



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 01:33 PM GMT

PDB ID : 3TZI  
Title : X-ray crystal structure of arachidonic acid bound in the cyclooxygenase channel of G533V murine COX-2  
Authors : Vecchio, A.J.; Malkowski, M.G.  
Deposited on : 2011-09-27  
Resolution : 2.15 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

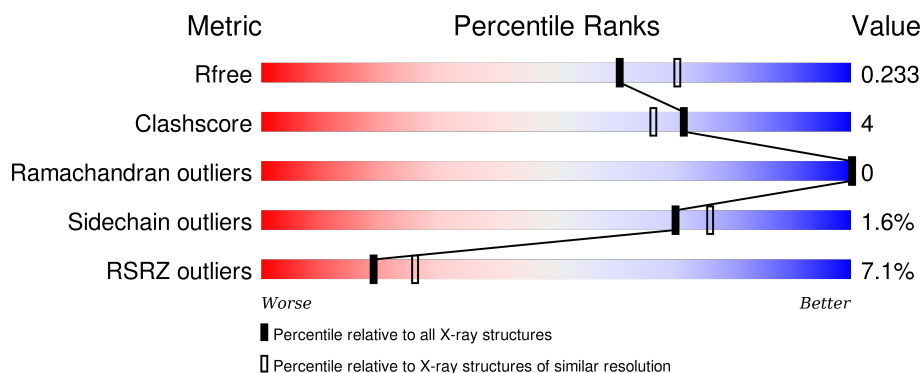
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1045 (2.16-2.16)
Clashscore	102246	1152 (2.16-2.16)
Ramachandran outliers	100387	1131 (2.16-2.16)
Sidechain outliers	100360	1131 (2.16-2.16)
RSRZ outliers	91569	1050 (2.16-2.16)

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	593	<div> <div>5%</div> <div>88%</div> <div>5% • 7%</div> </div>
1	B	593	<div> <div>8%</div> <div>87%</div> <div>6% • 7%</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	ACD	A	1	-	-	X	X
2	ACD	B	1	-	-	X	X
4	NAG	A	662	X	-	-	-
4	NAG	B	661	X	-	-	X
8	EDO	B	3	-	-	-	X
8	EDO	B	620	-	-	-	X
8	EDO	B	7	-	-	-	X
9	AKR	A	9	-	-	X	-

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 10096 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.

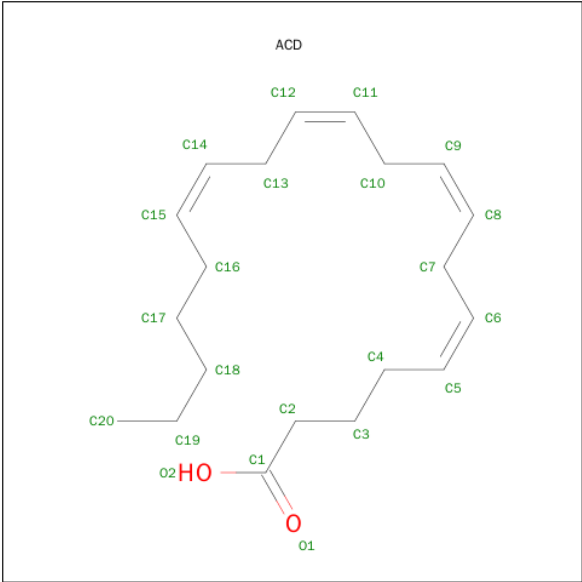
- Molecule 1 is a protein called Prostaglandin G/H synthase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	551	Total	C	N	O	S	0	5	0
			4468	2886	743	813	26			
1	B	552	Total	C	N	O	S	0	6	0
			4469	2892	743	809	25			

There are 18 discrepancies between the modelled and reference sequences:

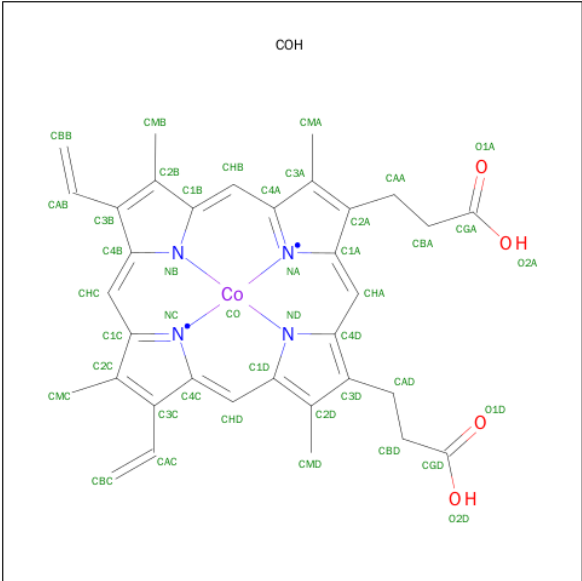
Chain	Residue	Modelled	Actual	Comment	Reference
A	27	ALA	-	EXPRESSION TAG	UNP Q05769
A	29	HIS	-	EXPRESSION TAG	UNP Q05769
A	30	HIS	-	EXPRESSION TAG	UNP Q05769
A	31	HIS	-	EXPRESSION TAG	UNP Q05769
A	32	HIS	-	EXPRESSION TAG	UNP Q05769
A	33	HIS	-	EXPRESSION TAG	UNP Q05769
A	34	HIS	-	EXPRESSION TAG	UNP Q05769
A	533	VAL	GLY	ENGINEERED MUTATION	UNP Q05769
A	594	ALA	ASN	ENGINEERED MUTATION	UNP Q05769
B	27	ALA	-	EXPRESSION TAG	UNP Q05769
B	29	HIS	-	EXPRESSION TAG	UNP Q05769
B	30	HIS	-	EXPRESSION TAG	UNP Q05769
B	31	HIS	-	EXPRESSION TAG	UNP Q05769
B	32	HIS	-	EXPRESSION TAG	UNP Q05769
B	33	HIS	-	EXPRESSION TAG	UNP Q05769
B	34	HIS	-	EXPRESSION TAG	UNP Q05769
B	533	VAL	GLY	ENGINEERED MUTATION	UNP Q05769
B	594	ALA	ASN	ENGINEERED MUTATION	UNP Q05769

- Molecule 2 is ARACHIDONIC ACID (three-letter code: ACD) (formula: C<sub>20</sub>H<sub>32</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			22	20	2		
2	B	1	Total	C	O	0	0
			22	20	2		

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING CO (three-letter code: COH) (formula: C<sub>34</sub>H<sub>32</sub>CoN<sub>4</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	Co	N	O	0	0
			43	34	1	4	4		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	Co	N	O	0	0
			43	34	1	4	4		

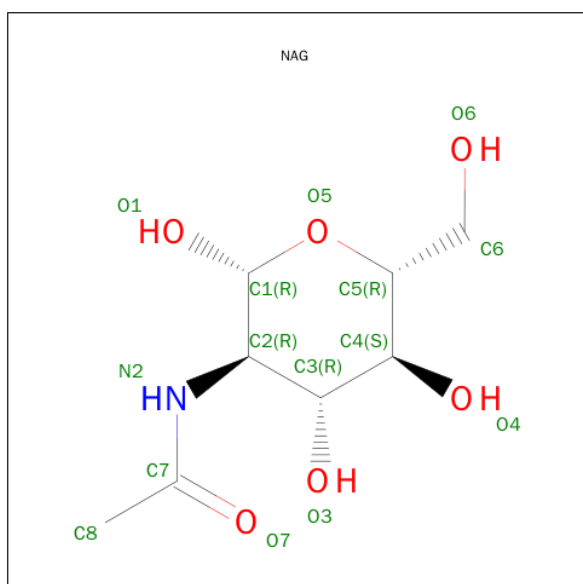
- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

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

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	3	Total	C	N	O	0	0
			39	22	2	15		

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



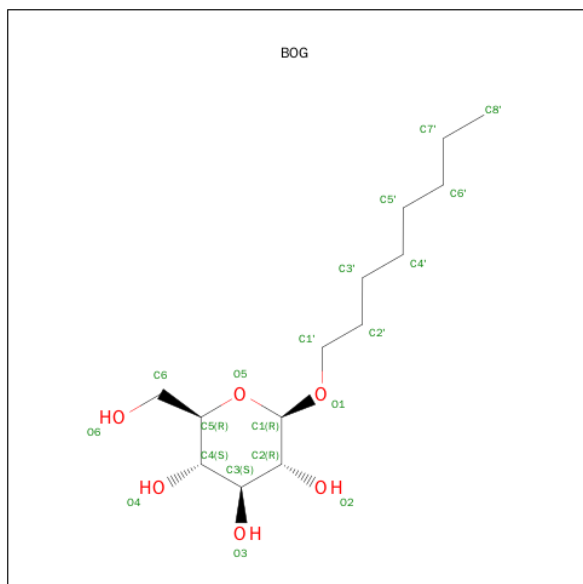
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

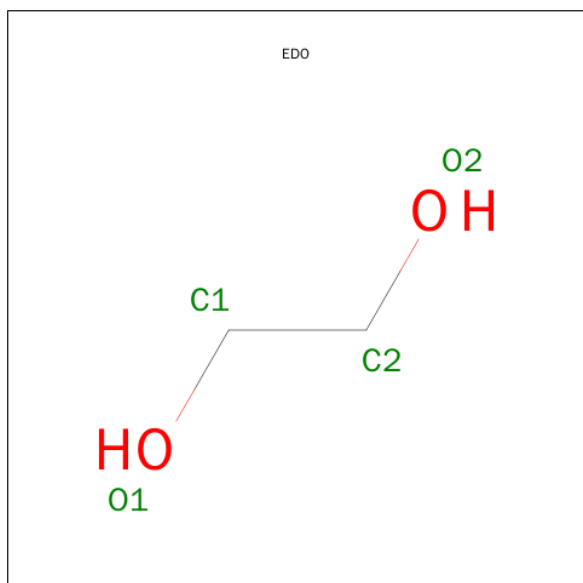
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is SUGAR (B-OCTYLGLUCOSIDE) (three-letter code: BOG) (formula:  $C_{14}H_{28}O_6$ ).



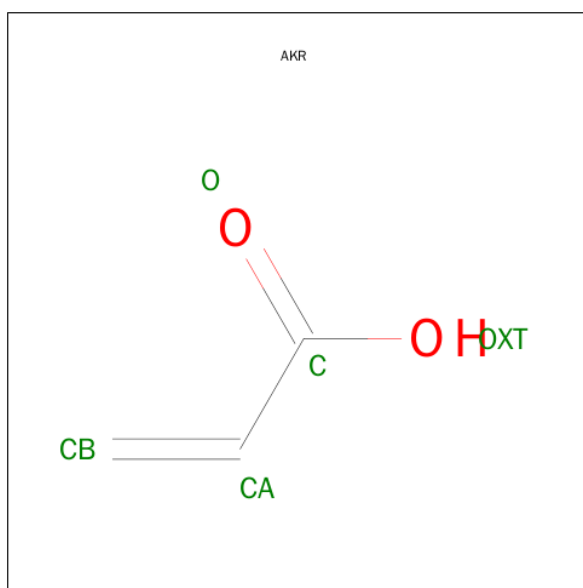
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			20	14	6		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0

- Molecule 9 is ACRYLIC ACID (three-letter code: AKR) (formula:  $C_3H_4O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total C O 5 3 2	0	0
9	A	1	Total C O 5 3 2	0	0

- Molecule 10 is water.

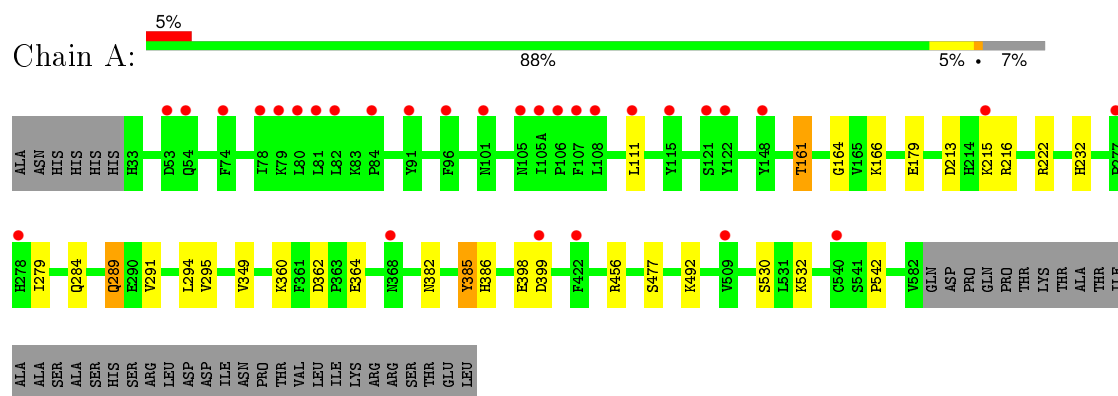


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	401	Total 401	O 401	0	0
10	B	415	Total 415	O 415	0	0

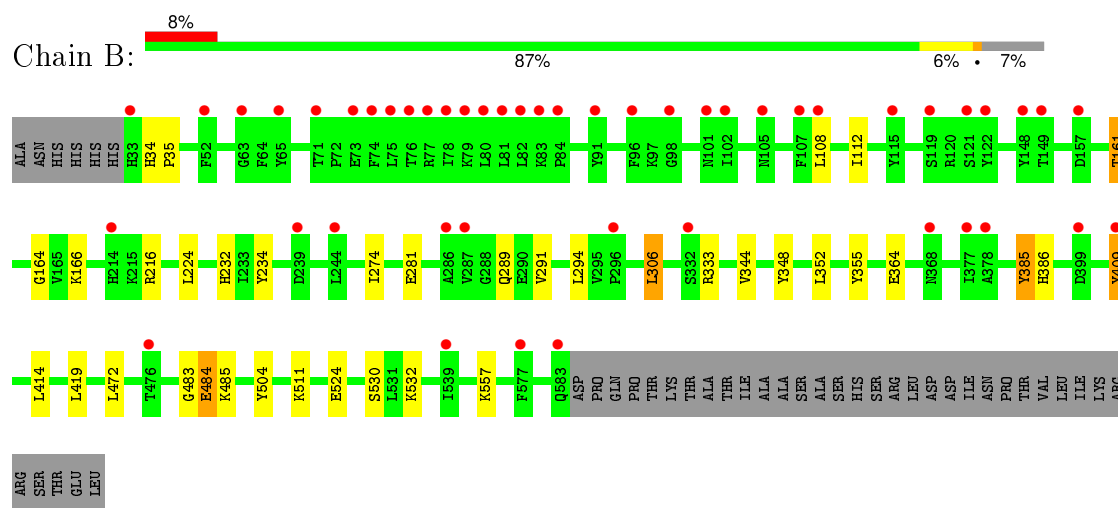
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Prostaglandin G/H synthase 2



#### • Molecule 1: Prostaglandin G/H synthase 2



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	122.27Å 133.51Å 181.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 2.15 19.99 – 2.15	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.99-2.15) 100.0 (19.99-2.15)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.84 (at 2.15Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, $R_{free}$	0.170 , 0.210 0.203 , 0.233	Depositor DCC
$R_{free}$ test set	4039 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.0	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 44.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 80416 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10096	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, ACD, AKR, EDO, COH, BOG, 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	0.49	0/4610	0.56	0/6260
1	B	0.49	0/4616	0.56	1/6272 (0.0%)
All	All	0.49	0/9226	0.56	1/12532 (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
4	A	1	0
4	B	1	0
All	All	2	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	306	LEU	CA-CB-CG	5.07	126.95	115.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	662	NAG	C1
4	B	661	NAG	C5

There are no planarity outliers.

## 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	A	4468	0	4308	31	0
1	B	4469	0	4286	30	0
2	A	22	0	31	10	0
2	B	22	0	31	10	0
3	A	43	0	30	0	0
3	B	43	0	30	4	0
4	A	28	0	25	0	0
4	B	56	0	50	3	0
5	A	39	0	34	4	0
6	A	14	0	13	0	0
6	B	14	0	13	0	0
7	A	20	0	28	1	0
8	A	8	0	12	1	0
8	B	24	0	36	2	0
9	A	10	0	6	4	0
10	A	401	0	0	4	0
10	B	415	0	0	4	0
All	All	10096	0	8933	78	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 4.

All (78) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:216:ARG:HG2	4:B:672:NAG:H81	1.53	0.90
1:A:530:SER:OG	2:A:1:ACD:H192	1.73	0.89
1:B:161:THR:HG21	1:B:166:LYS:O	1.78	0.81
1:A:161:THR:HG21	1:A:166:LYS:O	1.80	0.81
1:B:530:SER:OG	2:B:1:ACD:H182	1.82	0.80
2:B:1:ACD:H71	2:B:1:ACD:H11	1.65	0.78
1:A:492:LYS:HD2	9:A:9:AKR:HA1	1.67	0.76
1:A:216:ARG:HH11	5:A:672:NAG:H83	1.50	0.74
1:B:484:GLU:HG2	1:B:485:LYS:H	1.53	0.74
3:B:619:COH:HMB1	3:B:619:COH:HBB1	1.72	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:477:SER:HB2	9:A:9:AKR:HB2	1.75	0.69
1:A:216:ARG:NH1	5:A:672:NAG:H83	2.08	0.69
1:B:385:TYR:OH	2:B:1:ACD:H181	1.96	0.65
1:B:484:GLU:HG2	1:B:485:LYS:N	2.13	0.63
2:A:1:ACD:H11	2:A:1:ACD:H72	1.81	0.62
1:B:274:ILE:HD12	1:B:291:VAL:HG12	1.81	0.62
3:B:619:COH:HMB3	8:B:8:EDO:H11	1.84	0.60
1:B:216:ARG:HH11	4:B:672:NAG:H83	1.64	0.60
2:B:1:ACD:H71	2:B:1:ACD:C11	2.32	0.60
1:A:216:ARG:HG2	5:A:672:NAG:H81	1.85	0.58
1:B:161:THR:HG23	10:B:627:HOH:O	2.03	0.58
1:B:472:LEU:HD21	1:B:524:GLU:HG3	1.85	0.58
2:B:1:ACD:C12	2:B:1:ACD:H162	2.34	0.57
1:B:414:LEU:HD11	1:B:419:LEU:HD12	1.86	0.56
2:B:1:ACD:H162	2:B:1:ACD:C11	2.36	0.55
2:A:1:ACD:C12	2:A:1:ACD:H162	2.38	0.54
2:B:1:ACD:C12	2:B:1:ACD:C16	2.86	0.53
1:A:349:VAL:HG13	2:A:1:ACD:H101	1.89	0.53
1:B:355:TYR:OH	2:B:1:ACD:H6	2.08	0.53
1:A:456:ARG:HA	8:A:2:EDO:H12	1.91	0.53
1:B:294:LEU:HD22	1:B:409:TYR:CD2	2.44	0.52
2:A:1:ACD:H72	2:A:1:ACD:C11	2.38	0.50
1:B:483:GLY:HA3	1:B:511:LYS:HD3	1.94	0.50
1:A:385:TYR:CE1	2:A:1:ACD:H191	2.46	0.50
3:B:619:COH:HBC1	3:B:619:COH:HHD	1.93	0.50
1:A:222:ARG:NH1	10:A:974:HOH:O	2.45	0.49
1:A:294:LEU:HG	1:A:295:VAL:CG2	2.42	0.49
1:A:161:THR:CG2	1:A:166:LYS:O	2.58	0.49
1:A:294:LEU:HG	1:A:295:VAL:HG23	1.96	0.48
1:B:161:THR:HG22	1:B:164:GLY:H	1.78	0.47
1:A:289:GLN:OE1	1:A:291:VAL:HG22	2.14	0.47
1:B:216:ARG:NH1	4:B:672:NAG:H83	2.29	0.47
1:B:364:GLU:HB3	10:B:1008:HOH:O	2.15	0.47
3:B:619:COH:HMB1	3:B:619:COH:CBB	2.44	0.46
1:B:344:VAL:HA	1:B:348:TYR:HB3	1.96	0.46
1:B:294:LEU:HD22	1:B:409:TYR:HD2	1.80	0.46
1:B:108:LEU:O	1:B:112:ILE:HG12	2.15	0.46
1:B:234:TYR:CE2	1:B:333:ARG:HG3	2.51	0.45
1:A:477:SER:CB	9:A:9:AKR:HB2	2.45	0.45
1:A:385:TYR:HE1	2:A:1:ACD:H191	1.82	0.45
9:A:10:AKR:HB2	10:A:929:HOH:O	2.17	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:213:ASP:OD1	1:A:215:LYS:HG2	2.17	0.45
1:A:385:TYR:OH	2:A:1:ACD:H191	2.18	0.44
1:B:161:THR:CG2	1:B:166:LYS:O	2.59	0.44
2:A:1:ACD:C16	2:A:1:ACD:C12	2.96	0.44
1:A:216:ARG:HH11	5:A:672:NAG:C8	2.23	0.44
1:A:179:GLU:HB3	7:A:703:BOG:H3'2	2.01	0.43
1:B:484:GLU:CG	1:B:485:LYS:N	2.81	0.43
1:B:386:HIS:HB2	1:B:504[A]:TYR:CE1	2.53	0.43
2:A:1:ACD:H162	2:A:1:ACD:H12	2.00	0.43
1:B:532:LYS:HD2	1:B:532:LYS:HA	1.89	0.43
1:A:161:THR:HG22	1:A:164:GLY:N	2.34	0.43
1:B:161:THR:HG22	1:B:164:GLY:N	2.34	0.43
2:B:1:ACD:H131	2:B:1:ACD:H102	1.83	0.43
1:A:364:GLU:HB3	10:A:959:HOH:O	2.18	0.43
1:A:382:ASN:O	1:A:386:HIS:HD2	2.02	0.43
1:B:385:TYR:CE1	2:B:1:ACD:H181	2.54	0.42
1:A:360:LYS:HE2	1:A:362:ASP:HB2	2.01	0.42
1:B:557:LYS:HE2	10:B:943:HOH:O	2.19	0.42
1:A:398:GLU:HB3	1:A:399:ASP:H	1.65	0.42
1:A:532:LYS:HD2	1:A:532:LYS:HA	1.81	0.41
1:A:161:THR:HG22	1:A:164:GLY:H	1.85	0.41
8:B:7:EDO:H21	10:B:901:HOH:O	2.20	0.41
1:A:382:ASN:O	1:A:386:HIS:CD2	2.72	0.41
1:A:279:ILE:HG22	1:A:284:GLN:HG3	2.03	0.41
1:A:161:THR:HG23	10:A:25:HOH:O	2.20	0.41
1:B:34:HIS:HA	1:B:35:PRO:HD3	1.97	0.41
1:B:352:LEU:HA	1:B:352:LEU:HD23	1.90	0.41

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	554/593 (93%)	541 (98%)	13 (2%)	0	100	100
1	B	556/593 (94%)	544 (98%)	12 (2%)	0	100	100
All	All	1110/1186 (94%)	1085 (98%)	25 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	485/531 (91%)	479 (99%)	6 (1%)	78	83
1	B	480/531 (90%)	471 (98%)	9 (2%)	65	69
All	All	965/1062 (91%)	950 (98%)	15 (2%)	70	76

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	111	LEU
1	A	161	THR
1	A	232	HIS
1	A	289	GLN
1	A	385	TYR
1	A	542	PRO
1	B	161	THR
1	B	224	LEU
1	B	232	HIS
1	B	281	GLU
1	B	289	GLN
1	B	306	LEU
1	B	385	TYR
1	B	409	TYR
1	B	484	GLU

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



Mol	Chain	Res	Type
1	B	39	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 ⓘ

9 carbohydrates 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	$\# Z  > 2$
4	NAG	A	661	1,4	14,14,15	0.60	0	15,19,21	0.99	1 (6%)
4	NAG	A	662	4	14,14,15	0.80	0	15,19,21	0.63	0
5	NAG	A	671	1,5	14,14,15	0.50	0	15,19,21	1.49	2 (13%)
5	NAG	A	672	5	14,14,15	0.49	0	15,19,21	0.60	0
5	MAN	A	673	5	11,11,12	0.53	0	14,15,17	1.03	1 (7%)
4	NAG	B	661	1,4	14,14,15	1.49	1 (7%)	15,19,21	1.15	1 (6%)
4	NAG	B	662	4	14,14,15	0.78	1 (7%)	15,19,21	1.67	4 (26%)
4	NAG	B	671	1,4	14,14,15	0.51	0	15,19,21	1.23	1 (6%)
4	NAG	B	672	4	14,14,15	0.56	0	15,19,21	1.20	1 (6%)

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
4	NAG	A	661	1,4	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	662	4	1/1/5/7	0/6/23/26	0/1/1/1
5	NAG	A	671	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	672	5	-	0/6/23/26	0/1/1/1
5	MAN	A	673	5	-	0/2/19/22	0/1/1/1
4	NAG	B	661	1,4	1/1/5/7	0/6/23/26	0/1/1/1
4	NAG	B	662	4	-	0/6/23/26	0/1/1/1
4	NAG	B	671	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	672	4	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	661	NAG	O5-C1	-5.13	1.35	1.43
4	B	662	NAG	O5-C1	-2.08	1.40	1.43

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	662	NAG	C2-N2-C7	-3.41	118.65	123.04
4	B	662	NAG	C1-O5-C5	-3.01	108.42	112.25
5	A	671	NAG	C2-N2-C7	-2.82	119.42	123.04
4	A	661	NAG	C3-C4-C5	-2.06	106.60	110.20
4	B	662	NAG	O5-C5-C6	2.03	111.74	107.35
4	B	662	NAG	C4-C3-C2	3.01	115.90	111.23
5	A	673	MAN	C1-O5-C5	3.05	116.12	112.25
4	B	661	NAG	C3-C4-C5	3.10	115.59	110.20
4	B	672	NAG	C4-C3-C2	3.10	116.04	111.23
4	B	671	NAG	C1-O5-C5	3.58	116.79	112.25
5	A	671	NAG	C1-O5-C5	4.72	118.23	112.25

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	662	NAG	C1
4	B	661	NAG	C5

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	672	NAG	4	0
4	B	672	NAG	3	0

## 5.6 Ligand geometry

17 ligands 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	ACD	A	1	-	18,21,21	0.29	0	18,21,21	0.31	0
9	AKR	A	10	-	1,4,4	1.52	0	0,4,4	0.00	-
8	EDO	A	2	-	3,3,3	0.52	0	2,2,2	0.43	0
8	EDO	A	5	-	3,3,3	0.44	0	2,2,2	0.34	0
3	COH	A	619	1	28,50,50	2.25	8 (28%)	19,82,82	1.39	3 (15%)
6	NAG	A	681	1	14,14,15	0.49	0	15,19,21	0.91	1 (6%)
7	BOG	A	703	-	20,20,20	0.46	0	25,25,25	0.68	0
9	AKR	A	9	-	1,4,4	1.53	0	0,4,4	0.00	-
2	ACD	B	1	-	18,21,21	0.24	0	18,21,21	0.78	0
8	EDO	B	3	-	3,3,3	0.54	0	2,2,2	0.58	0
8	EDO	B	4	-	3,3,3	0.50	0	2,2,2	0.38	0
8	EDO	B	6	-	3,3,3	0.52	0	2,2,2	0.39	0
3	COH	B	619	1	28,50,50	2.22	9 (32%)	19,82,82	1.98	6 (31%)
8	EDO	B	620	-	3,3,3	0.55	0	2,2,2	0.34	0
6	NAG	B	681	1	14,14,15	0.49	0	15,19,21	1.35	1 (6%)
8	EDO	B	7	-	3,3,3	0.51	0	2,2,2	0.45	0
8	EDO	B	8	-	3,3,3	0.41	0	2,2,2	0.55	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	ACD	A	1	-	-	0/17/19/19	0/0/0/0
9	AKR	A	10	-	-	0/0/2/2	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	EDO	A	2	-	-	0/1/1/1	0/0/0/0
8	EDO	A	5	-	-	0/1/1/1	0/0/0/0
3	COH	A	619	1	-	0/6/54/54	0/0/8/8
6	NAG	A	681	1	-	0/6/23/26	0/1/1/1
7	BOG	A	703	-	-	0/11/31/31	0/1/1/1
9	AKR	A	9	-	-	0/0/2/2	0/0/0/0
2	ACD	B	1	-	-	0/17/19/19	0/0/0/0
8	EDO	B	3	-	-	0/1/1/1	0/0/0/0
8	EDO	B	4	-	-	0/1/1/1	0/0/0/0
8	EDO	B	6	-	-	0/1/1/1	0/0/0/0
3	COH	B	619	1	-	0/6/54/54	0/0/8/8
8	EDO	B	620	-	-	0/1/1/1	0/0/0/0
6	NAG	B	681	1	-	0/6/23/26	0/1/1/1
8	EDO	B	7	-	-	0/1/1/1	0/0/0/0
8	EDO	B	8	-	-	0/1/1/1	0/0/0/0

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	619	COH	C3B-C2B	-4.73	1.34	1.40
3	B	619	COH	C3C-C2C	-4.42	1.34	1.40
3	B	619	COH	C3B-C2B	-3.83	1.35	1.40
3	A	619	COH	C3C-C2C	-3.80	1.35	1.40
3	B	619	COH	CAA-C2A	2.06	1.56	1.52
3	A	619	COH	CO-ND	2.35	2.09	1.97
3	A	619	COH	CO-NA	2.37	2.09	1.97
3	B	619	COH	CO-ND	2.49	2.10	1.97
3	B	619	COH	CO-NA	2.60	2.11	1.97
3	B	619	COH	CO-NB	2.78	2.12	1.97
3	B	619	COH	C3C-CAC	3.35	1.54	1.47
3	A	619	COH	CO-NB	3.41	2.15	1.97
3	A	619	COH	C3B-CAB	3.57	1.55	1.47
3	A	619	COH	C3C-CAC	3.58	1.55	1.47
3	B	619	COH	C3B-CAB	4.03	1.56	1.47
3	B	619	COH	C3D-C2D	5.34	1.53	1.37
3	A	619	COH	C3D-C2D	5.43	1.53	1.37

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	619	COH	CBA-CAA-C2A	-4.69	104.12	112.53
3	B	619	COH	CBD-CAD-C3D	-4.30	104.82	112.53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	619	COH	C3C-CAC-CBC	-2.48	121.25	126.32
3	A	619	COH	C3C-CAC-CBC	-2.30	121.60	126.32
3	A	619	COH	CBA-CAA-C2A	-2.14	108.70	112.53
3	B	619	COH	CMB-C2B-C1B	-2.10	124.89	128.36
3	A	619	COH	CAA-C2A-C1A	-2.05	124.78	127.01
3	B	619	COH	C1D-ND-C4D	2.03	112.50	105.58
3	B	619	COH	CAA-CBA-CGA	2.67	117.65	112.75
6	A	681	NAG	C1-O5-C5	2.93	115.96	112.25
6	B	681	NAG	C1-O5-C5	4.45	117.90	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	ACD	10	0
9	A	10	AKR	1	0
8	A	2	EDO	1	0
7	A	703	BOG	1	0
9	A	9	AKR	3	0
2	B	1	ACD	10	0
3	B	619	COH	4	0
8	B	7	EDO	1	0
8	B	8	EDO	1	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	551/593 (92%)	0.30	30 (5%)	29 40	20, 29, 39, 47	0
1	B	552/593 (93%)	0.37	48 (8%)	13 19	17, 29, 39, 48	0
All	All	1103/1186 (93%)	0.33	78 (7%)	19 26	17, 29, 39, 48	0

All (78) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	81	LEU	6.8
1	B	81	LEU	6.6
1	B	74	PHE	6.5
1	B	82	LEU	6.4
1	B	91	TYR	5.3
1	A	74	PHE	5.1
1	A	121	SER	4.9
1	B	583	GLN	4.7
1	A	82	LEU	4.7
1	B	78	ILE	4.5
1	B	107	PHE	4.2
1	A	115	TYR	4.1
1	A	122	TYR	3.9
1	B	122	TYR	3.9
1	A	78	ILE	3.8
1	B	115	TYR	3.7
1	B	65	TYR	3.7
1	B	102	ILE	3.7
1	A	107	PHE	3.6
1	B	148	TYR	3.5
1	A	278	HIS	3.4
1	A	106	PRO	3.4
1	B	96	PHE	3.4
1	A	399	ASP	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	80	LEU	3.4
1	B	79	LYS	3.4
1	B	84	PRO	3.3
1	A	148	TYR	3.2
1	B	377	ILE	3.2
1	B	101	ASN	3.1
1	B	75	LEU	3.0
1	B	121	SER	3.0
1	B	368[A]	ASN	3.0
1	B	71	THR	2.9
1	B	73	GLU	2.9
1	B	149	THR	2.8
1	B	80	LEU	2.8
1	B	98	GLY	2.7
1	A	54	GLN	2.7
1	B	77	ARG	2.7
1	B	399[A]	ASP	2.7
1	B	52	PHE	2.7
1	A	368	ASN	2.6
1	B	287	VAL	2.6
1	B	33	HIS	2.6
1	A	215	LYS	2.6
1	B	476	THR	2.6
1	B	119	SER	2.5
1	B	577[A]	PHE	2.5
1	A	422	PHE	2.4
1	A	101	ASN	2.3
1	A	84	PRO	2.3
1	B	76	THR	2.3
1	A	108	LEU	2.3
1	B	244	LEU	2.3
1	B	108	LEU	2.3
1	B	409	TYR	2.3
1	B	157	ASP	2.3
1	A	91	TYR	2.2
1	A	105	ASN	2.2
1	B	239	ASP	2.2
1	B	105	ASN	2.2
1	B	296	PRO	2.2
1	B	332	SER	2.2
1	B	539	ILE	2.2
1	B	83	LYS	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	111	LEU	2.1
1	B	286	ALA	2.1
1	A	509	VAL	2.1
1	A	540	CYS	2.1
1	A	53	ASP	2.1
1	B	214	HIS	2.1
1	B	63	GLY	2.1
1	A	96	PHE	2.1
1	A	277	PRO	2.0
1	B	378	ALA	2.0
1	A	79	LYS	2.0
1	A	105(A)	ILE	2.0

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

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

## 6.3 Carbohydrates ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	B	661	14/15	0.78	0.31	3.40	44,49,53,59	0
4	NAG	B	671	14/15	0.94	0.10	-1.10	23,29,33,39	0
5	NAG	A	671	14/15	0.96	0.09	-1.46	26,32,34,39	0
4	NAG	B	672	14/15	0.83	0.28	-	44,46,49,50	0
5	NAG	A	672	14/15	0.88	0.15	-	44,46,51,57	0
4	NAG	A	662	14/15	0.78	0.45	-	60,63,65,65	0
4	NAG	B	662	14/15	0.83	0.48	-	64,67,68,68	0
5	MAN	A	673	11/12	0.82	0.42	-	61,64,65,66	0
4	NAG	A	661	14/15	0.88	0.23	-	40,47,51,56	0

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron



density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	EDO	B	7	4/4	0.91	0.33	35.86	40,40,41,42	0
8	EDO	B	3	4/4	0.90	0.27	6.81	41,41,42,43	0
2	ACD	B	1	22/22	0.82	0.26	3.86	52,58,64,65	0
2	ACD	A	1	22/22	0.82	0.25	3.35	53,57,64,64	0
8	EDO	B	620	4/4	0.94	0.20	2.24	26,32,32,34	0
6	NAG	B	681	14/15	0.90	0.23	1.98	42,49,51,52	0
8	EDO	B	4	4/4	0.80	0.27	1.94	48,49,51,51	0
8	EDO	A	5	4/4	0.93	0.15	1.64	36,37,37,38	0
7	BOG	A	703	20/20	0.93	0.14	1.31	34,38,42,43	0
9	AKR	A	9	5/5	0.91	0.19	1.25	53,53,54,54	0
6	NAG	A	681	14/15	0.87	0.26	1.24	38,44,48,49	0
9	AKR	A	10	5/5	0.89	0.21	0.71	47,47,48,48	0
8	EDO	B	8	4/4	0.79	0.18	0.53	52,52,52,53	0
3	COH	B	619	43/43	0.91	0.16	0.46	35,37,49,53	0
3	COH	A	619	43/43	0.91	0.16	0.39	35,39,51,56	0
8	EDO	A	2	4/4	0.94	0.13	-0.12	38,40,40,42	0
8	EDO	B	6	4/4	0.77	0.16	-	56,57,57,58	0

## 6.5 Other polymers

There are no such residues in this entry.