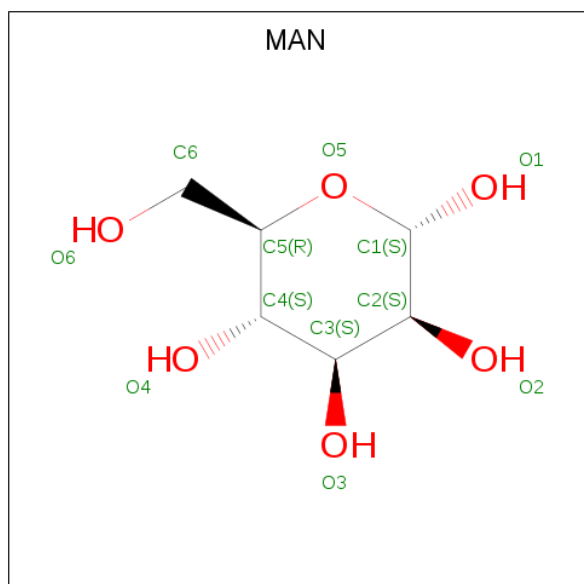
There are 12 unique types of molecules in this entry. The entry contains 5854 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 CHONDROITINASE AC.

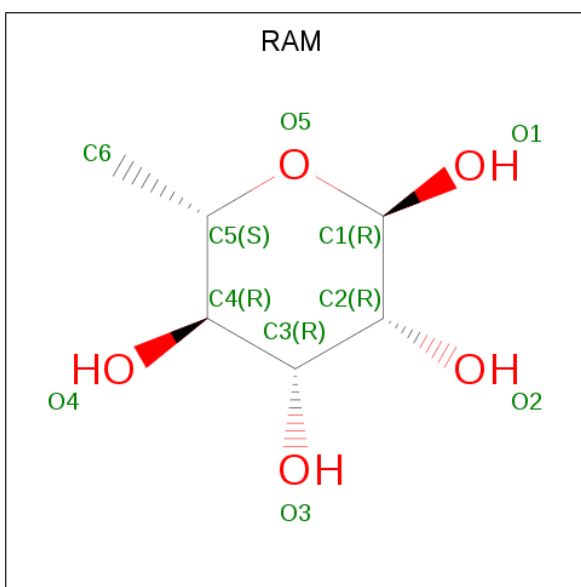
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	674	5381	3445	919	1003	14	0	0	0

- Molecule 2 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



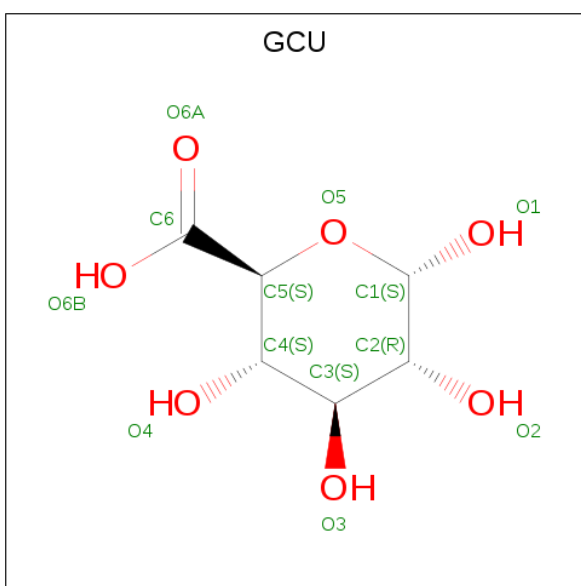
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	11	6	5	0	0
2	A	1	11	6	5	0	0

- Molecule 3 is ALPHA-L-RHAMNOSE (three-letter code: RAM) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>5</sub>).



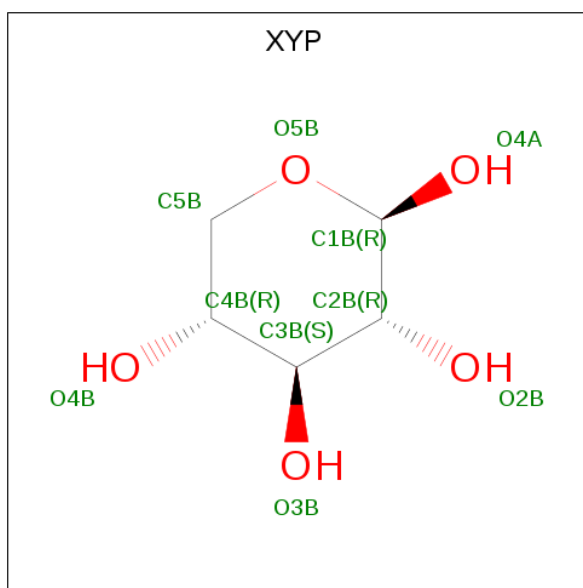
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	6	4		
3	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 4 is D-GLUCURONIC ACID (three-letter code: GCU) (formula:  $C_6H_{10}O_7$ ).



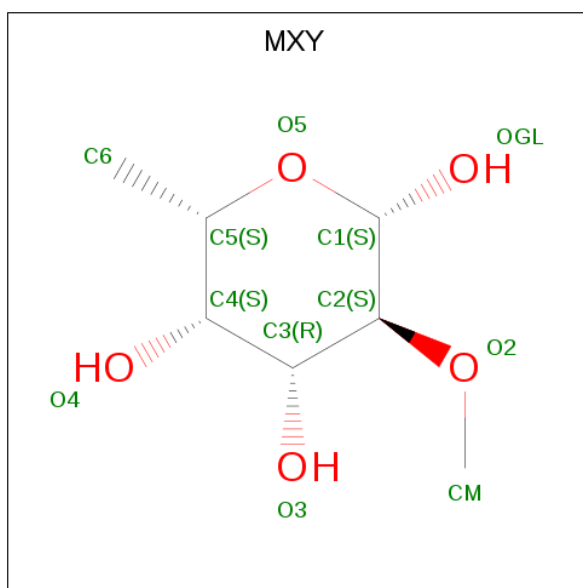
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			12	6	6		
4	A	1	Total	C	O	0	0
			12	6	6		

- Molecule 5 is BETA-D-XYLOPYRANOSE (three-letter code: XYP) (formula:  $C_5H_{10}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			9	5	4		

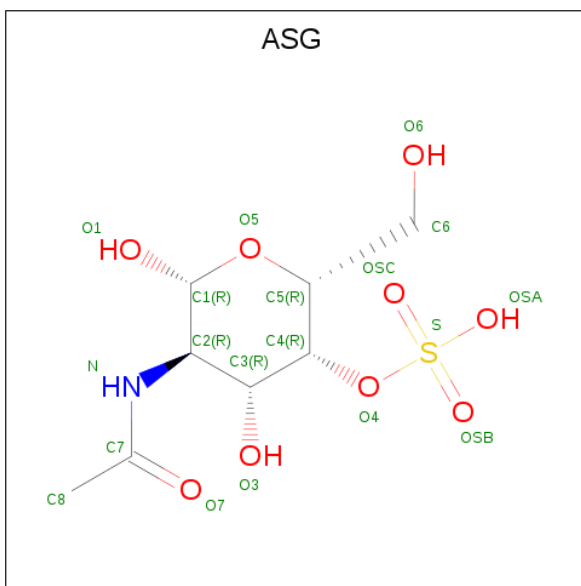
- Molecule 6 is 2-O-METHYL FUCOSE (three-letter code: MXY) (formula:  $C_7H_{14}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			11	7	4		

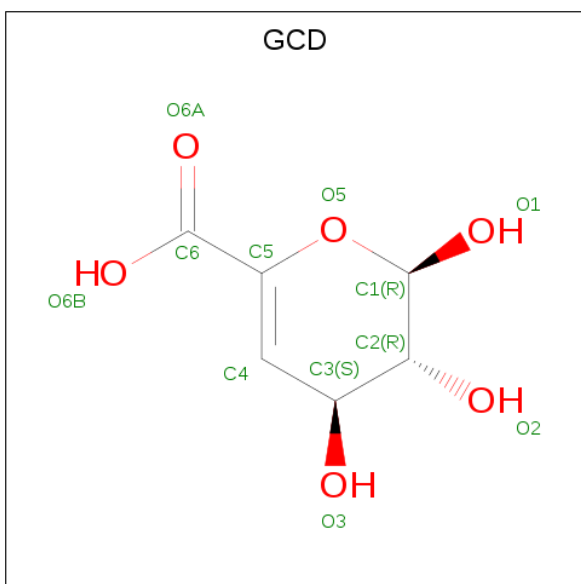
- Molecule 7 is 2-DEOXY-2-ACETAMIDO-BETA-D-GALACTOSE-4-SULFATE (three-

letter code: ASG) (formula:  $C_8H_{15}NO_9S$ ).



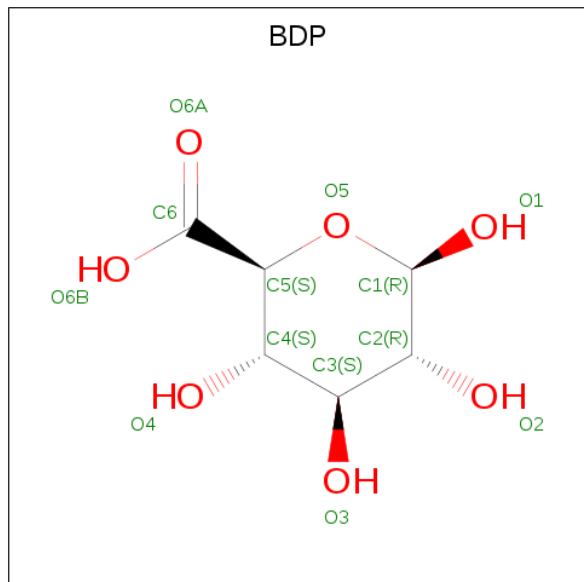
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	S	0	0
			18	8	1	8	1		

- Molecule 8 is 4,5-DEHYDRO-D-GLUCURONIC ACID (three-letter code: GCD) (formula:  $C_6H_8O_6$ ).



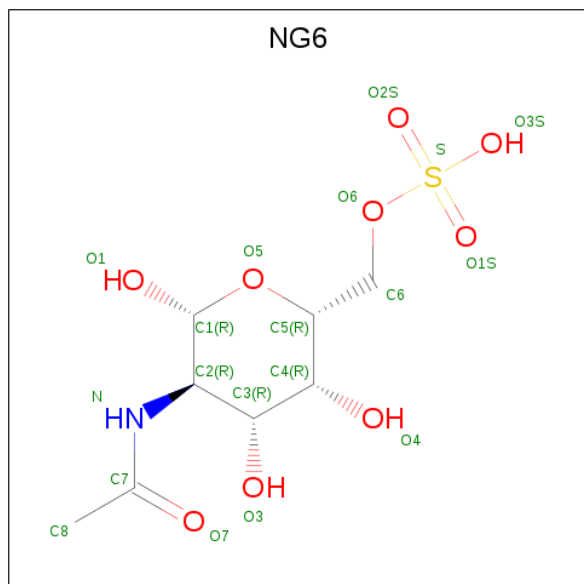
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 9 is BETA-D-GLUCOPYRANURONIC ACID (three-letter code: BDP) (formula:  $C_6H_{10}O_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			12	6	6		

- Molecule 10 is N-ACETYL-D-GALACTOSAMINE 6-SULFATE (three-letter code: NG6) (formula:  $C_8H_{15}NO_9S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
10	A	1	Total	C	N	O	S	0	0
			19	8	1	9	1		

- Molecule 11 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	1	Total 1	Ca 1	0	0

- Molecule 12 is water.

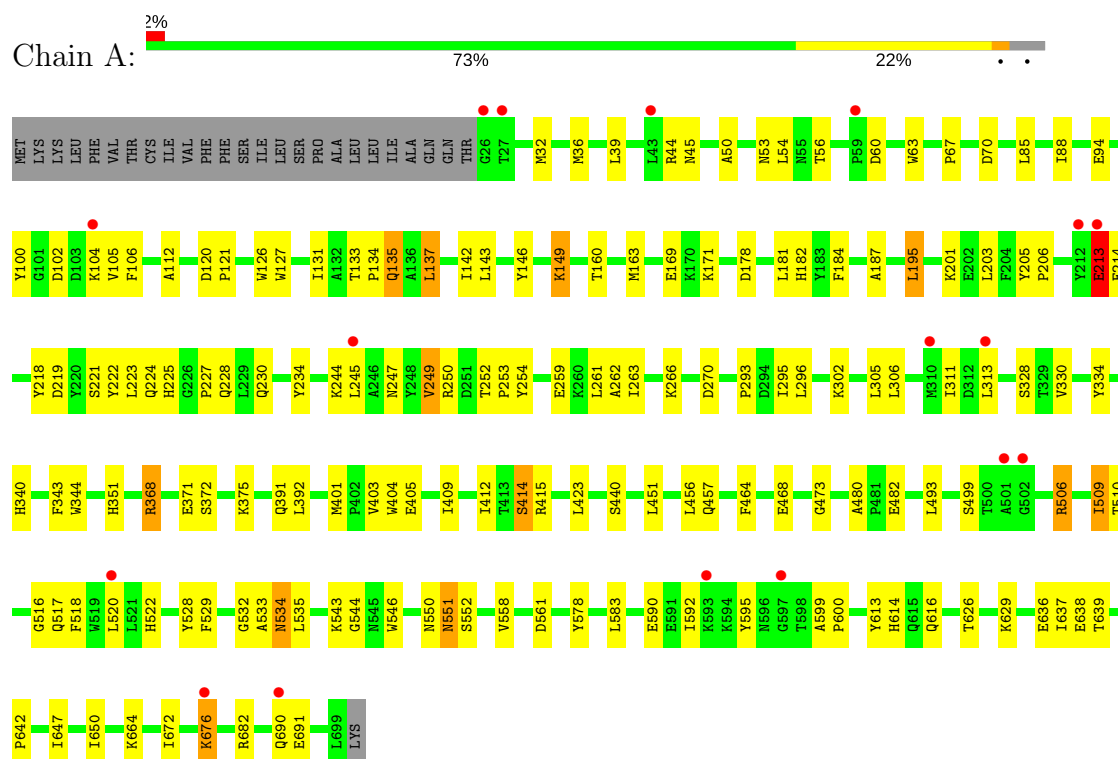
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	326	Total 326	O 326	0	0



### 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: CHONDROITINASE AC



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.90Å 86.90Å 192.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 20.03 – 2.29	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.30) 92.8 (20.03-2.29)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	16.73 (at 2.28Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.219 , 0.275 0.232 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.0	Xtriage
Anisotropy	0.671	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 40.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5854	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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: XYP, GCD, MXY, CA, RAM, NG6, GCU, BDP, ASG, MAN

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	0.35	0/5519	0.63	2/7485 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	405	GLU	N-CA-C	-6.23	94.18	111.00
1	A	213	GLU	N-CA-C	5.13	124.86	111.00

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	5381	0	5239	117	0
2	A	22	0	16	1	0
3	A	20	0	20	0	0
4	A	24	0	13	0	0
5	A	9	0	7	0	0
6	A	11	0	12	0	0
7	A	18	0	11	0	0
8	A	11	0	5	0	0
9	A	12	0	6	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	A	19	0	13	0	0
11	A	1	0	0	0	0
12	A	326	0	0	6	0
All	All	5854	0	5342	117	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 11.

The worst 5 of 117 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:213:GLU:HG3	1:A:227:PRO:HB3	1.39	1.05
1:A:203:LEU:HD13	1:A:245:LEU:HD22	1.64	0.80
1:A:247:ASN:HB2	1:A:311:ILE:HD11	1.65	0.77
1:A:187:ALA:HB2	1:A:195:LEU:HD13	1.70	0.73
1:A:54:LEU:HD13	1:A:88:ILE:HD13	1.76	0.66

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	672/700 (96%)	616 (92%)	52 (8%)	4 (1%)	27	33

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	213	GLU
1	A	551	ASN
1	A	214	GLU
1	A	249	VAL

### 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	563/598 (94%)	545 (97%)	18 (3%)	42 58

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	401	MET
1	A	404	TRP
1	A	534	ASN
1	A	306	LEU
1	A	368	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	79	ASN
1	A	211	HIS
1	A	534	ASN
1	A	540	GLN
1	A	551	ASN

### 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 13 ligands modelled in this entry, 1 is monoatomic - leaving 12 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$
2	MAN	A	701	1,3,4	11,11,12	0.54	0	15,15,17	0.62	0
3	RAM	A	702	2	9,10,11	1.53	1 (11%)	13,14,16	0.93	1 (7%)
4	GCU	A	703	2,5	9,12,13	1.51	1 (11%)	12,17,19	0.78	0
5	XYP	A	704	4,6	9,9,10	1.53	1 (11%)	10,12,14	0.67	0
6	MXY	A	705	5	10,11,12	1.11	1 (10%)	14,15,17	0.62	0
2	MAN	A	706	1,3,4	11,11,12	0.53	0	15,15,17	0.40	0
3	RAM	A	707	2	9,10,11	1.67	2 (22%)	13,14,16	0.71	0
4	GCU	A	708	2	9,12,13	1.80	2 (22%)	12,17,19	0.84	0
7	ASG	A	709	9,8	18,18,19	1.49	2 (11%)	18,26,28	1.03	2 (11%)
8	GCD	A	710	7	7,11,12	6.97	4 (57%)	7,15,17	3.97	3 (42%)
9	BDP	A	711	10,7	9,12,13	1.35	0	12,17,19	0.95	0
10	NG6	A	712	9	19,19,19	1.20	1 (5%)	26,28,28	0.97	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
2	MAN	A	701	1,3,4	-	0/2/19/22	0/1/1/1
3	RAM	A	702	2	-	0/0/17/20	0/1/1/1
4	GCU	A	703	2,5	-	0/0/21/24	0/1/1/1
5	XYP	A	704	4,6	-	0/0/14/17	0/1/1/1
6	MXY	A	705	5	-	0/2/19/22	0/1/1/1
2	MAN	A	706	1,3,4	-	0/2/19/22	0/1/1/1
3	RAM	A	707	2	-	0/0/17/20	0/1/1/1
4	GCU	A	708	2	-	0/0/21/24	0/1/1/1
7	ASG	A	709	9,8	-	0/11/28/31	0/1/1/1
8	GCD	A	710	7	-	0/0/17/20	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	BDP	A	711	10,7	-	0/0/21/24	0/1/1/1
10	NG6	A	712	9	-	0/10/30/30	0/1/1/1

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	A	712	NG6	O6-S	-4.78	1.44	1.56
7	A	709	ASG	O4-S	-4.39	1.44	1.56
4	A	708	GCU	O5-C1	2.27	1.47	1.43
3	A	707	RAM	O5-C1	2.29	1.47	1.43
6	A	705	MXV	C1-C2	2.40	1.55	1.51

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	710	GCD	O5-C5-C4	-9.03	117.05	124.84
8	A	710	GCD	O3-C3-C2	-3.66	103.38	109.54
8	A	710	GCD	C1-C2-C3	-3.18	105.63	109.66
7	A	709	ASG	O7-C7-C8	-2.11	118.25	122.07
3	A	702	RAM	C1-O5-C5	2.08	117.00	112.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	706	MAN	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	674/700 (96%)	-0.02	17 (2%) 57 65	23, 40, 64, 76	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	501	ALA	5.1
1	A	43	LEU	4.3
1	A	59	PRO	3.5
1	A	212	TYR	3.2
1	A	310	MET	3.1

### 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
6	MXY	A	705	11/12	0.73	0.36	78,80,82,83	0
8	GCD	A	710	11/12	0.83	0.22	59,62,64,65	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	XYP	A	704	9/10	0.84	0.19	64,67,71,75	0
3	RAM	A	702	10/11	0.85	0.40	70,72,74,74	0
9	BDP	A	711	12/13	0.86	0.18	58,61,64,64	0
4	GCU	A	708	12/13	0.88	0.22	55,58,62,64	0
4	GCU	A	703	12/13	0.89	0.12	44,53,55,59	0
10	NG6	A	712	19/19	0.90	0.18	64,68,70,70	0
2	MAN	A	701	11/12	0.91	0.14	53,57,62,66	0
2	MAN	A	706	11/12	0.93	0.13	44,47,51,51	0
7	ASG	A	709	18/19	0.94	0.15	48,51,55,56	0
3	RAM	A	707	10/11	0.95	0.16	46,50,52,54	0
11	CA	A	1801	1/1	0.98	0.04	28,28,28,28	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.