



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 08:00 PM GMT

PDB ID : 3IMM
Title : Crystal structure of Putative glycosyl hydrolase (YP_001301887.1) from Parabacteroides distasonis ATCC 8503 at 2.00 Å resolution
Authors : Joint Center for Structural Genomics (JCSG)
Deposited on : 2009-08-10
Resolution : 2.00 Å (reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

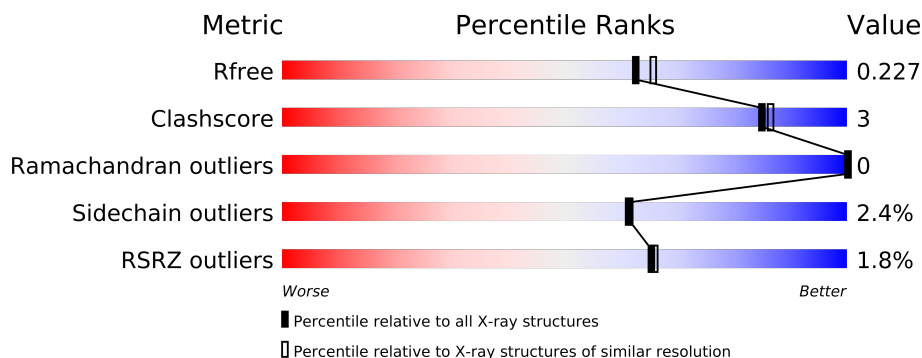
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.00 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}	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	201	
1	B	201	
1	C	201	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5199 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Putative secreted glycosylhydrolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	197	Total	C	N	O	S	Se	0	5	0
			1592	1014	267	302	4	5			
1	B	197	Total	C	N	O	S	Se	0	6	0
			1613	1030	271	303	4	5			
1	C	197	Total	C	N	O	S	Se	0	4	0
			1587	1014	265	299	4	5			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	leader sequence	UNP A6L9A1
B	0	GLY	-	leader sequence	UNP A6L9A1
C	0	GLY	-	leader sequence	UNP A6L9A1

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Na	0	0
			1	1		
2	A	1	Total	Na	0	0
			1	1		
2	C	1	Total	Na	0	0
			1	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C₄H₁₂NO₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	C	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	153	Total	O	0	0
			153	153		
5	B	152	Total	O	0	0
			152	152		
5	C	90	Total	O	0	0
			90	90		

i

- Molecule 1: Putative secreted glycosylhydrolase

GLY
ALA
ASP
ASN
N26
L74
C110
W113
Y119
V132
Q133
D134
M156
K163
M167
W172
F192
L195
L211
R215
L222

Chain B: 

[illegible]

Chain C: 

[illegible]

4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α , β , γ	45.40Å 45.40Å 241.07Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.11 – 2.00 28.10 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (28.11-2.00) 99.8 (28.10-2.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.33 (at 1.99Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.163 , 0.222 0.171 , 0.227	Depositor DCC
R_{free} test set	1876 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	26.5	Xtriage
Anisotropy	0.634	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 45.6	EDS
Estimated twinning fraction	0.007 for -h,-k,l 0.067 for h,-h-k,-l 0.039 for -k,-h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 37454 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5199	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, TRS, CL

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	0.75	1/1631 (0.1%)	0.73	1/2203 (0.0%)
1	B	0.72	0/1654	0.78	1/2233 (0.0%)
1	C	0.68	0/1626	0.74	0/2197
All	All	0.72	1/4911 (0.0%)	0.75	2/6633 (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 (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	110	CYS	CB-SG	-5.02	1.73	1.81

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	72	LEU	CA-CB-CG	5.58	128.13	115.30
1	A	215	ARG	CG-CD-NE	-5.19	100.90	111.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	205[A]	GLY	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1592	0	1521	7	0
1	B	1613	0	1550	13	0
1	C	1587	0	1515	7	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	B	1	0	0	0	0
4	C	8	0	12	0	0
5	A	153	0	0	3	0
5	B	152	0	0	0	0
5	C	90	0	0	2	0
All	All	5199	0	4598	26	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (26) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:106:ALA:HB1	1:B:113[A]:TRP:HD1	1.39	0.86
1:A:156[A]:MSE:HE2	1:A:163:LYS:HD2	1.71	0.72
1:B:106:ALA:CB	1:B:113[A]:TRP:HD1	2.12	0.60
1:C:93[A]:THR:HG22	5:C:283:HOH:O	2.02	0.59
1:B:36:LEU:HD21	1:B:53:LEU:HD13	1.90	0.53
1:A:132:VAL:HG13	5:A:327:HOH:O	2.08	0.53
1:C:180:ASP:C	1:C:180:ASP:OD1	2.49	0.52
1:A:167:MSE:HE2	1:A:172:TRP:CZ2	2.46	0.50
1:B:132:VAL:HG22	1:B:164:VAL:HA	1.93	0.50
1:A:192:PHE:CD1	1:A:195:LEU:HD22	2.48	0.49
1:B:146:MSE:HE3	1:B:159:LEU:HD13	1.94	0.49
1:B:57:LYS:HE3	1:B:59:GLU:OE2	2.13	0.48
1:A:211:LEU:HD22	5:A:342:HOH:O	2.14	0.48
1:B:102:GLU:OE2	1:B:204[A]:GLN:NE2	2.48	0.47
1:B:106:ALA:HB1	1:B:113[A]:TRP:CD1	2.31	0.46
1:C:151:ALA:HB2	1:C:222:LEU:CD1	2.45	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:77:LYS:HE3	1:C:143:TRP:CE2	2.52	0.45
1:A:156[B]:MSE:HE1	5:A:339:HOH:O	2.17	0.45
1:B:171:LYS:NZ	1:C:58:ASP:OD2	2.42	0.44
1:B:82:THR:HG21	1:B:212:ILE:HD12	2.01	0.43
1:C:31:LEU:O	5:C:248:HOH:O	2.21	0.42
1:A:167:MSE:HE2	1:A:172:TRP:CE2	2.55	0.42
1:C:73:ASP:CG	1:C:145[A]:HIS:HE2	2.24	0.41
1:B:73:ASP:OD1	1:B:145:HIS:CE1	2.73	0.41
1:B:125:ILE:HG12	1:B:165:THR:HG21	2.02	0.41
1:B:106:ALA:CB	1:B:113[A]:TRP:CD1	2.99	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	200/201 (100%)	193 (96%)	7 (4%)	0	100	100
1	B	201/201 (100%)	195 (97%)	6 (3%)	0	100	100
1	C	199/201 (99%)	195 (98%)	4 (2%)	0	100	100
All	All	600/603 (100%)	583 (97%)	17 (3%)	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	170/169 (101%)	167 (98%)	3 (2%)	71	73

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	173/169 (102%)	165 (95%)	8 (5%)	37	30
1	C	169/169 (100%)	167 (99%)	2 (1%)	82	84
All	All	512/507 (101%)	499 (98%)	13 (2%)	61	59

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	74	LEU
1	A	119	TYR
1	A	133	GLN
1	B	35	ASN
1	B	74	LEU
1	B	111	GLU
1	B	113[A]	TRP
1	B	113[B]	TRP
1	B	119	TYR
1	B	147	ARG
1	B	209	ASP
1	C	115	ASN
1	C	119	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	133	GLN
1	B	35	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	TRS	C	5	-	7,7,7	1.18	1 (14%)	9,9,9	1.16	1 (11%)

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	TRS	C	5	-	-	0/9/9/9	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	5	TRS	C-N	-2.95	1.46	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	5	TRS	C3-C-C1	-3.08	105.64	110.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 95th 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	197/201 (98%)	-0.46	2 (1%) 79 80	25, 30, 38, 55	0
1	B	197/201 (98%)	-0.35	4 (2%) 62 62	24, 30, 41, 57	0
1	C	197/201 (98%)	-0.15	5 (2%) 54 54	25, 31, 39, 55	0
All	All	591/603 (98%)	-0.32	11 (1%) 65 64	24, 30, 40, 57	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	114	GLY	7.0
1	B	115	ASN	5.6
1	B	113[A]	TRP	5.0
1	C	114	GLY	3.7
1	C	113	TRP	3.4
1	C	112	LYS	3.2
1	A	113	TRP	3.2
1	C	115	ASN	2.8
1	C	26	ASN	2.6
1	B	111	GLU	2.3
1	A	134	ASP	2.2

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 ⓘ

There are no carbohydrates in this entry.

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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	TRS	C	5	8/8	0.12	-0.32	28,37,39,42	0
2	NA	A	4	1/1	0.10	-0.76	39,39,39,39	0
2	NA	C	3	1/1	0.08	-2.06	37,37,37,37	0
2	NA	B	2	1/1	0.07	-2.60	30,30,30,30	0
3	CL	B	1	1/1	0.04	-3.94	28,28,28,28	0

6.5 Other polymers

There are no such residues in this entry.