

Full wwPDB X-ray Structure Validation Report (i)

Nov 5, 2023 – 09:55 pm GMT

PDB ID	:	10EW
Title	:	ATOMIC RESOLUTION STRUCTURE OF NATIVE ENDOTHIAPEPSIN
Authors	:	Coates, L.; Erskine, P.T.; Mall, S.; Gill, R.S.; Wood, S.P.; Myles, D.A.A.;
		Cooper, J.B.
Deposited on	:	2003-03-31
Resolution	:	0.90 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 A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

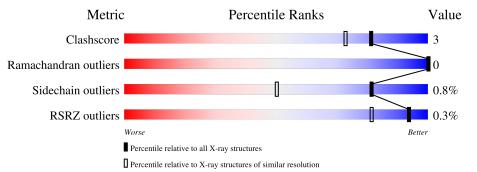
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as 541 be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 0.90 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
Clashscore	141614	1132(1.04-0.76)
Ramachandran outliers	138981	1055 (1.04-0.76)
Sidechain outliers	138945	1056 (1.04-0.76)
RSRZ outliers	127900	1028 (1.04-0.76)

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 >=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 <=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	А	329	93%	7%



$10\mathrm{EW}$

2 Entry composition (i)

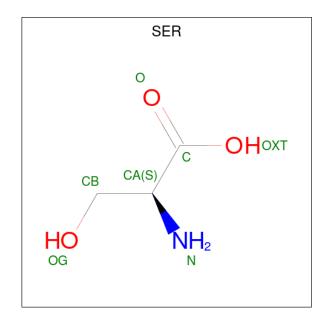
There are 6 unique types of molecules in this entry. The entry contains 3134 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 ENDOTHIAPEPSIN.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	329	Total 2480	C 1573	N 374	0 531	${ m S} { m 2}$	0	29	0

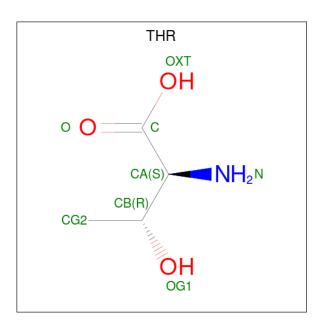
• Molecule 2 is SERINE (three-letter code: SER) (formula: C₃H₇NO₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	А	1	Total 6	С 3	N 1	O 2	0	0

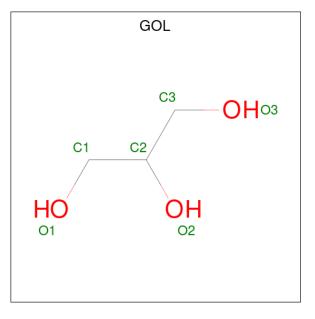
• Molecule 3 is THREONINE (three-letter code: THR) (formula: $C_4H_9NO_3$).





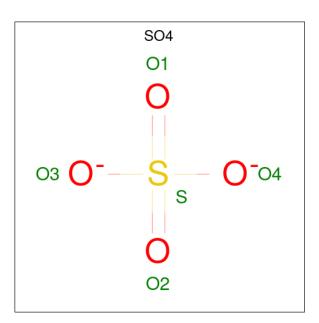
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	А	1	Total	С	Ν	0	0	0
0	11	Ĩ	8	4	1	3	U U	Ŭ

• Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	А	1	Total 6	С 3	O 3	0	0





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
5	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
5	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
5	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
5	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0

• Molecule 6 is water.

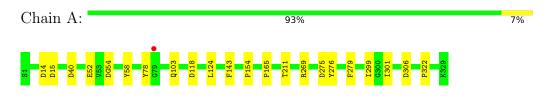
Μ	ol	Chain	Residues	Atoms		ZeroOcc	AltConf
(3	А	609	Total 609	O 609	0	0



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: ENDOTHIAPEPSIN





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	42.72Å 74.66Å 42.55Å	Deperitor
a, b, c, α , β , γ	90.00° 97.27° 90.00°	Depositor
Resolution (Å)	100.00 - 0.90	Depositor
Resolution (A)	42.38 - 0.91	EDS
% Data completeness	99.4 (100.00-0.90)	Depositor
(in resolution range)	92.3(42.38-0.91)	EDS
R _{merge}	0.11	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.96 (at 0.91 \text{\AA})$	Xtriage
Refinement program	SHELXL-97	Depositor
B B.	0.121 , 0.147	Depositor
R, R_{free}	0.127 , (Not available)	DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor $(Å^2)$	7.9	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.36 , 45.8	EDS
L-test for twinning ²	$< L > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.017 for l,-k,h	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	3134	wwPDB-VP
Average B, all atoms $(Å^2)$	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 9.01% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: GOL, SUI, 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).

Mal	Chain	Bo	nd lengths	Bond angles		
	Mol Chain		# Z > 5	RMSZ	# Z > 5	
1	А	0.77	2/2638~(0.1%)	1.03	8/3606~(0.2%)	

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	52	GLU	CD-OE2	6.32	1.32	1.25
1	А	52	GLU	CD-OE1	-5.30	1.19	1.25

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
1	А	275	ASP	CB-CG-OD2	-8.81	110.37	118.30
1	А	118	ASP	CB-CG-OD2	-7.51	111.55	118.30
1	А	276	TYR	CB-CG-CD1	6.64	124.99	121.00
1	А	58	TYR	CB-CG-CD2	6.59	124.95	121.00
1	А	269	ARG	NE-CZ-NH1	5.37	122.98	120.30
1	А	269	ARG	NE-CZ-NH2	-5.28	117.66	120.30
1	А	306	ASP	CB-CG-OD1	-5.17	113.65	118.30
1	А	14	ASP	CB-CG-OD1	5.03	122.82	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	А	2480	0	2367	12	0
2	А	6	0	4	1	0
3	А	8	0	7	0	0
4	А	6	0	8	0	0
5	А	25	0	0	0	0
6	А	609	0	0	9	0
All	All	3134	0	2386	13	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 3.

All (13) 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:279[A]:PHE:O	6:A:2539:HOH:O	1.63	1.15
1:A:322[B]:PRO:HG3	6:A:2573:HOH:O	1.82	0.79
1:A:322[B]:PRO:HD2	6:A:2426:HOH:O	1.84	0.78
1:A:165[B]:PRO:HG2	6:A:2383:HOH:O	1.91	0.70
1:A:211[B]:THR:HG21	6:A:2461:HOH:O	1.98	0.63
1:A:15:ASP:OD1	1:A:279[A]:PHE:HZ	1.91	0.54
1:A:154:PRO:HB3	6:A:2360:HOH:O	2.14	0.47
1:A:322[A]:PRO:HD2	6:A:2426:HOH:O	2.17	0.45
1:A:322[A]:PRO:HG2	6:A:2426:HOH:O	2.17	0.45
1:A:78:TYR:CZ	1:A:124:LEU:HD22	2.53	0.44
2:A:401:SER:N	6:A:2586:HOH:O	2.50	0.44
1:A:103:GLN:NE2	1:A:143:PHE:HA	2.32	0.44
1:A:299:ILE:HG22	1:A:301:ILE:HG22	2.03	0.40

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	А	354/329~(108%)	351 (99%)	3 (1%)	0	100 1	00

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	А	290/262~(111%)	289 (100%)	1 (0%)	92 67	

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type	
1	А	40	ASP	

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	31	ASN
1	А	103	GLN
1	А	146	ASN
1	А	169	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)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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	Type	Chain	Chain	Chain	Chain	Chain	Chain	Res	Link	Bo	ond leng	ths	B	ond ang	les
	Chain	i nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2						
1	SUI	А	54	1	10,11,12	1.36	1 (10%)	11,15,17	1.25	2 (18%)					

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
1	SUI	А	54	1	-	0/2/19/20	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	54	SUI	CB-CG	2.09	1.53	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	54	SUI	O-C-C2A	-2.59	118.56	126.39
1	А	54	SUI	CB-CG-N2	2.50	109.60	108.13

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

8 ligands are modelled in this entry.

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	Iol Type Chain Res		Link	B	Bond lengths			Bond angles		
	Type Cham	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2	
2	SER	А	401	-	4,5,6	0.53	0	$0,\!5,\!7$	-	-
5	SO4	А	502	-	4,4,4	0.72	0	$6,\!6,\!6$	1.36	0
5	SO4	А	504	-	4,4,4	0.24	0	$6,\!6,\!6$	0.49	0
3	THR	А	402	-	6,7,7	0.64	0	$7,\!9,\!9$	0.59	0
5	SO4	А	503	-	4,4,4	0.25	0	$6,\!6,\!6$	1.25	0
4	GOL	А	410	-	$5,\!5,\!5$	0.60	0	$5,\!5,\!5$	0.42	0
5	SO4	А	500	-	4,4,4	0.95	0	$6,\!6,\!6$	0.72	0
5	SO4	А	501	-	4,4,4	0.46	0	$6,\!6,\!6$	0.27	0

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	SER	А	401	-	-	2/2/4/6	-
3	THR	А	402	-	-	2/8/8/8	-
4	GOL	А	410	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	А	401	SER	N-CA-CB-OG
2	А	401	SER	C-CA-CB-OG
3	А	402	THR	OXT-C-CA-N
4	А	410	GOL	O1-C1-C2-C3
4	А	410	GOL	O1-C1-C2-O2
3	А	402	THR	N-CA-CB-OG1
4	А	410	GOL	O2-C2-C3-O3

There are no ring outliers.



1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	А	401	SER	1	0

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^{th} 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		$\mathbf{OWAB}(\mathbf{\AA}^2)$	Q < 0.9
1	А	328/329~(99%)	0.18	1 (0%) 9	94 81	6, 9, 17, 28	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ	
1	А	79	GLY	2.2	

6.2 Non-standard residues in protein, DNA, RNA chains (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^{th} 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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
1	SUI	А	54	11/12	0.98	0.06	$9,\!10,\!13,\!14$	0

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^{th} 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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
5	SO4	А	503	5/5	0.79	0.26	$8,\!12,\!18,\!22$	5

Continued on next page...



Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	SER	А	401	6/7	0.80	0.32	$12,\!13,\!16,\!32$	6
4	GOL	А	410	6/6	0.85	0.17	$16,\!17,\!27,\!35$	6
3	THR	А	402	8/8	0.87	0.31	$9,\!10,\!17,\!22$	8
5	SO4	А	500	5/5	0.91	0.15	16, 19, 23, 27	5
5	SO4	А	502	5/5	0.92	0.29	10, 14, 17, 20	5
5	SO4	А	504	5/5	0.92	0.19	12,16,27,28	5
5	SO4	А	501	5/5	0.98	0.09	8,9,10,12	5

Continued from previous page...

6.5 Other polymers (i)

There are no such residues in this entry.

