Full wwPDB X-ray Structure Validation Report (i)

Aug 26, 2024 – 05:58 pm BST

PDB ID	:	8QTY
Title	:	LytR LCP domain from Streptococcus dysgalactiae subs. dysgalactiae
Authors	:	Paquete-Ferreira, J.; Romao, M.J.; Santos-Silva, T.
Deposited on	:	2023-10-13
Resolution	:	2.80 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
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 (1)) were used in the production of this report:

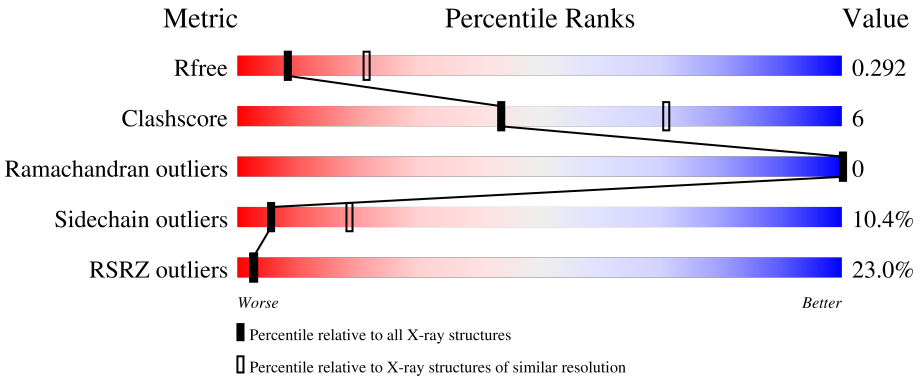
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.2

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 2.80 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}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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	305	20% 70% 16% • 11%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4216 atoms, of which 2100 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.

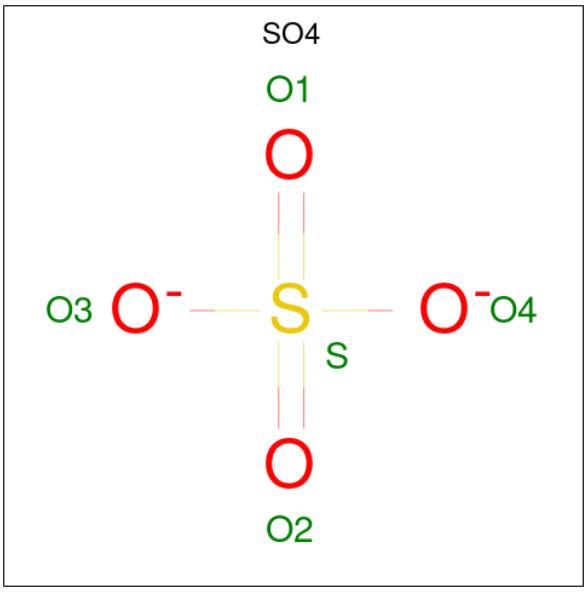
- Molecule 1 is a protein called Biofilm regulatory protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	270	4176	1294	2100	354	419	9	57	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	MET	-	initiating methionine	UNP A0A0A7LTX5
A	343	ALA	-	expression tag	UNP A0A0A7LTX5
A	344	LEU	-	expression tag	UNP A0A0A7LTX5
A	345	GLU	-	expression tag	UNP A0A0A7LTX5
A	346	HIS	-	expression tag	UNP A0A0A7LTX5
A	347	HIS	-	expression tag	UNP A0A0A7LTX5
A	348	HIS	-	expression tag	UNP A0A0A7LTX5
A	349	HIS	-	expression tag	UNP A0A0A7LTX5
A	350	HIS	-	expression tag	UNP A0A0A7LTX5
A	351	HIS	-	expression tag	UNP A0A0A7LTX5

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0

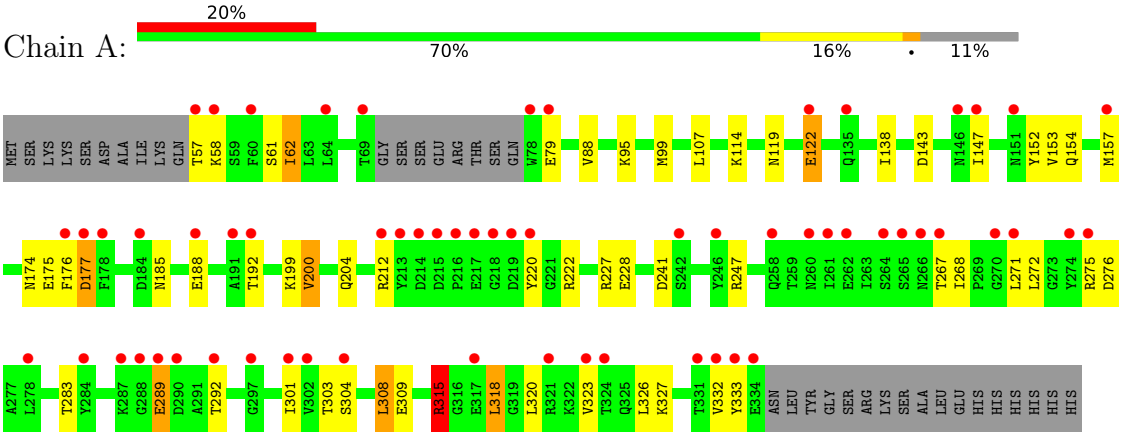
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	25	Total O 25 25	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 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: Biofilm regulatory protein



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α, β, γ	141.27Å 141.27Å 133.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.61 – 2.80 48.61 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.61-2.80) 99.8 (48.61-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.271 , 0.293 0.272 , 0.292	Depositor DCC
R_{free} test set	897 reflections (5.28%)	wwPDB-VP
Wilson B-factor (Å^2)	83.9	Xtriage
Anisotropy	0.176	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 51.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	4216	wwPDB-VP
Average B, all atoms (Å^2)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% 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: SO4

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	0.90	6/2099 (0.3%)	0.98	3/2839 (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	2

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	175	GLU	CD-OE1	11.05	1.37	1.25
1	A	289	GLU	CD-OE1	8.86	1.35	1.25
1	A	188	GLU	CD-OE2	7.89	1.34	1.25
1	A	188	GLU	CD-OE1	7.79	1.34	1.25
1	A	228	GLU	CD-OE1	6.97	1.33	1.25
1	A	79	GLU	CD-OE1	5.02	1.31	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	143	ASP	CB-CA-C	5.26	120.93	110.40
1	A	79	GLU	CB-CA-C	5.25	120.90	110.40
1	A	303	THR	OG1-CB-CG2	-5.10	98.28	110.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	275	ARG	Sidechain
1	A	315	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	2076	2100	2096	25	8
2	A	15	0	0	0	0
3	A	25	0	0	0	0
All	All	2116	2100	2096	25	8

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 6.

All (25) 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:61:SER:HB3	1:A:318:LEU:HD21	1.72	0.70
1:A:95:LYS:HD3	1:A:276:ASP:HB3	1.75	0.68
1:A:301:ILE:HB	1:A:333:TYR:CE1	2.30	0.67
1:A:247:ARG:HE	1:A:268:ILE:HD12	1.65	0.61
1:A:301:ILE:HG21	1:A:333:TYR:CD1	2.39	0.57
1:A:247:ARG:HG3	1:A:268:ILE:CD1	2.35	0.57
1:A:62:ILE:HG23	1:A:88:VAL:HB	1.88	0.55
1:A:114:LYS:HD3	1:A:119:ASN:ND2	2.22	0.54
1:A:147:ILE:HD12	1:A:315:ARG:HG3	1.91	0.52
1:A:62:ILE:HD11	1:A:153:VAL:HG23	1.93	0.49
1:A:308:LEU:HD13	1:A:326:LEU:HD13	1.95	0.49
1:A:99:MET:HE2	1:A:283:THR:HG22	1.95	0.48
1:A:332:VAL:HG12	1:A:333:TYR:N	2.30	0.46
1:A:308:LEU:O	1:A:309:GLU:C	2.54	0.45
1:A:318:LEU:HD12	1:A:318:LEU:HA	1.72	0.45
1:A:138:ILE:HD11	1:A:152:TYR:HB3	1.99	0.44
1:A:152:TYR:CE1	1:A:154:GLN:HG3	2.54	0.43
1:A:227:ARG:HD2	1:A:283:THR:HB	2.01	0.43
1:A:152:TYR:HE1	1:A:154:GLN:HG3	1.84	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:200:VAL:HG22	1:A:204:GLN:HB3	2.00	0.42
1:A:247:ARG:HG3	1:A:268:ILE:HD13	2.02	0.42
1:A:301:ILE:CG2	1:A:333:TYR:CE1	3.02	0.42
1:A:107:LEU:HA	1:A:107:LEU:HD12	1.81	0.41
1:A:267:THR:O	1:A:268:ILE:C	2.58	0.41
1:A:304:SER:HB3	1:A:332:VAL:HG11	2.02	0.41

All (8) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:177:ASP:OD2	1:A:220:TYR:CE2[7_555]	1.19	1.01
1:A:177:ASP:OD2	1:A:220:TYR:CZ[7_555]	1.29	0.91
1:A:177:ASP:OD2	1:A:220:TYR:OH[7_555]	1.63	0.57
1:A:177:ASP:OD2	1:A:220:TYR:HE2[7_555]	1.26	0.34
1:A:122:GLU:O	1:A:192:THR:HG23[3_555]	1.32	0.28
1:A:177:ASP:CG	1:A:220:TYR:CE2[7_555]	2.09	0.11
1:A:177:ASP:CG	1:A:220:TYR:HE2[7_555]	1.53	0.07
1:A:122:GLU:OE1	1:A:192:THR:HG22[3_555]	1.55	0.05

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	266/305 (87%)	257 (97%)	9 (3%)	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	230/261 (88%)	206 (90%)	24 (10%)	5 18

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	THR
1	A	58	LYS
1	A	62	ILE
1	A	122	GLU
1	A	157	MET
1	A	174	ASN
1	A	176	PHE
1	A	177	ASP
1	A	185	ASN
1	A	199	LYS
1	A	200	VAL
1	A	212	ARG
1	A	222	ARG
1	A	241	ASP
1	A	271	LEU
1	A	272	LEU
1	A	289	GLU
1	A	292	THR
1	A	308	LEU
1	A	315	ARG
1	A	318	LEU
1	A	320	LEU
1	A	323	VAL
1	A	327	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	231	GLN
1	A	266	ASN
1	A	300	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

3 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	SO4	A	401	-	4,4,4	0.57	0	6,6,6	0.37	0
2	SO4	A	403	-	4,4,4	1.05	0	6,6,6	0.48	0
2	SO4	A	402	-	4,4,4	0.37	0	6,6,6	0.73	0

There are no bond length outliers.

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 (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, 95^{th} 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	270/305 (88%)	1.44	62 (22%)	2 3	56, 79, 101, 116	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	213	TYR	7.8
1	A	214	ASP	6.9
1	A	333	TYR	6.8
1	A	177	ASP	6.7
1	A	301	ILE	5.9
1	A	216	PRO	5.9
1	A	290	ASP	5.7
1	A	219	ASP	5.4
1	A	270	GLY	4.8
1	A	220	TYR	4.7
1	A	218	GLY	4.6
1	A	289	GLU	4.5
1	A	334	GLU	4.3
1	A	69	THR	4.3
1	A	264	SER	4.0
1	A	122	GLU	4.0
1	A	266	ASN	3.8
1	A	57	THR	3.7
1	A	217	GLU	3.7
1	A	287	LYS	3.5
1	A	288	GLY	3.4
1	A	215	ASP	3.4
1	A	267	THR	3.3
1	A	78	TRP	3.2
1	A	79	GLU	3.2
1	A	58	LYS	3.2
1	A	184	ASP	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	262	GLU	3.2
1	A	212	ARG	3.1
1	A	188	GLU	3.1
1	A	317	GLU	3.1
1	A	176	PHE	3.1
1	A	324	THR	3.1
1	A	147	ILE	3.0
1	A	323	VAL	2.9
1	A	260	ASN	2.9
1	A	278	LEU	2.8
1	A	297	GLY	2.7
1	A	332	VAL	2.7
1	A	321	ARG	2.6
1	A	146	ASN	2.6
1	A	302	VAL	2.6
1	A	192	THR	2.5
1	A	178	PHE	2.5
1	A	271	LEU	2.5
1	A	64	LEU	2.5
1	A	261	ILE	2.5
1	A	274	TYR	2.5
1	A	242	SER	2.4
1	A	60	PHE	2.4
1	A	265	SER	2.3
1	A	135	GLN	2.3
1	A	292	THR	2.3
1	A	151	ASN	2.2
1	A	258	GLN	2.2
1	A	275	ARG	2.2
1	A	331	THR	2.1
1	A	191	ALA	2.1
1	A	304	SER	2.1
1	A	284	TYR	2.1
1	A	246	TYR	2.0
1	A	157	MET	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 monosaccharides in this entry.

6.4 Ligands (i)

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, 95^{th} 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(Å^2)	Q<0.9
2	SO4	A	401	5/5	0.79	0.22	87,88,90,90	0
2	SO4	A	402	5/5	0.84	0.13	120,120,124,127	0
2	SO4	A	403	5/5	0.84	0.24	79,81,82,84	0

6.5 Other polymers (i)

There are no such residues in this entry.