wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 07:39 PM EDT

PDB ID	:	4BZY
Title	:	Crystal structure of human glycogen branching enzyme (GBE1)
Authors	:	Froese, D.S.; Krojer, T.; Goubin, S.; Strain-Damerell, C.; Mahajan, P.; von Delft, F.; Burgess-Brown, N.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.; Yue, W.W.
Deposited on	:	2013-07-30
Resolution	:	2.75 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:

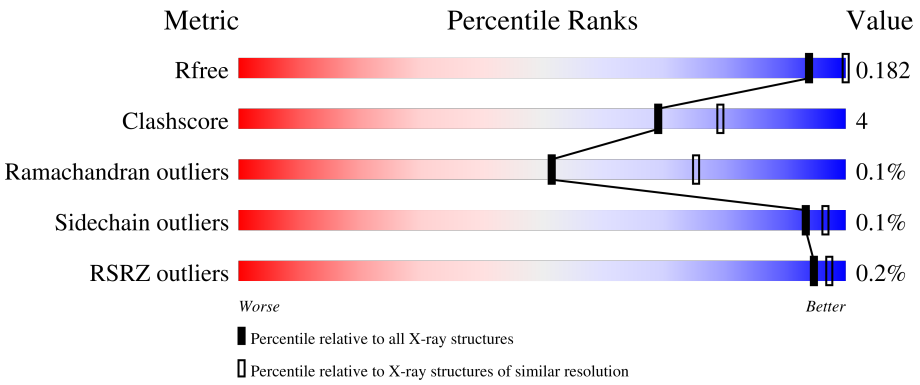
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	702	82% 9% 9%
1	B	702	84% 7% 8%
1	C	702	% 81% 9% 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 16084 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 1,4-ALPHA-GLUCAN-BRANCHING ENZYME.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	637	5142	3308	871	940	23	0	0	0
1	B	644	5234	3371	889	951	23	0	2	0
1	C	637	5098	3281	858	936	23	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	265	SER	THR	variant	UNP Q04446
A	334	VAL	ILE	variant	UNP Q04446
B	265	SER	THR	variant	UNP Q04446
B	334	VAL	ILE	variant	UNP Q04446
C	265	SER	THR	variant	UNP Q04446
C	334	VAL	ILE	variant	UNP Q04446

- Molecule 2 is water.

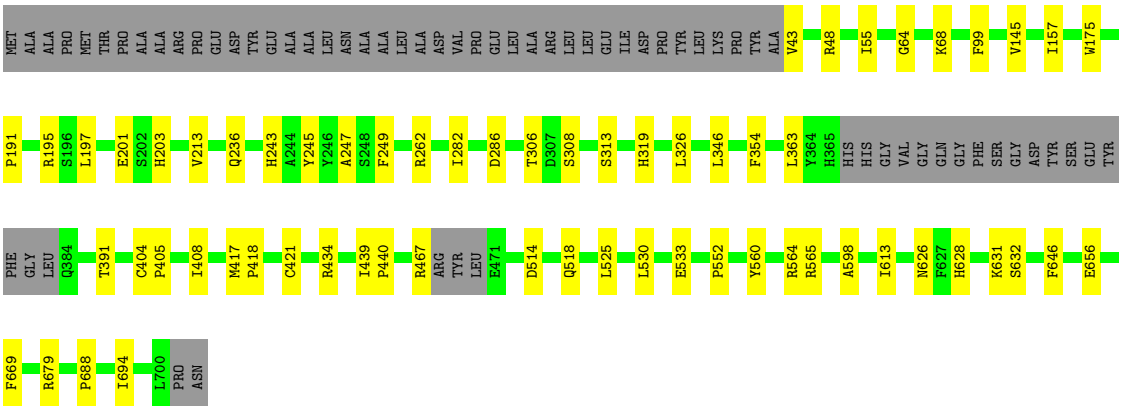
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	137	137	137	0	0
2	B	342	342	342	0	0
2	C	131	131	131	0	0

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.

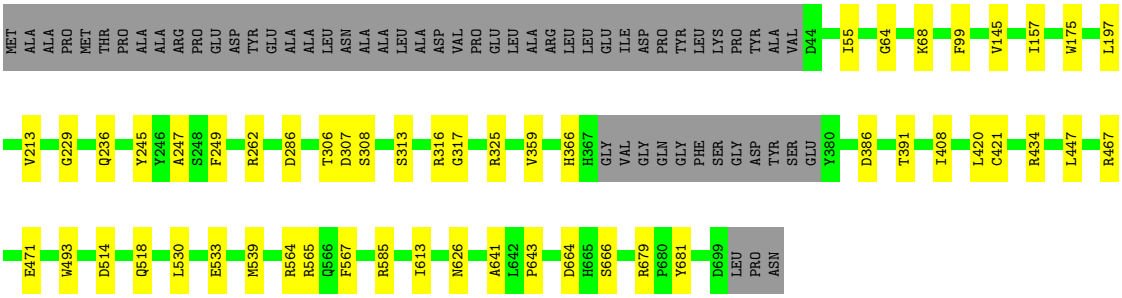
• Molecule 1: 1,4-ALPHA-GLUCAN-BRANCHING ENZYME

Chain A:



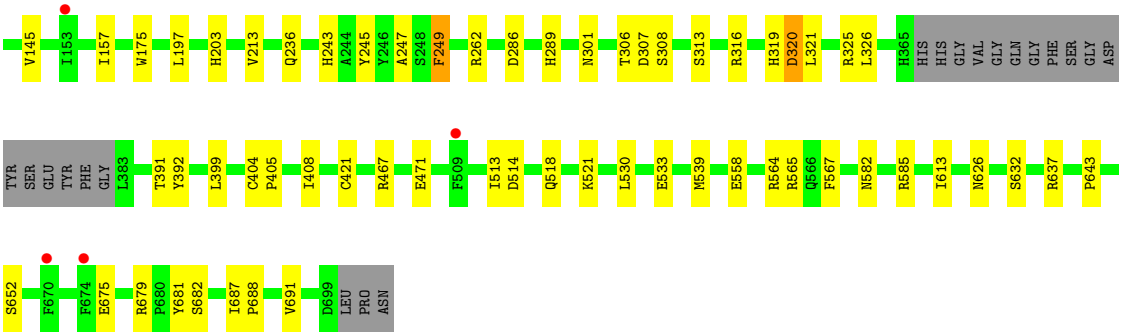
• Molecule 1: 1,4-ALPHA-GLUCAN-BRANCHING ENZYME

Chain B:



• Molecule 1: 1,4-ALPHA-GLUCAN-BRANCHING ENZYME

Chain C:


V145
I153
I157
W175
L197
H203
V213
Q236
H243
A244
Y245
Y246
A247
S248
F249
R262
D286
H289
N301
T306
D307
S308
S313
R316
H319
D320
L321
R325
L326
H365
HIS
HIS
GLY
VAL
GLY
GLN
GLY
PHE
SER
GLY
ASP
TYR
SER
GLU
TYR
PHE
GLY
L383
T391
Y392
L399
C404
P405
I408
C421
R467
E471
F509
I513
D514
Q518
K521
L530
E533
M539
E558
R564
R565
Q566
F567
N582
R585
I613
N626
S632
R637
P643
S652
F670
F674
E675
R679
P680
Y681
S682
I687
P688
V691
D699
LEU
PRO
ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	117.28Å 164.54Å 311.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.30 – 2.75 91.30 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.8 (91.30-2.75) 99.8 (91.30-2.75)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.90 (at 2.73Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.184 , 0.225 0.185 , 0.182	Depositor DCC
R_{free} test set	3936 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	48.8	Xtriage
Anisotropy	0.305	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 39.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	16084	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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	0.23	0/5297	0.40	0/7179
1	B	0.24	0/5400	0.41	0/7321
1	C	0.23	0/5257	0.42	0/7137
All	All	0.23	0/15954	0.41	0/21637

There are no bond length outliers.

There are no bond angle outliers.

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	5142	0	4853	36	0
1	B	5234	0	4932	32	0
1	C	5098	0	4751	43	0
2	A	137	0	0	0	0
2	B	342	0	0	1	0
2	C	131	0	0	2	0
All	All	16084	0	14536	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.

The worst 5 of 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:306:THR:HG22	1:B:308:SER:H	1.44	0.81
1:A:306:THR:HG22	1:A:308:SER:H	1.47	0.78
1:C:64:GLY:HA3	1:C:68:LYS:HD2	1.66	0.77
1:C:306:THR:HG22	1:C:308:SER:H	1.50	0.77
1:C:307:ASP:O	1:C:325:ARG:NH2	2.19	0.75

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	631/702 (90%)	619 (98%)	12 (2%)	0	100	100
1	B	642/702 (92%)	630 (98%)	12 (2%)	0	100	100
1	C	634/702 (90%)	621 (98%)	12 (2%)	1 (0%)	47	69
All	All	1907/2106 (91%)	1870 (98%)	36 (2%)	1 (0%)	51	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	320	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	536/599 (90%)	535 (100%)	1 (0%)	93	96
1	B	545/599 (91%)	545 (100%)	0	100	100
1	C	523/599 (87%)	522 (100%)	1 (0%)	93	96
All	All	1604/1797 (89%)	1602 (100%)	2 (0%)	93	96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	656	GLU
1	C	249	PHE

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

Mol	Chain	Res	Type
1	B	537	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	637/702 (90%)	-0.12	0 100 100	35, 49, 67, 95	0
1	B	644/702 (91%)	-0.13	0 100 100	18, 31, 54, 77	0
1	C	637/702 (90%)	0.04	4 (0%) 89 92	32, 52, 79, 98	0
All	All	1918/2106 (91%)	-0.07	4 (0%) 95 97	18, 46, 72, 98	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	509	PHE	3.0
1	C	153	ILE	2.6
1	C	674	PHE	2.3
1	C	670	PHE	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 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.