



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 08:51 PM GMT

PDB ID : 1MEE  
Title : THE COMPLEX BETWEEN THE SUBTILISIN FROM A MESOPHILIC BACTERIUM AND THE LEECH INHIBITOR EGLIN-C  
Authors : Dauter, Z.; Betzel, C.; Wilson, K.S.  
Deposited on : 1991-04-15  
Resolution : 2.00 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

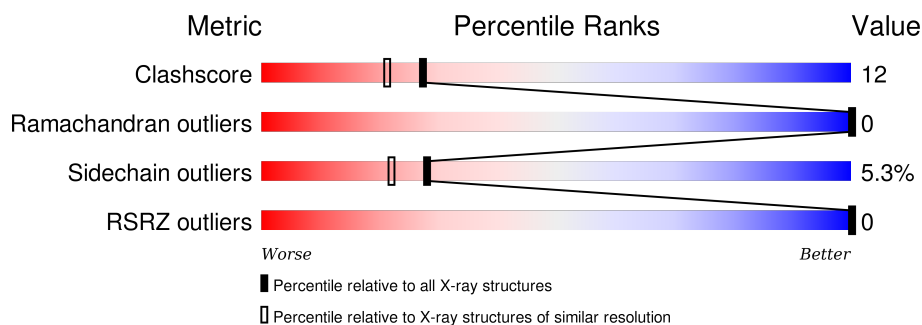
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



The reported resolution of this entry is 2.00 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	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	275	 75% 22% •
2	I	64	 66% 31% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CA	A	400	-	-	-	X
3	CA	A	401	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2792 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 MESENTERICOPEPTIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total	C	N	O	S	0	0	0
			1948	1210	335	399	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	88	ALA	SER	CONFLICT	UNP P07518
A	89	SER	ALA	CONFLICT	UNP P07518

- Molecule 2 is a protein called EGLIN C.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	I	64	Total	C	N	O	0	0	0
			530	345	91	94			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	33	ASN	ASP	CONFLICT	UNP P01051

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Ca	0	0
			2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	250	Total	O	0	0
			250	250		

*Continued on next page...*

*Continued from previous page...*

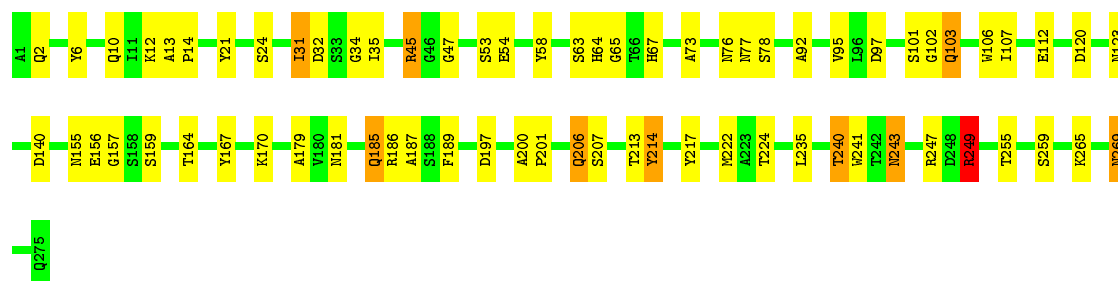
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	I	62	Total	O	0	0
			62	62		

### 3 Residue-property plots

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 ( $\text{RSRZ} > 2$ ). 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.

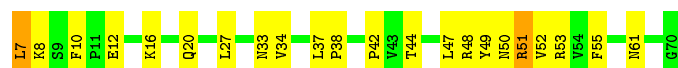
#### • Molecule 1: MESENTERICOPEPTIDASE

Chain A: 



#### • Molecule 2: EGLIN C

Chain I: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.98 Å 71.85 Å 48.34 Å 90.00° 110.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.00 21.36 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-2.00) 92.4 (21.36-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.26 (at 2.01 Å)	Xtriage
Refinement program	PROLSQ	Depositor
R, $R_{free}$	0.151 , (Not available) 0.140 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	14.3	Xtriage
Anisotropy	0.734	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 133.9	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 17300 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2792	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.72% 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

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

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	0.83	0/1987	1.63	26/2714 (1.0%)
2	I	0.85	0/548	1.61	4/749 (0.5%)
All	All	0.83	0/2535	1.62	30/3463 (0.9%)

There are no bond length outliers.

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	21	TYR	CB-CG-CD1	-11.27	114.24	121.00
1	A	45	ARG	CD-NE-CZ	10.62	138.47	123.60
2	I	51	ARG	NE-CZ-NH1	10.34	125.47	120.30
1	A	249	ARG	NE-CZ-NH2	-10.27	115.17	120.30
2	I	53	ARG	NE-CZ-NH1	-10.19	115.20	120.30
1	A	45	ARG	NE-CZ-NH1	9.69	125.14	120.30
1	A	197	ASP	CB-CG-OD1	-8.66	110.51	118.30
1	A	197	ASP	CB-CG-OD2	8.65	126.08	118.30
1	A	214	TYR	CB-CG-CD1	-7.91	116.25	121.00
1	A	247	ARG	NE-CZ-NH2	7.84	124.22	120.30
1	A	217	TYR	CA-CB-CG	7.52	127.69	113.40
2	I	53	ARG	CD-NE-CZ	-7.40	113.23	123.60
1	A	31	ILE	O-C-N	7.21	134.23	122.70
1	A	120	ASP	CB-CG-OD1	-7.07	111.93	118.30
1	A	97	ASP	CB-CG-OD1	7.06	124.66	118.30
1	A	179	ALA	N-CA-CB	-6.92	100.42	110.10
1	A	217	TYR	CB-CG-CD1	6.26	124.76	121.00
1	A	32	ASP	N-CA-CB	-6.16	99.52	110.60
1	A	249	ARG	NE-CZ-NH1	5.86	123.23	120.30
1	A	101	SER	CB-CA-C	5.73	120.99	110.10
1	A	21	TYR	CB-CG-CD2	5.71	124.43	121.00
1	A	73	ALA	N-CA-C	5.56	126.02	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	140	ASP	CB-CG-OD1	5.54	123.29	118.30
1	A	54	GLU	O-C-N	5.49	131.48	122.70
1	A	206	GLN	CA-CB-CG	5.47	125.44	113.40
1	A	73	ALA	N-CA-CB	-5.32	102.65	110.10
1	A	63	SER	N-CA-CB	5.28	118.42	110.50
1	A	45	ARG	NH1-CZ-NH2	-5.06	113.83	119.40
2	I	12	GLU	CG-CD-OE1	5.05	128.40	118.30
1	A	157	GLY	CA-C-O	5.04	129.68	120.60

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	1948	0	1901	46	1
2	I	530	0	511	15	0
3	A	2	0	0	0	0
4	A	250	0	0	11	1
4	I	62	0	0	1	0
All	All	2792	0	2412	58	1

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 12.

All (58) 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:112:GLU:OE1	4:A:677:HOH:O	2.09	0.69
2:I:7:LEU:HD21	2:I:47:LEU:CD2	2.26	0.66
1:A:269:ASN:C	1:A:269:ASN:HD22	2.01	0.64
1:A:240:THR:HB	4:A:510:HOH:O	2.00	0.62
2:I:7:LEU:HD21	2:I:47:LEU:HD22	1.83	0.60
1:A:53:SER:OG	4:A:707:HOH:O	2.17	0.59
1:A:243:ASN:N	1:A:243:ASN:HD22	2.02	0.57

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:13:ALA:N	1:A:14:PRO:CD	2.67	0.56
1:A:103:GLN:HG3	1:A:106:TRP:CE2	2.40	0.56
1:A:2:GLN:HE22	1:A:76:ASN:HD22	1.53	0.56
1:A:12:LYS:HE3	1:A:269:ASN:OD1	2.07	0.55
1:A:45:ARG:NH2	4:A:584:HOH:O	2.40	0.55
1:A:12:LYS:HE3	1:A:269:ASN:CG	2.27	0.54
1:A:12:LYS:CE	4:A:644:HOH:O	2.56	0.54
1:A:47:GLY:HA3	1:A:92:ALA:O	2.08	0.53
1:A:155:ASN:HA	1:A:189:PHE:O	2.09	0.52
1:A:186:ARG:HD3	4:A:684:HOH:O	2.10	0.52
1:A:181:ASN:HD21	1:A:185:GLN:HE21	1.59	0.51
1:A:67:HIS:CD2	1:A:207:SER:HB3	2.46	0.50
2:I:7:LEU:HD21	2:I:47:LEU:HD21	1.92	0.50
1:A:189:PHE:HB3	2:I:7:LEU:HD22	1.92	0.50
1:A:187:ALA:HB1	1:A:189:PHE:CE2	2.48	0.49
1:A:64:HIS:CD2	2:I:44:THR:HB	2.47	0.49
1:A:243:ASN:H	1:A:243:ASN:HD22	1.62	0.48
1:A:206:GLN:HG2	4:A:591:HOH:O	2.13	0.48
2:I:37:LEU:O	2:I:55:PHE:HA	2.14	0.48
1:A:200:ALA:HB1	1:A:201:PRO:HD2	1.95	0.47
1:A:6:TYR:HE1	1:A:10:GLN:HE22	1.61	0.47
1:A:156:GLU:HG3	4:A:683:HOH:O	2.14	0.47
2:I:48:ARG:HD2	2:I:51:ARG:NH1	2.30	0.47
1:A:167:TYR:CZ	1:A:170:LYS:HD3	2.49	0.46
2:I:49:TYR:CD1	2:I:49:TYR:N	2.83	0.46
1:A:35:ILE:O	1:A:58:TYR:HA	2.16	0.46
1:A:170:LYS:HD2	4:A:676:HOH:O	2.16	0.45
1:A:102:GLY:O	2:I:42:PRO:HD2	2.16	0.45
1:A:106:TRP:CD1	1:A:106:TRP:N	2.83	0.45
1:A:241:TRP:CZ3	1:A:249:ARG:HG2	2.52	0.45
1:A:255:THR:HG23	4:A:610:HOH:O	2.18	0.44
1:A:77:ASN:HB2	1:A:78:SER:H	1.61	0.44
2:I:37:LEU:HA	2:I:38:PRO:HD3	1.85	0.43
1:A:123:ASN:ND2	1:A:224:THR:HG22	2.33	0.43
2:I:27:LEU:HD13	4:I:667:HOH:O	2.18	0.43
1:A:214:TYR:N	1:A:214:TYR:CD1	2.86	0.43
1:A:13:ALA:N	1:A:14:PRO:HD3	2.34	0.43
1:A:269:ASN:C	1:A:269:ASN:ND2	2.70	0.43
1:A:103:GLN:O	1:A:107:ILE:HD13	2.18	0.42
1:A:31:ILE:HG22	1:A:95:VAL:HG21	2.00	0.42
1:A:12:LYS:HD2	1:A:269:ASN:HD21	1.85	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:200:ALA:HB1	1:A:201:PRO:CD	2.50	0.42
1:A:12:LYS:HE2	4:A:644:HOH:O	2.17	0.42
2:I:33:ASN:HB2	2:I:50:ASN:O	2.20	0.41
1:A:107:ILE:HD12	1:A:107:ILE:N	2.35	0.41
2:I:16:LYS:HB3	2:I:20:GLN:HB2	2.02	0.41
1:A:34:GLY:O	1:A:65:GLY:HA3	2.20	0.41
1:A:235:LEU:HD23	1:A:235:LEU:HA	1.92	0.41
2:I:8:LYS:HG2	2:I:10:PHE:CZ	2.56	0.41
1:A:156:GLU:HB2	1:A:164:THR:HB	2.04	0.40
2:I:34:VAL:HA	2:I:52:VAL:O	2.22	0.40

All (1) 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)
1:A:185:GLN:OE1	4:A:557:HOH:O[2_546]	2.13	0.07

## 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	273/275 (99%)	264 (97%)	9 (3%)	0	100	100
2	I	62/64 (97%)	61 (98%)	1 (2%)	0	100	100
All	All	335/339 (99%)	325 (97%)	10 (3%)	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	206/206 (100%)	194 (94%)	12 (6%)	25	19
2	I	59/59 (100%)	57 (97%)	2 (3%)	44	41
All	All	265/265 (100%)	251 (95%)	14 (5%)	28	22

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	SER
1	A	103	GLN
1	A	159	SER
1	A	185	GLN
1	A	213	THR
1	A	222	MET
1	A	240	THR
1	A	243	ASN
1	A	249	ARG
1	A	259	SER
1	A	265	LYS
1	A	269	ASN
2	I	7	LEU
2	I	61	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	10	GLN
1	A	43	ASN
1	A	59	GLN
1	A	76	ASN
1	A	109	ASN
1	A	123	ASN
1	A	184	ASN
1	A	185	GLN
1	A	243	ASN
1	A	269	ASN
2	I	28	HIS

### 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](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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 [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	275/275 (100%)	-0.77	0 100 100	6, 14, 37, 57	0
2	I	64/64 (100%)	-0.84	0 100 100	8, 17, 43, 65	0
All	All	339/339 (100%)	-0.78	0 100 100	6, 15, 40, 65	0

There are no RSRZ outliers to report.

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	CA	A	401	1/1	0.98	0.16	7.58	26,26,26,26	0
3	CA	A	400	1/1	1.00	0.09	2.46	9,9,9,9	0

## 6.5 Other polymers

There are no such residues in this entry.