Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2020 – 07:28 PM BST

PDB ID	:	4AZH
Title	:	Differential inhibition of the tandem GH20 catalytic modules in the pneumococcal exo-beta-D-N-acetylglucosaminidase, StrH
Authors	:	Pluvinage, B.; Stubbs, K.A.; Vocadlo, D.J.; Boraston, A.B.
Deposited on	:	2012-06-26
Resolution	:	2.22 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*
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:

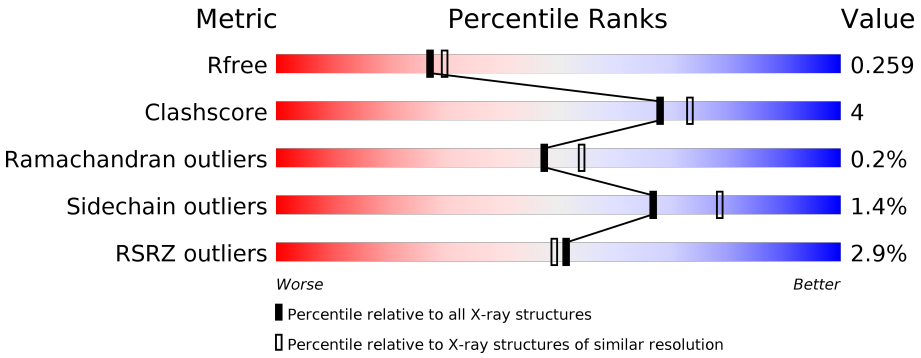
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.22 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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 >=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 <=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	443	4% 86% 9% 5%
1	B	443	3% 82% 12% 5%
1	C	443	2% 84% 9% 7%
1	D	443	2% 85% 7% • 7%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14690 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.

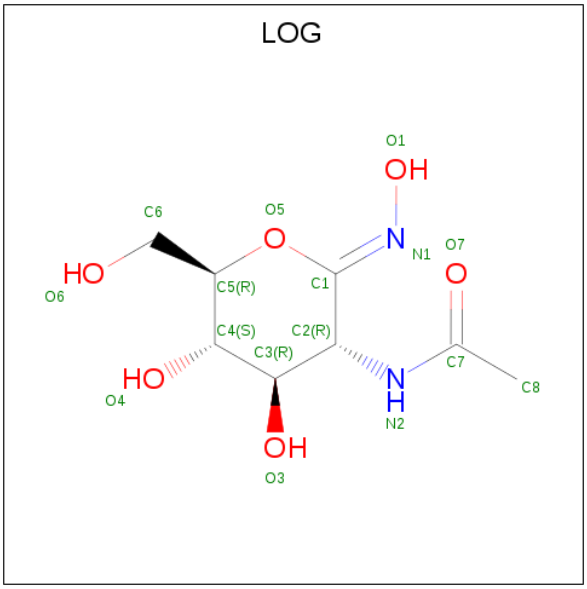
- Molecule 1 is a protein called BETA-N-ACETYLHEXOSAMINIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	421	Total	C	N	O	S	1	4	0
			3361	2152	548	649	12			
1	B	419	Total	C	N	O	S	3	3	0
			3346	2144	545	645	12			
1	C	414	Total	C	N	O	S	0	3	0
			3302	2118	537	635	12			
1	D	411	Total	C	N	O	S	7	2	0
			3276	2100	534	630	12			

There are 20 discrepancies between the modelled and reference sequences:

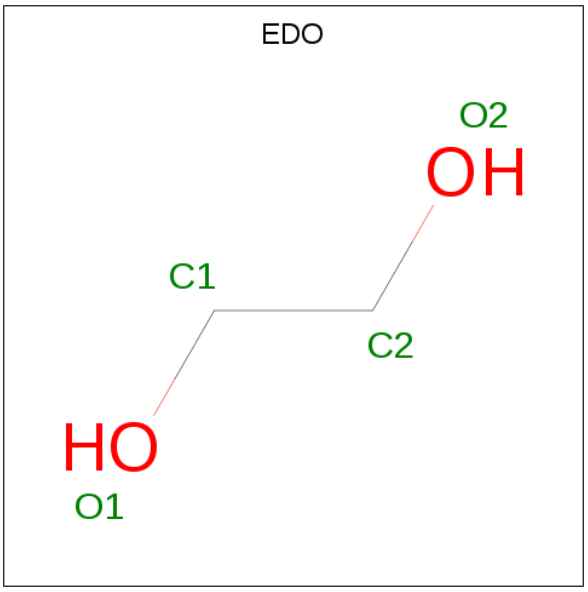
Chain	Residue	Modelled	Actual	Comment	Reference
A	622	ARG	-	expression tag	UNP P49610
A	623	GLY	-	expression tag	UNP P49610
A	624	SER	-	expression tag	UNP P49610
A	625	HIS	-	expression tag	UNP P49610
A	626	MET	-	expression tag	UNP P49610
B	622	ARG	-	expression tag	UNP P49610
B	623	GLY	-	expression tag	UNP P49610
B	624	SER	-	expression tag	UNP P49610
B	625	HIS	-	expression tag	UNP P49610
B	626	MET	-	expression tag	UNP P49610
C	622	ARG	-	expression tag	UNP P49610
C	623	GLY	-	expression tag	UNP P49610
C	624	SER	-	expression tag	UNP P49610
C	625	HIS	-	expression tag	UNP P49610
C	626	MET	-	expression tag	UNP P49610
D	622	ARG	-	expression tag	UNP P49610
D	623	GLY	-	expression tag	UNP P49610
D	624	SER	-	expression tag	UNP P49610
D	625	HIS	-	expression tag	UNP P49610
D	626	MET	-	expression tag	UNP P49610

- Molecule 2 is N-acetylglucosaminono-1,5-lactone (Z)-oxime (three-letter code: LOG) (formula: $C_8H_{14}N_2O_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			16	8	2	6		
2	B	1	Total	C	N	O	0	0
			16	8	2	6		
2	C	1	Total	C	N	O	0	0
			16	8	2	6		
2	D	1	Total	C	N	O	0	0
			16	8	2	6		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 4 is water.

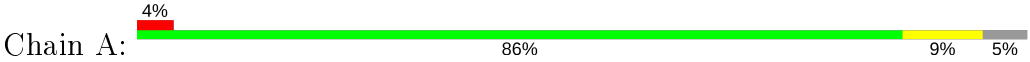
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	355	Total O 355 355	0	0
4	B	310	Total O 310 310	0	0
4	C	330	Total O 330 330	0	0
4	D	298	Total O 298 298	0	0

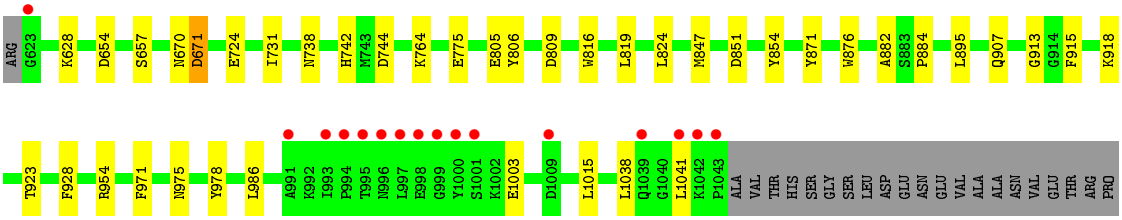
3 Residue-property plots (i)

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: BETA-N-ACETYLHEXOSAMINIDASE

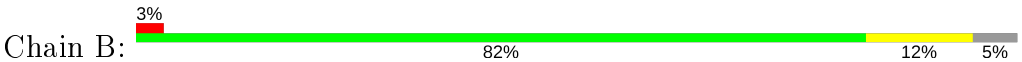
Chain A:

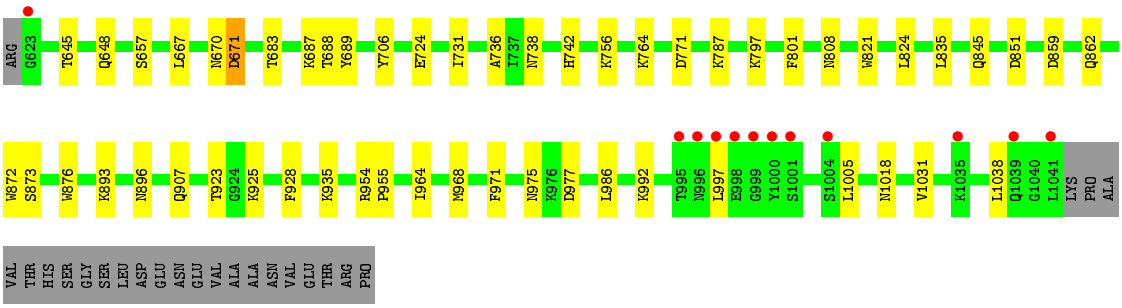




• Molecule 1: BETA-N-ACETYLHEXOSAMINIDASE

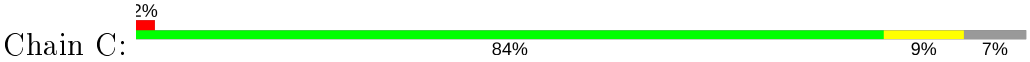
Chain B:

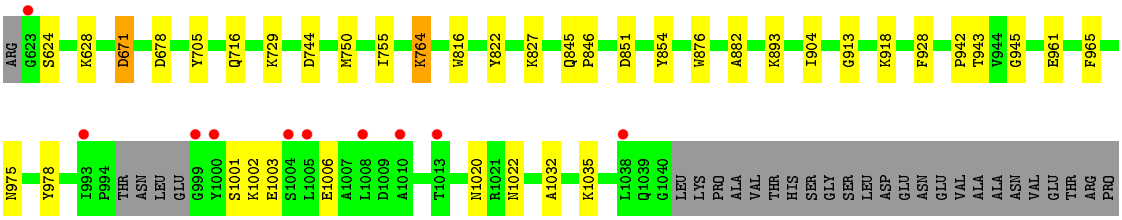




• Molecule 1: BETA-N-ACETYLHEXOSAMINIDASE

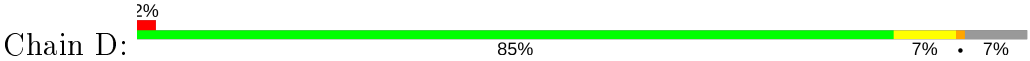
Chain C:

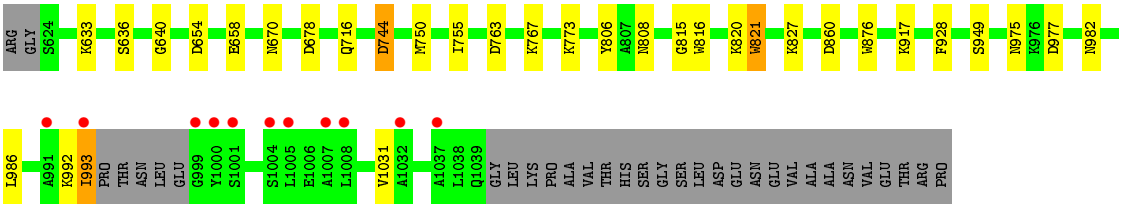




• Molecule 1: BETA-N-ACETYLHEXOSAMINIDASE

Chain D:




ARG
GLY
S624
K633
S636
G640
D654
E658
N670
D678
Q716
D744
M750
I755
D763
K767
K773
Y806
A807
N808
G815
W816
K820
W821
K827
D860
W876
K917
F928
S949
N975
K976
D977
N982
L986
A991
K992
I993
PRO
THR
ASN
LEU
GLU
G999
Y1000
S1001
S1004
L1005
E1006
A1007
L1008
V1031
A1032
A1037
L1038
Q1039
GLY
LEU
LYS
PRO
ALA
VAL
THR
HIS
SER
GLY
SER
LEU
ASP
GLU
ASN
GLU
VAL
ALA
ALA
ASN
VAL
GLU
THR
ARG
PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	67.51Å 115.97Å 133.09Å 90.00° 99.85° 90.00°	Depositor
Resolution (Å)	43.71 – 2.22 43.71 – 2.22	Depositor EDS
% Data completeness (in resolution range)	99.9 (43.71-2.22) 99.9 (43.71-2.22)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.87 (at 2.22Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.197 , 0.259 0.197 , 0.259	Depositor DCC
R_{free} test set	4975 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	15.0	Xtriage
Anisotropy	0.132	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 37.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	14690	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 61.64 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2715e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: LOG, EDO

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	0.67	2/3437 (0.1%)	0.71	0/4646
1	B	1.04	6/3424 (0.2%)	0.77	2/4627 (0.0%)
1	C	0.67	2/3379 (0.1%)	0.72	1/4565 (0.0%)
1	D	0.80	6/3352 (0.2%)	0.77	5/4528 (0.1%)
All	All	0.81	16/13592 (0.1%)	0.74	8/18366 (0.0%)

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	925	LYS	CE-NZ	-38.90	0.51	1.49
1	B	764	LYS	CE-NZ	-24.34	0.88	1.49
1	D	827	LYS	CE-NZ	-23.70	0.89	1.49
1	B	756	LYS	CE-NZ	-14.79	1.12	1.49
1	D	917	LYS	CG-CD	-13.40	1.06	1.52
1	D	767	LYS	CD-CE	-6.76	1.34	1.51
1	B	876	TRP	CD2-CE2	5.87	1.48	1.41
1	C	816	TRP	CD2-CE2	5.62	1.48	1.41
1	D	821	TRP	CD2-CE2	5.44	1.47	1.41
1	B	821	TRP	CD2-CE2	5.38	1.47	1.41
1	C	876	TRP	CD2-CE2	5.19	1.47	1.41
1	B	872	TRP	CD2-CE2	5.16	1.47	1.41
1	A	816	TRP	CD2-CE2	5.15	1.47	1.41
1	D	816	TRP	CD2-CE2	5.12	1.47	1.41
1	A	876	TRP	CD2-CE2	5.05	1.47	1.41
1	D	876	TRP	CD2-CE2	5.01	1.47	1.41

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	925	LYS	CD-CE-NZ	17.56	152.08	111.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	827	LYS	CD-CE-NZ	14.53	145.11	111.70
1	B	764	LYS	CD-CE-NZ	13.97	143.84	111.70
1	D	917	LYS	CB-CG-CD	9.18	135.47	111.60
1	D	773	LYS	CD-CE-NZ	-7.88	93.57	111.70
1	C	671	ASP	CB-CG-OD1	5.75	123.48	118.30
1	D	763	ASP	CB-CG-OD1	5.06	122.86	118.30
1	D	763	ASP	CB-CG-OD2	-5.06	113.75	118.30

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	3361	0	3259	23	0
1	B	3346	0	3265	33	0
1	C	3302	0	3218	26	0
1	D	3276	0	3187	14	0
2	A	16	0	14	2	0
2	B	16	0	14	0	0
2	C	16	0	14	0	0
2	D	16	0	14	0	0
3	A	16	0	24	2	0
3	B	8	0	12	5	0
3	C	24	0	36	2	0
4	A	355	0	0	5	0
4	B	310	0	0	9	0
4	C	330	0	0	6	0
4	D	298	0	0	1	0
All	All	14690	0	13057	99	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 (99) 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:C:1003:GLU:HA	1:C:1006:GLU:HG2	1.36	1.07
1:C:943:THR:HG22	1:C:945:GLY:H	1.33	0.93
1:D:820:LYS:HE3	4:D:3138:HOH:O	1.72	0.89
1:B:787:LYS:HE3	4:B:3047:HOH:O	1.74	0.87
1:D:678:ASP:H	1:D:716:GLN:HE21	1.22	0.86
1:C:943:THR:HG21	4:C:3260:HOH:O	1.76	0.86
1:B:648:GLN:HG2	3:B:2043:EDO:H22	1.64	0.79
1:C:678:ASP:H	1:C:716:GLN:NE2	1.81	0.78
1:D:678:ASP:H	1:D:716:GLN:NE2	1.82	0.78
1:C:678:ASP:H	1:C:716:GLN:HE21	1.32	0.77
1:A:907:GLN:OE1	1:A:954[A]:ARG:NH2	2.20	0.74
1:B:907:GLN:OE1	1:B:954[A]:ARG:NH2	2.22	0.72
1:D:750:MET:HB2	1:D:755:ILE:HD12	1.73	0.70
1:D:992:LYS:C	1:D:993:ILE:HG12	2.10	0.70
1:B:797:LYS:HE3	4:B:3178:HOH:O	1.95	0.66
1:C:827:LYS:HD3	4:C:3135:HOH:O	1.97	0.64
1:B:645:THR:H	1:B:648:GLN:HE21	1.48	0.61
1:C:854:TYR:O	1:C:882:ALA:HB2	2.00	0.61
1:B:687[B]:LYS:HE3	1:B:689:TYR:CE2	2.37	0.60
1:A:670:ASN:HD22	1:A:738:ASN:HD22	1.48	0.60
1:A:805:GLU:OE2	2:A:2044:LOG:H2	2.03	0.59
1:C:918:LYS:HD2	4:C:3254:HOH:O	2.03	0.57
1:B:667:LEU:HA	1:B:736:ALA:HB3	1.84	0.57
1:A:871:TYR:OH	1:A:884:PRO:HD3	2.06	0.56
1:A:670:ASN:ND2	1:A:738:ASN:HD22	2.03	0.55
1:B:893:LYS:HD3	4:B:3211:HOH:O	2.07	0.55
1:A:654[A]:ASP:OD2	4:A:3028:HOH:O	2.18	0.53
1:D:986:LEU:HD13	1:D:1031:VAL:HG22	1.89	0.53
1:B:935:LYS:HE3	4:B:3249:HOH:O	2.09	0.52
1:B:671:ASP:OD1	1:B:742:HIS:HD2	1.92	0.52
1:C:913:GLY:O	1:C:918:LYS:CE	2.58	0.51
1:A:847:MET:HE3	1:A:895:LEU:HD23	1.93	0.51
1:C:1032:ALA:O	1:C:1035:LYS:HB3	2.11	0.51
1:C:913:GLY:O	1:C:918:LYS:HE2	2.10	0.51
1:B:670:ASN:HD22	1:B:738:ASN:HD22	1.56	0.51
1:B:645:THR:H	1:B:648:GLN:NE2	2.10	0.50
1:D:636:SER:HB3	1:D:949:SER:HA	1.94	0.50
1:B:687[B]:LYS:HE2	4:B:3055:HOH:O	2.12	0.49
1:B:992:LYS:HB2	1:B:1038:LEU:HD21	1.94	0.49
1:A:871:TYR:CZ	1:A:884:PRO:HD3	2.48	0.49
1:C:904[B]:ILE:N	1:C:904[B]:ILE:CD1	2.74	0.49
1:A:824:LEU:HD13	3:A:2045:EDO:H11	1.94	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:928:PHE:HB2	1:A:975:ASN:ND2	2.27	0.49
1:C:729:LYS:HE3	4:C:3308:HOH:O	2.12	0.49
1:B:771:ASP:OD1	3:B:2044:EDO:H21	2.13	0.48
1:A:854:TYR:O	1:A:882:ALA:HB2	2.13	0.48
1:A:671:ASP:OD1	1:A:742:HIS:HD2	1.95	0.48
1:B:986:LEU:HD13	1:B:1031:VAL:HG22	1.95	0.48
1:B:824:LEU:HD13	3:B:2044:EDO:H22	1.96	0.48
1:B:928:PHE:HB2	1:B:975:ASN:ND2	2.29	0.48
1:A:913:GLY:O	1:A:918:LYS:HE2	2.13	0.47
1:B:787:LYS:HD3	4:B:3124:HOH:O	2.13	0.47
1:C:943:THR:HG22	1:C:945:GLY:N	2.16	0.47
1:B:964:ILE:O	1:B:968:MET:HG3	2.15	0.47
1:B:670:ASN:ND2	1:B:738:ASN:HD22	2.12	0.47
1:B:873:SER:HB2	4:B:3198:HOH:O	2.15	0.46
4:A:3307:HOH:O	3:B:2043:EDO:C2	2.63	0.46
1:B:787:LYS:HD2	1:B:787:LYS:HA	1.75	0.46
1:A:657:SER:HA	1:A:731:ILE:HD11	1.98	0.45
1:B:923:THR:HB	1:B:971:PHE:HA	1.97	0.45
1:B:657:SER:HA	1:B:731:ILE:HD11	1.98	0.45
1:B:801:PHE:HB2	1:B:835:LEU:HD13	1.98	0.45
1:C:750:MET:HB3	1:C:755:ILE:HB	1.99	0.45
1:B:859:ASP:HB2	1:B:862:GLN:HE22	1.82	0.45
1:D:640:GLY:HA3	1:D:670:ASN:HB3	1.99	0.44
1:A:775:GLU:HG2	4:A:3236:HOH:O	2.17	0.44
1:C:705:TYR:O	3:C:2046:EDO:H11	2.17	0.44
1:C:928:PHE:HB2	1:C:975:ASN:ND2	2.33	0.44
1:C:942:PRO:HD2	4:C:3232:HOH:O	2.18	0.44
1:C:764:LYS:HD2	4:C:3131:HOH:O	2.17	0.43
1:C:628:LYS:HG2	1:C:978:TYR:CE1	2.53	0.43
4:A:3307:HOH:O	3:B:2043:EDO:H22	2.17	0.43
2:A:2044:LOG:O7	2:A:2044:LOG:C1	2.66	0.43
1:A:923:THR:HB	1:A:971:PHE:HA	2.01	0.43
1:A:986:LEU:HD11	1:A:1015:LEU:HD13	2.01	0.43
1:C:904[B]:ILE:N	1:C:904[B]:ILE:HD12	2.34	0.43
1:C:961:GLU:HG2	1:C:965:PHE:CE2	2.53	0.43
1:A:764:LYS:HE3	1:A:809:ASP:O	2.19	0.43
1:D:928:PHE:HB2	1:D:975:ASN:ND2	2.33	0.42
1:B:706:TYR:CD2	1:B:955:PRO:HB2	2.54	0.42
1:A:724[B]:GLU:HG2	4:A:3218:HOH:O	2.19	0.42
1:B:997:LEU:HG	1:B:1005:LEU:HD11	2.00	0.42
1:C:1001:SER:OG	1:C:1002:LYS:N	2.52	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:822:TYR:CD1	3:C:2045:EDO:H11	2.54	0.42
1:A:806:TYR:CG	1:A:819:LEU:HD11	2.55	0.42
1:D:806:TYR:O	1:D:815:GLY:HA3	2.19	0.41
1:B:808:ASN:HA	1:B:808:ASN:HD22	1.69	0.41
1:B:845:GLN:NE2	4:B:3172:HOH:O	2.42	0.41
1:D:658:GLU:HG3	1:D:982:ASN:ND2	2.35	0.41
1:D:808:ASN:HD22	1:D:808:ASN:HA	1.69	0.41
3:A:2045:EDO:H22	3:A:2047:EDO:H11	2.01	0.41
1:A:1038:LEU:O	1:A:1041:LEU:HD12	2.21	0.41
1:C:845:GLN:NE2	1:C:846:PRO:HD2	2.35	0.41
1:B:787:LYS:CE	4:B:3047:HOH:O	2.50	0.40
1:D:744:ASP:OD1	1:D:744:ASP:N	2.53	0.40
1:A:907:GLN:HB3	1:A:915:PHE:HA	2.03	0.40
1:B:1018:ASN:HB3	1:D:821:TRP:CH2	2.56	0.40
1:C:1020:ASN:OD1	1:C:1022:ASN:HB2	2.21	0.40
1:A:628:LYS:HG2	1:A:978:TYR:CE1	2.56	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	423/443 (96%)	407 (96%)	15 (4%)	1 (0%)	47 54
1	B	420/443 (95%)	400 (95%)	18 (4%)	2 (0%)	29 30
1	C	413/443 (93%)	393 (95%)	19 (5%)	1 (0%)	47 54
1	D	409/443 (92%)	389 (95%)	20 (5%)	0	100 100
All	All	1665/1772 (94%)	1589 (95%)	72 (4%)	4 (0%)	47 54

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	671	ASP
1	B	896	ASN
1	B	671	ASP
1	C	671	ASP

5.3.2 Protein sidechains (i)

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	344/364 (94%)	341 (99%)	3 (1%)	78	87
1	B	345/364 (95%)	339 (98%)	6 (2%)	60	73
1	C	339/364 (93%)	334 (98%)	5 (2%)	65	76
1	D	336/364 (92%)	330 (98%)	6 (2%)	59	71
All	All	1364/1456 (94%)	1344 (98%)	20 (2%)	67	76

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	744	ASP
1	A	851	ASP
1	A	1003	GLU
1	B	683	THR
1	B	688	THR
1	B	724	GLU
1	B	851	ASP
1	B	977[A]	ASP
1	B	977[B]	ASP
1	C	624	SER
1	C	744	ASP
1	C	764	LYS
1	C	851	ASP
1	C	893	LYS
1	D	633	LYS
1	D	654	ASP
1	D	744	ASP
1	D	860	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	977	ASP
1	D	993	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (33) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	632	ASN
1	A	670	ASN
1	A	742	HIS
1	A	761	HIS
1	A	808	ASN
1	A	929	ASN
1	A	975	ASN
1	A	1033	ASN
1	A	1039	GLN
1	B	632	ASN
1	B	648	GLN
1	B	670	ASN
1	B	742	HIS
1	B	808	ASN
1	B	975	ASN
1	B	1033	ASN
1	C	670	ASN
1	C	716	GLN
1	C	742	HIS
1	C	761	HIS
1	C	814	GLN
1	C	929	ASN
1	C	975	ASN
1	C	1018	ASN
1	C	1033	ASN
1	D	670	ASN
1	D	716	GLN
1	D	742	HIS
1	D	761	HIS
1	D	808	ASN
1	D	922	ASN
1	D	929	ASN
1	D	975	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 monosaccharides in this entry.

5.6 Ligand geometry

16 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	LOG	D	2040	-	14,16,16	1.63	1 (7%)	14,22,22	1.22	1 (7%)
3	EDO	A	2048	-	3,3,3	0.47	0	2,2,2	0.35	0
3	EDO	C	2047	-	3,3,3	0.56	0	2,2,2	0.17	0
3	EDO	C	2043	-	3,3,3	0.45	0	2,2,2	0.36	0
2	LOG	C	2041	-	14,16,16	1.56	3 (21%)	14,22,22	1.38	3 (21%)
3	EDO	C	2044	-	3,3,3	0.39	0	2,2,2	0.49	0
3	EDO	C	2045	-	3,3,3	0.39	0	2,2,2	0.50	0
3	EDO	B	2043	-	3,3,3	0.40	0	2,2,2	0.38	0
3	EDO	A	2045	-	3,3,3	0.46	0	2,2,2	0.33	0
3	EDO	B	2044	-	3,3,3	0.53	0	2,2,2	0.21	0
2	LOG	B	2042	-	14,16,16	1.29	2 (14%)	14,22,22	1.31	3 (21%)
3	EDO	C	2046	-	3,3,3	0.36	0	2,2,2	0.69	0
3	EDO	A	2046	-	3,3,3	0.51	0	2,2,2	0.74	0
3	EDO	C	2042	-	3,3,3	0.59	0	2,2,2	0.18	0
3	EDO	A	2047	-	3,3,3	0.39	0	2,2,2	0.54	0
2	LOG	A	2044	-	14,16,16	1.83	3 (21%)	14,22,22	1.29	2 (14%)

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	LOG	D	2040	-	-	1/6/28/28	0/1/1/1
3	EDO	A	2048	-	-	0/1/1/1	-
3	EDO	C	2047	-	-	1/1/1/1	-
3	EDO	C	2043	-	-	1/1/1/1	-
2	LOG	C	2041	-	-	1/6/28/28	0/1/1/1
3	EDO	C	2044	-	-	0/1/1/1	-
3	EDO	C	2045	-	-	1/1/1/1	-
3	EDO	B	2043	-	-	1/1/1/1	-
3	EDO	A	2045	-	-	1/1/1/1	-
3	EDO	B	2044	-	-	1/1/1/1	-
2	LOG	B	2042	-	-	2/6/28/28	0/1/1/1
3	EDO	C	2046	-	-	0/1/1/1	-
3	EDO	A	2046	-	-	0/1/1/1	-
3	EDO	C	2042	-	-	1/1/1/1	-
3	EDO	A	2047	-	-	1/1/1/1	-
2	LOG	A	2044	-	-	1/6/28/28	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2044	LOG	O5-C5	-5.17	1.39	1.46
2	D	2040	LOG	O5-C5	-4.98	1.39	1.46
2	C	2041	LOG	O5-C5	-3.70	1.41	1.46
2	C	2041	LOG	O5-C1	2.82	1.41	1.36
2	A	2044	LOG	O3-C3	2.80	1.49	1.43
2	B	2042	LOG	O5-C5	-2.35	1.43	1.46
2	B	2042	LOG	C8-C7	2.16	1.55	1.50
2	C	2041	LOG	C8-C7	2.15	1.55	1.50
2	A	2044	LOG	O1-N1	-2.05	1.34	1.40

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2040	LOG	O5-C1-C2	2.91	126.33	118.10
2	A	2044	LOG	O5-C1-C2	2.83	126.10	118.10
2	C	2041	LOG	O5-C1-C2	2.64	125.56	118.10
2	B	2042	LOG	O5-C1-C2	2.57	125.34	118.10
2	C	2041	LOG	C3-C4-C5	-2.51	105.76	110.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2042	LOG	C3-C2-N2	-2.45	108.45	112.28
2	C	2041	LOG	C3-C2-N2	-2.21	108.82	112.28
2	B	2042	LOG	C3-C4-C5	-2.20	106.32	110.24
2	A	2044	LOG	C3-C2-N2	-2.09	109.01	112.28

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	2045	EDO	O1-C1-C2-O2
3	C	2047	EDO	O1-C1-C2-O2
3	C	2043	EDO	O1-C1-C2-O2
3	B	2043	EDO	O1-C1-C2-O2
3	B	2044	EDO	O1-C1-C2-O2
3	C	2042	EDO	O1-C1-C2-O2
3	A	2047	EDO	O1-C1-C2-O2
3	A	2045	EDO	O1-C1-C2-O2
2	D	2040	LOG	C1-C2-N2-C7
2	C	2041	LOG	C1-C2-N2-C7
2	B	2042	LOG	C1-C2-N2-C7
2	A	2044	LOG	C1-C2-N2-C7
2	B	2042	LOG	O5-C5-C6-O6

There are no ring outliers.

7 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	2045	EDO	1	0
3	B	2043	EDO	3	0
3	A	2045	EDO	2	0
3	B	2044	EDO	2	0
3	C	2046	EDO	1	0
3	A	2047	EDO	1	0
2	A	2044	LOG	2	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^{th} 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^2)	Q<0.9
1	A	421/443 (95%)	-0.33	16 (3%)	40	38	6, 13, 46, 94	12 (2%)
1	B	419/443 (94%)	-0.34	12 (2%)	51	49	9, 16, 44, 85	11 (2%)
1	C	414/443 (93%)	-0.38	10 (2%)	59	57	7, 13, 46, 84	15 (3%)
1	D	411/443 (92%)	-0.38	11 (2%)	54	52	9, 16, 46, 77	13 (3%)
All	All	1665/1772 (93%)	-0.36	49 (2%)	51	49	6, 14, 46, 94	51 (3%)

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	997	LEU	5.7
1	B	997	LEU	5.6
1	C	1004	SER	5.4
1	A	1043	PRO	4.9
1	D	1000	TYR	4.6
1	A	994	PRO	4.3
1	B	995	THR	4.0
1	B	996	ASN	3.7
1	A	996	ASN	3.6
1	C	1005	LEU	3.6
1	A	995	THR	3.5
1	A	1042	LYS	3.4
1	A	999	GLY	3.3
1	C	1000	TYR	3.2
1	A	993	ILE	3.1
1	A	623	GLY	3.1
1	A	1000	TYR	3.1
1	D	1037	ALA	3.1
1	B	1000	TYR	3.0
1	D	1004	SER	3.0
1	B	1035	LYS	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	1008	LEU	2.9
1	B	998	GLU	2.8
1	A	1041	LEU	2.8
1	C	993	ILE	2.8
1	D	991	ALA	2.8
1	A	1039	GLN	2.7
1	C	999	GLY	2.7
1	D	1005	LEU	2.6
1	D	1008	LEU	2.6
1	D	1032	ALA	2.6
1	D	999	GLY	2.5
1	D	1001	SER	2.5
1	C	1013	THR	2.5
1	B	1001	SER	2.4
1	B	999	GLY	2.4
1	B	1039	GLN	2.4
1	C	623	GLY	2.4
1	B	1041	LEU	2.3
1	C	1010	ALA	2.3
1	B	1004	SER	2.3
1	A	1009	ASP	2.3
1	A	991	ALA	2.3
1	A	998	GLU	2.2
1	D	1007	ALA	2.1
1	A	1001	SER	2.1
1	B	623	GLY	2.1
1	C	1038	LEU	2.0
1	D	993	ILE	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides 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^{th} 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^2)	Q<0.9
3	EDO	A	2048	4/4	0.62	0.35	81,84,87,91	0
3	EDO	C	2047	4/4	0.82	0.24	38,42,42,44	0
3	EDO	C	2044	4/4	0.88	0.22	31,32,33,34	0
3	EDO	C	2046	4/4	0.89	0.23	33,34,34,38	0
3	EDO	C	2042	4/4	0.90	0.13	23,23,25,26	0
3	EDO	C	2045	4/4	0.91	0.15	36,38,39,42	0
3	EDO	A	2047	4/4	0.92	0.14	31,33,35,35	0
3	EDO	A	2046	4/4	0.94	0.12	27,28,28,28	0
3	EDO	B	2043	4/4	0.95	0.19	23,24,25,25	0
3	EDO	C	2043	4/4	0.96	0.11	27,28,28,29	0
2	LOG	D	2040	16/16	0.96	0.10	13,14,17,19	1
3	EDO	A	2045	4/4	0.96	0.13	20,22,22,23	0
3	EDO	B	2044	4/4	0.96	0.14	23,23,25,26	0
2	LOG	C	2041	16/16	0.97	0.12	10,10,12,14	1
2	LOG	B	2042	16/16	0.97	0.10	12,13,14,16	1
2	LOG	A	2044	16/16	0.97	0.10	8,8,9,10	1

6.5 Other polymers

There are no such residues in this entry.