



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 08:08 PM GMT

PDB ID : 1IXC  
Title : Crystal structure of CbnR, a LysR family transcriptional regulator  
Authors : Muraoka, S.; Okumura, R.; Ogawa, N.; Miyashita, K.; Senda, T.  
Deposited on : 2002-06-18  
Resolution : 2.20 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

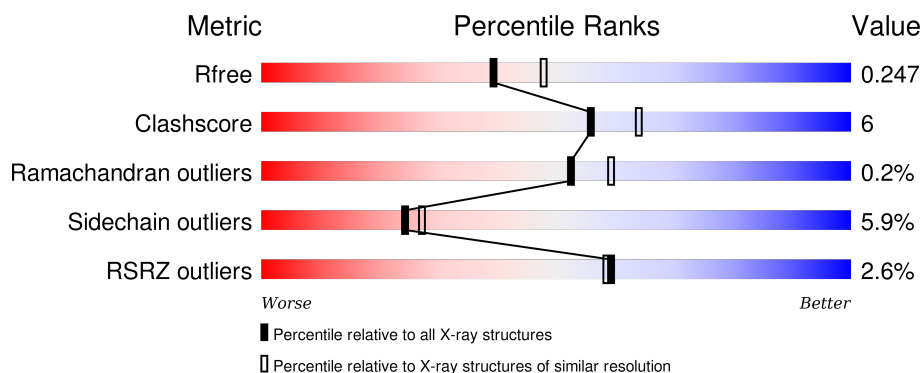
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	 2% 84% 12% • •
1	B	294	 3% 81% 16% • •

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4564 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
1	A	289	Total	C	N	O	S	Se	0	0	0
			2217	1412	405	393	2	5			
1	B	289	Total	C	N	O	S	Se	0	0	0
			2219	1413	409	390	2	5			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
A	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
A	36	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
A	127	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
A	234	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
B	18	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
B	36	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
B	127	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7
B	234	MSE	MET	MODIFIED RESIDUE	UNP Q9WXC7

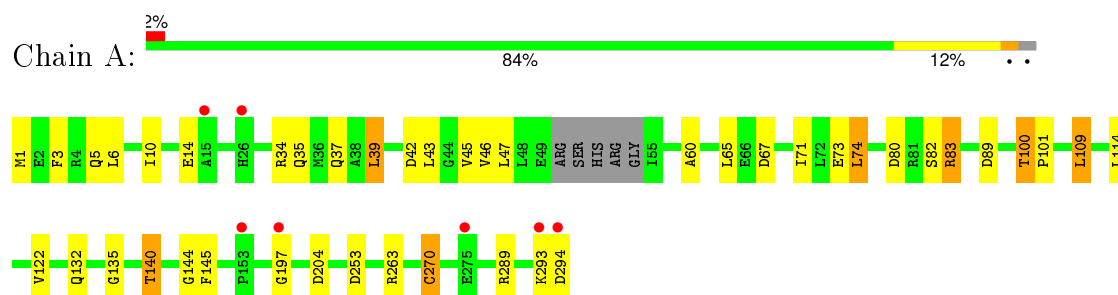
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	88	Total	O	0	0
			88	88		
2	B	40	Total	O	0	0
			40	40		

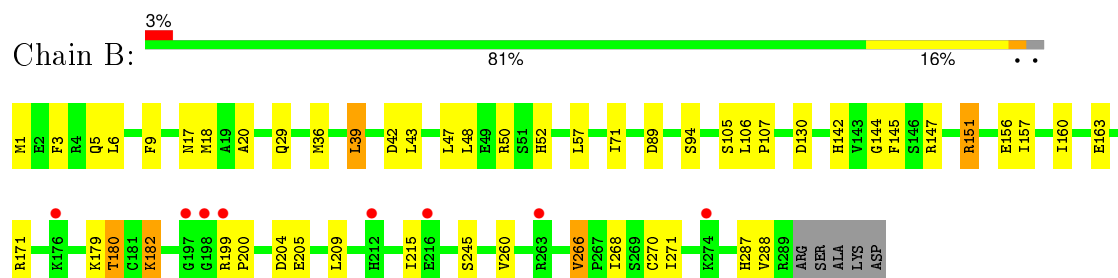
### 3 Residue-property plots [i](#)

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.

#### • Molecule 1: LysR-type regulatory protein



#### • Molecule 1: LysR-type regulatory protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.22Å 100.86Å 87.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.71 – 2.20 35.61 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (87.71-2.20) 100.0 (35.61-2.20)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.54 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5	Depositor
R, $R_{free}$	0.218 , 0.245 0.219 , 0.247	Depositor DCC
$R_{free}$ test set	1676 reflections (5.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.4	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 45.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 32756 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4564	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.47% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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	0.51	0/2255	0.80	6/3043 (0.2%)
1	B	0.48	0/2259	0.74	4/3051 (0.1%)
All	All	0.49	0/4514	0.77	10/6094 (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	A	0	1

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	42	ASP	CB-CG-OD2	7.06	124.65	118.30
1	A	89	ASP	CB-CG-OD2	5.97	123.67	118.30
1	B	89	ASP	CB-CG-OD2	5.94	123.65	118.30
1	A	197	GLY	N-CA-C	5.87	127.77	113.10
1	A	42	ASP	CB-CG-OD2	5.73	123.46	118.30
1	A	67	ASP	CB-CG-OD2	5.55	123.30	118.30
1	B	204	ASP	CB-CG-OD2	5.45	123.21	118.30
1	A	80	ASP	CB-CG-OD2	5.37	123.14	118.30
1	B	130	ASP	CB-CG-OD2	5.37	123.13	118.30
1	A	253	ASP	CB-CG-OD2	5.05	122.84	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	35	GLN	Mainchain

## 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	2217	0	2288	28	0
1	B	2219	0	2290	33	0
2	A	88	0	0	1	2
2	B	40	0	0	0	0
All	All	4564	0	4578	56	2

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 (56) 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:1:MSE:HE1	1:A:71:ILE:HD13	1.50	0.94
1:B:163:GLU:OE1	1:B:268:ILE:HD11	1.82	0.80
1:B:17:ASN:HD22	1:B:20:ALA:H	1.38	0.72
1:A:1:MSE:HE3	1:A:6:LEU:HD21	1.74	0.68
1:B:1:MSE:HE3	1:B:6:LEU:HD21	1.76	0.66
1:B:1:MSE:HE3	1:B:6:LEU:HD11	1.78	0.64
1:B:147:ARG:NH1	1:B:268:ILE:HG12	2.13	0.64
1:A:1:MSE:HE1	1:A:71:ILE:CD1	2.26	0.62
1:B:147:ARG:HH12	1:B:268:ILE:HG12	1.66	0.61
1:B:9:PHE:N	1:B:36:MSE:HE1	2.15	0.61
1:B:147:ARG:HD3	1:B:266:VAL:HG13	1.82	0.61
1:B:94:SER:H	1:B:142:HIS:CD2	2.19	0.60
1:B:17:ASN:ND2	1:B:20:ALA:H	2.00	0.60
1:B:151:ARG:HB3	1:B:151:ARG:CZ	2.30	0.59
1:B:1:MSE:HE1	1:B:71:ILE:HD13	1.85	0.59
1:A:5:GLN:HB3	1:A:39:LEU:HG	1.83	0.59
1:B:94:SER:H	1:B:142:HIS:HD2	1.51	0.59
1:B:1:MSE:CE	1:B:6:LEU:HD21	2.35	0.55
1:A:3:PHE:CZ	1:B:1:MSE:HE2	2.42	0.54
1:B:106:LEU:HB3	1:B:107:PRO:HD3	1.90	0.54
1:A:74:LEU:HD13	1:B:71:ILE:HG13	1.91	0.53
1:A:45:VAL:HG11	1:A:60:ALA:HB1	1.88	0.53
1:A:204:ASP:HA	2:A:347:HOH:O	2.06	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:135:GLY:HA2	1:A:140:THR:CG2	2.39	0.52
1:A:293:LYS:O	1:A:294:ASP:CB	2.58	0.52
1:A:132:GLN:NE2	1:A:144:GLY:HA3	2.25	0.52
1:B:5:GLN:HB3	1:B:39:LEU:HG	1.91	0.51
1:A:1:MSE:HE2	1:B:3:PHE:CZ	2.46	0.51
1:A:45:VAL:HG11	1:A:60:ALA:CB	2.41	0.51
1:A:293:LYS:O	1:A:294:ASP:HB2	2.11	0.50
1:B:160:ILE:HG22	1:B:288:VAL:HG13	1.93	0.50
1:A:135:GLY:HA2	1:A:140:THR:HG23	1.94	0.49
1:B:182:LYS:NZ	1:B:215:ILE:HD11	2.28	0.49
1:A:45:VAL:HG12	1:A:46:VAL:N	2.28	0.48
1:A:1:MSE:HE3	1:A:6:LEU:CD2	2.42	0.48
1:A:1:MSE:CE	1:A:6:LEU:HD21	2.43	0.48
1:B:180:THR:HG23	1:B:260:VAL:CG2	2.43	0.47
1:B:199:ARG:O	1:B:200:PRO:C	2.53	0.47
1:A:74:LEU:HD13	1:B:71:ILE:CG1	2.46	0.46
1:B:205:GLU:O	1:B:209:LEU:HG	2.15	0.46
1:B:180:THR:HG23	1:B:260:VAL:HG23	1.96	0.46
1:A:100:THR:N	1:A:101:PRO:CD	2.79	0.46
1:A:109:LEU:C	1:A:109:LEU:HD23	2.36	0.45
1:A:39:LEU:HD13	1:A:47:LEU:CD1	2.46	0.45
1:B:144:GLY:O	1:B:270:CYS:HA	2.17	0.44
1:A:83:ARG:HD3	1:B:43:LEU:O	2.18	0.43
1:A:145:PHE:CE2	1:A:270:CYS:HB3	2.54	0.43
1:A:1:MSE:CE	1:A:71:ILE:HD13	2.35	0.42
1:B:145:PHE:CE2	1:B:270:CYS:HB3	2.54	0.42
1:B:105:SER:CB	1:B:268:ILE:CD1	2.98	0.42
1:B:157:ILE:HG12	1:B:271:ILE:HG22	2.02	0.42
1:A:10:ILE:O	1:A:14:GLU:HG3	2.20	0.41
1:A:114:LEU:HD21	1:A:122:VAL:HG23	2.03	0.41
1:A:39:LEU:HD22	1:A:43:LEU:CD1	2.51	0.41
1:B:48:LEU:HD23	1:B:57:LEU:CD2	2.51	0.41
1:B:39:LEU:HD13	1:B:47:LEU:CD1	2.51	0.40

All (2) 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)
2:A:347:HOH:O	2:A:347:HOH:O[2_555]	1.33	0.87
2:A:356:HOH:O	2:A:356:HOH:O[2_555]	1.91	0.29



## 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	285/294 (97%)	276 (97%)	8 (3%)	1 (0%)	39	42
1	B	287/294 (98%)	275 (96%)	12 (4%)	0	100	100
All	All	572/588 (97%)	551 (96%)	20 (4%)	1 (0%)	52	59

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	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	227/226 (100%)	214 (94%)	13 (6%)	25	29
1	B	227/226 (100%)	213 (94%)	14 (6%)	23	25
All	All	454/452 (100%)	427 (94%)	27 (6%)	24	27

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	34	ARG
1	A	37	GLN
1	A	39	LEU
1	A	65	LEU
1	A	73	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	74	LEU
1	A	82	SER
1	A	83	ARG
1	A	109	LEU
1	A	140	THR
1	A	263	ARG
1	A	270	CYS
1	A	289	ARG
1	B	18	MSE
1	B	29	GLN
1	B	39	LEU
1	B	50	ARG
1	B	52	HIS
1	B	151	ARG
1	B	156	GLU
1	B	171	ARG
1	B	179	LYS
1	B	180	THR
1	B	182	LYS
1	B	245	SER
1	B	266	VAL
1	B	287	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	132	GLN
1	A	277	GLN
1	B	5	GLN
1	B	17	ASN
1	B	29	GLN
1	B	132	GLN
1	B	142	HIS
1	B	170	HIS
1	B	287	HIS

### 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	284/294 (96%)	0.18	7 (2%) 61 60	2, 12, 30, 66	0
1	B	284/294 (96%)	0.18	8 (2%) 56 55	7, 19, 28, 39	0
All	All	568/588 (96%)	0.18	15 (2%) 59 58	2, 16, 29, 66	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	198	GLY	7.3
1	A	294	ASP	5.3
1	A	293	LYS	3.7
1	B	216	GLU	3.5
1	A	197	GLY	3.0
1	B	197	GLY	2.9
1	B	212	HIS	2.8
1	A	275	GLU	2.6
1	A	15	ALA	2.6
1	B	199	ARG	2.5
1	B	176	LYS	2.5
1	B	263	ARG	2.3
1	B	274	LYS	2.3
1	A	26	HIS	2.2
1	A	153	PRO	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 carbohydrates 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.