



wwPDB EM Validation Summary Report ⓘ

Sep 2, 2025 – 01:28 AM JST

PDB ID : 9K21 / pdb_00009k21
EMDB ID : EMD-61987
Title : Cryo-EM structure of human taurine transporter TauT bound with homotaurine in an inward-open state
Authors : Du, B.; Yan, K.
Deposited on : 2024-10-16
Resolution : 3.45 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp>
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:

EMDB validation analysis : **FAILED**
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4-5-2 with Phenix2.0rc1
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.45.1

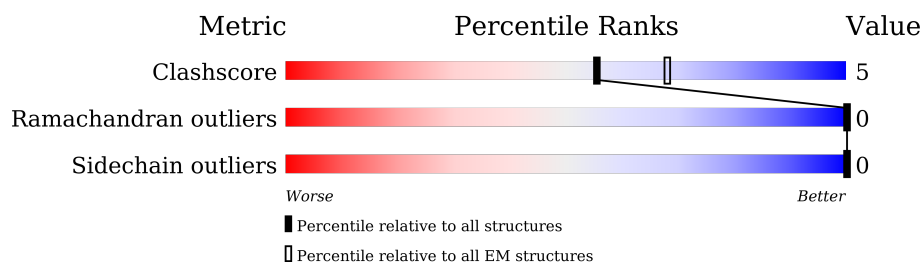
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.45 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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	H	118	100%
2	L	138	69% 9% 22%
3	A	651	73% 8% 19%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	A20	A	702	-	X	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5489 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Fragment antigen binding heavy chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	H	118	Total	C	N	O	S	0	0
			825	527	139	156	3		

- Molecule 2 is a protein called Fragment antigen binding light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	L	107	Total	C	N	O	S	0	0
			723	460	121	139	3		

- Molecule 3 is a protein called Sodium- and chloride-dependent taurine transporter.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	A	529	Total	C	N	O	S	0	0
			3932	2635	599	668	30		

There are 43 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	319	HIS	LYS	engineered mutation	UNP P31641
A	320	ASN	TYR	engineered mutation	UNP P31641
A	322	VAL	SER	engineered mutation	UNP P31641
A	486	ASN	ASP	engineered mutation	UNP P31641
A	487	ARG	ASN	engineered mutation	UNP P31641
A	488	PHE	LEU	engineered mutation	UNP P31641
A	489	SER	TYR	engineered mutation	UNP P31641
A	490	GLU	ASP	engineered mutation	UNP P31641
A	491	ASP	GLY	engineered mutation	UNP P31641
A	493	ARG	GLU	engineered mutation	UNP P31641
A	498	PHE	TYR	engineered mutation	UNP P31641
A	499	PRO	ARG	engineered mutation	UNP P31641
A	621	GLY	-	expression tag	UNP P31641
A	622	SER	-	expression tag	UNP P31641
A	623	ASP	-	expression tag	UNP P31641

Continued on next page...

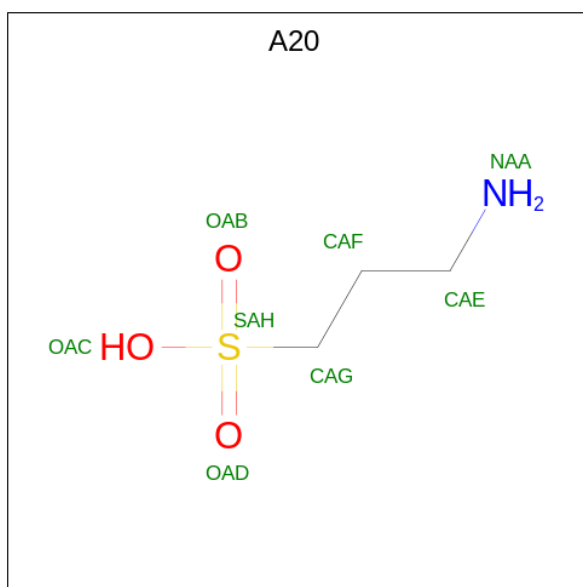
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	624	GLU	-	expression tag	UNP P31641
A	625	VAL	-	expression tag	UNP P31641
A	626	ASP	-	expression tag	UNP P31641
A	627	ALA	-	expression tag	UNP P31641
A	628	GLY	-	expression tag	UNP P31641
A	629	SER	-	expression tag	UNP P31641
A	630	HIS	-	expression tag	UNP P31641
A	631	HIS	-	expression tag	UNP P31641
A	632	HIS	-	expression tag	UNP P31641
A	633	HIS	-	expression tag	UNP P31641
A	634	HIS	-	expression tag	UNP P31641
A	635	HIS	-	expression tag	UNP P31641
A	636	HIS	-	expression tag	UNP P31641
A	637	HIS	-	expression tag	UNP P31641
A	638	HIS	-	expression tag	UNP P31641
A	639	HIS	-	expression tag	UNP P31641
A	640	GLY	-	expression tag	UNP P31641
A	641	SER	-	expression tag	UNP P31641
A	642	VAL	-	expression tag	UNP P31641
A	643	GLU	-	expression tag	UNP P31641
A	644	ASP	-	expression tag	UNP P31641
A	645	TYR	-	expression tag	UNP P31641
A	646	LYS	-	expression tag	UNP P31641
A	647	ASP	-	expression tag	UNP P31641
A	648	ASP	-	expression tag	UNP P31641
A	649	ASP	-	expression tag	UNP P31641
A	650	ASP	-	expression tag	UNP P31641
A	651	LYS	-	expression tag	UNP P31641

- Molecule 4 is CHLORIDE ION (CCD ID: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	AltConf
4	A	1	Total Cl 1 1	0

- Molecule 5 is 3-aminopropane-1-sulfonic acid (CCD ID: A20) (formula: C₃H₉NO₃S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	S	
5	A	1	8	3	1	3	1	0

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	164725	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: A20, 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	H	0.14	0/844	0.40	0/1153
2	L	0.15	0/742	0.36	0/1021
3	A	0.31	0/4059	0.60	0/5572
All	All	0.28	0/5645	0.54	0/7746

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
3	A	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	302	TYR	Sidechain
3	A	532	TYR	Sidechain

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	H	825	0	724	0	0
2	L	723	0	608	9	0
3	A	3932	0	3696	41	0
4	A	1	0	0	0	0
5	A	8	0	9	7	0
All	All	5489	0	5037	50	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 5.

The worst 5 of 50 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)
3:A:303:ALA:HB2	5:A:702:A20:CAE	1.89	1.02
3:A:303:ALA:HB2	5:A:702:A20:H6	1.36	1.02
3:A:303:ALA:CB	5:A:702:A20:H6	2.04	0.88
3:A:63:ASN:ND2	3:A:333:ASN:OD1	2.18	0.76
2:L:31:SER:O	2:L:52:THR:OG1	2.14	0.66

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 PDB entries followed by that with respect to all EM entries.

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	H	116/118 (98%)	113 (97%)	3 (3%)	0	100	100
2	L	105/138 (76%)	101 (96%)	4 (4%)	0	100	100
3	A	525/651 (81%)	512 (98%)	13 (2%)	0	100	100
All	All	746/907 (82%)	726 (97%)	20 (3%)	0	100	100

There are no Ramachandran outliers to report.

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 PDB entries followed by that with respect to all EM entries.

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	H	72/98 (74%)	72 (100%)	0	100	100
2	L	61/116 (53%)	61 (100%)	0	100	100
3	A	383/561 (68%)	383 (100%)	0	100	100
All	All	516/775 (67%)	516 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	74	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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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
5	A20	A	702	-	7,7,7	0.92	0	9,9,9	4.50	5 (55%)

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
5	A20	A	702	-	-	4/5/5/5	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	702	A20	CAF-CAG-SAH	-8.74	99.85	113.25
5	A	702	A20	OAB-SAH-OAD	-5.51	94.87	113.95
5	A	702	A20	OAD-SAH-CAG	5.31	113.31	106.92
5	A	702	A20	OAC-SAH-CAG	5.17	114.13	105.77
5	A	702	A20	OAB-SAH-CAG	3.80	111.49	106.92

There are no chirality outliers.

All (4) torsion outliers are listed below:

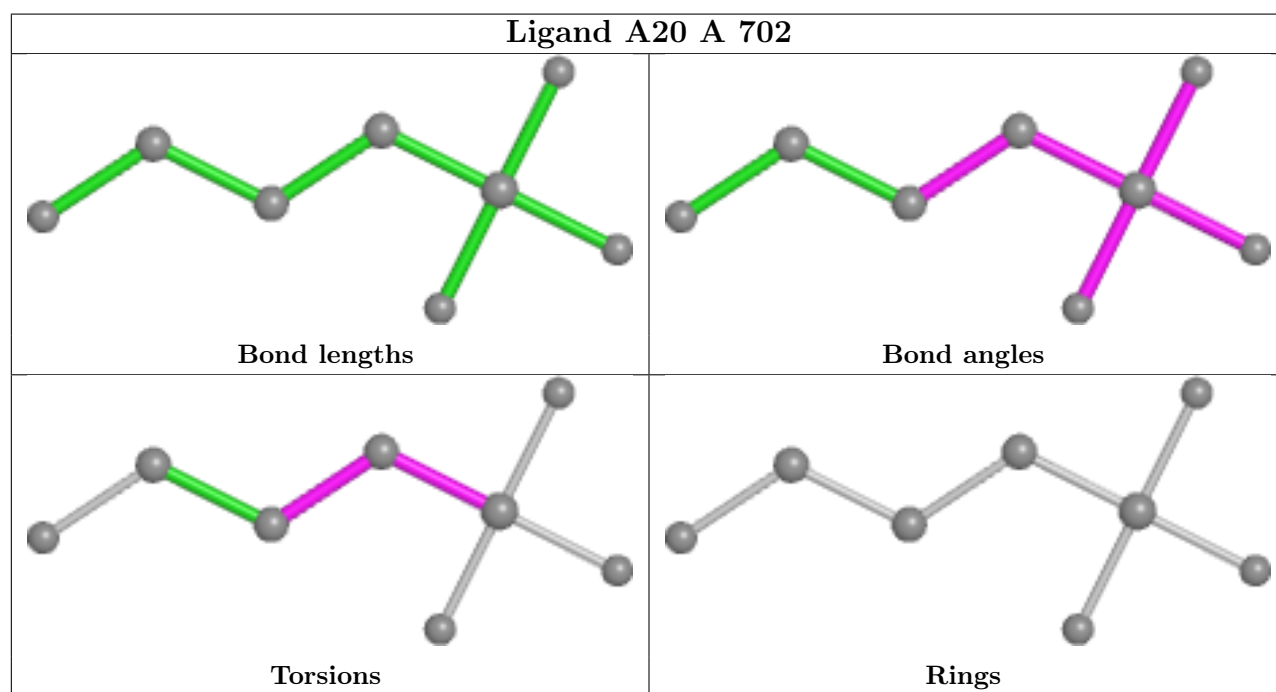
Mol	Chain	Res	Type	Atoms
5	A	702	A20	CAF-CAG-SAH-OAC
5	A	702	A20	CAE-CAF-CAG-SAH
5	A	702	A20	CAF-CAG-SAH-OAD
5	A	702	A20	CAF-CAG-SAH-OAB

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	702	A20	7	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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.