

Full wwPDB X-ray Structure Validation Report (i)

May 31, 2021 – 04:03 pm BST

PDB ID	:	6THA
Title	:	Crystal structure of human sugar transporter GLUT1 (SLC2A1) in the inward
		conformation
Authors	:	Pedersen, B.P.; Paulsen, P.A.; Custodio, T.F.
Deposited on	:	2019-11-19
Resolution	:	2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity	÷	4.02b-467
Mogul		1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.19
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
$\operatorname{CCP4}$:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.19

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries},{ m resolution\ range}({ m \AA}))$		
R_{free}	130704	3907 (2.40-2.40)		
Clashscore	141614	4398 (2.40-2.40)		
Ramachandran outliers	138981	4318 (2.40-2.40)		
Sidechain outliers	138945	4319 (2.40-2.40)		
RSRZ outliers	127900	3811(2.40-2.40)		

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 <=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			
			21%			
1	А	496	77%	13%	·	10%



$6 \mathrm{THA}$

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3556 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 Solute carrier family 2, facilitated glucose transporter member 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	А	448	Total 3478	C 2299	N 563	O 594	S 22	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	493	LEU	-	expression tag	UNP P11166
А	494	VAL	-	expression tag	UNP P11166
A	495	PRO	-	expression tag	UNP P11166
А	496	ARG	-	expression tag	UNP P11166

• Molecule 2 is nonyl beta-D-glucopyranoside (three-letter code: BNG) (formula: $C_{15}H_{30}O_6$) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	1	Total C O 21 15 6	0	0
2	А	1	Total C O 21 15 6	0	0

• Molecule 3 is 3,6,9,12,15,18-HEXAOXAICOSANE-1,20-DIOL (three-letter code: P33) (formula: $C_{14}H_{30}O_8$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	А	1	Total 22	C 14	O 8	0	0

• Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	Total Cl 1 1	0	0

• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	13	Total O 13 13	0	0



3 Residue-property plots (i)

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: Solute carrier family 2, facilitated glucose transporter member 1





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	120.00Å 102.18 Å 69.00 Å	Depositor
a, b, c, α , β , γ	90.00° 98.89° 90.00°	Depositor
$\mathbf{B}_{\mathrm{ascolution}}(\mathbf{\hat{A}})$	41.66 - 2.40	Depositor
Resolution (A)	48.78 - 2.40	EDS
% Data completeness	99.8 (41.66-2.40)	Depositor
(in resolution range)	99.9 (48.78-2.40)	EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.03 (at 2.39 \text{\AA})$	Xtriage
Refinement program	PHENIX dev-2614	Depositor
B B.	0.204 , 0.229	Depositor
n, n_{free}	0.205 , 0.232	DCC
R_{free} test set	1583 reflections (4.92%)	wwPDB-VP
Wilson B-factor $(Å^2)$	78.6	Xtriage
Anisotropy	0.412	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.30 , 90.1	EDS
L-test for twinning ²	$ < L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3556	wwPDB-VP
Average B, all atoms $(Å^2)$	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.78% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: CL, P33, BNG

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).

Mal	Chain	Bond	lengths	Bond angles		
		RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.29	0/3561	0.43	0/4836	

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	А	3478	0	3601	39	0
2	А	42	0	60	1	0
3	А	22	0	30	1	0
4	А	1	0	0	0	0
5	А	13	0	0	0	0
All	All	3556	0	3691	39	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.

All (39) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.



Atom 1	Atom 2	Interatomic	Clash	
Atom-1	At0111-2	distance (Å)	overlap (Å)	
1:A:278:LEU:HD12	1:A:390:ILE:HG12	1.68	0.75	
1:A:426:GLU:HA	1:A:433:VAL:HG21	1.80	0.64	
1:A:145:GLY:O	1:A:153:ARG:NH2	2.30	0.63	
1:A:334:ARG:NH1	1:A:448:THR:O	2.37	0.57	
1:A:146:GLU:HG2	1:A:212:ARG:HD3	1.87	0.57	
1:A:340:GLY:HA3	1:A:383:PRO:HD3	1.88	0.55	
1:A:324:SER:HA	1:A:327:VAL:HG12	1.88	0.53	
1:A:290:VAL:HG22	1:A:372:ILE:HD13	1.91	0.53	
1:A:274:ILE:HG23	1:A:452:VAL:HG22	1.91	0.53	
1:A:434:PHE:O	1:A:438:THR:HG23	2.08	0.53	
1:A:220:GLU:OE1	1:A:223:ARG:HD3	2.10	0.52	
1:A:354:ALA:HB2	1:A:367:LEU:HD23	1.90	0.52	
1:A:147:VAL:O	1:A:232:ARG:NH2	2.40	0.52	
1:A:430:GLY:O	1:A:433:VAL:HG23	2.10	0.50	
1:A:150:THR:OG1	1:A:243:GLU:OE1	2.29	0.50	
1:A:149:PRO:HG2	1:A:152:LEU:HD12	1.94	0.49	
1:A:49:VAL:HG12	1:A:55:SER:HA	1.96	0.48	
1:A:140:VAL:HB	1:A:141:PRO:HD3	1.97	0.47	
1:A:361:LEU:HD22	1:A:363:TRP:CZ2	2.49	0.47	
1:A:87:VAL:CG2	1:A:92:ARG:HG3	2.44	0.47	
1:A:385:PRO:HA	2:A:501:BNG:H4'1	1.98	0.46	
1:A:334:ARG:NH1	1:A:452:VAL:O	2.49	0.46	
1:A:274:ILE:HG22	1:A:390:ILE:HD13	1.98	0.45	
1:A:343:GLY:HA3	1:A:378:PHE:CD1	2.51	0.45	
1:A:106:SER:OG	1:A:129:ILE:HB	2.17	0.45	
1:A:388:TRP:CZ2	3:A:503:P33:H142	2.51	0.45	
1:A:320:PHE:HA	1:A:323:VAL:HB	2.00	0.44	
1:A:336:LEU:HB2	1:A:386:ILE:HD11	2.00	0.43	
1:A:285:SER:HB2	1:A:441:LEU:HD21	1.99	0.43	
1:A:180:MET:HA	1:A:185:LEU:HB2	2.00	0.42	
1:A:278:LEU:HD12	1:A:390:ILE:CG1	2.46	0.42	
1:A:392:ALA:HA	1:A:400:ARG:HD3	2.02	0.42	
1:A:429:CYS:HB2	1:A:433:VAL:HG22	2.01	0.42	
1:A:426:GLU:HA	1:A:433:VAL:CG2	2.48	0.41	
1:A:360:GLN:OE1	1:A:360:GLN:N	2.50	0.41	
1:A:294:SER:O	1:A:297:ILE:HG13	2.20	0.41	
1:A:290:VAL:O	1:A:294:SER:HB3	2.20	0.41	
1:A:435:ILE:O	1:A:439:VAL:HG23	2.20	0.41	
1:A:249:ARG:O	1:A:253:ARG:HG3	2.21	0.40	

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	А	446/496~(90%)	433~(97%)	12 (3%)	1 (0%)	47 62

All (1) Ramachandran outliers are listed below:

Mol	Chain	\mathbf{Res}	Type
1	А	304	GLN

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	А	376/416~(90%)	368~(98%)	8 (2%)	53 72	

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	81	PHE
1	А	147	VAL
1	А	153	ARG
1	А	178	SER
1	А	280	LEU
1	А	361	LEU
1	А	372	ILE
1	А	394	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.



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)

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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 Turn	Tune	Chain	Dog	Tink	Bo	Bond lengths			Bond angles		
WIOI	туре	Ullalli	nes		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2	
3	P33	A	503	-	21,21,21	0.20	0	20,20,20	0.15	0	
2	BNG	А	502	-	21,21,21	0.12	0	26,26,26	0.32	0	
2	BNG	А	501	-	21,21,21	0.19	0	26,26,26	0.21	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
3	P33	А	503	-	-	9/19/19/19	-
2	BNG	А	502	-	-	1/12/32/32	0/1/1/1
2	BNG	A	501	-	-	1/12/32/32	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.



There are no chirality outliers.

Mol	Chain	\mathbf{Res}	Type	Atoms
3	А	503	P33	C11-C12-O13-C14
2	А	501	BNG	C2'-C1'-O1-C1
2	А	502	BNG	C4'-C5'-C6'-C7'
3	А	503	P33	O13-C14-C15-O16
3	А	503	P33	C5-C6-O7-C8
3	А	503	P33	C18-C17-O16-C15
3	А	503	P33	C21-C20-O19-C18
3	А	503	P33	C17-C18-O19-C20
3	А	503	P33	O10-C11-C12-O13
3	А	503	P33	O16-C17-C18-O19
3	А	503	P33	O4-C5-C6-O7

All (11) torsion outliers are listed below:

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	А	503	P33	1	0
2	А	501	BNG	1	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 then 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.



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^{th} 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	$OWAB(Å^2)$	Q<0.9
1	А	448/496~(90%)	1.02	106 (23%) 0 0	66, 101, 152, 188	0

All (106) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	450	PHE	8.5
1	А	452	VAL	7.0
1	А	449	TYR	6.7
1	А	152	LEU	6.3
1	А	326	PHE	6.2
1	А	455	THR	6.2
1	А	360	GLN	5.7
1	А	13	MET	5.6
1	А	8	LEU	5.5
1	А	9	THR	5.1
1	А	237	VAL	5.1
1	А	330	ARG	4.9
1	А	133	CYS	4.9
1	А	451	LYS	4.7
1	А	101	LEU	4.5
1	А	129	ILE	4.3
1	А	132	TYR	4.3
1	А	54	GLU	4.2
1	А	363	TRP	4.1
1	А	136	THR	4.0
1	А	432	TYR	4.0
1	А	334	ARG	3.9
1	А	454	GLU	3.9
1	А	273	LEU	3.9
1	А	102	LEU	3.9
1	А	135	LEU	3.8
1	А	242	GLN	3.8

Continued on next page...



67	THA
~ -	

Mol	Chain	Res	Type	RSRZ	
1	А	103	ALA	3.8	
1	А	185 LEU		3.8	
1	А	100 ASN		3.6	
1	А	331	ALA	3.6	
1	А	105	VAL	3.5	
1	А	361	LEU	3.4	
1	А	389	PHE	3.4	
1	А	357	LEU	3.4	
1	А	241	LEU	3.4	
1	А	49	VAL	3.4	
1	А	239	HIS	3.4	
1	А	234	THR	3.4	
1	А	99	MET	3.3	
1	А	178	SER	3.3	
1	А	447	PHE	3.3	
1	А	156	LEU	3.3	
1	А	390	ILE	3.3	
1	А	151	ALA	3.2	
1	А	238	THR	3.2	
1	А	249	ARG	3.2	
1	А	236	ASP	3.2	
1	А	12	LEU	3.1	
1	А	53	GLY	3.1	
1	А	453	PRO	3.1	
1	А	225	LYS	3.1	
1	А	333	ARG	3.0	
1	А	52	TYR	3.0	
1	А	80	SER	3.0	
1	А	134	GLY	3.0	
1	А	104	PHE	3.0	
1	А	137	THR	2.9	
1	А	362	PRO	2.8	
1	A	149	PRO	2.7	
1	А	260	LEU	2.7	
1	А	274	ILE	2.7	
1	А	243	GLU	2.6	
1	А	393	GLU	2.6	
1	А	98	MET	2.6	
1	А	79	GLY	2.6	
1	А	358	LEU	2.6	
1	А	440	LEU	2.6	
1	А	183	LYS	2.6	

Continued from previous page...

Continued on next page...



Continued from previous page						
Mol	Chain	Res	Type	RSRZ		
1	А	250	GLN	2.6		
1	А	206	PHE	2.6		
1	А	26	PHE	2.6		
1	А	106	SER	2.6		
1	А	270	GLN	2.5		
1	А	14	LEU	2.5		
1	А	328	VAL	2.5		
1	А	325	LEU	2.5		
1	А	179	ILE	2.4		
1	А	329	GLU	2.4		
1	А	29	ASN	2.4		
1	А	96	MET	2.4		
1	А	83	VAL	2.4		
1	А	128	ILE	2.4		
1	А	81	PHE	2.4		
1	А	107	ALA	2.4		
1	А	246	GLU	2.4		
1	А	443	LEU	2.4		
1	А	117	LYS	2.3		
1	А	229	LYS	2.3		
1	А	10	GLY	2.3		
1	А	253	ARG	2.3		
1	А	359	GLU	2.3		
1	А	139	PHE	2.3		
1	А	240	ASP	2.3		
1	А	215	LEU	2.2		
1	A	394	LEU	2.2		
1	A	97	LEU	2.2		
1	А	353	ILE	2.2		
1	А	72	PHE	2.2		
1	A	130	GLY	2.2		
1	A	277	VAL	2.1		
1	A	$\overline{50}$	HIS	2.1		
1	A	448	THR	2.1		
1	A	245	LYS	2.1		
1	A	22	GLY	2.0		
1	A	138	GLY	2.0		

6.2Non-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 monosaccharides 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^{th} 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	$\mathbf{B} ext{-factors}(\mathbf{A}^2)$	Q<0.9
2	BNG	А	502	21/21	0.86	0.26	$94,\!133,\!140,\!145$	0
4	CL	А	504	1/1	0.87	0.26	$106,\!106,\!106,\!106$	0
3	P33	А	503	22/22	0.90	0.36	73,111,142,144	0
2	BNG	А	501	21/21	0.91	0.36	85,98,123,128	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.













6.5 Other polymers (i)

There are no such residues in this entry.

