



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 8, 2023 – 12:52 am GMT

PDB ID : 2W70  
Title : Crystal structure of Biotin carboxylase from E. coli in complex with the amino-thiazole-pyrimidine fragment  
Authors : Mochalkin, I.; Miller, J.R.  
Deposited on : 2008-12-19  
Resolution : 1.77 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
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

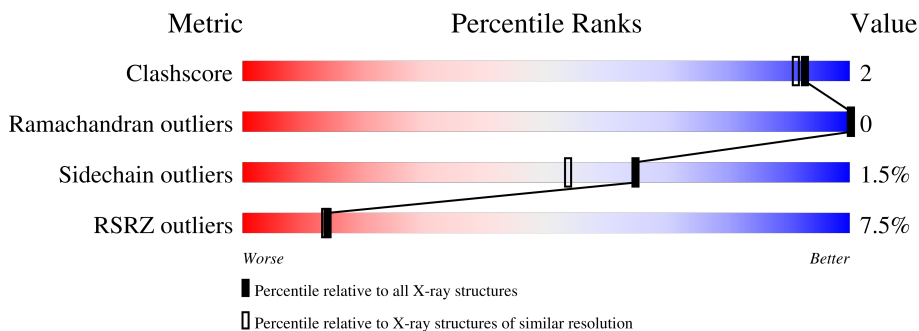
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.77 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))
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)
RSRZ outliers	127900	9032 (1.80-1.76)

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  $\leq 5\%$ . 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	449	 2% 96%
1	B	449	 13% 95%

## 2 Entry composition [i](#)

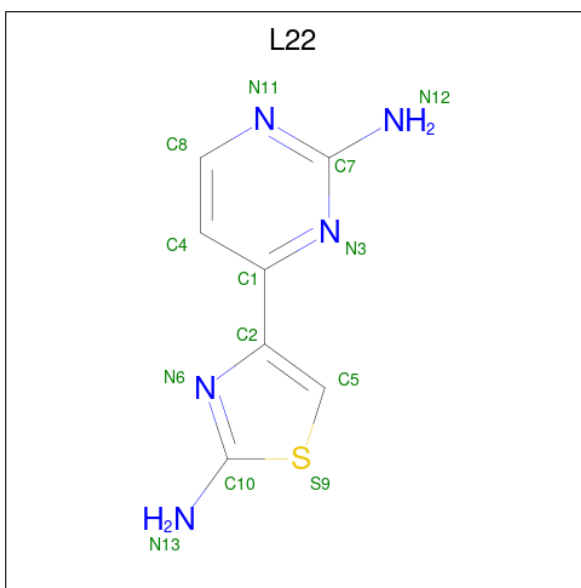
There are 4 unique types of molecules in this entry. The entry contains 7863 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 BIOTIN CARBOXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	446	Total 3444	C 2170	N 614	O 638	S 22	0	4	0
1	B	446	Total 3449	C 2172	N 617	O 638	S 22	0	4	0

- Molecule 2 is 4-(2-amino-1,3-thiazol-4-yl)pyrimidin-2-amine (three-letter code: L22) (formula: C<sub>7</sub>H<sub>7</sub>N<sub>5</sub>S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	S		
2	A	1	Total 13	C 7	N 5	S 1	0	0
2	B	1	Total 13	C 7	N 5	S 1	0	0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0

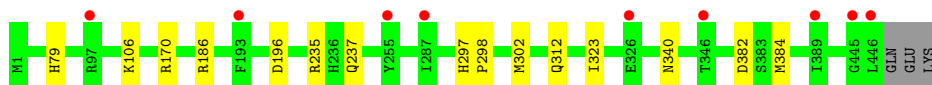
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	538	Total O 538 538	0	0
4	B	404	Total O 404 404	0	0

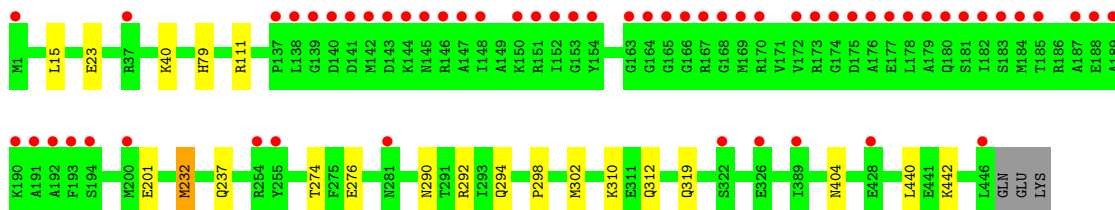
### 3 Residue-property plots [i](#)

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: BIOTIN CARBOXYLASE



- Molecule 1: BIOTIN CARBOXYLASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.23Å 106.22Å 122.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.77 29.33 – 1.77	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-1.77) 93.6 (29.33-1.77)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.32 (at 1.77Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.174 , 0.202 0.192 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.1	Xtrriage
Anisotropy	0.289	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7863	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.91% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL, L22

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	0.39	0/3527	0.57	0/4759
1	B	0.38	0/3532	0.56	0/4766
All	All	0.39	0/7059	0.57	0/9525

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	2	0

There are no bond length outliers.

There are no bond angle outliers.

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	72[A]	ILE	CB
1	A	72[B]	ILE	CB

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	3444	0	3476	5	0
1	B	3449	0	3479	16	0
2	A	13	0	7	0	0
2	B	13	0	7	0	0
3	A	1	0	0	0	0
3	B	1	0	0	1	0
4	A	538	0	0	0	0
4	B	404	0	0	3	0
All	All	7863	0	6969	21	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.

All (21) 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:232:MET:HE3	4:B:2390:HOH:O	1.81	0.80
1:B:232:MET:HE2	1:B:440:LEU:HB2	1.82	0.62
1:B:232:MET:CE	1:B:232:MET:HA	2.31	0.60
1:A:340:ASN:HD22	1:A:384:MET:HA	1.70	0.56
1:B:232:MET:HE2	1:B:232:MET:HA	1.86	0.56
1:B:292:ARG:NH1	3:B:1448:CL:CL	2.76	0.55
1:A:79:HIS:HE2	1:A:312:GLN:NE2	2.06	0.54
1:A:186:ARG:HG2	1:A:196:ASP:HB2	1.90	0.53
1:B:276:GLU:CD	1:B:290:ASN:HD21	2.14	0.51
1:B:79:HIS:HE2	1:B:312:GLN:NE2	2.09	0.50
1:B:232:MET:HE1	1:B:440:LEU:N	2.29	0.48
1:B:298:PRO:O	1:B:302:MET:HG2	2.13	0.47
1:B:319:GLN:NE2	4:B:2259:HOH:O	2.48	0.46
1:B:274:THR:HG21	1:B:294:GLN:HE22	1.81	0.46
1:B:23:GLU:OE1	1:B:310:LYS:NZ	2.40	0.46
1:B:232:MET:CE	1:B:440:LEU:HB2	2.45	0.45
1:B:15:LEU:CD2	1:B:40:LYS:HD2	2.48	0.44
1:A:298:PRO:O	1:A:302:MET:HG2	2.19	0.43
1:A:297:HIS:N	1:A:298:PRO:CD	2.83	0.41
1:B:15:LEU:HD22	1:B:40:LYS:HD2	2.02	0.41
1:B:404:ASN:ND2	4:B:2351:HOH:O	2.50	0.41

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	448/449 (100%)	440 (98%)	8 (2%)	0	100	100
1	B	448/449 (100%)	436 (97%)	12 (3%)	0	100	100
All	All	896/898 (100%)	876 (98%)	20 (2%)	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	362/361 (100%)	355 (98%)	7 (2%)	57	43
1	B	362/361 (100%)	356 (98%)	6 (2%)	60	48
All	All	724/722 (100%)	711 (98%)	13 (2%)	65	45

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	106	LYS
1	A	170	ARG
1	A	235	ARG
1	A	237	GLN
1	A	323	ILE
1	A	382[A]	ASP
1	A	382[B]	ASP
1	B	111	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	201	GLU
1	B	232	MET
1	B	237[A]	GLN
1	B	237[B]	GLN
1	B	442	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	180	GLN
1	A	312	GLN
1	A	319	GLN
1	A	340	ASN
1	A	404	ASN
1	A	431	GLN
1	B	94	GLN
1	B	206	ASN
1	B	290	ASN
1	B	294	GLN
1	B	312	GLN
1	B	319	GLN
1	B	404	ASN

### 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](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	L22	B	1447	-	10,14,14	1.06	1 (10%)	10,19,19	3.48	4 (40%)
2	L22	A	1447	-	10,14,14	1.02	0	10,19,19	4.35	5 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	L22	B	1447	-	-	0/0/4/4	0/2/2/2
2	L22	A	1447	-	-	0/0/4/4	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1447	L22	C4-C8	2.04	1.39	1.36

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1447	L22	C2-C5-S9	-11.11	109.43	112.53
2	B	1447	L22	C2-C5-S9	-8.23	110.23	112.53
2	A	1447	L22	C8-N11-C7	4.57	120.92	116.24
2	A	1447	L22	N11-C7-N3	-4.18	120.63	125.70
2	A	1447	L22	C4-C8-N11	-3.95	119.92	123.81
2	B	1447	L22	C8-N11-C7	3.92	120.25	116.24
2	B	1447	L22	C4-C8-N11	-3.77	120.09	123.81
2	B	1447	L22	N11-C7-N3	-3.73	121.17	125.70
2	A	1447	L22	N12-C7-N11	2.29	119.76	117.44

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

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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<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	446/449 (99%)	0.15	9 (2%) 65 65	16, 21, 28, 36	0
1	B	446/449 (99%)	0.71	58 (13%) 3 3	16, 21, 57, 60	0
All	All	892/898 (99%)	0.43	67 (7%) 14 13	16, 21, 46, 60	0

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	167	ARG	9.0
1	B	168	GLY	8.9
1	B	193	PHE	8.8
1	B	142	MET	8.0
1	B	166	GLY	7.3
1	B	143	ASP	6.4
1	B	180	GLN	6.4
1	B	187	ALA	6.3
1	B	150	LYS	6.2
1	B	192	ALA	6.1
1	B	154	TYR	5.9
1	B	194	SER	5.8
1	B	184	MET	5.7
1	B	146	ARG	5.6
1	B	191	ALA	5.0
1	B	177	GLU	4.8
1	B	141	ASP	4.8
1	B	151	ARG	4.7
1	B	164	GLY	4.7
1	B	152	ILE	4.6
1	B	140	ASP	4.5
1	B	181	SER	4.5
1	A	446	LEU	4.3
1	B	176	ALA	4.2

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	144	LYS	4.1
1	B	190	LYS	4.1
1	B	183	SER	4.0
1	B	188	GLU	3.8
1	B	185	THR	3.8
1	B	147	ALA	3.7
1	B	189	ALA	3.7
1	B	139	GLY	3.7
1	B	148	ILE	3.7
1	B	165	GLY	3.7
1	B	163	GLY	3.5
1	B	179	ALA	3.5
1	B	254	ARG	3.5
1	B	174	GLY	3.4
1	B	170	ARG	3.3
1	B	37	ARG	3.2
1	A	445	GLY	3.1
1	B	255	TYR	3.0
1	B	281	ASN	3.0
1	B	446	LEU	3.0
1	B	173	ARG	2.7
1	B	153	GLY	2.7
1	B	172	VAL	2.7
1	B	1	MET	2.7
1	B	178	LEU	2.7
1	B	145	ASN	2.7
1	A	326	GLU	2.6
1	B	326	GLU	2.6
1	B	200	MET	2.6
1	B	175	ASP	2.5
1	B	182	ILE	2.5
1	B	322	SER	2.5
1	A	389	ILE	2.5
1	B	137	PRO	2.4
1	A	346	THR	2.4
1	B	169	MET	2.4
1	A	193	PHE	2.3
1	A	97	ARG	2.3
1	A	287	ILE	2.2
1	B	389	ILE	2.2
1	B	428	GLU	2.2
1	A	255	TYR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	138	LEU	2.1

## 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](#)

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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	L22	B	1447	13/13	0.88	0.11	35,36,37,39	0
2	L22	A	1447	13/13	0.94	0.11	20,22,25,29	0
3	CL	B	1448	1/1	0.98	0.06	21,21,21,21	0
3	CL	A	1448	1/1	1.00	0.02	17,17,17,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.