WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Oct 13, 2024 – 10:09 am BST

PDB ID	:	1OGO
Title	:	Dex49A from Penicillium minioluteum complex with isomaltose
Authors	:	Larsson, A.M.; Stahlberg, J.; Jones, T.A.
Deposited on	:	2003-05-08
Resolution	:	1.65 Å(reported)

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

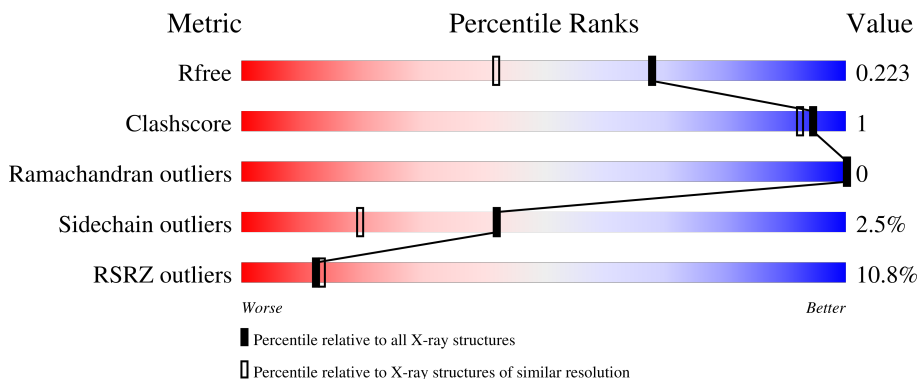|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
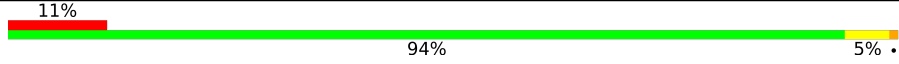*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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	Percentile Ranks	Value
Rfree		0.223
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		2.5%
RSRZ outliers		10.8%

Worse ... *Better*

▊ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	2328 (1.66-1.66)
Clashscore	180529	2515 (1.66-1.66)
Ramachandran outliers	177936	2475 (1.66-1.66)
Sidechain outliers	177891	2475 (1.66-1.66)
RSRZ outliers	164620	2328 (1.66-1.66)

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	X	574	11% / 94% / 5% •
2	A	2	50% / 50%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4929 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 DEXTRANASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	572	Total 4398	C 2782	N 735	O 863	S 18	0	1	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	1	HIS	MET	engineered mutation	UNP P48845
X	5	ALA	ASN	engineered mutation	UNP P48845
X	537	ALA	ASN	engineered mutation	UNP P48845
X	540	ALA	ASN	engineered mutation	UNP P48845
X	543	ILE	VAL	engineered mutation	UNP P48845

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-6)-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	A	2	Total 23	C 12	O 11	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	X	508	Total 508	O 508	0	0

3 Residue-property plots ⓘ

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: DEXTRANASE
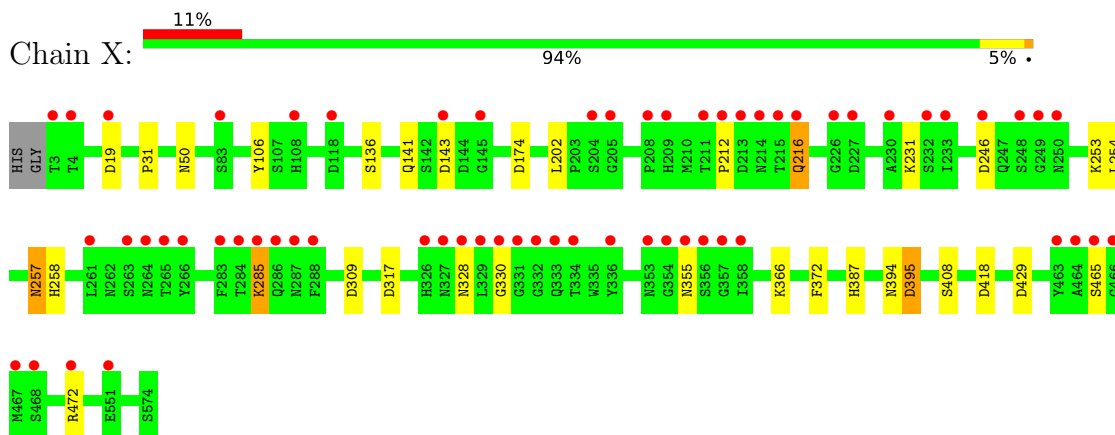
• Molecule 2: alpha-D-glucopyranose-(1-6)-beta-D-glucopyranose

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	115.41Å 103.77Å 49.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.73 – 1.65 25.73 – 1.65	Depositor EDS
% Data completeness (in resolution range)	99.1 (25.73-1.65) 99.0 (25.73-1.65)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.32 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.1.19	Depositor
R, R_{free}	0.188 , 0.215 0.199 , 0.223	Depositor DCC
R_{free} test set	3613 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å²)	17.5	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.41 , 42.6	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4929	wwPDB-VP
Average B, all atoms (Å²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.96% 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: BGC, GLC

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	X	0.45	0/4523	0.74	7/6180 (0.1%)

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	X	0	1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	X	317	ASP	CB-CG-OD2	6.46	124.11	118.30
1	X	19	ASP	CB-CG-OD1	6.45	124.10	118.30
1	X	309	ASP	CB-CG-OD2	5.83	123.54	118.30
1	X	418	ASP	CB-CG-OD2	5.55	123.30	118.30
1	X	174	ASP	CB-CG-OD2	5.54	123.29	118.30
1	X	395	ASP	CB-CG-OD2	5.31	123.08	118.30
1	X	143	ASP	CB-CG-OD2	5.19	122.97	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	X	328	ASN	Peptide

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	X	4398	0	4153	10	0
2	A	23	0	21	0	0
3	X	508	0	0	2	0
All	All	4929	0	4174	10	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 1.

All (10) 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:X:246:ASP:OD2	1:X:253:LYS:NZ	2.13	0.81
1:X:285:LYS:NZ	3:X:2320:HOH:O	2.30	0.59
1:X:136:SER:OG	3:X:2187:HOH:O	2.17	0.57
1:X:212:PRO:O	1:X:231:LYS:NZ	2.45	0.49
1:X:31:PRO:HG3	1:X:254:LEU:HD11	1.96	0.47
1:X:372:PHE:O	1:X:394:ASN:HB3	2.16	0.45
1:X:387:HIS:HA	1:X:408:SER:O	2.17	0.44
1:X:330:GLY:O	1:X:355:ASN:ND2	2.51	0.43
1:X:257:ASN:HD22	1:X:258:HIS:H	1.68	0.42
1:X:216:GLN:HE21	1:X:216:GLN:HA	1.85	0.41

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	X	571/574 (100%)	543 (95%)	28 (5%)	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	X	485/485 (100%)	473 (98%)	12 (2%)	42 19

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	50	ASN
1	X	106	TYR
1	X	141	GLN
1	X	202	LEU
1	X	216	GLN
1	X	257	ASN
1	X	285	LYS
1	X	366	LYS
1	X	395	ASP
1	X	429	ASP
1	X	465	SER
1	X	472	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	50	ASN
1	X	51	ASN
1	X	53	HIS
1	X	156	ASN
1	X	216	GLN
1	X	257	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	X	264	ASN
1	X	287	ASN
1	X	300	ASN
1	X	304	GLN
1	X	333	GLN
1	X	343	ASN
1	X	355	ASN
1	X	374	GLN
1	X	562	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 ⓘ

2 monosaccharides 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	BGC	A	1	2	12,12,12	0.51	0	17,17,17	0.71	0
2	GLC	A	2	2	11,11,12	0.60	0	15,15,17	0.94	1 (6%)

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BGC	A	1	2	-	0/2/22/22	0/1/1/1
2	GLC	A	2	2	-	0/2/19/22	0/1/1/1

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	A	2	GLC	C1-O5-C5	2.87	116.08	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain A

Bond lengths

Bond angles

Torsions

Rings
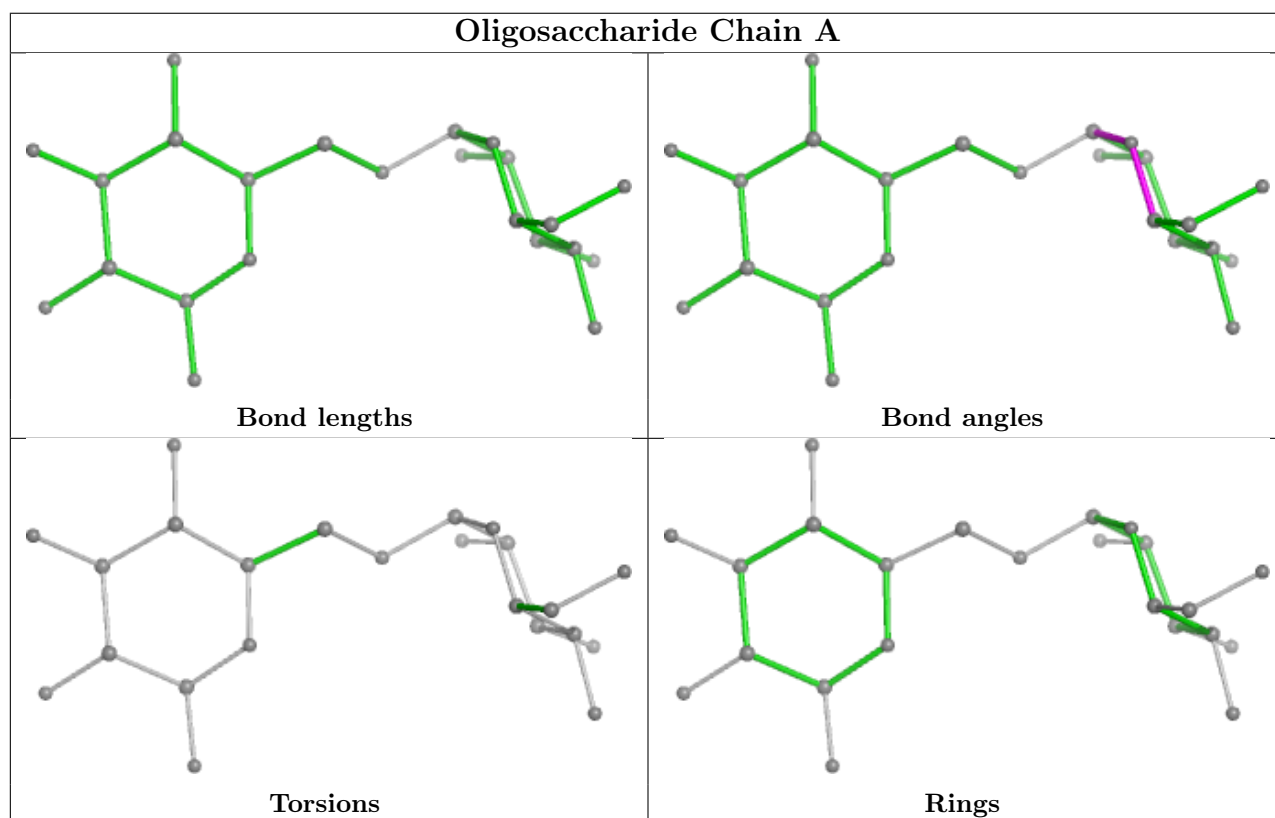
5.6 Ligand geometry (i)

There are no ligands in this entry.

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(Å²)	Q<0.9
1	X	572/574 (99%)	0.47	62 (10%)	12	13	8, 21, 35, 45	1 (0%)

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	329	LEU	7.4
1	X	3	THR	6.0
1	X	214	ASN	5.0
1	X	465	SER	4.2
1	X	330	GLY	4.1
1	X	354	GLY	4.0
1	X	332	GLY	3.8
1	X	230	ALA	3.8
1	X	4	THR	3.8
1	X	209	HIS	3.8
1	X	213	ASP	3.7
1	X	249	GLY	3.6
1	X	551	GLU	3.5
1	X	143	ASP	3.4
1	X	466	GLY	3.4
1	X	464	ALA	3.2
1	X	463	TYR	3.2
1	X	263	SER	3.2
1	X	248	SER	3.1
1	X	331	GLY	3.1
1	X	357	GLY	3.1
1	X	205	GLY	3.1
1	X	233	ILE	3.0
1	X	355	ASN	3.0
1	X	264	ASN	3.0
1	X	208	PRO	2.8
1	X	265	THR	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	X	472	ARG	2.8
1	X	467	MET	2.7
1	X	211	THR	2.7
1	X	328	ASN	2.7
1	X	266	TYR	2.7
1	X	215	THR	2.6
1	X	334	THR	2.6
1	X	284	THR	2.6
1	X	204	SER	2.6
1	X	356	SER	2.6
1	X	212	PRO	2.5
1	X	288	PHE	2.5
1	X	327	ASN	2.5
1	X	468	SER	2.5
1	X	226	GLY	2.5
1	X	19	ASP	2.5
1	X	287	ASN	2.4
1	X	336	TYR	2.4
1	X	118	ASP	2.4
1	X	283	PHE	2.4
1	X	83	SER	2.4
1	X	227	ASP	2.4
1	X	250	ASN	2.3
1	X	285	LYS	2.3
1	X	353	ASN	2.3
1	X	108	HIS	2.3
1	X	232	SER	2.3
1	X	246	ASP	2.2
1	X	286	GLN	2.2
1	X	145	GLY	2.1
1	X	216	GLN	2.1
1	X	326	HIS	2.0
1	X	333	GLN	2.0
1	X	358	ILE	2.0
1	X	261	LEU	2.0

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 ⓘ

SUGAR-RSR INFOmissingINFO

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.