



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 23, 2026 – 01:56 pm GMT

PDB ID : 9IC8 / pdb\_00009ic8  
Title : Structure of BstEII with DNA substrate  
Authors : Rety, S.; Xi, X.G.; Zhang, J.H.  
Deposited on : 2025-02-14  
Resolution : 2.72 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 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 : 4-5-2 with Phenix2.0  
Xtrriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.48.1

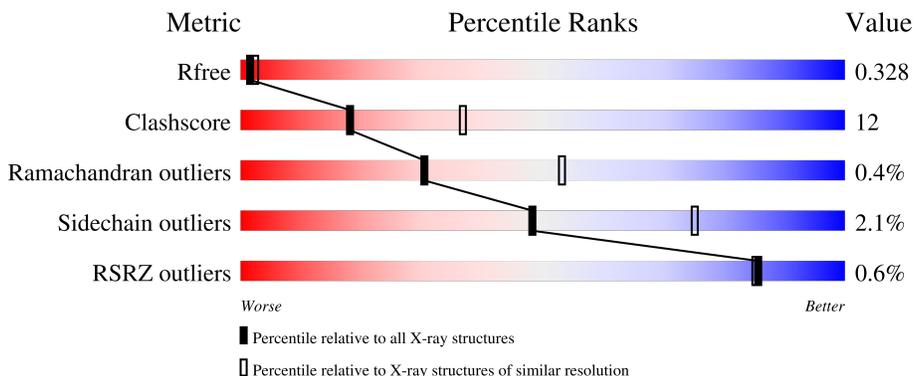
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.72 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))
$R_{free}$	164625	4050 (2.74-2.70)
Clashscore	180529	4439 (2.74-2.70)
Ramachandran outliers	177936	4374 (2.74-2.70)
Sidechain outliers	177891	4375 (2.74-2.70)
RSRZ outliers	164620	4050 (2.74-2.70)

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	288	 68% 31% .
1	B	288	 71% 28% .
1	C	288	 71% 27% ..
1	D	288	 2% 69% 27% ..

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 9428 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 BstEII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	286	Total 2357	C 1515	N 398	O 437	S 7	0	0	0
1	B	286	Total 2357	C 1515	N 398	O 437	S 7	0	0	0
1	C	286	Total 2357	C 1515	N 398	O 437	S 7	0	0	0
1	D	286	Total 2357	C 1515	N 398	O 437	S 7	0	0	0





- Molecule 1: BstEII



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	37.97Å 203.99Å 75.77Å 90.00° 91.76° 90.00°	Depositor
Resolution (Å)	101.99 – 2.72 101.99 – 2.72	Depositor EDS
% Data completeness (in resolution range)	98.6 (101.99-2.72) 89.8 (101.99-2.72)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.38 (at 2.73Å)	Xtrriage
Refinement program	PHENIX 1.21.2_5419	Depositor
R, $R_{free}$	0.275 , 0.325 0.278 , 0.328	Depositor DCC
$R_{free}$ test set	1490 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	62.5	Xtrriage
Anisotropy	0.357	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 99.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.118 for h,-k,-l	Xtrriage
Reported twinning fraction	0.210 for h,-k,-l	Depositor
Outliers	0 of 30279 reflections	Xtrriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	9428	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	88.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.68% 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	0.12	0/2410	0.38	0/3249
1	B	0.13	0/2410	0.38	0/3249
1	C	0.12	0/2410	0.39	0/3249
1	D	0.15	0/2410	0.51	0/3249
All	All	0.13	0/9640	0.42	0/12996

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	D	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	185	PHE	Peptide
1	D	190	ARG	Peptide
1	D	24	TYR	Peptide
1	D	253	MET	Peptide

### 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	A	2357	0	2349	60	3
1	B	2357	0	2349	55	2
1	C	2357	0	2349	55	1
1	D	2357	0	2349	59	0
All	All	9428	0	9396	225	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 12.

The worst 5 of 225 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:D:63:LYS:HB2	1:D:64:ASP:HA	1.56	0.87
1:C:11:VAL:O	1:C:13:ALA:N	2.14	0.81
1:A:238:LYS:NZ	1:A:284:TRP:O	2.22	0.73
1:C:264:GLY:HA3	1:C:268:VAL:HG12	1.71	0.73
1:B:174:LEU:HD11	1:B:229:TYR:CZ	2.25	0.71

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:A:206:ASP:OD2	1:B:83:SER:OG[1_655]	2.01	0.19
1:A:83:SER:OG	1:B:206:ASP:OD2[1_655]	2.13	0.07
1:A:14:LYS:NZ	1:C:99:GLU:O[2_555]	2.17	0.03

## 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	284/288 (99%)	282 (99%)	0	2 (1%)	19 40
1	B	284/288 (99%)	283 (100%)	0	1 (0%)	30 53

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	284/288 (99%)	281 (99%)	1 (0%)	2 (1%)	19	40
1	D	284/288 (99%)	282 (99%)	2 (1%)	0	100	100
All	All	1136/1152 (99%)	1128 (99%)	3 (0%)	5 (0%)	30	53

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	21	SER
1	A	206	ASP
1	B	206	ASP
1	C	12	ALA
1	C	21	SER

### 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	255/257 (99%)	252 (99%)	3 (1%)	67	85
1	B	255/257 (99%)	251 (98%)	4 (2%)	58	81
1	C	255/257 (99%)	252 (99%)	3 (1%)	67	85
1	D	255/257 (99%)	244 (96%)	11 (4%)	25	50
All	All	1020/1028 (99%)	999 (98%)	21 (2%)	48	75

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	192	ASP
1	D	213	ASP
1	D	288	LEU
1	D	270	VAL
1	D	200	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	119	HIS
1	D	58	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 oligosaccharides 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	286/288 (99%)	-0.49	0 <b>100</b>   <b>100</b>	38, 68, 140, 212	0
1	B	286/288 (99%)	-0.16	1 (0%) <b>90</b>   <b>90</b>	37, 93, 179, 249	0
1	C	286/288 (99%)	-0.40	1 (0%) <b>90</b>   <b>90</b>	34, 81, 158, 231	0
1	D	286/288 (99%)	-0.23	5 (1%) <b>69</b>   <b>69</b>	38, 95, 172, 381	0
All	All	1144/1152 (99%)	-0.32	7 (0%) <b>85</b>   <b>85</b>	34, 81, 165, 381	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	231	ALA	3.4
1	D	50	SER	3.2
1	D	188	TYR	3.1
1	B	226	ILE	2.5
1	D	198	VAL	2.2

### 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 oligosaccharides 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.