wwPDB X-ray Structure Validation Summary Report (i)

Oct 18, 2023 – 01:49 PM EDT

PDB ID	:	1IZ1
Title	:	CRYSTAL STRUCTURE OF CBNR, A LYSR FAMILY TRANSCRIPTIONAL REGULATOR
Authors	:	Muraoka, S.; Okumura, R.; Ogawa, N.; Miyashita, K.; Senda, T.
Deposited on	:	2002-09-18
Resolution	:	2.50 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 (1)) were used in the production of this report:

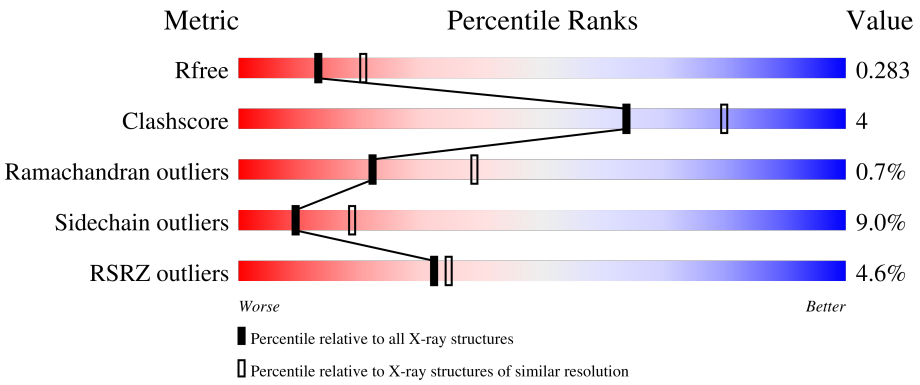
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	294	84% 14% •••
1	B	294	3% 82% 17% •
1	P	294	3% 84% 13% • •
1	Q	294	12% 78% 18% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9097 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 LysR-type regulatory protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	292	2241	1425	415	394	7	0	0	0
1	B	294	2259	1435	418	399	7	0	0	0
1	P	292	2241	1425	415	394	7	0	0	0
1	Q	292	2241	1425	415	394	7	0	0	0

- Molecule 2 is water.

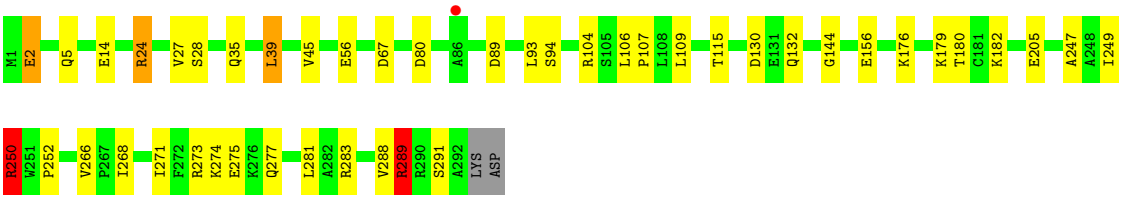
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	34	34	34	0	0
2	B	30	30	30	0	0
2	P	27	27	27	0	0
2	Q	24	24	24	0	0

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.

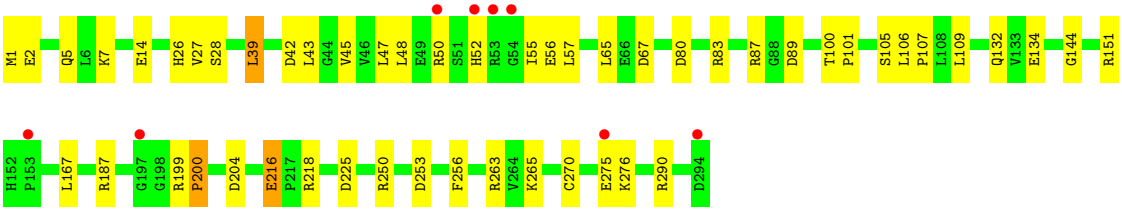
• Molecule 1: LysR-type regulatory protein

Chain A:



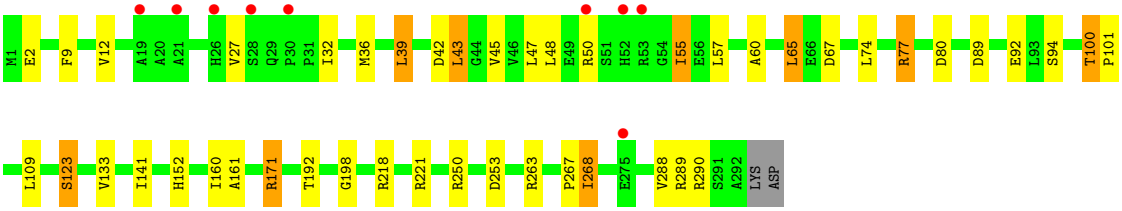
• Molecule 1: LysR-type regulatory protein

Chain B:



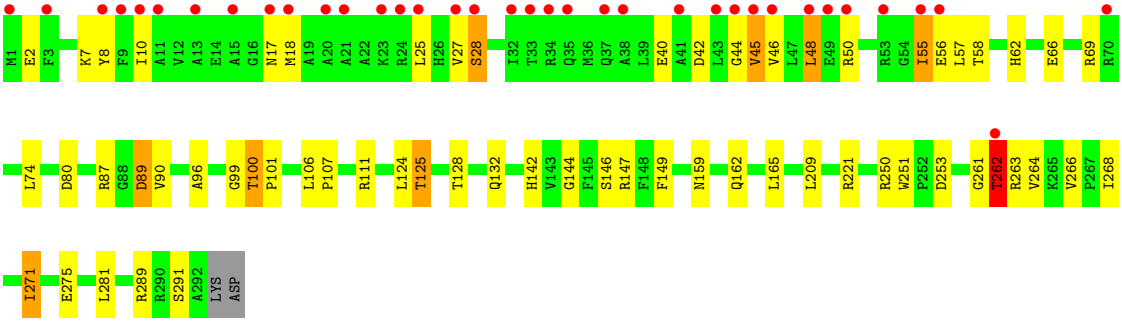
• Molecule 1: LysR-type regulatory protein

Chain P:



• Molecule 1: LysR-type regulatory protein

Chain Q:


M1
E2
F3
K7
Y8
F9
I10
A11
V12
A13
E14
A15
G16
N17
M18
A19
A20
A21
A22
K23
R24
L25
H26
V27
S28
I32
T33
R34
Q35
M36
Q37
A38
L39
E40
A41
D42
L43
G44
V45
V46
L47
L48
E49
R50
R53
G54
I55
E56
L57
T58
H62
E66
R69
R70
L74
D80
R87
G88
D89
V90
A96
G99
T100
P101
L106
P107
R111
L124
T125
T128
Q132
H142
V143
G144
F145
S146
R147
F148
F149
N159
Q162
L165
L209
R221
R250
W251
P252
D253
G261
T262
R263
V264
K265
V266
P267
I268
I271
E275
L281
R289
R290
S291
A292
LYS
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.27Å 124.48Å 166.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.50 43.49 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.8 (50.00-2.50) 97.8 (43.49-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.1.19	Depositor
R, R_{free}	0.222 , 0.286 0.224 , 0.283	Depositor DCC
R_{free} test set	2374 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	32.6	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 34.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	9097	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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)

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	0.65	0/2285	0.87	5/3092 (0.2%)
1	B	0.64	1/2303 (0.0%)	0.81	6/3114 (0.2%)
1	P	0.65	0/2285	0.82	7/3092 (0.2%)
1	Q	0.63	0/2285	0.82	4/3092 (0.1%)
All	All	0.64	1/9158 (0.0%)	0.83	22/12390 (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	P	0	1
1	Q	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	216	GLU	CD-OE2	5.70	1.31	1.25

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	P	80	ASP	CB-CG-OD2	7.13	124.71	118.30
1	A	67	ASP	CB-CG-OD2	6.92	124.53	118.30
1	B	67	ASP	CB-CG-OD2	6.73	124.35	118.30
1	B	80	ASP	CB-CG-OD2	6.72	124.35	118.30
1	Q	89	ASP	CB-CG-OD2	6.61	124.25	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	P	198	GLY	Peptide
1	Q	261	GLY	Peptide

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	2241	0	2313	15	0
1	B	2259	0	2330	14	0
1	P	2241	0	2313	21	0
1	Q	2241	0	2313	27	0
2	A	34	0	0	0	0
2	B	30	0	0	2	0
2	P	27	0	0	0	0
2	Q	24	0	0	0	0
All	All	9097	0	9269	76	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 4.

The worst 5 of 76 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:Q:262:THR:O	1:Q:264:VAL:HG23	1.70	0.90
1:Q:149:PHE:CE1	1:Q:271:ILE:HD13	2.23	0.73
1:Q:159:ASN:HD21	1:Q:162:GLN:HE21	1.39	0.71
1:A:247:ALA:O	1:A:250:ARG:NH2	2.27	0.68
1:B:83:ARG:NH1	2:B:301:HOH:O	2.28	0.66

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	290/294 (99%)	278 (96%)	10 (3%)	2 (1%)	22	39
1	B	292/294 (99%)	275 (94%)	16 (6%)	1 (0%)	41	61
1	P	290/294 (99%)	273 (94%)	15 (5%)	2 (1%)	22	39
1	Q	290/294 (99%)	268 (92%)	19 (7%)	3 (1%)	15	28
All	All	1162/1176 (99%)	1094 (94%)	60 (5%)	8 (1%)	22	39

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	Q	28	SER
1	A	104	ARG
1	A	289	ARG
1	P	55	ILE
1	Q	100	THR

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	229/231 (99%)	207 (90%)	22 (10%)	8	16
1	B	231/231 (100%)	208 (90%)	23 (10%)	7	15
1	P	229/231 (99%)	214 (93%)	15 (7%)	16	32
1	Q	229/231 (99%)	206 (90%)	23 (10%)	7	15
All	All	918/924 (99%)	835 (91%)	83 (9%)	9	19

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	268	ILE
1	Q	111	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	P	290	ARG
1	Q	48	LEU
1	Q	221	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	Q	37	GLN
1	Q	162	GLN
1	Q	277	GLN
1	Q	173	GLN
1	B	37	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)

There are no ligands in this entry.

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)

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	292/294 (99%)	-0.26	1 (0%)	94	94	13, 25, 43, 52	0
1	B	294/294 (100%)	-0.02	8 (2%)	54	58	14, 28, 48, 72	0
1	P	292/294 (99%)	0.04	9 (3%)	49	52	11, 29, 66, 74	0
1	Q	292/294 (99%)	0.24	36 (12%)	4	3	14, 29, 78, 83	0
All	All	1170/1176 (99%)	0.00	54 (4%)	32	34	11, 27, 67, 83	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	Q	50	ARG	4.8
1	Q	55	ILE	4.6
1	P	52	HIS	4.1
1	B	53	ARG	4.1
1	Q	23	LYS	4.1

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.