



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 09:22 PM UTC

PDB ID : 2WIT / pdb\_00002wit  
Title : CRYSTAL STRUCTURE OF THE SODIUM-COUPLED GLYCINE BETAINE SYMPORTER BETP FROM CORYNEBACTERIUM GLUTAMICUM WITH BOUND SUBSTRATE  
Authors : Ressler, S.; Terwisscha Van Scheltinga, A.C.; Vonrhein, C.; Ott, V.; Ziegler, C.  
Deposited on : 2009-05-17  
Resolution : 3.35 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  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
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) : 2.49

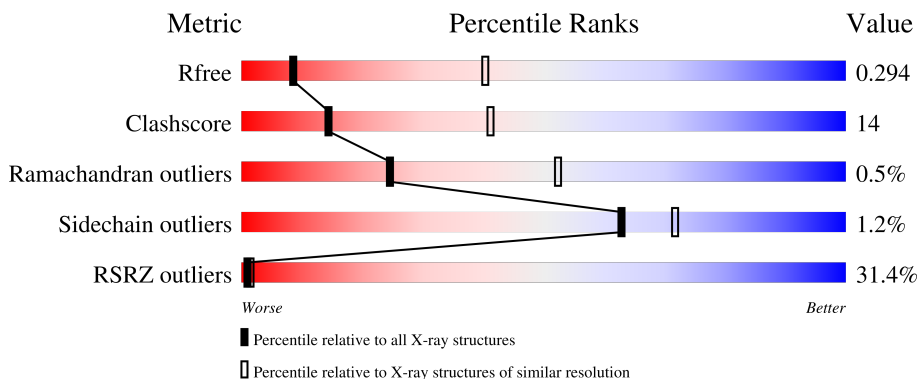
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.35 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))
$R_{free}$	180053	1099 (3.40-3.32)
Clashscore	190562	1116 (3.40-3.32)
Ramachandran outliers	187476	1101 (3.40-3.32)
Sidechain outliers	187428	1101 (3.40-3.32)
RSRZ outliers	180081	1099 (3.40-3.32)

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	566	
1	B	566	
1	C	566	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11737 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.

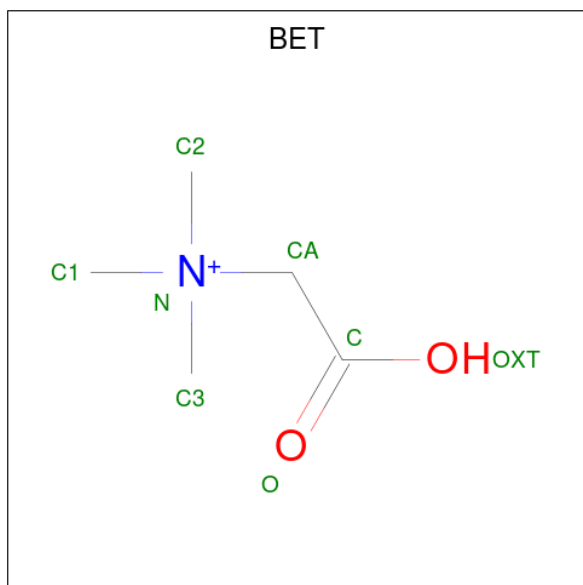
- Molecule 1 is a protein called GLYCINE BETAINES TRANSPORTER BETP.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	531	4057	2659	676	706	1	15	0	0	0
1	B	500	3783	2497	606	664	1	15	0	0	0
1	C	510	3873	2551	629	677	1	15	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	44	ALA	GLU	engineered mutation	UNP P54582
A	45	ALA	GLU	engineered mutation	UNP P54582
A	46	ALA	GLU	engineered mutation	UNP P54582
B	44	ALA	GLU	engineered mutation	UNP P54582
B	45	ALA	GLU	engineered mutation	UNP P54582
B	46	ALA	GLU	engineered mutation	UNP P54582
C	44	ALA	GLU	engineered mutation	UNP P54582
C	45	ALA	GLU	engineered mutation	UNP P54582
C	46	ALA	GLU	engineered mutation	UNP P54582

- Molecule 2 is TRIMETHYL GLYCINE (CCD ID: BET) (formula: C<sub>5</sub>H<sub>12</sub>NO<sub>2</sub>).

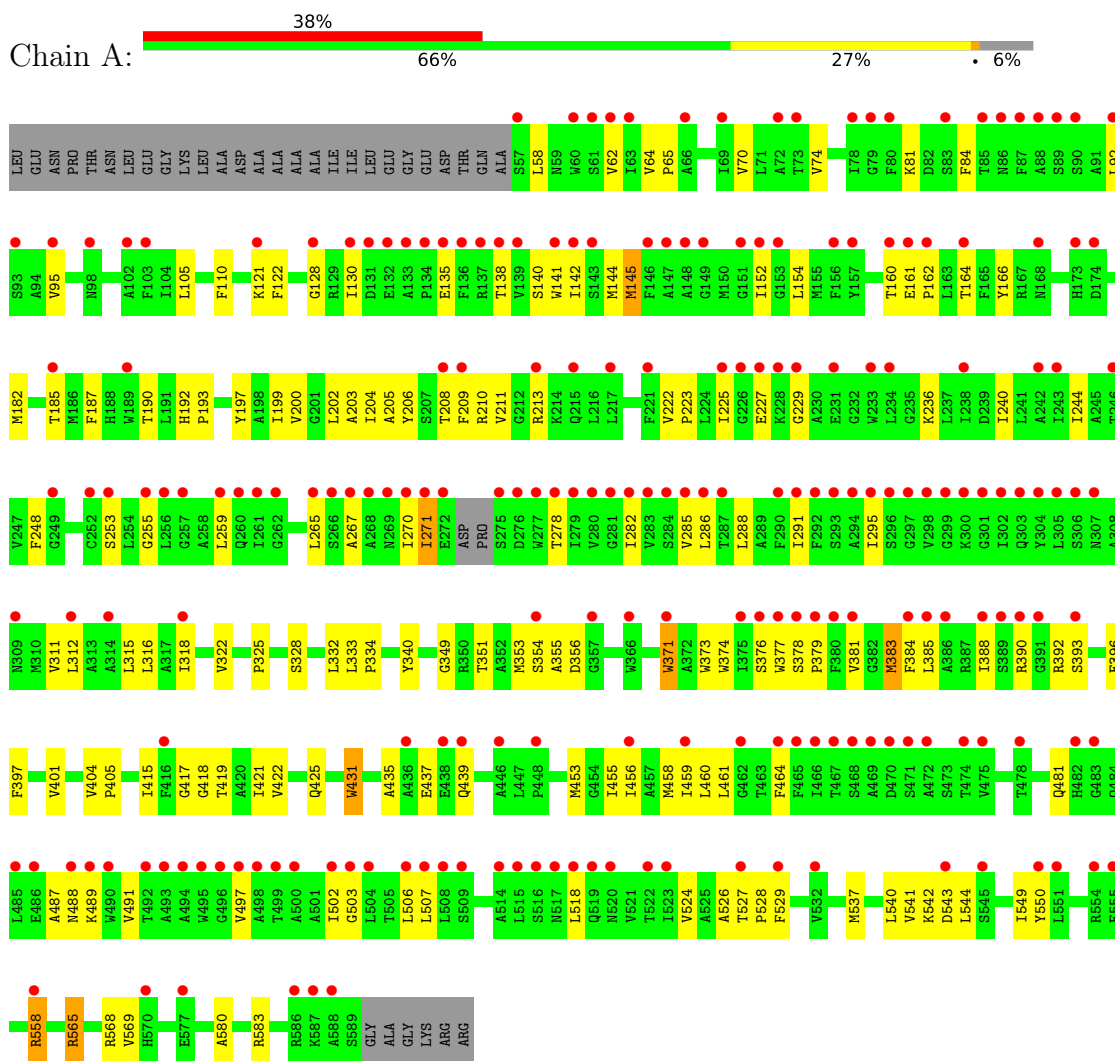


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total 8	5	1	2	0	0
2	B	1	Total 8	5	1	2	0	0
2	C	1	Total 8	5	1	2	0	0

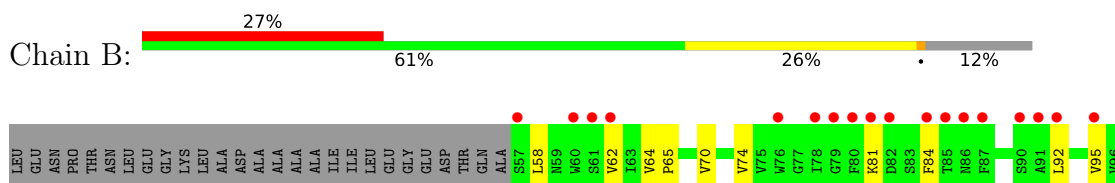
### 3 Residue-property plots

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: GLYCINE BETAINÉ TRANSPORTER BETP



#### • Molecule 1: GLYCINE BETAINÉ TRANSPORTER BETP





A561	•
R562	•
L563	•
A564	•
R565	•
F566	•
R567	•
R568	•
VAL	
HIS	
ASN	
GLU	
HIS	
ARG	
LYS	
ARG	
GLU	
LEU	
ALA	
ALA	
LYS	
ARG	
ARG	
ARG	
GLU	
ARG	
LYS	
ALA	
SER	
GLY	
ALA	
GLY	
LYS	
ARG	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.09Å 129.42Å 182.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.47 – 3.35 39.47 – 3.35	Depositor EDS
% Data completeness (in resolution range)	90.4 (39.47-3.35) 90.6 (39.47-3.35)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 3.32Å)	Xtrriage
Refinement program	BUSTER-TNT 2.7.0	Depositor
R, $R_{free}$	0.257 , 0.265 (Not available) , 0.294	Depositor DCC
$R_{free}$ test set	1886 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.7	Xtrriage
Anisotropy	0.286	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 44.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.76	EDS
Total number of atoms	11737	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% 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](#)

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

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	0/4142	0.68	0/5609
1	B	0.29	0/3865	0.67	0/5245
1	C	0.31	0/3956	0.68	0/5365
All	All	0.29	0/11963	0.68	0/16219

There are no bond length outliers.

There are no bond angle outliers.

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	A	4057	0	4104	119	0
1	B	3783	0	3815	110	0
1	C	3873	0	3909	110	0
2	A	8	0	11	4	0
2	B	8	0	11	3	0
2	C	8	0	11	0	0
All	All	11737	0	11861	332	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 14.

The worst 5 of 332 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:558:ARG:HH11	1:A:558:ARG:HG2	1.36	0.91
1:A:565:ARG:HG3	1:A:565:ARG:HH11	1.42	0.83
1:B:548:VAL:HG13	1:B:549:ILE:H	1.43	0.83
1:B:270:ILE:HG23	1:B:271:ILE:HG23	1.64	0.79
1:B:334:PRO:HG3	1:C:105:LEU:HB2	1.64	0.78

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	A	527/566 (93%)	486 (92%)	39 (7%)	2 (0%)	30 58
1	B	496/566 (88%)	454 (92%)	39 (8%)	3 (1%)	21 49
1	C	506/566 (89%)	466 (92%)	38 (8%)	2 (0%)	30 58
All	All	1529/1698 (90%)	1406 (92%)	116 (8%)	7 (0%)	24 52

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	225	ILE
1	B	225	ILE
1	B	547	ASP
1	C	225	ILE
1	A	431	TRP

### 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	415/424 (98%)	409 (99%)	6 (1%)	59	69
1	B	389/424 (92%)	385 (99%)	4 (1%)	68	74
1	C	397/424 (94%)	392 (99%)	5 (1%)	61	70
All	All	1201/1272 (94%)	1186 (99%)	15 (1%)	63	72

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	271	ILE
1	C	383	MSE
1	B	371	TRP
1	C	568	ARG
1	C	271	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	346	GLN
1	C	484	GLN
1	C	556	GLN
1	C	520	ASN
1	B	309	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

3 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	BET	A	1001	-	7,7,7	1.18	0	10,10,10	1.61	1 (10%)
2	BET	C	1001	-	7,7,7	1.17	0	10,10,10	1.54	1 (10%)
2	BET	B	1001	-	7,7,7	1.16	0	10,10,10	1.55	1 (10%)

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	BET	A	1001	-	-	2/5/5/5	-
2	BET	C	1001	-	-	2/5/5/5	-
2	BET	B	1001	-	-	2/5/5/5	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	BET	C-CA-N	-4.86	109.73	116.31
2	B	1001	BET	C-CA-N	-4.63	110.06	116.31
2	C	1001	BET	C-CA-N	-4.61	110.08	116.31

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1001	BET	O-C-CA-N
2	A	1001	BET	OXT-C-CA-N
2	B	1001	BET	O-C-CA-N

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	B	1001	BET	OXT-C-CA-N
2	C	1001	BET	O-C-CA-N

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	BET	4	0
2	B	1001	BET	3	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<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	516/566 (91%)	2.02	216 (41%) 0 0	13, 76, 164, 193	0
1	B	485/566 (85%)	1.61	154 (31%) 1 1	5, 64, 113, 211	0
1	C	495/566 (87%)	1.21	100 (20%) 3 4	5, 40, 104, 192	0
All	All	1496/1698 (88%)	1.62	470 (31%) 1 1	5, 60, 138, 211	0

The worst 5 of 470 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	131	ASP	11.0
1	A	499	THR	8.3
1	A	302	ILE	8.2
1	A	271	ILE	8.1
1	A	270	ILE	7.9

### 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 oligosaccharides 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<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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	BET	A	1001	8/8	0.86	0.28	3,108,128,151	0
2	BET	C	1001	8/8	0.90	0.25	30,78,154,165	0
2	BET	B	1001	8/8	0.91	0.30	18,69,117,138	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.