



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 11:37 AM BST

PDB ID : 2JFE  
Title : The crystal structure of human cytosolic beta-glucosidase  
Authors : Czjzek, M.; Tribolo, S.; Berrin, J.G.; Kroon, P.A.; Juge, N.  
Deposited on : 2007-01-31  
Resolution : 2.70 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 : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

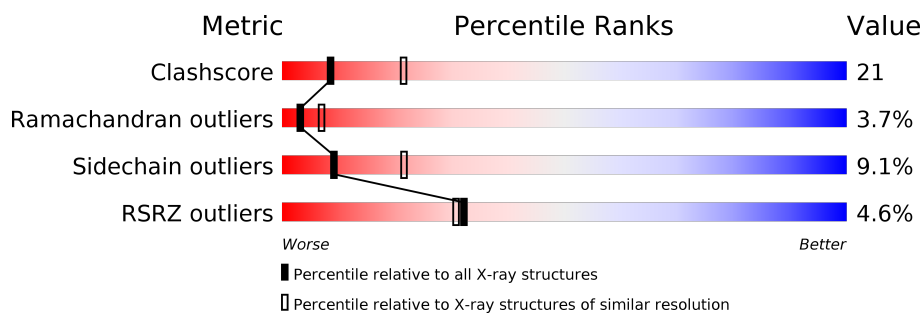
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	X	469	
2	A	2	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3911 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 CYTOSOLIC BETA-GLUCOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	459	Total	C	N	O	S	0	0	0
			3725	2417	608	689	11			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	341	ILE	LYS	conflict	UNP Q9H227

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	A	2	Total	C	N	O	0	0	0
			28	16	2	10			

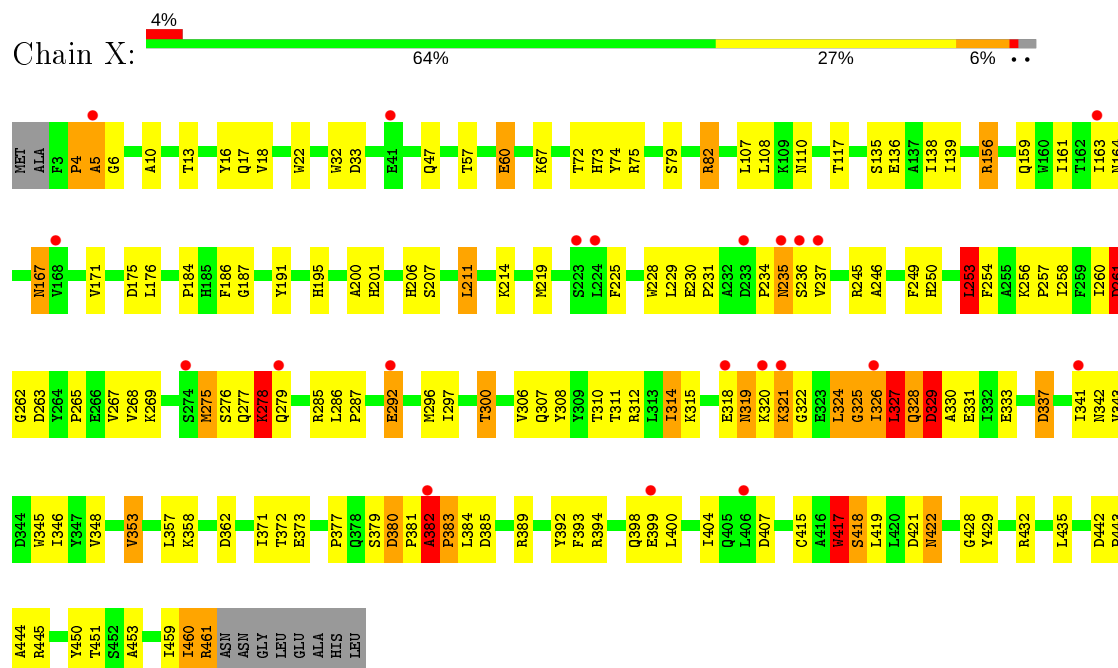
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	X	158	Total	O	0	0
			158	158		

### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

#### • Molecule 1: CYTOSOLIC BETA-GLUCOSIDASE



#### • Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.48Å 90.48Å 127.92Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.64 – 2.70 42.65 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (42.64-2.70) 100.0 (42.65-2.70)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.28 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.197 , 0.280 0.203 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.0	Xtriage
Anisotropy	0.394	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 61.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3911	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.80% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality

### 5.1 Standard geometry

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

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	X	0.67	0/3841	0.79	2/5221 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	X	0	5

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	X	253	LEU	CA-CB-CG	7.53	132.62	115.30
1	X	82	ARG	NE-CZ-NH1	5.13	122.87	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	X	277	GLN	Peptide
1	X	329	ASP	Peptide
1	X	380	ASP	Peptide
1	X	382	ALA	Peptide
1	X	417	TRP	Peptide

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	X	3725	0	3567	152	0
2	A	28	0	25	0	0
3	X	158	0	0	24	1
All	All	3911	0	3592	152	1

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 21.

All (152) 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:X:275:MET:O	1:X:278:LYS:HB2	1.23	1.28
1:X:6:GLY:HA2	3:X:2004:HOH:O	1.30	1.22
1:X:275:MET:O	1:X:278:LYS:CB	1.91	1.17
1:X:278:LYS:HG2	1:X:279:GLN:N	1.66	1.07
1:X:321:LYS:HA	3:X:2117:HOH:O	1.61	1.01
1:X:326:ILE:HA	3:X:2075:HOH:O	1.62	0.99
1:X:382:ALA:H	1:X:383:PRO:HD3	1.31	0.96
1:X:399:GLU:HB3	3:X:2127:HOH:O	1.67	0.94
1:X:75:ARG:HH22	1:X:164:ASN:ND2	1.72	0.87
1:X:278:LYS:CG	1:X:279:GLN:N	2.34	0.87
1:X:308:TYR:CD2	1:X:353:VAL:HG22	2.10	0.86
1:X:206:HIS:HE1	1:X:296:MET:O	1.56	0.86
1:X:325:GLY:O	1:X:327:LEU:N	2.10	0.85
1:X:258:ILE:HG23	1:X:300:THR:HG22	1.56	0.85
1:X:308:TYR:CE2	1:X:353:VAL:HG22	2.14	0.82
1:X:278:LYS:CG	1:X:279:GLN:H	1.93	0.81
1:X:47:GLN:HE22	1:X:429:TYR:H	1.27	0.81
1:X:67:LYS:HD3	1:X:110:ASN:ND2	1.96	0.81
1:X:383:PRO:HB2	1:X:450:TYR:HE2	1.45	0.79
1:X:10:ALA:HB2	1:X:73:HIS:HB2	1.65	0.79
1:X:382:ALA:H	1:X:383:PRO:CD	1.95	0.78
1:X:267:VAL:HG23	3:X:2092:HOH:O	1.83	0.78
1:X:175:ASP:O	1:X:184:PRO:O	2.03	0.75
1:X:333:GLU:HB3	3:X:2121:HOH:O	1.85	0.75

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:326:ILE:C	1:X:328:GLN:H	1.91	0.74
1:X:275:MET:O	1:X:278:LYS:HB3	1.87	0.73
1:X:300:THR:HG23	1:X:300:THR:O	1.88	0.72
1:X:229:LEU:HD22	1:X:314:ILE:HD12	1.71	0.72
1:X:382:ALA:N	1:X:383:PRO:HD3	2.06	0.71
1:X:278:LYS:HG2	1:X:279:GLN:H	1.53	0.70
1:X:75:ARG:NH2	1:X:164:ASN:HD22	1.90	0.70
1:X:325:GLY:C	1:X:327:LEU:H	1.94	0.70
1:X:75:ARG:HH22	1:X:164:ASN:HD22	1.40	0.69
1:X:417:TRP:CZ3	3:X:2147:HOH:O	2.45	0.69
1:X:325:GLY:C	1:X:327:LEU:N	2.45	0.68
1:X:322:GLY:O	3:X:2118:HOH:O	2.12	0.68
1:X:342:ASN:O	1:X:346:ILE:O	2.11	0.68
1:X:353:VAL:HB	1:X:399:GLU:OE1	1.94	0.68
1:X:383:PRO:CB	1:X:450:TYR:HE2	2.06	0.68
1:X:383:PRO:HB2	1:X:450:TYR:CE2	2.29	0.67
1:X:60:GLU:CD	1:X:60:GLU:H	1.96	0.67
1:X:161:ILE:HD11	1:X:372:THR:HG21	1.76	0.67
1:X:292:GLU:HG2	3:X:2105:HOH:O	1.94	0.67
1:X:442:ASP:O	1:X:444:ALA:O	2.13	0.67
1:X:13:THR:HG23	3:X:2039:HOH:O	1.95	0.65
1:X:17:GLN:O	1:X:422:ASN:HB2	1.97	0.65
1:X:319:ASN:H	1:X:331:GLU:HG2	1.61	0.64
1:X:47:GLN:NE2	1:X:429:TYR:H	1.94	0.64
1:X:256:LYS:O	1:X:260:ILE:O	2.16	0.63
1:X:260:ILE:O	1:X:262:GLY:N	2.31	0.63
1:X:312:ARG:NH2	1:X:345:TRP:O	2.34	0.61
1:X:257:PRO:O	1:X:262:GLY:HA2	1.99	0.60
1:X:167:ASN:HD21	1:X:225:PHE:H	1.50	0.59
1:X:206:HIS:CE1	1:X:296:MET:O	2.48	0.59
1:X:326:ILE:O	1:X:328:GLN:N	2.36	0.59
1:X:246:ALA:O	1:X:250:HIS:HD2	1.86	0.59
1:X:135:SER:O	1:X:138:ILE:HG22	2.03	0.58
1:X:389:ARG:HD3	1:X:393:PHE:HE1	1.67	0.58
1:X:108:LEU:HD11	1:X:156:ARG:HG3	1.86	0.57
1:X:72:THR:OG1	1:X:73:HIS:HD2	1.88	0.57
1:X:383:PRO:HG3	3:X:2135:HOH:O	2.05	0.56
1:X:373:GLU:HG2	1:X:417:TRP:CD1	2.41	0.56
1:X:253:LEU:HD22	1:X:268:VAL:HG11	1.87	0.56
1:X:234:PRO:C	1:X:236:SER:H	2.09	0.56
1:X:357:LEU:HD21	1:X:371:ILE:HD11	1.89	0.55

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:300:THR:CG2	1:X:300:THR:O	2.54	0.55
1:X:318:GLU:OE1	1:X:320:LYS:HE3	2.05	0.55
1:X:321:LYS:HD2	3:X:2116:HOH:O	2.05	0.55
1:X:310:THR:HG22	1:X:311:THR:N	2.22	0.55
1:X:310:THR:HG22	1:X:311:THR:O	2.06	0.55
1:X:382:ALA:N	3:X:2135:HOH:O	2.39	0.55
1:X:461:ARG:HB2	1:X:461:ARG:NH1	2.22	0.54
1:X:4:PRO:HD3	1:X:404:ILE:HD13	1.90	0.54
1:X:117:THR:HA	1:X:161:ILE:O	2.08	0.54
1:X:79:SER:HB3	1:X:82:ARG:HB3	1.89	0.54
1:X:417:TRP:HZ3	3:X:2147:HOH:O	1.87	0.54
1:X:307:GLN:HB3	1:X:373:GLU:HB2	1.88	0.54
1:X:47:GLN:HE21	1:X:428:GLY:HA3	1.73	0.54
1:X:443:PRO:C	1:X:444:ALA:O	2.45	0.53
1:X:326:ILE:HG22	1:X:326:ILE:O	2.09	0.53
1:X:228:TRP:O	1:X:311:THR:HA	2.10	0.52
1:X:75:ARG:NH2	1:X:164:ASN:ND2	2.45	0.52
1:X:297:ILE:O	1:X:300:THR:HB	2.09	0.52
1:X:75:ARG:NH1	1:X:373:GLU:HG3	2.24	0.52
1:X:358:LYS:HE2	1:X:362:ASP:OD2	2.09	0.52
1:X:191:TYR:HE2	3:X:2100:HOH:O	1.92	0.52
1:X:385:ASP:HA	1:X:451:THR:OG1	2.10	0.51
1:X:320:LYS:HG2	3:X:2115:HOH:O	2.12	0.50
1:X:234:PRO:O	1:X:236:SER:N	2.44	0.50
1:X:47:GLN:NE2	1:X:428:GLY:HA3	2.27	0.50
1:X:278:LYS:HE2	1:X:285:ARG:NH1	2.27	0.50
1:X:171:VAL:CG1	1:X:176:LEU:HD23	2.42	0.49
1:X:167:ASN:HB3	1:X:254:PHE:CE2	2.47	0.49
1:X:319:ASN:N	1:X:331:GLU:HG2	2.28	0.49
1:X:326:ILE:C	1:X:328:GLN:N	2.63	0.49
1:X:256:LYS:HB3	1:X:257:PRO:HD3	1.94	0.48
1:X:399:GLU:CB	3:X:2127:HOH:O	2.39	0.48
1:X:310:THR:HG22	1:X:311:THR:H	1.78	0.48
1:X:330:ALA:N	3:X:2120:HOH:O	2.46	0.48
1:X:187:GLY:HA3	1:X:325:GLY:HA3	1.96	0.48
1:X:394:ARG:NH1	3:X:2142:HOH:O	2.47	0.47
1:X:249:PHE:CZ	1:X:330:ALA:HB2	2.50	0.47
1:X:373:GLU:HG2	1:X:417:TRP:HD1	1.79	0.47
1:X:394:ARG:NH1	3:X:2141:HOH:O	2.48	0.47
1:X:228:TRP:O	1:X:311:THR:HG23	2.15	0.47
1:X:346:ILE:HD13	1:X:377:PRO:HG3	1.95	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:13:THR:OG1	1:X:18:VAL:HG21	2.15	0.47
1:X:16:TYR:CE2	1:X:32:TRP:HB3	2.50	0.46
1:X:459:ILE:O	1:X:461:ARG:N	2.48	0.46
1:X:75:ARG:HH12	1:X:373:GLU:HG3	1.80	0.46
1:X:337:ASP:C	1:X:337:ASP:OD1	2.54	0.46
1:X:163:ILE:HD12	1:X:201:HIS:HB2	1.98	0.46
1:X:279:GLN:HG2	3:X:2097:HOH:O	2.16	0.45
1:X:4:PRO:HG3	1:X:404:ILE:HG21	1.97	0.45
1:X:195:HIS:CD2	1:X:287:PRO:HB2	2.52	0.45
1:X:461:ARG:HH11	1:X:461:ARG:H	1.63	0.45
1:X:260:ILE:HG22	1:X:261:ASP:HB3	1.97	0.45
1:X:320:LYS:O	1:X:321:LYS:HB2	2.17	0.44
1:X:159:GLN:HE22	1:X:219:MET:HE3	1.82	0.44
1:X:421:ASP:CG	1:X:432:ARG:HB3	2.37	0.44
1:X:329:ASP:C	1:X:331:GLU:H	2.21	0.44
1:X:207:SER:O	1:X:211:LEU:HB2	2.18	0.44
1:X:256:LYS:HD3	1:X:265:PRO:HA	1.99	0.44
1:X:230:GLU:HA	1:X:231:PRO:HD2	1.81	0.43
1:X:435:LEU:O	1:X:453:ALA:HB2	2.19	0.43
1:X:263:ASP:OD1	1:X:269:LYS:NZ	2.50	0.43
1:X:79:SER:HB3	1:X:82:ARG:CB	2.48	0.43
1:X:159:GLN:HE22	1:X:219:MET:CE	2.31	0.43
1:X:171:VAL:HG13	1:X:176:LEU:HD23	2.01	0.43
1:X:13:THR:OG1	1:X:18:VAL:HG11	2.19	0.43
1:X:341:ILE:HG21	1:X:392:TYR:OH	2.19	0.43
1:X:139:ILE:HG23	1:X:200:ALA:HA	2.00	0.42
1:X:459:ILE:O	1:X:461:ARG:HD3	2.20	0.42
1:X:167:ASN:HD22	1:X:167:ASN:N	2.16	0.42
1:X:379:SER:HB2	3:X:2134:HOH:O	2.19	0.42
1:X:421:ASP:OD2	1:X:432:ARG:HB3	2.20	0.42
1:X:10:ALA:O	1:X:415:CYS:HA	2.21	0.41
1:X:108:LEU:CD1	1:X:156:ARG:HG3	2.48	0.41
1:X:278:LYS:HG2	1:X:279:GLN:CA	2.46	0.41
1:X:186:PHE:O	1:X:324:LEU:HD22	2.21	0.41
1:X:326:ILE:CA	3:X:2075:HOH:O	2.41	0.41
1:X:460:ILE:O	1:X:461:ARG:O	2.39	0.41
1:X:417:TRP:HH2	1:X:422:ASN:HD21	1.67	0.41
1:X:22:TRP:HB2	1:X:33:ASP:OD2	2.21	0.41
1:X:74:TYR:HD2	1:X:107:LEU:HD11	1.86	0.41
1:X:417:TRP:CH2	3:X:2147:HOH:O	2.73	0.41
1:X:230:GLU:OE1	1:X:231:PRO:HD2	2.21	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:13:THR:HG21	1:X:74:TYR:OH	2.21	0.40
1:X:276:SER:OG	1:X:285:ARG:HG3	2.21	0.40
1:X:167:ASN:H	1:X:167:ASN:HD22	1.70	0.40
1:X:245:ARG:HG3	1:X:267:VAL:HG11	2.04	0.40
1:X:4:PRO:HB2	1:X:5:ALA:H	1.72	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:X:2141:HOH:O	3:X:2141:HOH:O[6_765]	2.15	0.05

## 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	X	457/469 (97%)	410 (90%)	30 (7%)	17 (4%)	<b>3</b> <b>7</b>

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	5	ALA
1	X	235	ASN
1	X	326	ILE
1	X	381	PRO
1	X	382	ALA
1	X	418	SER
1	X	261	ASP
1	X	278	LYS
1	X	325	GLY
1	X	327	LEU
1	X	384	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	X	343	VAL
1	X	383	PRO
1	X	4	PRO
1	X	460	ILE
1	X	321	LYS
1	X	348	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	X	394/402 (98%)	358 (91%)	36 (9%)	9 21

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

Mol	Chain	Res	Type
1	X	57	THR
1	X	60	GLU
1	X	136	GLU
1	X	156	ARG
1	X	167	ASN
1	X	211	LEU
1	X	214	LYS
1	X	235	ASN
1	X	237	VAL
1	X	253	LEU
1	X	261	ASP
1	X	275	MET
1	X	278	LYS
1	X	286	LEU
1	X	292	GLU
1	X	300	THR
1	X	306	VAL
1	X	314	ILE
1	X	315	LYS
1	X	319	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	X	324	LEU
1	X	327	LEU
1	X	328	GLN
1	X	329	ASP
1	X	337	ASP
1	X	353	VAL
1	X	380	ASP
1	X	398	GLN
1	X	400	LEU
1	X	407	ASP
1	X	417	TRP
1	X	418	SER
1	X	419	LEU
1	X	422	ASN
1	X	445	ARG
1	X	461	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	38	GLN
1	X	47	GLN
1	X	73	HIS
1	X	110	ASN
1	X	164	ASN
1	X	167	ASN
1	X	206	HIS
1	X	216	GLN
1	X	240	GLN
1	X	250	HIS
1	X	317	GLN
1	X	395	GLN
1	X	405	GLN
1	X	422	ASN
1	X	426	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 ⓘ

2 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
2	NAG	A	1	1,2	14,14,15	0.67	0	17,19,21	1.07	1 (5%)
2	NAG	A	2	2	14,14,15	0.76	0	17,19,21	0.92	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	NAG	A	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	A	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	NAG	O5-C5-C6	3.14	112.13	107.20

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1	NAG	C4-C5-C6-O6
2	A	2	NAG	C8-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	A	2	NAG	O7-C7-N2-C2
2	A	1	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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	X	459/469 (97%)	0.24	21 (4%) 32 31	19, 36, 58, 66	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	326	ILE	5.0
1	X	320	LYS	4.7
1	X	235	ASN	4.1
1	X	321	LYS	3.4
1	X	406	LEU	3.4
1	X	382	ALA	3.2
1	X	237	VAL	2.8
1	X	318	GLU	2.7
1	X	224	LEU	2.6
1	X	233	ASP	2.6
1	X	274	SER	2.5
1	X	292	GLU	2.4
1	X	399	GLU	2.3
1	X	163	ILE	2.2
1	X	41	GLU	2.1
1	X	341	ILE	2.1
1	X	5	ALA	2.1
1	X	223	SER	2.1
1	X	236	SER	2.1
1	X	279	GLN	2.1
1	X	168	VAL	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](#)

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( $\text{\AA}^2$ )	Q<0.9
2	NAG	A	2	14/15	0.89	0.34	42,47,53,53	0
2	NAG	A	1	14/15	0.92	0.31	39,41,44,45	0

### 6.4 Ligands [i](#)

There are no ligands in this entry.

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

There are no such residues in this entry.