



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 09:04 PM GMT

PDB ID : 4ZM6  
Title : A unique GCN5-related glucosamine N-acetyltransferase region exist in the fungal multi-domain GH3 beta-N-acetylglucosaminidase  
Authors : Qin, Z.; Xiao, Y.; Yang, X.; Jiang, Z.; Yang, S.; Mesters, J.R.  
Deposited on : 2015-05-02  
Resolution : 2.80 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

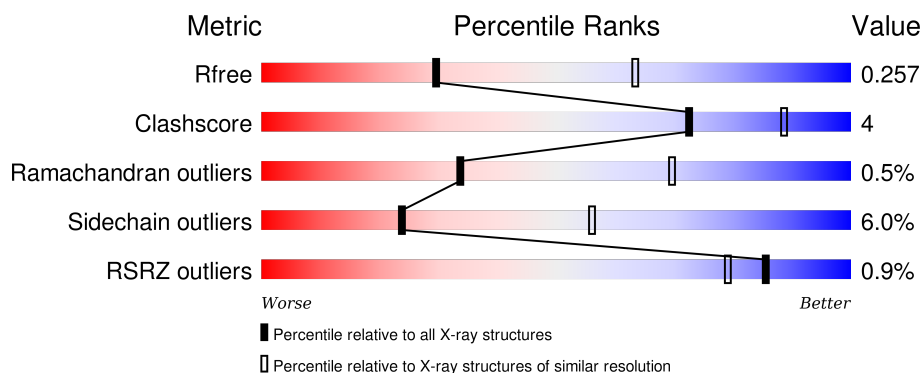
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*



The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	858	 86% 11% ..
1	B	858	 86% 11% ..

## 2 Entry composition [i](#)

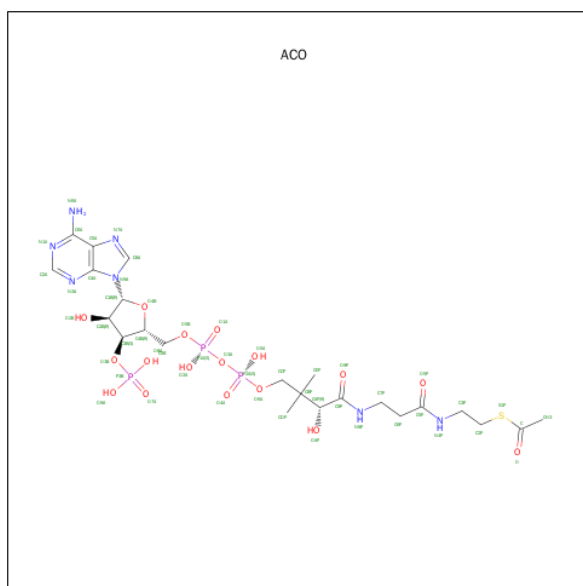
There are 4 unique types of molecules in this entry. The entry contains 13467 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 N-acetyl-beta-D glucosaminidase.

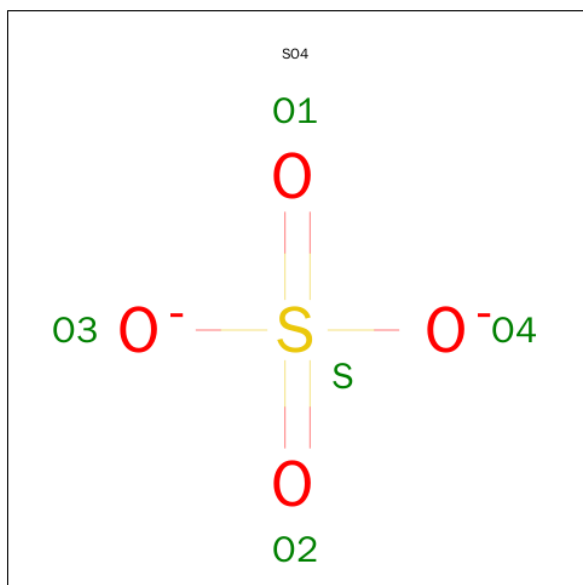
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	844	Total	C	N	O	S	0	2	0
			6628	4206	1159	1236	27			
1	B	844	Total	C	N	O	S	0	0	0
			6608	4199	1155	1228	26			

- Molecule 2 is ACETYL COENZYME \*A (three-letter code: ACO) (formula: C<sub>23</sub>H<sub>38</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P S	0	0
			51	23	7	17	3 1		
2	B	1	Total	C	N	O	P S	0	0
			51	23	7	17	3 1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

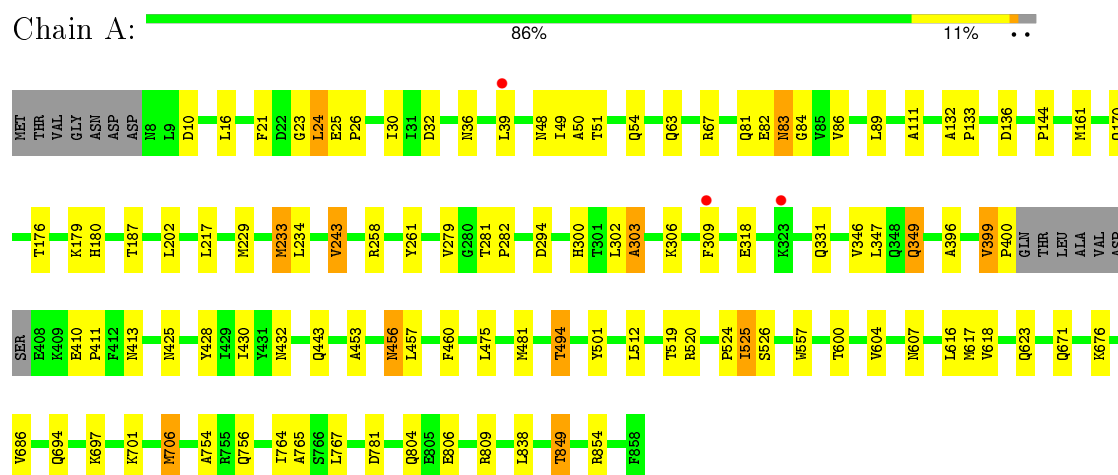
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	70	Total	O	0	0
			70	70		
4	B	44	Total	O	0	0
			44	44		

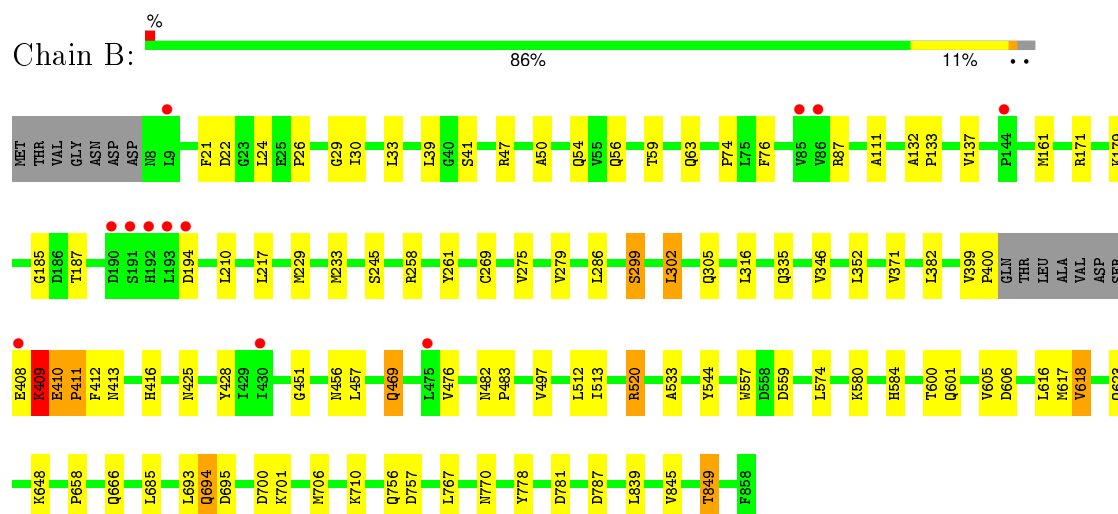
### 3 Residue-property plots

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: N-acetyl-beta-D glucosaminidase



- Molecule 1: N-acetyl-beta-D glucosaminidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	245.03 Å   245.03 Å   94.52 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	86.63 – 2.80 86.63 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.0 (86.63-2.80) 99.0 (86.63-2.80)	Depositor EDS
$R_{merge}$	0.01	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.50 (at 2.82 Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.229   ,   0.254 0.233   ,   0.257	Depositor DCC
$R_{free}$ test set	3440 reflections (5.14%)	DCC
Wilson B-factor (Å <sup>2</sup> )	62.7	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 46.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 70415 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13467	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: ACO, SO4

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	0.34	1/6772 (0.0%)	0.53	1/9194 (0.0%)
1	B	0.32	1/6750 (0.0%)	0.52	1/9160 (0.0%)
All	All	0.33	2/13522 (0.0%)	0.53	2/18354 (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	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	411	PRO	N-CD	5.53	1.55	1.47
1	A	411	PRO	N-CD	5.25	1.55	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	410	GLU	C-N-CD	5.57	140.10	128.40
1	B	410	GLU	C-N-CD	5.21	139.35	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	757	ASP	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	A	6628	0	6523	51	0
1	B	6608	0	6541	59	0
2	A	51	0	34	2	0
2	B	51	0	34	1	0
3	A	5	0	0	0	0
3	B	10	0	0	0	0
4	A	70	0	0	2	0
4	B	44	0	0	1	0
All	All	13467	0	13132	111	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 (111) 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:399:VAL:CG1	1:B:400:PRO:HD2	1.15	1.62
1:B:399:VAL:CG1	1:B:400:PRO:CD	2.02	1.36
1:B:399:VAL:HG13	1:B:400:PRO:CD	1.59	1.30
1:A:399:VAL:HG13	1:A:400:PRO:HD2	1.38	1.05
1:B:399:VAL:HG12	1:B:400:PRO:HD2	1.02	1.01
1:B:408:GLU:N	1:B:408:GLU:OE2	1.95	1.00
1:B:399:VAL:HG12	1:B:400:PRO:CD	1.76	0.99
1:B:399:VAL:HG13	1:B:400:PRO:HD2	0.96	0.94
1:B:399:VAL:HG13	1:B:400:PRO:HD3	1.60	0.80
1:B:269:CYS:SG	1:B:299:SER:OG	2.40	0.79
1:B:63:GLN:HG2	1:B:346:VAL:HG12	1.70	0.72
1:B:409:LYS:HD3	1:B:410:GLU:H	1.55	0.70
1:B:839:LEU:HD11	1:B:849:THR:HG22	1.78	0.66
1:A:413:ASN:OD1	1:A:428:TYR:OH	2.15	0.62
1:A:399:VAL:HG13	1:A:400:PRO:CD	2.24	0.62
1:B:483:PRO:HB3	1:B:497:VAL:HG21	1.82	0.61
1:B:409:LYS:CD	1:B:410:GLU:H	2.12	0.61
1:A:519:THR:HG22	4:A:1054:HOH:O	2.00	0.60
1:B:410:GLU:HB3	1:B:413:ASN:OD1	2.03	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:399:VAL:HG12	1:B:400:PRO:N	2.18	0.59
1:A:756:GLN:HE22	1:A:809:ARG:HH22	1.50	0.59
1:B:409:LYS:CG	1:B:410:GLU:H	2.14	0.58
1:A:671:GLN:HE22	1:A:854:ARG:HE	1.52	0.57
1:A:136:ASP:OD2	1:A:180:HIS:HD2	1.87	0.57
1:B:408:GLU:O	1:B:409:LYS:HB2	2.04	0.56
1:B:839:LEU:CD1	1:B:849:THR:HG22	2.35	0.56
1:A:21:PHE:CD2	1:A:30:ILE:HD13	2.41	0.55
1:B:63:GLN:CG	1:B:346:VAL:HG12	2.36	0.55
1:A:111:ALA:HB2	1:A:161:MET:CE	2.37	0.54
1:A:706:MET:HE3	1:A:706:MET:HA	1.90	0.53
1:A:24:LEU:HD23	1:A:50:ALA:HB2	1.90	0.53
1:B:469:GLN:HA	1:B:469:GLN:HE21	1.74	0.52
1:B:710:LYS:O	1:B:756:GLN:O	2.28	0.52
1:A:50:ALA:HB3	1:A:54:GLN:HB3	1.91	0.52
1:A:217:LEU:HD22	1:A:261:TYR:CD1	2.45	0.52
1:B:409:LYS:CG	1:B:410:GLU:N	2.72	0.52
1:B:21:PHE:CE1	1:B:30:ILE:HD13	2.46	0.51
1:B:74:PRO:HB3	1:B:346:VAL:HG11	1.92	0.51
1:B:21:PHE:CD1	1:B:30:ILE:HD13	2.46	0.50
1:A:132:ALA:HB1	1:A:133:PRO:HA	1.94	0.50
1:A:16:LEU:HG	1:A:294:ASP:O	2.11	0.50
1:A:63:GLN:HG2	1:A:346:VAL:HG12	1.93	0.50
1:B:217:LEU:HD22	1:B:261:TYR:CD1	2.46	0.49
1:B:648:LYS:NZ	1:B:658:PRO:O	2.46	0.49
1:B:111:ALA:HB2	1:B:161:MET:CE	2.42	0.49
1:B:580:LYS:O	1:B:601:GLN:HB2	2.13	0.49
1:B:409:LYS:O	1:B:411:PRO:HD3	2.13	0.48
1:A:170:GLN:NE2	1:A:176:THR:OG1	2.47	0.48
1:A:302:LEU:HD12	1:A:303:ALA:N	2.28	0.48
1:A:618:VAL:HG22	1:A:623:GLN:HG2	1.96	0.48
1:B:520:ARG:NH2	2:B:901:ACO:O2B	2.47	0.48
1:A:399:VAL:CG1	1:A:400:PRO:HD2	2.25	0.48
1:B:29:GLY:HA3	1:B:302:LEU:HD21	1.95	0.48
1:B:416:HIS:CG	1:B:428:TYR:CD1	3.02	0.47
1:B:574:LEU:HD11	1:B:778:TYR:HA	1.96	0.47
1:B:382:LEU:HD22	1:B:513:ILE:HA	1.97	0.47
1:A:475:LEU:O	1:A:494:THR:HG23	2.15	0.47
1:A:302:LEU:O	1:A:303:ALA:HB2	2.15	0.46
1:A:767:LEU:C	1:A:767:LEU:HD12	2.35	0.46
1:B:63:GLN:HG2	1:B:346:VAL:CG1	2.43	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:412:PHE:CE1	1:B:451:GLY:HA3	2.50	0.46
1:A:81:GLN:HE21	1:A:83:ASN:ND2	2.13	0.46
1:A:604:VAL:CG1	1:A:607:ASN:HA	2.46	0.46
1:A:281:THR:N	1:A:282:PRO:CD	2.78	0.46
1:B:408:GLU:O	1:B:409:LYS:CB	2.64	0.46
1:B:618:VAL:HG22	1:B:623:GLN:HG2	1.97	0.45
1:A:754:ALA:HB3	1:A:765:ALA:HB3	1.98	0.45
1:A:24:LEU:CD2	1:A:50:ALA:HB2	2.46	0.45
1:A:557:TRP:HA	1:A:617:MET:CE	2.46	0.45
1:A:179:LYS:HB3	1:A:229:MET:HB3	1.97	0.45
1:A:23:GLY:O	1:A:48:ASN:HA	2.17	0.45
1:B:533:ALA:HB2	4:B:1044:HOH:O	2.14	0.45
1:B:557:TRP:HA	1:B:617:MET:CE	2.47	0.45
1:B:26:PRO:HA	1:B:30:ILE:HD11	1.99	0.44
1:A:82:GLU:O	1:A:83:ASN:HB2	2.17	0.44
1:A:302:LEU:HD12	1:A:303:ALA:H	1.81	0.44
1:A:849:THR:HG23	4:A:1043:HOH:O	2.18	0.44
1:B:411:PRO:C	1:B:413:ASN:H	2.20	0.44
1:A:302:LEU:HD11	1:A:306:LYS:HG2	2.00	0.44
1:A:764:ILE:HB	1:A:806:GLU:HG3	2.00	0.43
1:B:693:LEU:O	1:B:694:GLN:C	2.56	0.43
2:A:901:ACO:O2B	2:A:901:ACO:O8A	2.28	0.43
1:B:409:LYS:HG2	1:B:410:GLU:N	2.33	0.42
1:A:671:GLN:NE2	1:A:854:ARG:HE	2.16	0.42
1:A:26:PRO:HA	1:A:30:ILE:HD11	2.01	0.42
1:A:111:ALA:CB	1:A:161:MET:HE3	2.50	0.42
1:A:302:LEU:O	1:A:303:ALA:CB	2.67	0.42
1:B:132:ALA:HB1	1:B:133:PRO:HA	2.00	0.42
1:B:24:LEU:CD2	1:B:50:ALA:HB2	2.49	0.42
1:B:137:VAL:O	1:B:185:GLY:HA3	2.20	0.42
1:B:767:LEU:HD12	1:B:767:LEU:C	2.39	0.42
1:B:56:GLN:NE2	1:B:352:LEU:HD21	2.36	0.41
1:A:432:ASN:HA	1:A:460:PHE:CE2	2.55	0.41
1:B:413:ASN:O	1:B:416:HIS:HB3	2.19	0.41
1:A:676:LYS:NZ	2:A:901:ACO:O9A	2.44	0.41
1:B:179:LYS:HB3	1:B:229:MET:HB3	2.02	0.41
1:B:59:THR:O	1:B:63:GLN:HB2	2.21	0.41
1:A:453:ALA:HB1	1:A:481:MET:SD	2.60	0.41
1:A:81:GLN:HE21	1:A:83:ASN:HD22	1.68	0.41
1:B:544:TYR:HB3	1:B:584:HIS:HB2	2.03	0.41
1:A:32:ASP:OD1	1:A:36:ASN:ND2	2.54	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:233:MET:N	1:A:243:VAL:HG22	2.35	0.41
1:B:275:VAL:O	1:B:279:VAL:O	2.39	0.41
1:B:33:LEU:HD11	1:B:305:GLN:HE21	1.86	0.41
1:A:67:ARG:CD	1:A:347:LEU:HD11	2.51	0.41
1:A:396:ALA:O	1:A:430:ILE:HA	2.21	0.41
1:A:524:PRO:O	1:A:525:ILE:HD12	2.21	0.41
1:B:24:LEU:HD23	1:B:50:ALA:HB2	2.03	0.40
1:A:49:ILE:HG22	1:A:49:ILE:O	2.21	0.40
1:A:25:GLU:HG2	1:A:26:PRO:HD2	2.03	0.40
1:B:482:ASN:HD22	1:B:482:ASN:N	2.20	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	842/858 (98%)	795 (94%)	40 (5%)	7 (1%)	24	58
1	B	840/858 (98%)	797 (95%)	41 (5%)	2 (0%)	52	84
All	All	1682/1716 (98%)	1592 (95%)	81 (5%)	9 (0%)	34	69

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	409	LYS
1	A	83	ASN
1	A	300	HIS
1	A	303	ALA
1	A	349	GLN
1	A	456	ASN
1	B	371	VAL
1	A	144	PRO

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	84	GLY

### 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	704/731 (96%)	665 (94%)	39 (6%)	27	59
1	B	702/731 (96%)	657 (94%)	45 (6%)	22	52
All	All	1406/1462 (96%)	1322 (94%)	84 (6%)	24	56

All (84) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	ASP
1	A	24	LEU
1	A	39	LEU
1	A	51	THR
1	A	86	VAL
1	A	89	LEU
1	A	187	THR
1	A	202	LEU
1	A	233	MET
1	A	234	LEU
1	A	243	VAL
1	A	258	ARG
1	A	279	VAL
1	A	309	PHE
1	A	318	GLU
1	A	331	GLN
1	A	349	GLN
1	A	399	VAL
1	A	425	ASN
1	A	443	GLN
1	A	456	ASN
1	A	457	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	494	THR
1	A	501	TYR
1	A	512	LEU
1	A	520	ARG
1	A	525	ILE
1	A	526	SER
1	A	600	THR
1	A	616	LEU
1	A	686	VAL
1	A	694	GLN
1	A	697	LYS
1	A	701	LYS
1	A	706	MET
1	A	781	ASP
1	A	804	GLN
1	A	838	LEU
1	A	849	THR
1	B	22	ASP
1	B	39	LEU
1	B	41	SER
1	B	47	ARG
1	B	54	GLN
1	B	76	PHE
1	B	87	ARG
1	B	171	ARG
1	B	187	THR
1	B	194	ASP
1	B	210	LEU
1	B	233	MET
1	B	245	SER
1	B	258	ARG
1	B	286	LEU
1	B	299	SER
1	B	302	LEU
1	B	316	LEU
1	B	335	GLN
1	B	409	LYS
1	B	425	ASN
1	B	456	ASN
1	B	457	LEU
1	B	469	GLN
1	B	476	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	512	LEU
1	B	520	ARG
1	B	559	ASP
1	B	600	THR
1	B	605	VAL
1	B	606	ASP
1	B	616	LEU
1	B	618	VAL
1	B	666	GLN
1	B	685	LEU
1	B	694	GLN
1	B	695	ASP
1	B	700	ASP
1	B	701	LYS
1	B	706	MET
1	B	770	ASN
1	B	781	ASP
1	B	787	ASP
1	B	845	VAL
1	B	849	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (42) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	36	ASN
1	A	60	HIS
1	A	64	GLN
1	A	83	ASN
1	A	138	ASN
1	A	153	GLN
1	A	170	GLN
1	A	180	HIS
1	A	305	GLN
1	A	331	GLN
1	A	339	GLN
1	A	349	GLN
1	A	425	ASN
1	A	432	ASN
1	A	440	GLN
1	A	443	GLN
1	A	456	ASN
1	A	467	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	469	GLN
1	A	486	GLN
1	A	584	HIS
1	A	671	GLN
1	A	756	GLN
1	A	770	ASN
1	B	56	GLN
1	B	138	ASN
1	B	153	GLN
1	B	170	GLN
1	B	184	HIS
1	B	305	GLN
1	B	339	GLN
1	B	425	ASN
1	B	440	GLN
1	B	456	ASN
1	B	467	GLN
1	B	469	GLN
1	B	482	ASN
1	B	486	GLN
1	B	584	HIS
1	B	610	HIS
1	B	671	GLN
1	B	756	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 ⓘ

5 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	ACO	A	901	-	44,53,53	0.88	1 (2%)	54,79,79	1.58	6 (11%)
3	SO4	A	902	-	4,4,4	0.36	0	6,6,6	0.19	0
2	ACO	B	901	-	44,53,53	0.86	1 (2%)	54,79,79	1.46	3 (5%)
3	SO4	B	902	-	4,4,4	0.41	0	6,6,6	0.14	0
3	SO4	B	903	-	4,4,4	0.36	0	6,6,6	0.10	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	ACO	A	901	-	-	0/47/67/67	0/3/3/3
3	SO4	A	902	-	-	0/0/0/0	0/0/0/0
2	ACO	B	901	-	-	0/47/67/67	0/3/3/3
3	SO4	B	902	-	-	0/0/0/0	0/0/0/0
3	SO4	B	903	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	901	ACO	C5A-C4A	3.13	1.47	1.40
2	A	901	ACO	C5A-C4A	3.32	1.48	1.40

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	901	ACO	N3A-C2A-N1A	-7.81	122.73	128.87
2	A	901	ACO	N3A-C2A-N1A	-7.28	123.15	128.87
2	A	901	ACO	O5P-C5P-C6P	-3.82	115.33	121.97
2	B	901	ACO	C3P-C2P-S1P	-2.19	105.65	111.47
2	A	901	ACO	N6A-C6A-N1A	2.04	121.93	118.52
2	A	901	ACO	P3B-O3B-C3B	2.36	127.60	121.56
2	B	901	ACO	O4B-C1B-N9A	2.71	113.23	108.11

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	901	ACO	C3B-C2B-C1B	2.73	106.01	100.06
2	A	901	ACO	C6P-C5P-N4P	3.50	122.54	116.46

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	901	ACO	2	0
2	B	901	ACO	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	844/858 (98%)	-0.03	3 (0%) 93 90	31, 60, 110, 130	0
1	B	844/858 (98%)	0.04	12 (1%) 78 69	41, 65, 96, 125	0
All	All	1688/1716 (98%)	0.01	15 (0%) 85 79	31, 63, 104, 130	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	85	VAL	3.1
1	B	194	ASP	3.1
1	B	193	LEU	3.0
1	B	475	LEU	3.0
1	B	408	GLU	3.0
1	B	192	HIS	2.9
1	B	190	ASP	2.6
1	A	309	PHE	2.6
1	B	9	LEU	2.4
1	A	323	LYS	2.4
1	B	144	PRO	2.4
1	B	86	VAL	2.3
1	A	39	LEU	2.3
1	B	191	SER	2.2
1	B	430	ILE	2.1

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ACO	A	901	51/51	0.95	0.21	0.78	34,41,67,68	51
3	SO4	B	903	5/5	0.76	0.18	0.25	130,132,136,136	0
2	ACO	B	901	51/51	0.94	0.21	0.17	46,54,76,80	51
3	SO4	B	902	5/5	0.95	0.12	-1.44	83,86,88,90	0
3	SO4	A	902	5/5	0.96	0.12	-1.86	78,80,84,84	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.