



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 08:59 am BST

PDB ID : 2H8L  
Title : Crystal structure of the bb' fragment of ERp57  
Authors : Kozlov, G.; Schrag, J.D.; Cygler, M.; Gehring, K.  
Deposited on : 2006-06-07  
Resolution : 2.00 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 : 2.11  
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) : 2.11

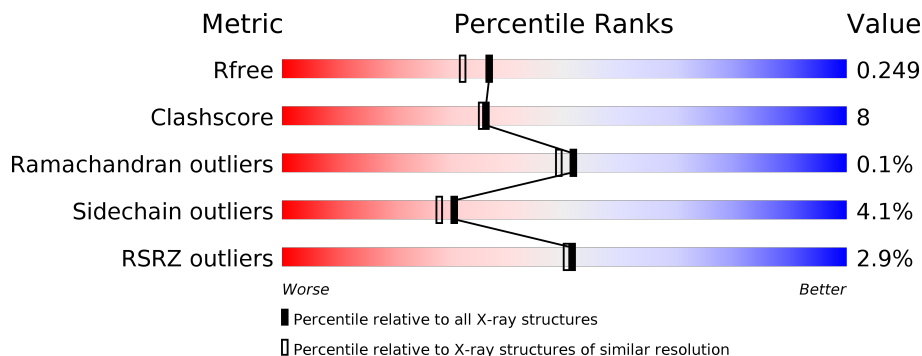
# 1 Overall quality at a glance

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))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 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	252	 3% 76% 13% • 8%
1	B	252	 5% 69% 19% • 9%
1	C	252	 74% 15% • 8%

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 6087 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 Protein disulfide-isomerase A3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	232	1883	1198	318	361	1	5	0	0	0
1	B	230	1868	1187	316	359	1	5	0	0	0
1	C	232	1883	1198	318	361	1	5	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	129	GLY	-	CLONING ARTIFACT	UNP P30101
A	130	PRO	-	CLONING ARTIFACT	UNP P30101
A	131	LEU	-	CLONING ARTIFACT	UNP P30101
A	132	GLY	-	CLONING ARTIFACT	UNP P30101
A	133	SER	-	CLONING ARTIFACT	UNP P30101
A	227	MSE	MET	MODIFIED RESIDUE	UNP P30101
A	247	MSE	MET	MODIFIED RESIDUE	UNP P30101
A	284	MSE	MET	MODIFIED RESIDUE	UNP P30101
A	285	MSE	MET	MODIFIED RESIDUE	UNP P30101
A	338	MSE	MET	MODIFIED RESIDUE	UNP P30101
A	377	ALA	-	CLONING ARTIFACT	UNP P30101
A	378	ALA	-	CLONING ARTIFACT	UNP P30101
A	379	ALA	-	CLONING ARTIFACT	UNP P30101
A	380	SER	-	CLONING ARTIFACT	UNP P30101
B	129	GLY	-	CLONING ARTIFACT	UNP P30101
B	130	PRO	-	CLONING ARTIFACT	UNP P30101
B	131	LEU	-	CLONING ARTIFACT	UNP P30101
B	132	GLY	-	CLONING ARTIFACT	UNP P30101
B	133	SER	-	CLONING ARTIFACT	UNP P30101
B	227	MSE	MET	MODIFIED RESIDUE	UNP P30101
B	247	MSE	MET	MODIFIED RESIDUE	UNP P30101
B	284	MSE	MET	MODIFIED RESIDUE	UNP P30101
B	285	MSE	MET	MODIFIED RESIDUE	UNP P30101

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	338	MSE	MET	MODIFIED RESIDUE	UNP P30101
B	377	ALA	-	CLONING ARTIFACT	UNP P30101
B	378	ALA	-	CLONING ARTIFACT	UNP P30101
B	379	ALA	-	CLONING ARTIFACT	UNP P30101
B	380	SER	-	CLONING ARTIFACT	UNP P30101
C	129	GLY	-	CLONING ARTIFACT	UNP P30101
C	130	PRO	-	CLONING ARTIFACT	UNP P30101
C	131	LEU	-	CLONING ARTIFACT	UNP P30101
C	132	GLY	-	CLONING ARTIFACT	UNP P30101
C	133	SER	-	CLONING ARTIFACT	UNP P30101
C	227	MSE	MET	MODIFIED RESIDUE	UNP P30101
C	247	MSE	MET	MODIFIED RESIDUE	UNP P30101
C	284	MSE	MET	MODIFIED RESIDUE	UNP P30101
C	285	MSE	MET	MODIFIED RESIDUE	UNP P30101
C	338	MSE	MET	MODIFIED RESIDUE	UNP P30101
C	377	ALA	-	CLONING ARTIFACT	UNP P30101
C	378	ALA	-	CLONING ARTIFACT	UNP P30101
C	379	ALA	-	CLONING ARTIFACT	UNP P30101
C	380	SER	-	CLONING ARTIFACT	UNP P30101

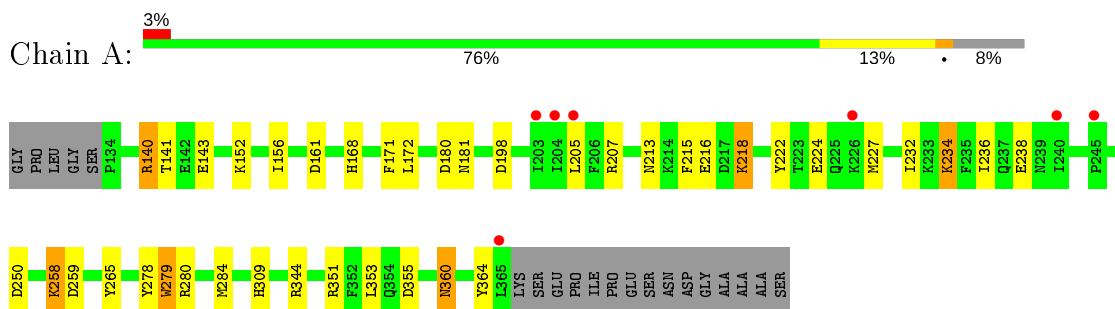
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	184	Total O 184 184	0	0
2	B	129	Total O 129 129	0	0
2	C	140	Total O 140 140	0	0

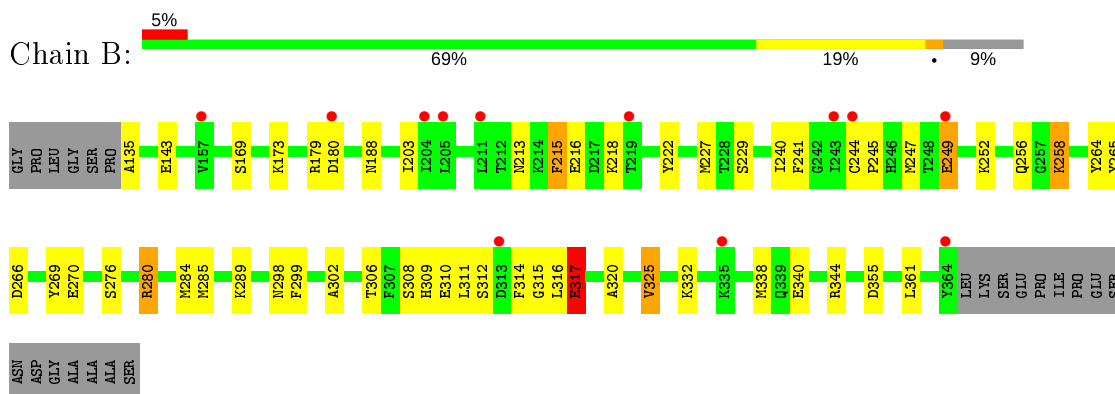
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

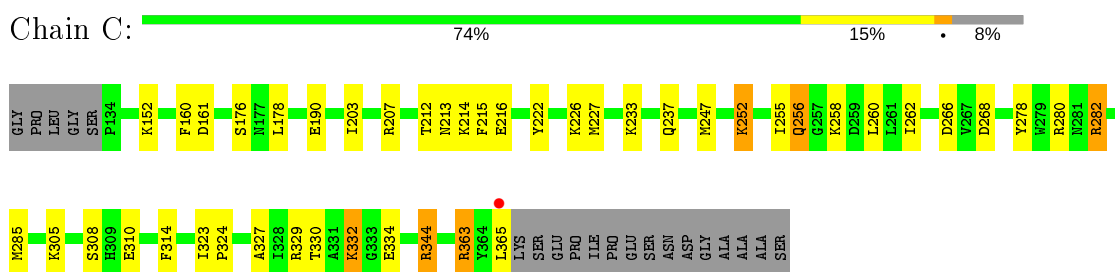
- Molecule 1: Protein disulfide-isomerase A3



- Molecule 1: Protein disulfide-isomerase A3



- Molecule 1: Protein disulfide-isomerase A3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.34Å 62.41Å 99.29Å 90.00° 98.40° 90.00°	Depositor
Resolution (Å)	32.70 – 2.00 32.74 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (32.70-2.00) 99.7 (32.74-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.85 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.194 , 0.250 0.193 , 0.249	Depositor DCC
$R_{free}$ test set	3162 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.1	Xtrriage
Anisotropy	0.058	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 42.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6087	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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](#)

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	1.15	3/1918 (0.2%)	0.98	8/2566 (0.3%)
1	B	1.49	16/1902 (0.8%)	1.04	6/2544 (0.2%)
1	C	1.04	2/1918 (0.1%)	1.08	12/2566 (0.5%)
All	All	1.24	21/5738 (0.4%)	1.04	26/7676 (0.3%)

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 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	317	GLU	CG-CD	27.32	1.93	1.51
1	B	317	GLU	CD-OE1	18.53	1.46	1.25
1	B	312	SER	CB-OG	11.18	1.56	1.42
1	B	312	SER	C-N	11.10	1.59	1.34
1	B	314	PHE	C-N	9.42	1.50	1.33

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	344	ARG	NE-CZ-NH2	-15.50	112.55	120.30
1	C	207	ARG	NE-CZ-NH1	13.21	126.91	120.30
1	C	207	ARG	NE-CZ-NH2	-11.97	114.31	120.30
1	C	344	ARG	NE-CZ-NH1	11.47	126.04	120.30
1	C	266	ASP	CB-CG-OD1	7.51	125.06	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	364	TYR	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	1883	0	1829	31	1
1	B	1868	0	1810	31	1
1	C	1883	0	1829	28	2
2	A	184	0	0	9	1
2	B	129	0	0	2	0
2	C	140	0	0	4	1
All	All	6087	0	5468	89	3

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 8.

The worst 5 of 89 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:317:GLU:CD	1:B:317:GLU:CG	1.92	1.37
1:A:227:MSE:HE3	1:A:232:ILE:HD11	1.40	1.03
1:A:218:LYS:HE3	2:A:418:HOH:O	1.57	1.02
1:A:140:ARG:HH11	1:A:140:ARG:HG3	1.26	1.00
1:A:258:LYS:HG3	2:A:506:HOH:O	1.65	0.96

All (3) 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:C:282:ARG:NH1	2:A:487:HOH:O[1_455]	2.09	0.11
1:B:320:ALA:CB	2:C:516:HOH:O[1_545]	2.15	0.05
1:A:198:ASP:OD1	1:C:282:ARG:NH2[1_655]	2.18	0.02



## 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	230/252 (91%)	224 (97%)	6 (3%)	0	100	100
1	B	228/252 (90%)	214 (94%)	13 (6%)	1 (0%)	34	30
1	C	230/252 (91%)	220 (96%)	10 (4%)	0	100	100
All	All	688/756 (91%)	658 (96%)	29 (4%)	1 (0%)	51	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	332	LYS

### 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	202/211 (96%)	197 (98%)	5 (2%)	47	49
1	B	200/211 (95%)	191 (96%)	9 (4%)	27	24
1	C	202/211 (96%)	191 (95%)	11 (5%)	22	18
All	All	604/633 (95%)	579 (96%)	25 (4%)	30	28

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	308	SER
1	B	325	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	332	LYS
1	B	317	GLU
1	C	152	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	194	ASN
1	B	213	ASN
1	B	309	HIS
1	B	177	ASN
1	C	213	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 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	227/252 (90%)	0.12	7 (3%) 49 48	35, 42, 51, 59	0
1	B	225/252 (89%)	0.40	12 (5%) 26 25	32, 42, 51, 54	0
1	C	227/252 (90%)	-0.09	1 (0%) 92 92	36, 42, 51, 54	0
All	All	679/756 (89%)	0.14	20 (2%) 51 50	32, 42, 51, 59	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	365	LEU	4.9
1	A	205	LEU	3.5
1	B	335	LYS	2.9
1	B	204	ILE	2.9
1	B	180	ASP	2.9

### 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

There are no such residues in this entry.