wwPDB X-ray Structure Validation Summary Report (i)

Jan 24, 2023 – 01:41 AM EST

PDB ID	:	3H3L
Title	:	Crystal structure of PUTATIVE SUGAR HYDROLASE (YP_001304206.1) from Parabacteroides distasonis ATCC 8503 at 1.59 A resolution
Authors	:	Joint Center for Structural Genomics (JCSG)
Deposited on	:	2009-04-16
Resolution	:	1.59 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

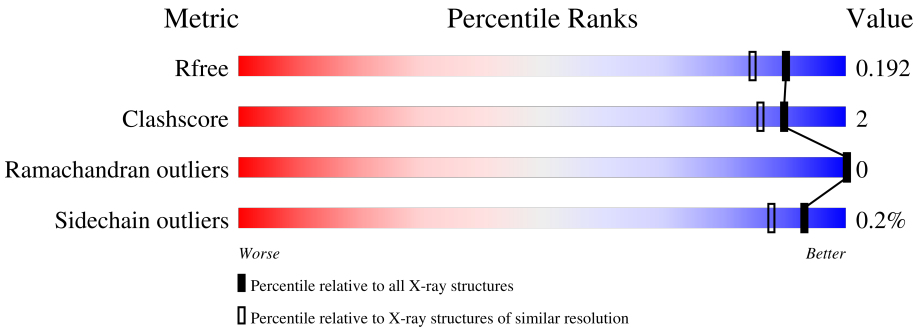
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 1.59 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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 of 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%

Mol	Chain	Length	Quality of chain
1	A	241	93%
1	B	241	92% 5%
1	C	241	88% 5% 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6134 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.

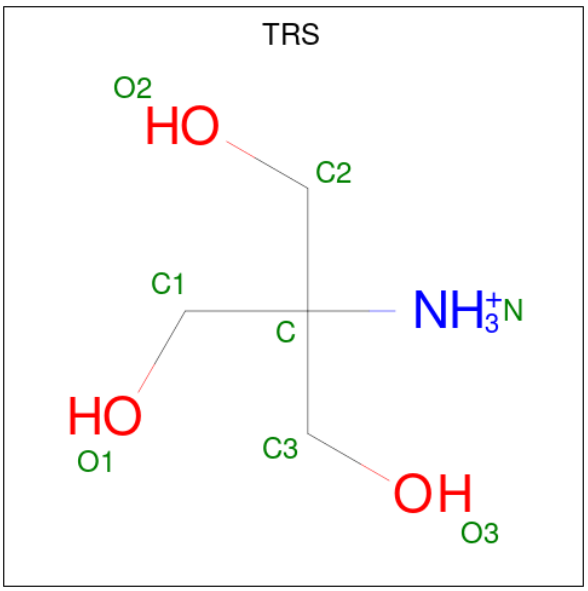
- Molecule 1 is a protein called PUTATIVE SUGAR HYDROLASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	232	Total	C	N	O	S	Se	0	11	0
			1899	1220	314	350	6	9			
1	B	229	Total	C	N	O	S	Se	0	13	0
			1885	1216	309	346	6	8			
1	C	223	Total	C	N	O	S	Se	0	13	0
			1838	1184	297	341	6	10			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP A6LFX0
B	0	GLY	-	expression tag	UNP A6LFX0
C	0	GLY	-	expression tag	UNP A6LFX0

- Molecule 2 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	178	Total	O	0	0
			178	178		
3	B	169	Total	O	0	0
			169	169		
3	C	141	Total	O	0	0
			141	141		

3 Residue-property plots (i)

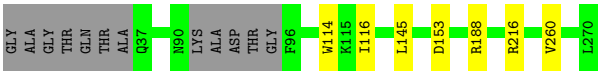
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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: PUTATIVE SUGAR HYDROLASE

Chain A:

• Molecule 1: PUTATIVE SUGAR HYDROLASE

Chain B:



• Molecule 1: PUTATIVE SUGAR HYDROLASE

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	117.26Å 117.26Å 118.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.31 – 1.59 29.31 – 1.59	Depositor EDS
% Data completeness (in resolution range)	96.6 (29.31-1.59) 96.6 (29.31-1.59)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 1.59Å)	Xtriage
Refinement program	REFMAC 5.2.0019, PHENIX	Depositor
R, R_{free}	0.170 , 0.185 0.181 , 0.192	Depositor DCC
R_{free} test set	5379 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	21.8	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.019 for -h,l,k 0.013 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6134	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% of the height of the origin peak. No significant pseudotranslation is detected.*

[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: TRS

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	0.58	0/1974	0.78	2/2659 (0.1%)
1	B	0.58	0/1968	0.74	3/2650 (0.1%)
1	C	0.60	0/1921	0.78	0/2589
All	All	0.58	0/5863	0.77	5/7898 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	188[A]	ARG	NE-CZ-NH1	6.50	123.55	120.30
1	A	188[B]	ARG	NE-CZ-NH1	6.50	123.55	120.30
1	B	216	ARG	NE-CZ-NH1	5.13	122.86	120.30
1	B	188[A]	ARG	NE-CZ-NH1	5.08	122.84	120.30
1	B	188[B]	ARG	NE-CZ-NH1	5.08	122.84	120.30

There are no chirality outliers.

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	1899	0	1819	5	0
1	B	1885	0	1817	2	0
1	C	1838	0	1755	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	8	0	12	0	0
2	B	8	0	12	0	0
2	C	8	0	12	0	0
3	A	178	0	0	2	0
3	B	169	0	0	0	0
3	C	141	0	0	1	0
All	All	6134	0	5427	20	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 2.

The worst 5 of 20 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:121[A]:ASN:HB2	1:C:151[A]:MSE:HE3	1.54	0.89
1:C:195[A]:ARG:HG2	1:C:209[A]:GLN:OE1	1.74	0.87
1:C:122:SER:O	3:C:420:HOH:O	2.13	0.67
1:C:146:LEU:HD11	1:C:151[A]:MSE:HG2	1.84	0.60
1:C:146:LEU:CD1	1:C:151[A]:MSE:HG2	2.36	0.55

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	240/241 (100%)	229 (95%)	11 (5%)	0	100	100
1	B	238/241 (99%)	231 (97%)	7 (3%)	0	100	100
1	C	232/241 (96%)	226 (97%)	6 (3%)	0	100	100
All	All	710/723 (98%)	686 (97%)	24 (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	200/197 (102%)	200 (100%)	0	100 100
1	B	200/197 (102%)	199 (100%)	1 (0%)	88 80
1	C	196/197 (100%)	196 (100%)	0	100 100
All	All	596/591 (101%)	595 (100%)	1 (0%)	93 88

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	153	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

3 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	TRS	C	1	-	7,7,7	0.28	0	9,9,9	1.10	1 (11%)
2	TRS	B	3	-	7,7,7	0.34	0	9,9,9	0.82	0
2	TRS	A	2	-	7,7,7	0.26	0	9,9,9	0.86	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	TRS	C	1	-	-	2/9/9/9	-
2	TRS	B	3	-	-	0/9/9/9	-
2	TRS	A	2	-	-	0/9/9/9	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1	TRS	O3-C3-C	2.11	117.70	111.00

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	TRS	N-C-C1-O1
2	C	1	TRS	C3-C-C1-O1

There are no ring outliers.

No monomer is involved in short contacts.

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.