



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 10:13 PM GMT

PDB ID : 2BZA  
Title : BOVINE PANCREAS BETA-TRYPSIN IN COMPLEX WITH BENZYLAMINE  
Authors : Ota, N.; Stroupe, C.; Ferreira-Da-Silva, J.M.S.; Shah, S.S.; Mares-Guia, M.; Brunger, A.T.  
Deposited on : 1999-03-16  
Resolution : 1.90 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

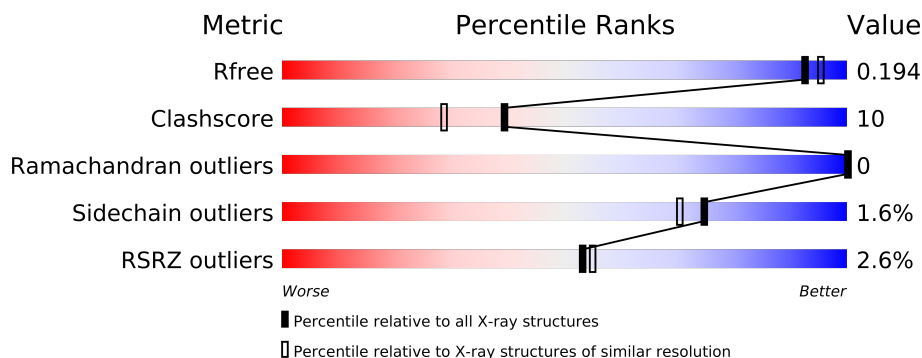
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance


The reported resolution of this entry is 1.90 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	3684 (1.90-1.90)
Clashscore	79885	4465 (1.90-1.90)
Ramachandran outliers	78287	4413 (1.90-1.90)
Sidechain outliers	78261	4414 (1.90-1.90)
RSRZ outliers	66119	3686 (1.90-1.90)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	223	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 1747 atoms, of which 0 are hydrogen and 0 are deuterium.

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 (TRYPSIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	0	0	0
			1629	1012	279	324	14			

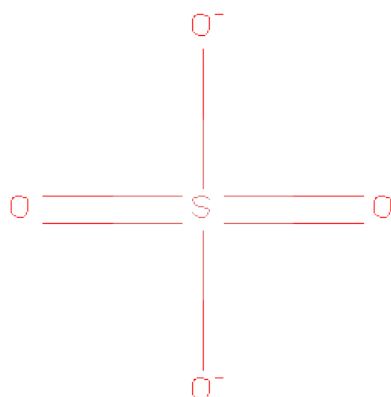
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

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

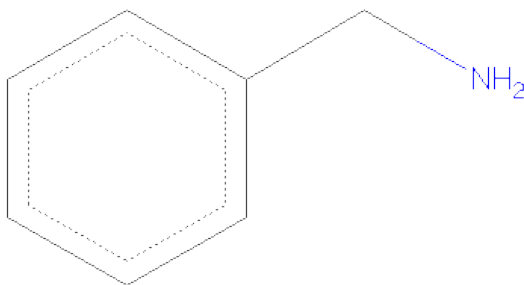
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Cl	0	0
			1	1		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is BENZYLAMINE (three-letter code: ABN) (formula:  $C_7H_9N$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	N	0	0
			8	7	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	103	Total	O	0	0
			103	103		

i

- Molecule 1: PROTEIN (TRYPSIN)

Q192	P198	L209	V213	A221A	Q221	K222	M223	Y228	S236	Q240	M245	I16	V31	F41	S45	L46	I47	N48	S49	Q50	M51	V52	M53	A56	Y59	K60	D64	V65	R66	L67	G78	N79	Y90	P124	T125	S127	Q135	M143	T144	K145	S146	S147	G148	T149	P152	P161	D165	K169	Y184	L185	E186	K188	G194
------	------	------	------	-------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.40Å 63.72Å 69.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.00 – 1.90 30.40 – 1.90	Depositor EDS
% Data completeness (in resolution range)	92.1 (32.00-1.90) 92.3 (30.40-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.14 (at 1.89Å)	Xtriage
Refinement program	CNS 0.3	Depositor
R, $R_{free}$	0.175 , 0.195 0.175 , 0.194	Depositor DCC
$R_{free}$ test set	2061 reflections (9.84%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.8	Xtriage
Anisotropy	0.010	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 40.8	EDS
Estimated twinning fraction	0.028 for k,h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 20943 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1747	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.39% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, SO4, ABN, CL

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	0.28	0/1660	0.63	0/2250

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1629	0	1588	31	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	5	0	0	0	0
5	A	8	0	9	0	0
6	A	103	0	0	0	0
All	All	1747	0	1597	31	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 10.

All (31) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:64:GLN:HE21	1:A:66:ARG:HE	1.20	0.88
1:A:48:ASN:C	1:A:48:ASN:HD22	1.91	0.73
1:A:64:GLN:NE2	1:A:66:ARG:HE	1.89	0.69
1:A:48:ASN:ND2	1:A:51:TRP:H	1.91	0.69
1:A:45:SER:OG	1:A:198:PRO:HB3	1.94	0.67
1:A:143:ASN:HD21	1:A:149:THR:HA	1.61	0.65
1:A:48:ASN:HD21	1:A:51:TRP:N	2.01	0.58
1:A:78:GLY:O	1:A:79:ASN:HB2	2.07	0.55
1:A:48:ASN:ND2	1:A:51:TRP:N	2.54	0.55
1:A:191:CYS:SG	1:A:192:GLN:NE2	2.80	0.54
1:A:165:ASP:OD1	1:A:169:LYS:HE3	2.09	0.53
1:A:48:ASN:ND2	1:A:50:GLN:H	2.07	0.52
1:A:64:GLN:HE22	1:A:66:ARG:HH21	1.56	0.52
1:A:135:GLN:NE2	1:A:161:PRO:HB3	2.25	0.52
1:A:124:PRO:HD3	1:A:209:LEU:O	2.10	0.52
1:A:48:ASN:C	1:A:48:ASN:ND2	2.62	0.50
1:A:144:THR:HG23	1:A:152:PRO:HD3	1.95	0.49
1:A:222:LYS:HG2	1:A:223:ASN:ND2	2.30	0.47
1:A:64:GLN:NE2	1:A:66:ARG:HH21	2.14	0.45
1:A:184:TYR:HB3	1:A:186:GLU:OE1	2.17	0.45
1:A:41:PHE:CZ	1:A:60:LYS:HD3	2.54	0.43
1:A:184:TYR:CE1	1:A:188:LYS:HE3	2.53	0.43
1:A:213:VAL:HA	1:A:228:TYR:CD2	2.54	0.42
1:A:236:SER:O	1:A:240:GLN:HG3	2.20	0.42
1:A:56:ALA:HB1	1:A:90:VAL:HG13	2.02	0.42
1:A:135:GLN:HE22	1:A:161:PRO:HB3	1.86	0.41
1:A:41:PHE:CE2	1:A:60:LYS:HD3	2.56	0.41
1:A:146:SER:OG	1:A:221(A):ALA:HB3	2.20	0.40
1:A:47:ILE:HD13	1:A:53:VAL:CG2	2.51	0.40
1:A:59:TYR:C	1:A:59:TYR:CD1	2.94	0.40
1:A:31:VAL:HG12	1:A:67:LEU:CD2	2.51	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	221/223 (99%)	216 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	184/184 (100%)	181 (98%)	3 (2%)	75	70

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	ASN
1	A	161	PRO
1	A	192	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	GLN
1	A	48	ASN
1	A	50	GLN
1	A	64	GLN
1	A	101	ASN
1	A	135	GLN
1	A	192	GLN
1	A	210	GLN
1	A	223	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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	$\# Z  > 2$
4	SO4	A	702	-	4,4,4	0.22	0	6,6,6	0.10	0
5	ABN	A	703	-	8,8,8	1.82	1 (12%)	9,9,9	0.66	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
4	SO4	A	702	-	-	0/0/0/0	0/0/0/0
5	ABN	A	703	-	-	0/2/2/2	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	703	ABN	C-C1	-3.72	1.41	1.52

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	223/223 (100%)	0.18	6 (2%) 52 53	12, 19, 39, 72	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	147	SER	7.4
1	A	148	GLY	5.6
1	A	149	THR	4.3
1	A	146	SER	3.7
1	A	145	LYS	2.3
1	A	127	SER	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	A	2	1/1	0.13	-	46,46,46,46	0
5	ABN	A	703	8/8	0.15	-	18,21,22,22	0
4	SO4	A	702	5/5	0.15	-	42,44,49,49	0
2	CA	A	1	1/1	0.04	-	17,17,17,17	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.