



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 11, 2023 – 06:48 pm GMT

PDB ID : 2WIT
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.; Vonnrhein, 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

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 : 4.02b-467
Mogul : 1.8.4, CSD as541be (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

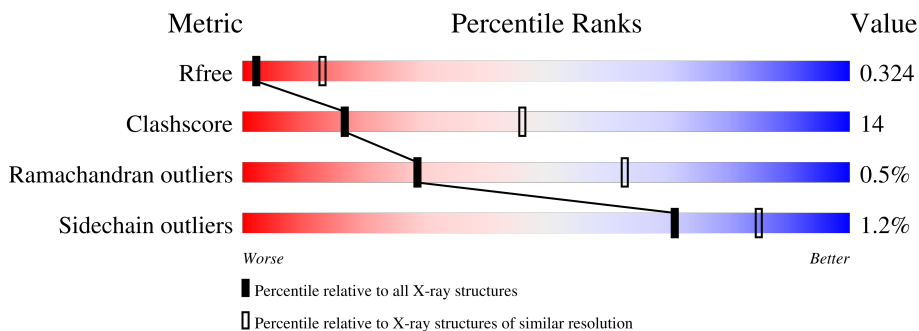
1 Overall quality at a glance

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}	130704	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	566	66% (green), 27% (yellow), 6% (grey)
1	B	566	61% (green), 26% (yellow), 12% (grey)
1	C	566	62% (green), 27% (yellow), 10% (grey)

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.

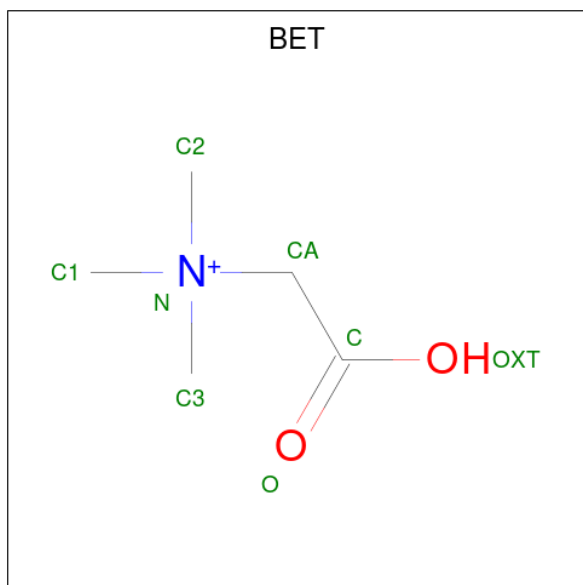
- Molecule 1 is a protein called GLYCINE BETAINES TRANSPORTER BETP.

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 (three-letter code: BET) (formula: C₅H₁₂NO₂).



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

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	118.09Å 129.42Å 182.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.47 – 3.35 39.37 – 3.35	Depositor EDS
% Data completeness (in resolution range)	90.4 (39.47-3.35) 90.6 (39.37-3.35)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.46 (at 3.32Å)	Xtrriage
Refinement program	BUSTER-TNT 2.7.0	Depositor
R, R_{free}	0.257 , 0.265 0.310 , 0.324	Depositor DCC
R_{free} test set	1886 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	46.7	Xtrriage
Anisotropy	0.286	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 49.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.74	EDS
Total number of atoms	11737	wwPDB-VP
Average B, all atoms (Å ²)	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.*

¹Intensities estimated from amplitudes.

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.

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	0.23	0/4142	0.34	0/5609
1	B	0.24	0/3865	0.33	0/5245
1	C	0.25	0/3956	0.34	0/5365
All	All	0.24	0/11963	0.34	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%)	34 68
1	B	496/566 (88%)	454 (92%)	39 (8%)	3 (1%)	25 59
1	C	506/566 (89%)	466 (92%)	38 (8%)	2 (0%)	34 68
All	All	1529/1698 (90%)	1406 (92%)	116 (8%)	7 (0%)	29 63

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%)	410 (99%)	5 (1%)	71	85
1	B	389/424 (92%)	385 (99%)	4 (1%)	76	87
1	C	397/424 (94%)	392 (99%)	5 (1%)	69	84
All	All	1201/1272 (94%)	1187 (99%)	14 (1%)	71	85

5 of 14 residues with a non-rotameric sidechain are listed below:

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	484	GLN
1	C	556	GLN
1	C	520	ASN
1	B	484	GLN
1	C	481	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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	1.19	0	10,10,10	1.59	1 (10%)
2	BET	B	1001	-	7,7,7	1.18	0	10,10,10	1.53	1 (10%)
2	BET	C	1001	-	7,7,7	1.18	0	10,10,10	1.53	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	B	1001	-	-	2/5/5/5	-
2	BET	C	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	-4.80	109.73	116.34
2	B	1001	BET	C-CA-N	-4.57	110.06	116.34
2	C	1001	BET	C-CA-N	-4.55	110.08	116.34

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.