



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:04 AM GMT

PDB ID : 2P90  
Title : The crystal structure of a protein of unknown function from *Corynebacterium glutamicum* ATCC 13032  
Authors : Zhang, R.; Duggan, E.; Gu, M.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2007-03-23  
Resolution : 2.35 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  
<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

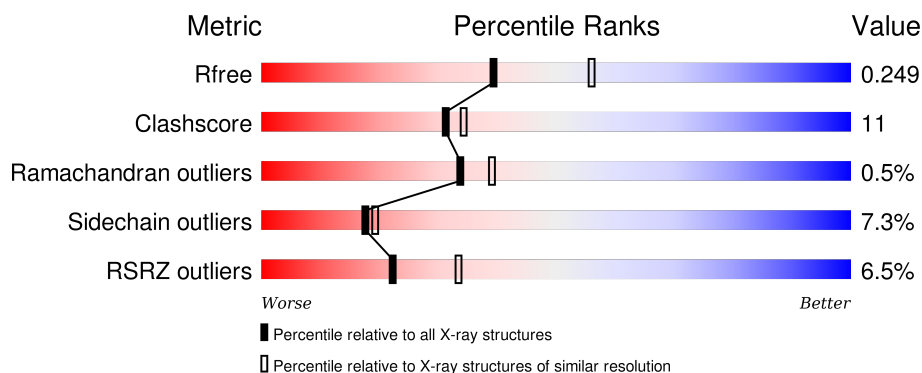
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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	1352 (2.38-2.34)
Clashscore	102246	1456 (2.38-2.34)
Ramachandran outliers	100387	1435 (2.38-2.34)
Sidechain outliers	100360	1436 (2.38-2.34)
RSRZ outliers	91569	1358 (2.38-2.34)

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	319	<div> <div>9%</div> <div>68%13%•16%</div> </div>
1	B	319	<div> <div>3%</div> <div>65%15%5%15%</div> </div>
1	C	319	<div> <div>4%</div> <div>64%15%•17%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6595 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 Hypothetical protein Cgl1923.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	269	Total	C	N	O	S	0	0	0
			2092	1303	364	416	9			
1	B	271	Total	C	N	O	S	0	0	0
			2109	1314	369	417	9			
1	C	266	Total	C	N	O	S	0	0	0
			2068	1289	361	409	9			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	125	Total	O	0	0
			125	125		
2	B	107	Total	O	0	0
			107	107		
2	C	94	Total	O	0	0
			94	94		



ALA

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.83Å 113.35Å 117.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	81.65 – 2.35 44.40 – 2.34	Depositor EDS
% Data completeness (in resolution range)	99.8 (81.65-2.35) 99.4 (44.40-2.34)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.193 , 0.251 0.192 , 0.249	Depositor DCC
$R_{free}$ test set	2420 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	44.8	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.8	EDS
Estimated twinning fraction	0.013 for -h,l,k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 47914 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6595	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% 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	0.78	2/2132 (0.1%)	0.86	6/2900 (0.2%)
1	B	0.80	4/2151 (0.2%)	0.84	5/2927 (0.2%)
1	C	0.79	0/2107	0.95	10/2864 (0.3%)
All	All	0.79	6/6390 (0.1%)	0.89	21/8691 (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

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	43	GLU	CD-OE1	7.27	1.33	1.25
1	A	43	GLU	CD-OE1	6.94	1.33	1.25
1	B	43	GLU	CG-CD	6.46	1.61	1.51
1	B	269	GLU	CG-CD	6.21	1.61	1.51
1	A	43	GLU	CG-CD	6.08	1.61	1.51

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	72	ARG	NE-CZ-NH2	-13.89	113.35	120.30
1	C	73	ARG	NE-CZ-NH2	-12.49	114.06	120.30
1	C	115	ARG	NE-CZ-NH2	-12.29	114.15	120.30
1	A	115	ARG	NE-CZ-NH2	-11.02	114.79	120.30
1	C	73	ARG	NE-CZ-NH1	10.07	125.33	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	171	THR	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	2092	0	2048	45	0
1	B	2109	0	2064	53	0
1	C	2068	0	2028	51	0
2	A	125	0	0	13	0
2	B	107	0	0	14	0
2	C	94	0	0	17	0
All	All	6595	0	6140	134	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 11.

The worst 5 of 134 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:33:GLN:HG2	2:B:403:HOH:O	1.32	1.23
1:C:152:VAL:HG13	1:C:167:VAL:HG21	1.39	1.03
1:A:106:MET:SD	2:A:433:HOH:O	2.19	0.99
1:C:90:ASN:HB2	2:C:333:HOH:O	1.59	0.98
1:C:152:VAL:HG13	1:C:167:VAL:CG2	1.97	0.95

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	267/319 (84%)	257 (96%)	9 (3%)	1 (0%)	39	46
1	B	269/319 (84%)	259 (96%)	8 (3%)	2 (1%)	26	29
1	C	262/319 (82%)	251 (96%)	10 (4%)	1 (0%)	39	46
All	All	798/957 (83%)	767 (96%)	27 (3%)	4 (0%)	34	39

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	21	GLN
1	C	252	SER
1	A	268	SER
1	B	167	VAL

### 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	234/276 (85%)	217 (93%)	17 (7%)	17	19
1	B	236/276 (86%)	219 (93%)	17 (7%)	18	19
1	C	231/276 (84%)	214 (93%)	17 (7%)	17	19
All	All	701/828 (85%)	650 (93%)	51 (7%)	17	19

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	161	ASP
1	B	241	VAL
1	C	245	LEU
1	B	192	LYS
1	B	245	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	134	ASN
1	B	156	HIS
1	C	48	HIS
1	B	81	ASN
1	C	21	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 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	269/319 (84%)	0.55	28 (10%) <b>8</b> <b>14</b>	33, 44, 96, 122	0
1	B	271/319 (84%)	0.17	10 (3%) 45 59	34, 49, 71, 86	0
1	C	266/319 (83%)	0.17	14 (5%) 30 45	31, 45, 90, 102	0
All	All	806/957 (84%)	0.29	52 (6%) <b>22</b> <b>33</b>	31, 46, 89, 122	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	273	TYR	7.6
1	B	169	LEU	7.5
1	A	270	LEU	6.5
1	A	254	GLU	5.5
1	A	249	THR	5.4

### 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.