wwPDB X-ray Structure Validation Summary Report (i)

Dec 13, 2023 – 12:54 am GMT

PDB ID	:	2Y5Y
Title	:	Crystal structure of LacY in complex with an affinity inactivator
Authors	:	Chaptal, V.; Kwon, S.; Sawaya, M.R.; Guan, L.; Kaback, H.R.; Abramson, J.
Deposited on	:	2011-01-19
Resolution	:	3.38 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 (i)) were used in the production of this report:

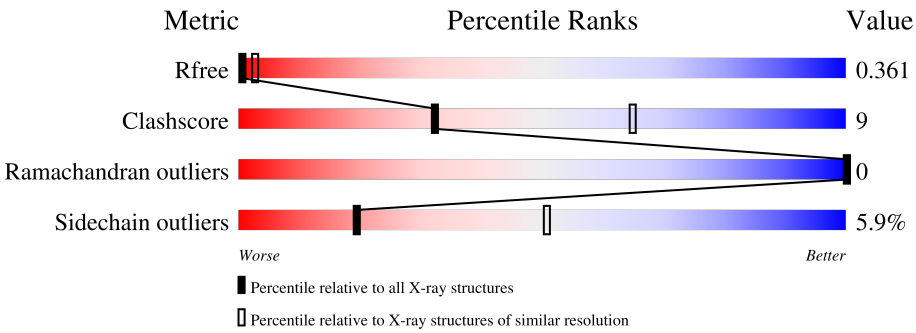
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.38 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	1691 (3.46-3.30)
Clashscore	141614	1762 (3.46-3.30)
Ramachandran outliers	138981	1732 (3.46-3.30)
Sidechain outliers	138945	1731 (3.46-3.30)

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	423	71% 18% • 9%
1	B	423	73% 19% • 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6209 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 LACTOSE PERMEASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	386	3065	2086	466	498	15	0	0	0
1	B	394	3112	2114	475	508	15	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	117	SER	CYS	engineered mutation	UNP P02920
A	122	CYS	ALA	engineered mutation	UNP P02920
A	148	ALA	CYS	engineered mutation	UNP P02920
A	154	VAL	CYS	engineered mutation	UNP P02920
A	176	MET	CYS	engineered mutation	UNP P02920
A	234	SER	CYS	engineered mutation	UNP P02920
A	333	SER	CYS	engineered mutation	UNP P02920
A	353	ALA	CYS	engineered mutation	UNP P02920
A	355	ALA	CYS	engineered mutation	UNP P02920
A	418	HIS	-	expression tag	UNP P02920
A	419	HIS	-	expression tag	UNP P02920
A	420	HIS	-	expression tag	UNP P02920
A	421	HIS	-	expression tag	UNP P02920
A	422	HIS	-	expression tag	UNP P02920
A	423	HIS	-	expression tag	UNP P02920
B	117	SER	CYS	engineered mutation	UNP P02920
B	122	CYS	ALA	engineered mutation	UNP P02920
B	148	ALA	CYS	engineered mutation	UNP P02920
B	154	VAL	CYS	engineered mutation	UNP P02920
B	176	MET	CYS	engineered mutation	UNP P02920
B	234	SER	CYS	engineered mutation	UNP P02920
B	333	SER	CYS	engineered mutation	UNP P02920
B	353	ALA	CYS	engineered mutation	UNP P02920
B	355	ALA	CYS	engineered mutation	UNP P02920
B	418	HIS	-	expression tag	UNP P02920

Continued on next page...

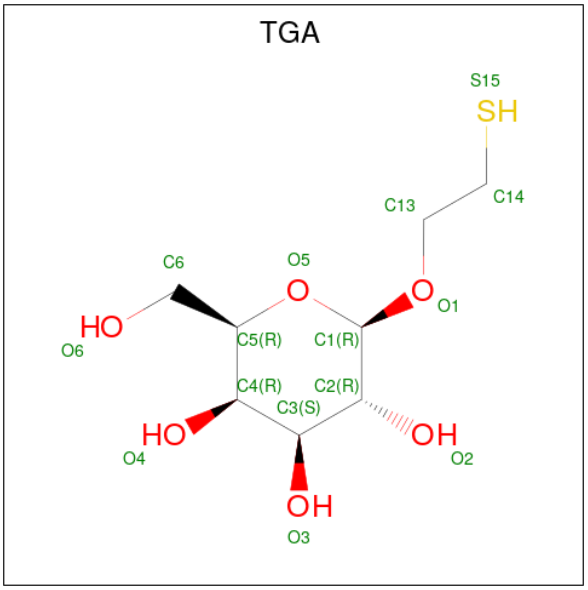
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	419	HIS	-	expression tag	UNP P02920
B	420	HIS	-	expression tag	UNP P02920
B	421	HIS	-	expression tag	UNP P02920
B	422	HIS	-	expression tag	UNP P02920
B	423	HIS	-	expression tag	UNP P02920

- Molecule 2 is BARIUM ION (three-letter code: BA) (formula: Ba).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ba		
2	A	1	1	1	0	0
			Total	Ba		
2	B	1	1	1	0	0

- Molecule 3 is 2-sulfanylethyl beta-D-galactopyranoside (three-letter code: TGA) (formula: $C_8H_{16}O_6S$).



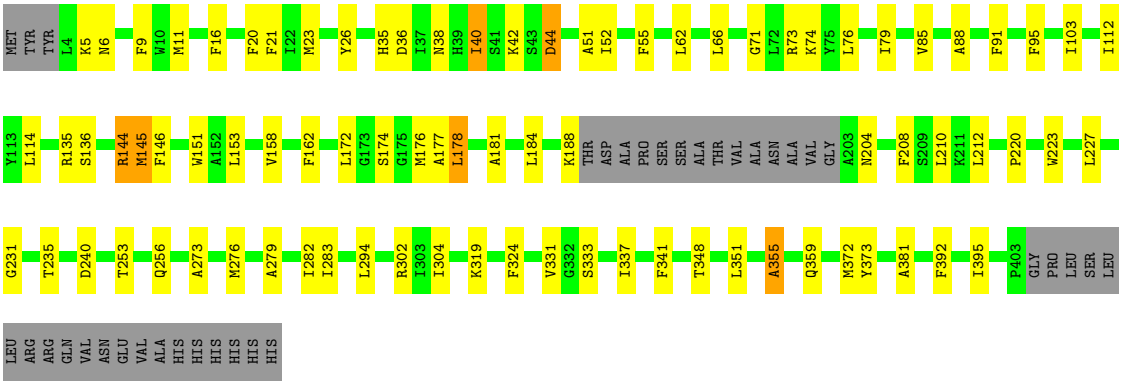
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
3	A	1	15	8	6	1	0	0
			Total	C	O	S		
3	B	1	15	8	6	1	0	0

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. 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.

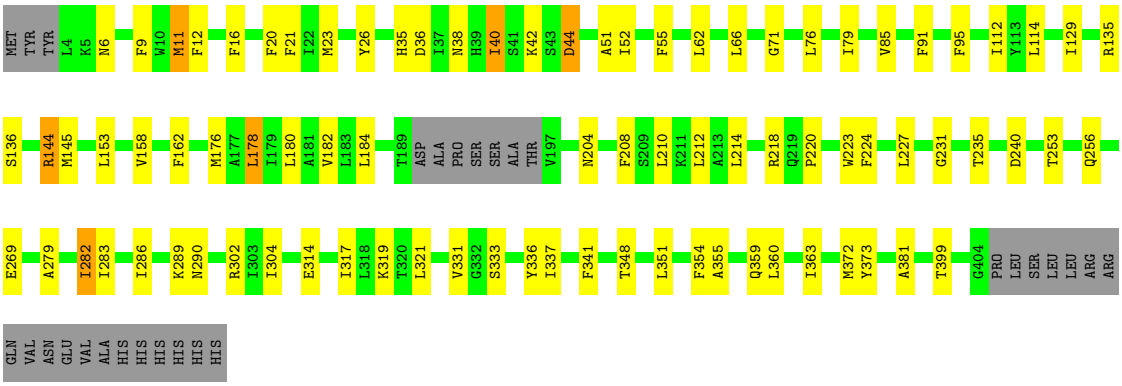
• Molecule 1: LACTOSE PERMEASE

Chain A:



• Molecule 1: LACTOSE PERMEASE

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	102.56Å 127.84Å 189.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.28 – 3.38 45.12 – 3.38	Depositor EDS
% Data completeness (in resolution range)	(Not available) (28.28-3.38) 98.4 (45.12-3.38)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 3.40Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, R_{free}	0.286 , 0.303 0.349 , 0.361	Depositor DCC
R_{free} test set	1760 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	105.2	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.20 , 73.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.37$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	6209	wwPDB-VP
Average B, all atoms (Å^2)	163.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: BA, TGA

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	0.43	0/3157	0.73	7/4273 (0.2%)
1	B	0.43	0/3204	0.72	5/4338 (0.1%)
All	All	0.43	0/6361	0.72	12/8611 (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	A	0	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	135	ARG	CB-CA-C	8.80	127.99	110.40
1	A	135	ARG	CB-CA-C	8.18	126.76	110.40
1	B	136	SER	N-CA-CB	-7.85	98.72	110.50
1	A	36	ASP	CB-CA-C	7.69	125.79	110.40
1	A	136	SER	N-CA-CB	-7.32	99.53	110.50

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	144	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	144	ARG	Sidechain

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	3065	0	3112	52	0
1	B	3112	0	3159	58	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	15	0	15	1	0
3	B	15	0	15	3	0
All	All	6209	0	6301	110	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 9.

The worst 5 of 110 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:55:PHE:HB2	1:B:112:ILE:HG21	1.41	1.02
1:A:55:PHE:HB2	1:A:112:ILE:HG21	1.41	0.99
1:A:55:PHE:CD2	1:A:112:ILE:HG22	2.10	0.87
1:B:55:PHE:CD2	1:B:112:ILE:HG22	2.12	0.85
1:B:253:THR:OG1	1:B:256:GLN:HG2	1.77	0.84

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	382/423 (90%)	359 (94%)	23 (6%)	0	100	100
1	B	390/423 (92%)	365 (94%)	25 (6%)	0	100	100
All	All	772/846 (91%)	724 (94%)	48 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	320/350 (91%)	301 (94%)	19 (6%)	19	51
1	B	324/350 (93%)	305 (94%)	19 (6%)	19	51
All	All	644/700 (92%)	606 (94%)	38 (6%)	19	51

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	210	LEU
1	B	317	ILE
1	B	212	LEU
1	B	282	ILE
1	B	351	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	137	ASN
1	B	241	GLN
1	B	272	ASN
1	A	290	ASN
1	A	379	GLN

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)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
3	TGA	A	701	1	15,15,15	0.93	1 (6%)	20,20,20	1.07	2 (10%)
3	TGA	B	801	1	15,15,15	0.78	0	20,20,20	1.18	3 (15%)

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
3	TGA	A	701	1	-	1/6/26/26	0/1/1/1
3	TGA	B	801	1	-	4/6/26/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	701	TGA	O1-C1	2.71	1.44	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	801	TGA	O5-C5-C6	3.29	114.61	106.44
3	B	801	TGA	C13-O1-C1	-2.96	108.93	113.84
3	A	701	TGA	O5-C5-C6	2.47	112.58	106.44
3	A	701	TGA	O5-C1-C2	2.47	115.57	110.35
3	B	801	TGA	C1-O5-C5	-2.02	109.73	113.69

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	801	TGA	O1-C13-C14-S15
3	B	801	TGA	C4-C5-C6-O6
3	B	801	TGA	O5-C5-C6-O6
3	A	701	TGA	O1-C13-C14-S15
3	B	801	TGA	C14-C13-O1-C1

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	701	TGA	1	0
3	B	801	TGA	3	0

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)

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

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

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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