wwPDB X-ray Structure Validation Summary Report (i)

Aug 21, 2023 – 08:05 PM EDT

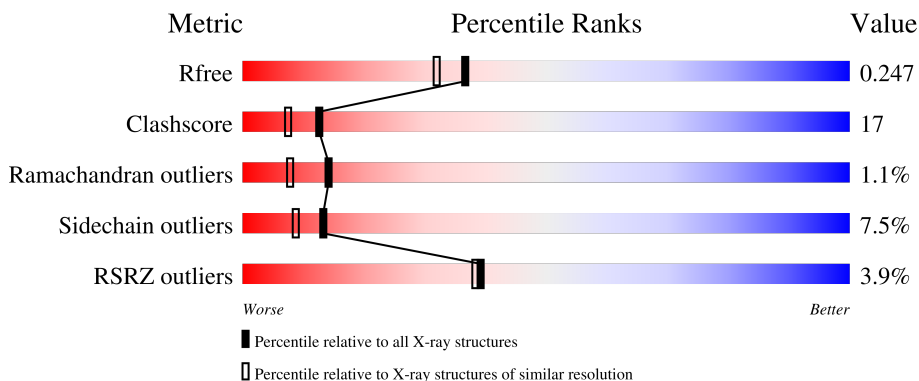|		
PDB ID	:	2P9H
Title	:	High resolution structure of the Lactose Repressor bound to IPTG
Authors	:	Daber, R.
Deposited on	:	2007-03-26
Resolution	:	2.00 Å(reported)

---

1 Overall quality at a glance ⓘ

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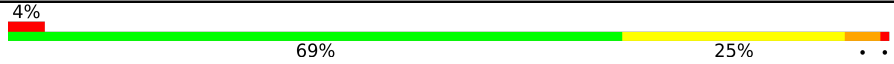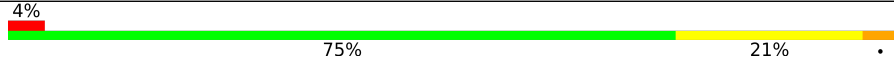*X-RAY DIFFRACTION*

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	Percentile Ranks	Value
Rfree		0.247
Clashscore		17
Ramachandran outliers		1.1%
Sidechain outliers		7.5%
RSRZ outliers		3.9%

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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 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	269	4% 69% 25% • •
1	B	269	4% 75% 21% •

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4794 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 operon repressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	269	Total	C	N	O	S	0	0	0
			2013	1260	356	387	10			
1	B	269	Total	C	N	O	S	0	0	0
			2013	1260	356	387	10			

- Molecule 2 is 1-methylethyl 1-thio-beta-D-galactopyranoside (three-letter code: IPT) (formula: $C_9H_{18}O_5S$).
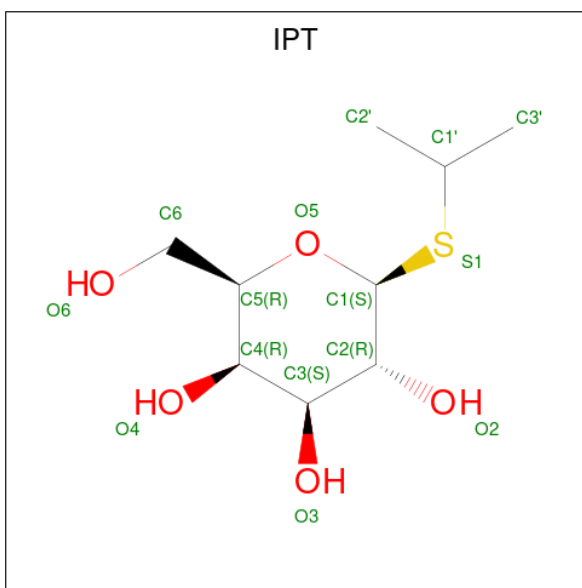
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			15	9	5	1		
2	B	1	Total	C	O	S	0	0
			15	9	5	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	353	Total 353	O 353	0	0
3	B	385	Total 385	O 385	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: Lactose operon repressor

Chain A:

● Molecule 1: Lactose operon repressor

Chain B:
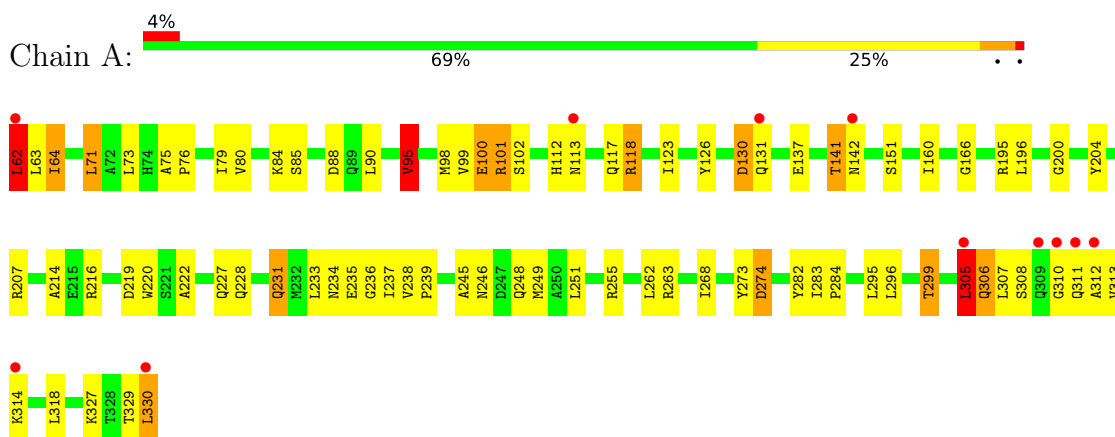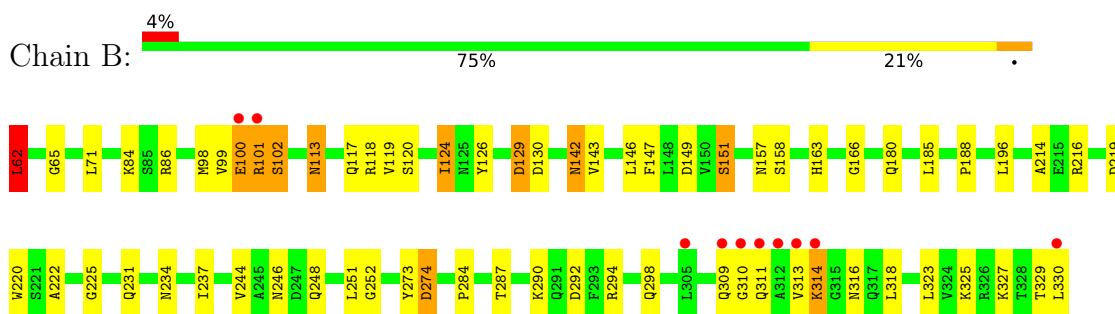
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	77.12Å 77.12Å 210.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.04 – 2.00 34.04 – 2.00	Depositor EDS
% Data completeness (in resolution range)	83.6 (34.04-2.00) 83.6 (34.04-2.00)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.90 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.185 , 0.246 0.188 , 0.247	Depositor DCC
R_{free} test set	1868 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	28.1	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 52.0	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	4794	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.49 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.9424e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: IPT

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	0.78	0/2040	0.90	7/2773 (0.3%)
1	B	0.79	1/2040 (0.0%)	0.82	3/2773 (0.1%)
All	All	0.78	1/4080 (0.0%)	0.86	10/5546 (0.2%)

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	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	113	ASN	CB-CG	5.21	1.63	1.51

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	255	ARG	NE-CZ-NH1	7.95	124.27	120.30
1	A	255	ARG	NE-CZ-NH2	-7.80	116.40	120.30
1	A	62	LEU	CA-CB-CG	7.04	131.50	115.30
1	A	305	LEU	CA-CB-CG	5.91	128.89	115.30
1	A	195	ARG	NE-CZ-NH1	-5.80	117.40	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	99	VAL	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	A	2013	0	2058	85	0
1	B	2013	0	2058	61	0
2	A	15	0	18	0	0
2	B	15	0	18	1	0
3	A	353	0	0	30	0
3	B	385	0	0	20	0
All	All	4794	0	4152	137	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 17.

The worst 5 of 137 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:227:GLN:HG3	3:A:1339:HOH:O	1.20	1.31
1:A:234:ASN:HB2	3:A:1316:HOH:O	1.17	1.27
1:B:129:ASP:HB2	3:B:1370:HOH:O	1.31	1.22
1:A:113:ASN:HB3	3:A:1091:HOH:O	1.38	1.18
1:B:113:ASN:HB2	3:B:1184:HOH:O	1.46	1.16

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	267/269 (99%)	253 (95%)	11 (4%)	3 (1%)	14	8
1	B	267/269 (99%)	254 (95%)	10 (4%)	3 (1%)	14	8
All	All	534/538 (99%)	507 (95%)	21 (4%)	6 (1%)	14	8

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	309	GLN
1	A	274	ASP
1	B	274	ASP
1	A	141	THR
1	A	310	GLY

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	220/220 (100%)	201 (91%)	19 (9%)	10	6
1	B	220/220 (100%)	206 (94%)	14 (6%)	17	13
All	All	440/440 (100%)	407 (92%)	33 (8%)	13	9

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	146	LEU
1	B	151	SER
1	B	314	LYS
1	A	299	THR
1	A	231	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	311	GLN
1	B	125	ASN
1	B	316	ASN
1	B	291	GLN
1	B	306	GLN

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 ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

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	IPT	B	999	-	14,15,15	1.20	1 (7%)	18,21,21	1.14	2 (11%)
2	IPT	A	998	-	14,15,15	1.39	1 (7%)	18,21,21	1.39	2 (11%)

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	IPT	B	999	-	-	0/6/26/26	0/1/1/1
2	IPT	A	998	-	-	2/6/26/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	998	IPT	O5-C1	-4.10	1.36	1.42
2	B	999	IPT	O2-C2	2.91	1.49	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	998	IPT	C2-C1-S1	-3.86	105.36	111.30
2	A	998	IPT	C1-O5-C5	-2.71	107.58	112.58
2	B	999	IPT	C2-C1-S1	-2.56	107.35	111.30
2	B	999	IPT	C1-O5-C5	-2.35	108.26	112.58

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	998	IPT	C4-C5-C6-O6
2	A	998	IPT	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	999	IPT	1	0

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	269/269 (100%)	-0.20	11 (4%)	37	36	18, 28, 46, 70	0
1	B	269/269 (100%)	-0.08	10 (3%)	41	41	19, 29, 45, 71	0
All	All	538/538 (100%)	-0.14	21 (3%)	39	38	18, 28, 46, 71	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	311	GLN	5.9
1	A	330	LEU	4.6
1	B	311	GLN	4.6
1	A	312	ALA	4.1
1	B	305	LEU	3.8

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
2	IPT	B	999	15/15	0.96	0.16	18,20,23,23	0
2	IPT	A	998	15/15	0.98	0.10	14,19,22,23	0

6.5   Other polymers ⓘ

There are no such residues in this entry.