wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:37 PM GMT

PDB ID	:	1PV7
Title	:	Crystal structure of lactose permease with TDG
Authors	:	Abramson, J.; Smirnova, I.; Kasho, V.; Verner, G.; Kaback, H.R.; Iwata, S.
Deposited on	:	2003-06-26
Resolution	:	3.60 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

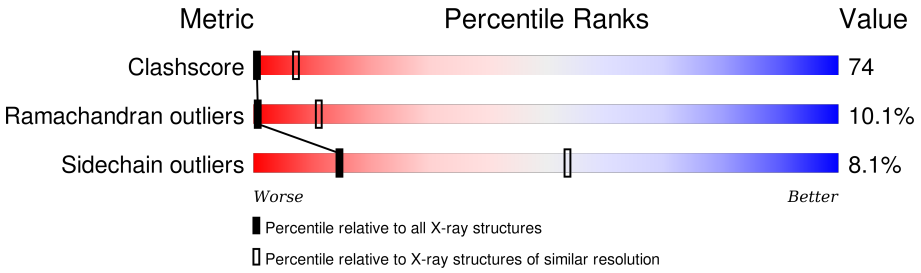
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.60 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))
Clashscore	102246	1010 (3.74-3.46)
Ramachandran outliers	100387	1007 (3.76-3.44)
Sidechain outliers	100360	1007 (3.76-3.44)

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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	417	23% 63% 14%
1	B	417	24% 62% 13%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6626 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.

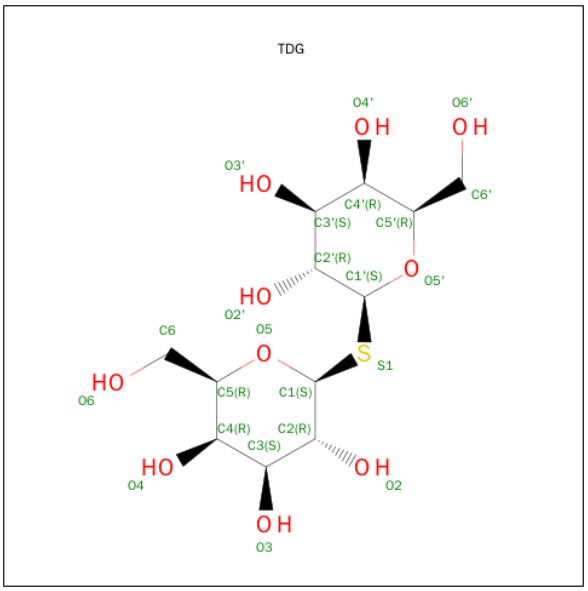
- Molecule 1 is a protein called Lactose permease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	417	Total	C	N	O	S	0	0	0
			3290	2222	506	541	21			
1	B	417	Total	C	N	O	S	0	0	0
			3290	2222	506	541	21			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	154	GLY	CYS	ENGINEERED	UNP P02920
B	154	GLY	CYS	ENGINEERED	UNP P02920

- Molecule 2 is THIODIGALACTOSIDE (three-letter code: TDG) (formula: $C_{12}H_{22}O_{10}S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			23	12	10	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	B	1	Total	C	O	S	0	0
			23	12	10	1		

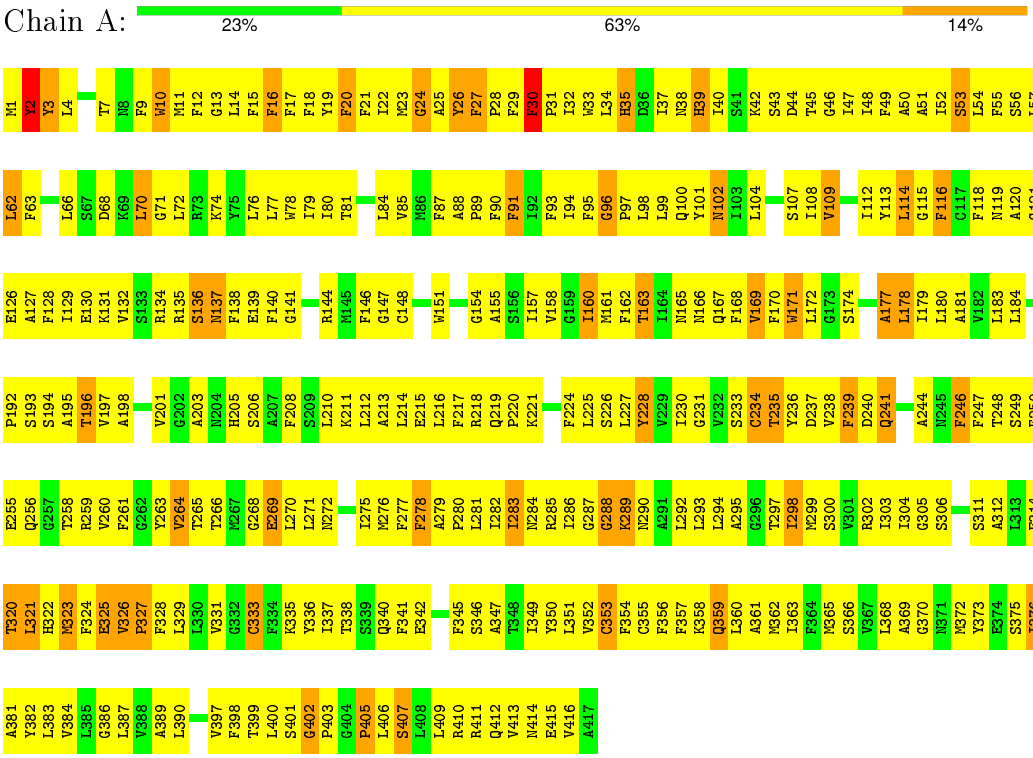
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

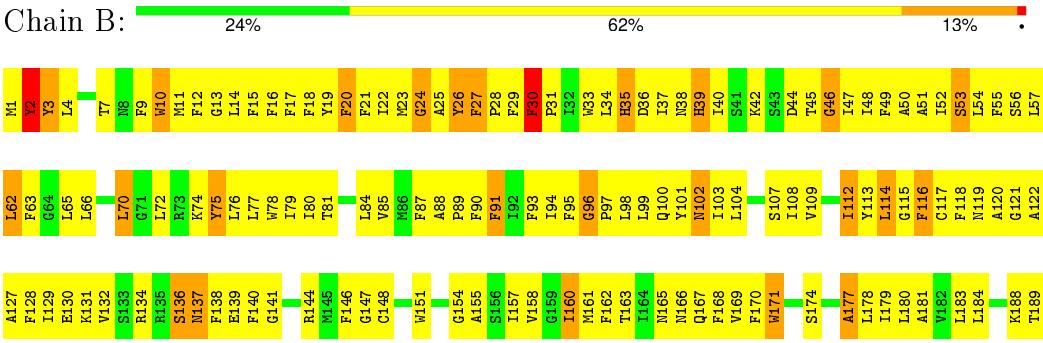
- Molecule 1: Lactose permease

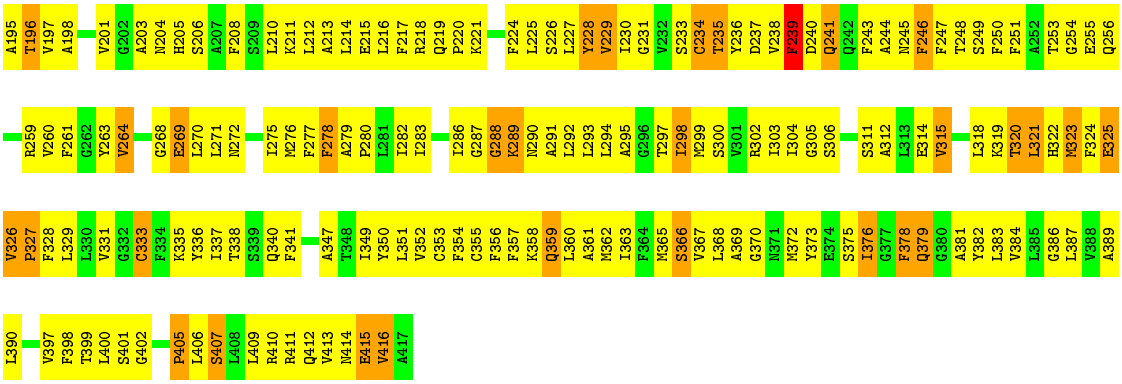
Chain A:



- Molecule 1: Lactose permease

Chain B:




A195 T196 V197 A198 V201 G202 A203 N204 H205 S206 A207 F208 S209 L210 K211 L212 A213 L214 E215 L216 F217 R218 Q219 P220 K221 F224 L225 S226 L227 Y228 V229 I230 G231 V232 S233 C234 T235 Y236 D237 V238 F239 D240 Q241 Q242 F243 A244 N245 F246 F247 T248 S249 F250 F251 A252 T253 G254 E255 Q256
R259 V260 F261 G262 Y263 V264 G268 E269 L270 L271 N272 I275 M276 F277 F278 A279 P280 L281 I282 I283 I286 G287 G288 K289 N290 A291 L292 L293 L294 A295 G296 T297 I298 M299 S300 V301 R302 I303 I304 G305 S306 S311 A312 L313 E314 V315 L318 K319 T320 L321 H322 M323 F324 E325
V326 P327 F328 L329 L330 V331 G332 C333 F334 K335 Y336 I337 T338 S339 Q340 F341 A347 T348 I349 Y350 L351 V352 C353 F354 C355 F356 F357 K358 Q359 L360 A361 M362 I363 F364 M365 S366 V367 L368 A369 G370 N371 M372 Y373 E374 S375 I376 G377 F378 Q379 G380 A381 Y382 L383 V384 L385 G386 L387 V388 A389
I390 V397 F398 T399 L400 S401 G402 P405 L406 S407 I408 L409 R410 R411 Q412 V413 N414 E415 V416 A417

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	101.35Å 125.84Å 188.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	4.00 – 3.60	Depositor
% Data completeness (in resolution range)	(Not available) (4.00-3.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.271 , 0.296	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6626	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	0.69	1/3387 (0.0%)	0.84	2/4588 (0.0%)
1	B	0.67	1/3387 (0.0%)	0.84	2/4588 (0.0%)
All	All	0.68	2/6774 (0.0%)	0.84	4/9176 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	315	VAL	CB-CG1	-5.82	1.40	1.52
1	B	315	VAL	CB-CG1	-5.50	1.41	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	70	LEU	CA-CB-CG	5.74	128.50	115.30
1	B	70	LEU	CA-CB-CG	5.67	128.34	115.30
1	B	114	LEU	N-CA-C	-5.12	97.18	111.00
1	A	114	LEU	N-CA-C	-5.00	97.49	111.00

There are no chirality outliers.

There are no planarity outliers.

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	A	3290	0	3333	504	0
1	B	3290	0	3333	490	0
2	A	23	0	22	0	0
2	B	23	0	22	0	0
All	All	6626	0	6710	990	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 74.

The worst 5 of 990 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:30:PHE:HB3	1:A:31:PRO:HD3	1.29	1.10
1:B:87:PHE:HB3	1:B:174:SER:HB2	1.36	1.08
1:B:256:GLN:OE1	1:B:259:ARG:HD2	1.54	1.07
1:B:30:PHE:HB3	1:B:31:PRO:HD3	1.31	1.07
1:A:293:LEU:HD13	1:A:397:VAL:HG22	1.36	1.06

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	415/417 (100%)	261 (63%)	114 (28%)	40 (10%)	1 12
1	B	415/417 (100%)	261 (63%)	110 (26%)	44 (11%)	0 10
All	All	830/834 (100%)	522 (63%)	224 (27%)	84 (10%)	1 11

5 of 84 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	102	ASN
1	A	160	ILE
1	A	196	THR
1	A	407	SER

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	345/345 (100%)	317 (92%)	28 (8%)	15 54
1	B	345/345 (100%)	317 (92%)	28 (8%)	15 54
All	All	690/690 (100%)	634 (92%)	56 (8%)	15 54

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	333	CYS
1	B	20	PHE
1	B	333	CYS
1	A	353	CYS
1	B	2	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	379	GLN
1	B	8	ASN
1	B	290	ASN
1	A	371	ASN
1	B	340	GLN

5.3.3 RNA

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

2 ligands 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	TDG	A	500	-	24,24,24	2.05	7 (29%)	32,35,35	0.82	0
2	TDG	B	1500	-	24,24,24	1.84	7 (29%)	32,35,35	0.81	0

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
2	TDG	A	500	-	-	0/8/48/48	0/2/2/2
2	TDG	B	1500	-	-	0/8/48/48	0/2/2/2

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1500	TDG	C4-C5	2.29	1.57	1.53
2	A	500	TDG	C1'-C2'	2.32	1.57	1.53
2	A	500	TDG	C3-C2	2.37	1.58	1.52
2	B	1500	TDG	O5'-C1'	2.56	1.46	1.42
2	B	1500	TDG	C3-C2	2.62	1.59	1.52

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.