wwPDB X-ray Structure Validation Summary Report (i)

Feb 12, 2024 – 04:51 AM EST

PDB ID	:	3GTN
Title	:	Crystal Structure of XynC from Bacillus subtilis 168
Authors	:	St John, F.J.; Hurlbert, J.C.; Pozharski, E.
Deposited on	:	2009-03-27
Resolution	:	2.68 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:

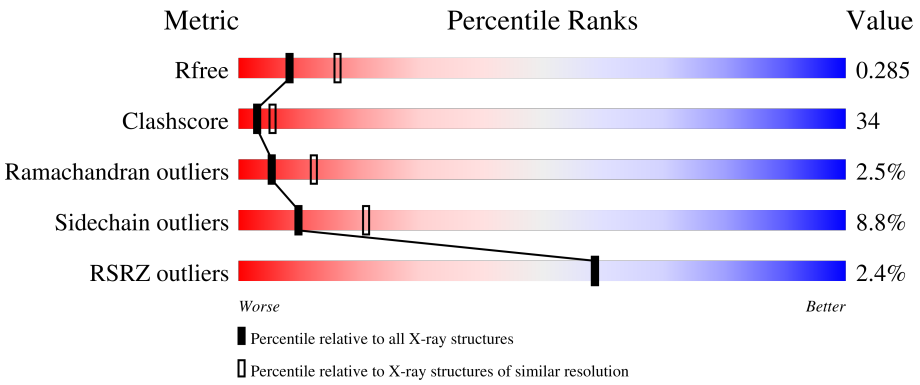
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.68 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	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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% 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	401	% 50% 43% 5% ••
1	B	401	4% 40% 50% 8% •

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 6306 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 Glucuronoxylanase xynC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	395	3158	1997	564	588	9	0	0	0
1	B	394	3148	1991	561	587	9	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

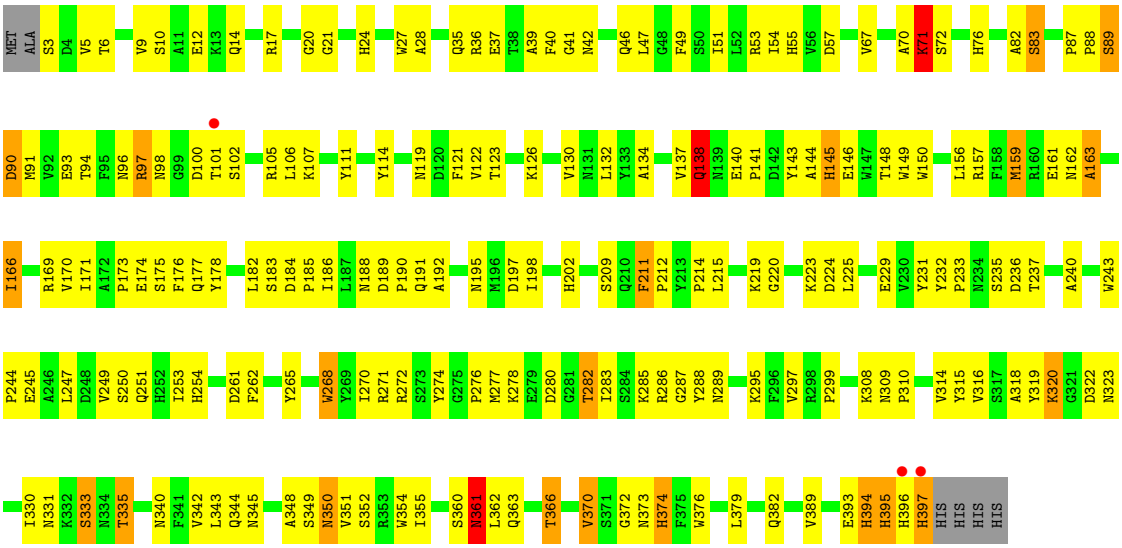
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q45070
A	392	LEU	-	expression tag	UNP Q45070
A	393	GLU	-	expression tag	UNP Q45070
A	394	HIS	-	expression tag	UNP Q45070
A	395	HIS	-	expression tag	UNP Q45070
A	396	HIS	-	expression tag	UNP Q45070
A	397	HIS	-	expression tag	UNP Q45070
A	398	HIS	-	expression tag	UNP Q45070
A	399	HIS	-	expression tag	UNP Q45070
A	400	HIS	-	expression tag	UNP Q45070
A	401	HIS	-	expression tag	UNP Q45070
B	1	MET	-	initiating methionine	UNP Q45070
B	392	LEU	-	expression tag	UNP Q45070
B	393	GLU	-	expression tag	UNP Q45070
B	394	HIS	-	expression tag	UNP Q45070
B	395	HIS	-	expression tag	UNP Q45070
B	396	HIS	-	expression tag	UNP Q45070
B	397	HIS	-	expression tag	UNP Q45070
B	398	HIS	-	expression tag	UNP Q45070
B	399	HIS	-	expression tag	UNP Q45070
B	400	HIS	-	expression tag	UNP Q45070
B	401	HIS	-	expression tag	UNP Q45070

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.

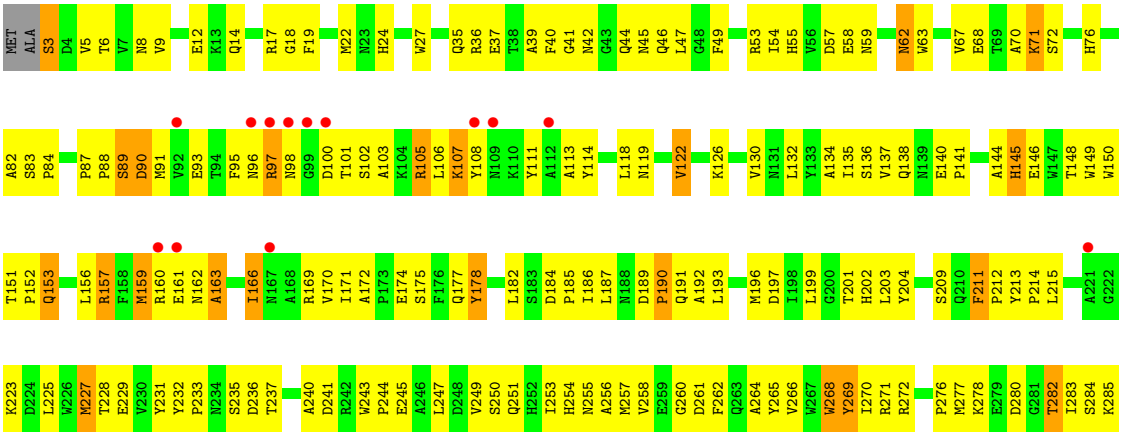
- Molecule 1: Glucuronoxylanase xynC

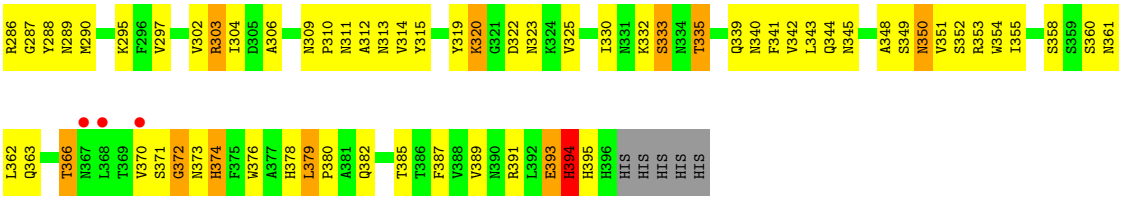
Chain A:



- Molecule 1: Glucuronoxylanase xynC

Chain B:




R286
G287
Y288
N289
M290
K295
F296
V297
V302
R303
I304
D305
A306
N309
P310
N311
A312
N313
V314
Y315
Y319
K320
G321
D322
N323
K324
V325
I330
N331
K332
S333
N334
T335
Q339
N340
F341
V342
L343
Q344
N345
A348
S349
N350
V351
S352
R353
W354
I355
S358
S359
S360
N361
L362
Q363
T366
N367
L368
T369
V370
S371
G372
N373
H374
F375
W376
A377
H378
L379
P380
A381
Q382
T385
T386
F387
V388
V389
N390
R391
L392
E393
H394
H395
H396
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	61.34Å 82.11Å 96.65Å 90.00° 104.70° 90.00°	Depositor
Resolution (Å)	33.33 – 2.68 33.33 – 2.68	Depositor EDS
% Data completeness (in resolution range)	95.7 (33.33-2.68) 95.7 (33.33-2.68)	Depositor EDS
R_{merge}	0.24	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.59 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.236 , 0.287 0.238 , 0.285	Depositor DCC
R_{free} test set	1281 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	30.6	Xtriage
Anisotropy	0.226	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 34.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6306	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.97% 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 (i)

5.1 Standard geometry (i)

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	0.78	6/3249 (0.2%)	0.82	1/4425 (0.0%)
1	B	0.80	4/3238 (0.1%)	0.87	5/4410 (0.1%)
All	All	0.79	10/6487 (0.2%)	0.85	6/8835 (0.1%)

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	211	PHE	CE2-CZ	-6.60	1.24	1.37
1	B	153	GLN	CD-NE2	-6.00	1.17	1.32
1	A	211	PHE	CG-CD1	-5.89	1.29	1.38
1	A	211	PHE	CG-CD2	-5.44	1.30	1.38
1	B	211	PHE	CG-CD2	-5.24	1.30	1.38

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	157	ARG	NE-CZ-NH1	-10.16	115.22	120.30
1	B	157	ARG	NE-CZ-NH2	9.22	124.91	120.30
1	A	157	ARG	NE-CZ-NH2	-6.18	117.21	120.30
1	B	17	ARG	NE-CZ-NH2	-5.97	117.32	120.30
1	B	303	ARG	NE-CZ-NH2	-5.83	117.38	120.30

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	3158	0	3003	181	1
1	B	3148	0	2996	238	1
All	All	6306	0	5999	419	2

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 34.

The worst 5 of 419 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:A:138:GLN:HG2	1:A:150:TRP:NE1	1.63	1.12
1:B:35:GLN:HG2	1:B:277:MET:HE3	1.30	1.09
1:B:134:ALA:HB1	1:B:171:ILE:CD1	1.83	1.08
1:B:394:HIS:O	1:B:394:HIS:ND1	1.89	1.05
1:B:288:TYR:CE2	1:B:361:ASN:ND2	2.23	1.05

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:37:GLU:OE2	1:B:209:SER:OG[2_546]	1.98	0.22
1:A:37:GLU:OE2	1:A:209:SER:OG[2_555]	2.18	0.02

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	393/401 (98%)	353 (90%)	31 (8%)	9 (2%)	6 14
1	B	392/401 (98%)	337 (86%)	44 (11%)	11 (3%)	5 10
All	All	785/802 (98%)	690 (88%)	75 (10%)	20 (2%)	5 12

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	268	TRP
1	B	166	ILE
1	B	268	TRP
1	B	394	HIS
1	A	71	LYS

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	336/341 (98%)	306 (91%)	30 (9%)	9 20
1	B	335/341 (98%)	306 (91%)	29 (9%)	10 21
All	All	671/682 (98%)	612 (91%)	59 (9%)	10 21

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	395	HIS
1	B	374	HIS
1	B	101	THR
1	B	366	THR
1	B	335	THR

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

Mol	Chain	Res	Type
1	A	340	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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, 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	395/401 (98%)	0.01	3 (0%)	86	86	13, 25, 41, 66	0
1	B	394/401 (98%)	0.17	16 (4%)	37	35	13, 25, 40, 50	0
All	All	789/802 (98%)	0.09	19 (2%)	59	59	13, 25, 41, 66	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	396	HIS	5.1
1	B	99	GLY	3.5
1	B	92	VAL	3.5
1	B	108	TYR	3.4
1	B	98	ASN	3.3

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.