



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 12:04 AM GMT

PDB ID : 1ZCU
Title : apo form of the 162S mutant of glycogenin
Authors : Hurley, T.D.; Stout, S.L.; Miner, E.; Zhou, J.; Roach, P.J.
Deposited on : 2005-04-13
Resolution : 2.00 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

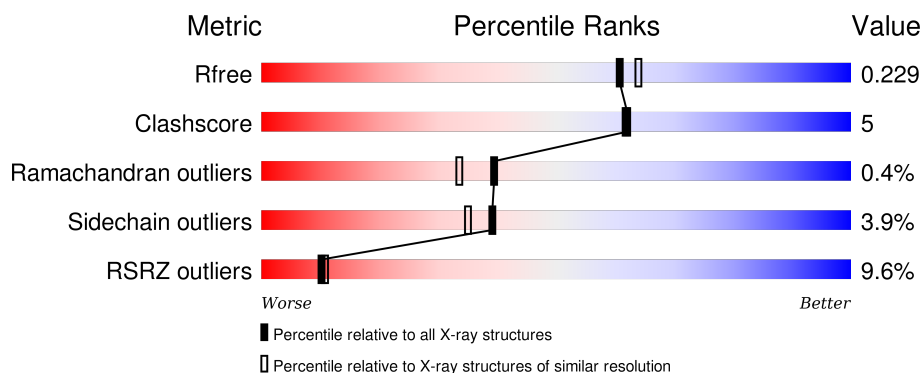
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

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}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (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. 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 $\leq 5\%$. 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	353	<div> <div>7%</div> <div>63%</div> <div>9%</div> <div>26%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2152 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 Glycogenin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	0	0
			2061	1327	338	388	8			

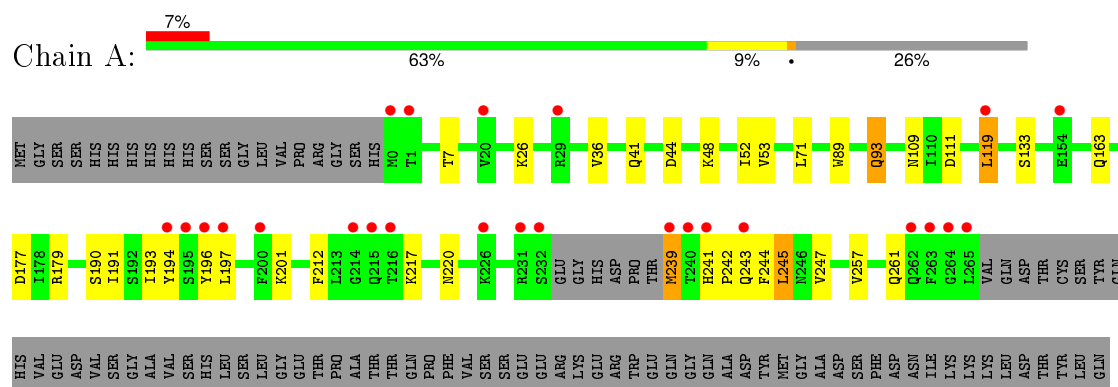
There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	CLONING ARTIFACT	UNP P13280
A	-19	GLY	-	CLONING ARTIFACT	UNP P13280
A	-18	SER	-	CLONING ARTIFACT	UNP P13280
A	-17	SER	-	CLONING ARTIFACT	UNP P13280
A	-16	HIS	-	CLONING ARTIFACT	UNP P13280
A	-15	HIS	-	CLONING ARTIFACT	UNP P13280
A	-14	HIS	-	CLONING ARTIFACT	UNP P13280
A	-13	HIS	-	CLONING ARTIFACT	UNP P13280
A	-12	HIS	-	CLONING ARTIFACT	UNP P13280
A	-11	HIS	-	CLONING ARTIFACT	UNP P13280
A	-10	SER	-	CLONING ARTIFACT	UNP P13280
A	-9	SER	-	CLONING ARTIFACT	UNP P13280
A	-8	GLY	-	CLONING ARTIFACT	UNP P13280
A	-7	LEU	-	CLONING ARTIFACT	UNP P13280
A	-6	VAL	-	CLONING ARTIFACT	UNP P13280
A	-5	PRO	-	CLONING ARTIFACT	UNP P13280
A	-4	ARG	-	CLONING ARTIFACT	UNP P13280
A	-3	GLY	-	CLONING ARTIFACT	UNP P13280
A	-2	SER	-	CLONING ARTIFACT	UNP P13280
A	-1	HIS	-	CLONING ARTIFACT	UNP P13280
A	162	SER	ASP	ENGINEERED	UNP P13280

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	91	Total	O	0	0
			91	91		

- Molecule 1: Glycogenin-1



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	58.95Å 104.78Å 121.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.00 24.06 – 2.00	Depositor EDS
% Data completeness (in resolution range)	94.5 (25.00-2.00) 94.8 (24.06-2.00)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.40 (at 2.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.216 , 0.229 0.217 , 0.229	Depositor DCC
R_{free} test set	1233 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	32.5	Xtriage
Anisotropy	0.730	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 41.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 24452 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2152	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.39	0/2114	0.62	0/2881

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	2061	0	2022	22	0
2	A	91	0	0	0	0
All	All	2152	0	2022	22	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 5.

All (22) 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:93:GLN:H	1:A:93:GLN:HE21	1.15	0.90
1:A:93:GLN:N	1:A:93:GLN:HE21	1.91	0.68
1:A:93:GLN:H	1:A:93:GLN:NE2	1.92	0.65
1:A:239:MET:HE3	1:A:245:LEU:HD23	1.79	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:197:LEU:HD11	1:A:241:HIS:CE1	2.40	0.56
1:A:190:SER:HA	1:A:193:ILE:HD12	1.88	0.54
1:A:191:ILE:HA	1:A:194:TYR:CE2	2.44	0.53
1:A:193:ILE:HD13	1:A:244:PHE:CE1	2.44	0.53
1:A:217:LYS:H	1:A:220:ASN:HD22	1.58	0.51
1:A:44:ASP:O	1:A:48:LYS:HG3	2.13	0.47
1:A:212:PHE:CE2	1:A:245:LEU:HD13	2.49	0.47
1:A:243:GLN:O	1:A:247:VAL:HG23	2.15	0.46
1:A:257:VAL:O	1:A:261:GLN:HG2	2.15	0.46
1:A:109:ASN:ND2	1:A:111:ASP:H	2.14	0.46
1:A:7:THR:HG23	1:A:36:VAL:HG23	1.98	0.45
1:A:197:LEU:HG	1:A:201:LYS:HE3	2.00	0.44
1:A:26:LYS:HE2	1:A:53:VAL:O	2.18	0.43
1:A:239:MET:HE1	1:A:242:PRO:HB3	2.01	0.42
1:A:133:SER:H	1:A:163:GLN:HE21	1.68	0.41
1:A:48:LYS:O	1:A:52:ILE:HD13	2.20	0.41
1:A:119:LEU:HA	1:A:119:LEU:HD22	1.69	0.41
1:A:197:LEU:HD11	1:A:241:HIS:HE1	1.86	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	256/353 (72%)	250 (98%)	5 (2%)	1 (0%)	39	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	177	ASP

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	230 / 311 (74%)	221 (96%)	9 (4%)	39 35

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	GLN
1	A	71	LEU
1	A	89	TRP
1	A	93	GLN
1	A	119	LEU
1	A	179	ARG
1	A	196	TYR
1	A	239	MET
1	A	245	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	93	GLN
1	A	109	ASN
1	A	139	GLN
1	A	163	GLN
1	A	206	ASN
1	A	220	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 [i](#)

There are no carbohydrates 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

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	260/353 (73%)	0.44	25 (9%) 10 11	24, 37, 69, 83	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	215	GLN	5.9
1	A	240	THR	5.8
1	A	194	TYR	5.1
1	A	263	PHE	4.7
1	A	0	MET	4.6
1	A	232	SER	4.1
1	A	195	SER	4.1
1	A	265	LEU	3.9
1	A	214	GLY	3.8
1	A	231	ARG	3.3
1	A	241	HIS	3.2
1	A	216	THR	3.1
1	A	239	MET	3.1
1	A	262	GLN	3.0
1	A	154	GLU	2.9
1	A	1	THR	2.9
1	A	29	ARG	2.9
1	A	226	LYS	2.8
1	A	197	LEU	2.7
1	A	196	TYR	2.6
1	A	119	LEU	2.5
1	A	264	GLY	2.5
1	A	20	VAL	2.3
1	A	200	PHE	2.2
1	A	243	GLN	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 carbohydrates 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.