



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2021 – 03:41 PM EDT

PDB ID : 2020  
Title : Crystal structure of transcription regulator CcpA of *Lactococcus lactis*  
Authors : Loll, B.; Kowalczyk, M.; Alings, C.; Chieduch, A.; Bardowski, J.; Saenger, W.; Biesiadka, J.  
Deposited on : 2006-11-29  
Resolution : 1.90 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](mailto: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 ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
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) : 2.23.2

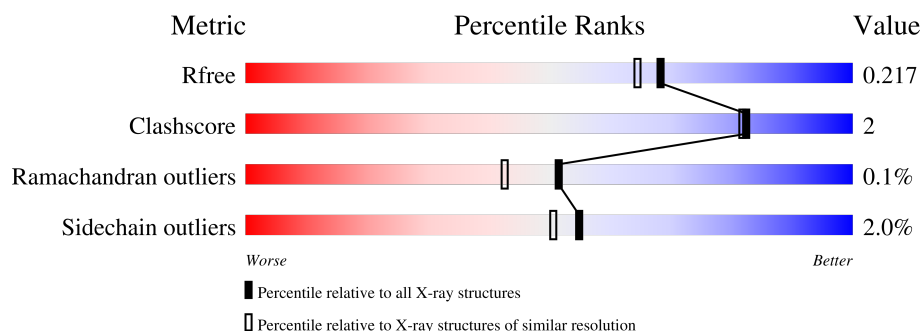
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)

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  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	332	
1	B	332	
1	C	332	
1	D	332	
1	E	332	
1	F	332	
1	G	332	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	H	332	 A horizontal bar chart showing the quality of chain H. The bar is divided into three segments: a green segment representing 78%, a yellow segment representing 1%, and a grey segment representing 18%. A small black dot is located at the end of the yellow segment.

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 18486 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 Catabolite control protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	271	Total	C	N	O	S	0	6	0
			2135	1348	348	430	9			
1	C	270	Total	C	N	O	S	0	2	0
			2106	1327	342	428	9			
1	D	270	Total	C	N	O	S	0	4	0
			2120	1335	347	429	9			
1	B	275	Total	C	N	O	S	0	11	0
			2176	1370	353	444	9			
1	E	271	Total	C	N	O	S	0	11	0
			2151	1358	349	435	9			
1	F	272	Total	C	N	O	S	0	7	0
			2154	1357	356	432	9			
1	G	271	Total	C	N	O	S	0	2	0
			2116	1334	346	427	9			
1	H	272	Total	C	N	O	S	0	6	0
			2152	1355	358	430	9			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	89	ALA	ASP	engineered mutation	UNP Q9CF33
C	89	ALA	ASP	engineered mutation	UNP Q9CF33
D	89	ALA	ASP	engineered mutation	UNP Q9CF33
B	89	ALA	ASP	engineered mutation	UNP Q9CF33
E	89	ALA	ASP	engineered mutation	UNP Q9CF33
F	89	ALA	ASP	engineered mutation	UNP Q9CF33
G	89	ALA	ASP	engineered mutation	UNP Q9CF33
H	89	ALA	ASP	engineered mutation	UNP Q9CF33

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		
2	G	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total	Cl	0	0
			1	1		

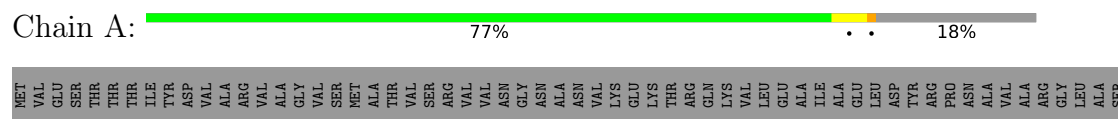
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	244	Total 244	O 244	0	0
4	C	17	Total 17	O 17	0	0
4	D	92	Total 92	O 92	0	0
4	B	217	Total 217	O 217	0	0
4	E	297	Total 297	O 297	0	0
4	F	255	Total 255	O 255	0	0
4	G	137	Total 137	O 137	0	0
4	H	76	Total 76	O 76	0	0

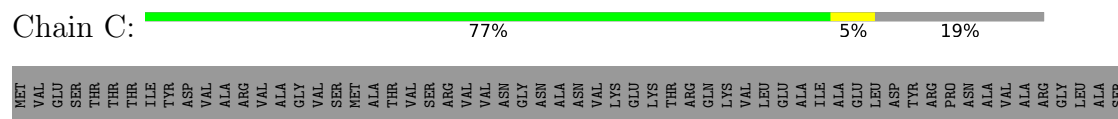
### 3 Residue-property plots [i](#)

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. 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. 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.

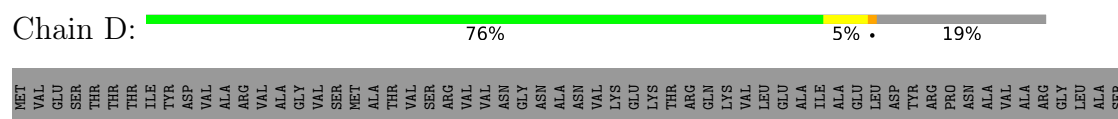
- Molecule 1: Catabolite control protein A



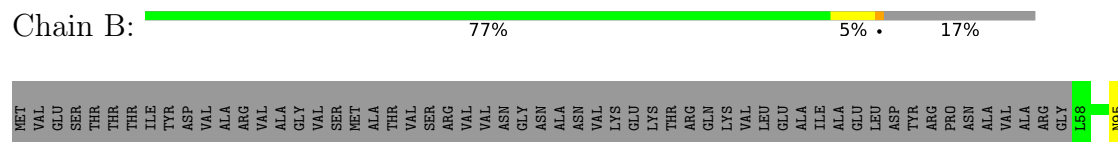
- Molecule 1: Catabolite control protein A



- Molecule 1: Catabolite control protein A



- Molecule 1: Catabolite control protein A



- Molecule 1: Catabolite control protein A





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.76Å 74.27Å 160.30Å 90.00° 102.36° 90.00°	Depositor
Resolution (Å)	19.90 – 1.90 19.87 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.4 (19.90-1.90) 99.5 (19.87-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.89 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.176 , 0.211 0.185 , 0.217	Depositor DCC
$R_{free}$ test set	10565 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.2	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 55.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	18486	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, 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	0.68	0/2188	0.77	3/2960 (0.1%)
1	B	0.62	0/2248	0.73	2/3040 (0.1%)
1	C	0.39	0/2142	0.54	0/2898
1	D	0.46	0/2164	0.61	0/2927
1	E	0.77	2/2223 (0.1%)	0.83	4/3005 (0.1%)
1	F	0.67	0/2210	0.75	3/2985 (0.1%)
1	G	0.58	0/2152	0.63	0/2911
1	H	0.48	0/2204	0.59	0/2976
All	All	0.60	2/17531 (0.0%)	0.69	12/23702 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	226	GLU	CD-OE1	5.18	1.31	1.25
1	E	226	GLU	CG-CD	5.17	1.59	1.51

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	83	ARG	NE-CZ-NH1	10.68	125.64	120.30
1	E	234	ARG	NE-CZ-NH2	-9.72	115.44	120.30
1	F	234	ARG	NE-CZ-NH2	-9.24	115.68	120.30
1	A	83	ARG	NE-CZ-NH2	-8.35	116.12	120.30
1	B	238	ARG	NE-CZ-NH1	7.85	124.23	120.30

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	2135	0	2148	10	0
1	B	2176	0	2178	16	0
1	C	2106	0	2116	4	0
1	D	2120	0	2130	12	0
1	E	2151	0	2164	23	0
1	F	2154	0	2172	9	0
1	G	2116	0	2134	7	0
1	H	2152	0	2178	9	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
2	E	5	0	0	0	0
2	F	10	0	0	0	0
2	G	5	0	0	0	0
3	E	1	0	0	0	0
4	A	244	0	0	2	0
4	B	217	0	0	3	0
4	C	17	0	0	0	0
4	D	92	0	0	4	0
4	E	297	0	0	9	0
4	F	255	0	0	1	0
4	G	137	0	0	1	0
4	H	76	0	0	0	0
All	All	18486	0	17220	85	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 2.

The worst 5 of 85 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:B:226[A]:GLU:OE1	4:B:2079:HOH:O	1.65	1.14
1:D:119[A]:GLN:CD	4:D:415:HOH:O	1.90	1.08
1:E:303[A]:ARG:NE	4:E:2302:HOH:O	1.88	1.06
1:E:86[A]:ASP:OD1	4:E:2295:HOH:O	1.73	1.05
1:E:303[A]:ARG:CZ	4:E:2302:HOH:O	2.05	1.04

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	275/332 (83%)	271 (98%)	3 (1%)	1 (0%)	34	24
1	B	284/332 (86%)	277 (98%)	6 (2%)	1 (0%)	34	24
1	C	270/332 (81%)	263 (97%)	7 (3%)	0	100	100
1	D	272/332 (82%)	265 (97%)	7 (3%)	0	100	100
1	E	280/332 (84%)	276 (99%)	4 (1%)	0	100	100
1	F	277/332 (83%)	272 (98%)	4 (1%)	1 (0%)	34	24
1	G	271/332 (82%)	267 (98%)	4 (2%)	0	100	100
1	H	276/332 (83%)	271 (98%)	5 (2%)	0	100	100
All	All	2205/2656 (83%)	2162 (98%)	40 (2%)	3 (0%)	51	42

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	315	GLU
1	A	315	GLU
1	B	315	GLU

### 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	243/286 (85%)	240 (99%)	3 (1%)	71	70
1	B	250/286 (87%)	244 (98%)	6 (2%)	49	43
1	C	238/286 (83%)	229 (96%)	9 (4%)	33	24
1	D	240/286 (84%)	234 (98%)	6 (2%)	47	41
1	E	247/286 (86%)	242 (98%)	5 (2%)	55	51
1	F	244/286 (85%)	240 (98%)	4 (2%)	62	60
1	G	239/286 (84%)	235 (98%)	4 (2%)	60	57
1	H	244/286 (85%)	241 (99%)	3 (1%)	71	70
All	All	1945/2288 (85%)	1905 (98%)	40 (2%)	55	48

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	234	ARG
1	G	237	GLU
1	F	165	HIS
1	F	313	ASP
1	H	185	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	210	ASN
1	F	103	ASN
1	H	247	HIS
1	F	95	ASN
1	F	165	HIS

### 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 [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 for Mogul analysis.

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	E	2004	-	4,4,4	0.24	0	6,6,6	0.24	0
2	SO4	A	2007	-	4,4,4	0.19	0	6,6,6	0.30	0
2	SO4	G	2008	-	4,4,4	0.18	0	6,6,6	0.23	0
2	SO4	A	2006	-	4,4,4	0.18	0	6,6,6	0.10	0
2	SO4	B	2003	-	4,4,4	0.17	0	6,6,6	0.14	0
2	SO4	B	2001	-	4,4,4	0.24	0	6,6,6	0.71	0
2	SO4	F	2005	-	4,4,4	0.12	0	6,6,6	0.19	0
2	SO4	F	2002	-	4,4,4	0.11	0	6,6,6	0.49	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.