



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 09:16 pm GMT

PDB ID : 2HZL
Title : Crystal structures of a sodium-alpha-keto acid binding subunit from a TRAP transporter in its closed forms
Authors : Gonin, S.; Arnoux, P.; Pierru, B.; Alonso, B.; Sabaty, M.; Pignol, D.
Deposited on : 2006-08-09
Resolution : 1.40 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

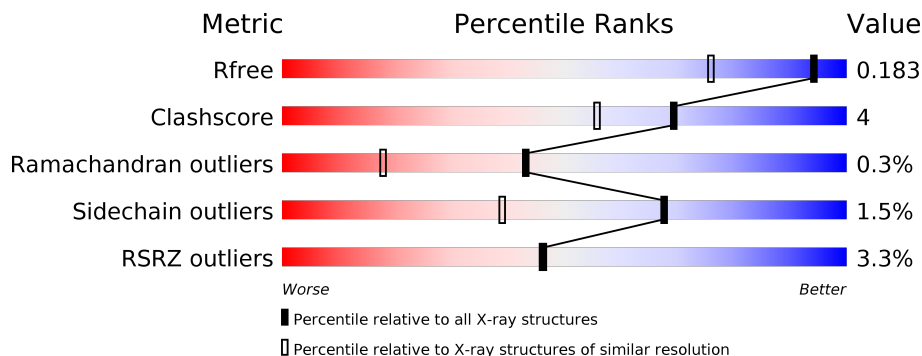
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	1450 (1.40-1.40)
Clashscore	122126	1541 (1.40-1.40)
Ramachandran outliers	120053	1500 (1.40-1.40)
Sidechain outliers	120020	1499 (1.40-1.40)
RSRZ outliers	108989	1412 (1.40-1.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	365	 3% 83% 8% 8%
1	B	365	 3% 85% 7% 8%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5795 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.

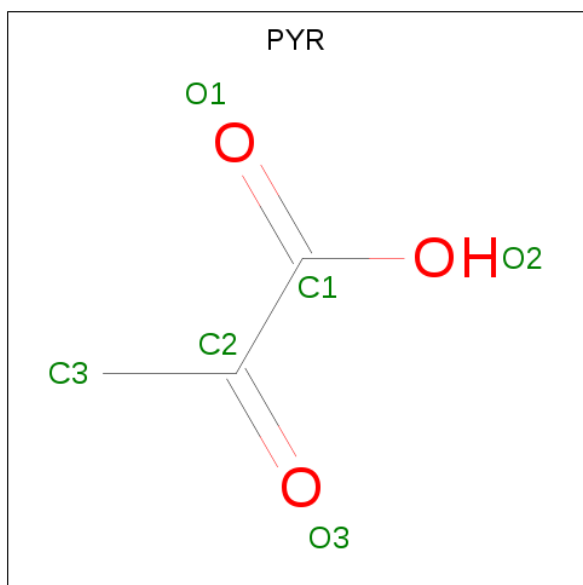
- Molecule 1 is a protein called TRAP-T family sorbitol/mannitol transporter, periplasmic binding protein, SmoM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	334	Total 2632	C 1701	N 430	O 487	S 14	0	4	0
1	B	337	Total 2672	C 1727	N 437	O 494	S 14	0	6	0

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Na 1	0	0
2	A	1	Total 1	Na 1	0	0

- Molecule 3 is PYRUVIC ACID (three-letter code: PYR) (formula: C₃H₄O₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0

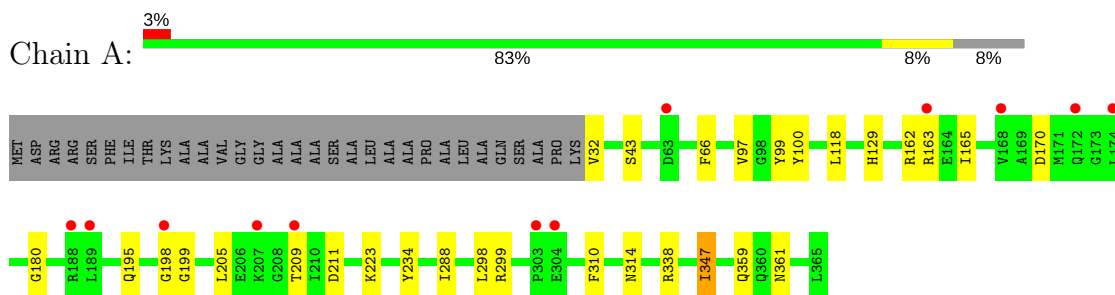
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	222	Total O 222 222	0	0
4	B	249	Total O 249 249	0	0

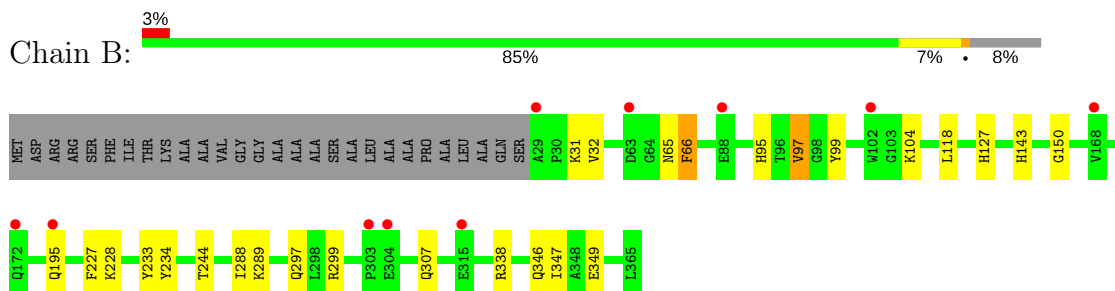
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: TRAP-T family sorbitol/mannitol transporter, periplasmic binding protein, SmoM



- Molecule 1: TRAP-T family sorbitol/mannitol transporter, periplasmic binding protein, SmoM



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	118.02Å 78.06Å 95.14Å 90.00° 124.98° 90.00°	Depositor
Resolution (Å)	39.01 – 1.40 39.03 – 1.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (39.01-1.40) 98.9 (39.03-1.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	79.12 (at 1.40Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.173 , 0.184 0.172 , 0.183	Depositor DCC
R_{free} test set	6908 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	9.4	Xtrriage
Anisotropy	0.068	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 49.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5795	wwPDB-VP
Average B, all atoms (Å ²)	11.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.03% 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

5.1 Standard geometry

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

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	0.39	0/2723	0.55	0/3703
1	B	0.39	0/2758	0.56	0/3751
All	All	0.39	0/5481	0.55	0/7454

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

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	2632	0	2521	25	0
1	B	2672	0	2561	23	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	6	0	3	0	0
3	B	12	0	6	0	0
4	A	222	0	0	3	0
4	B	249	0	0	6	0
All	All	5795	0	5091	42	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 4.

All (42) 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:B:349:GLU:HG3	4:B:747:HOH:O	1.60	1.00
1:A:180:GLY:H	1:A:195:GLN:HE21	1.27	0.83
1:B:297:GLN:HE21	1:B:299:ARG:HH22	1.32	0.76
1:A:359:GLN:NE2	1:B:307:GLN:HE22	1.87	0.73
1:A:198:GLY:HA2	4:A:823:HOH:O	1.89	0.72
1:A:359:GLN:HE22	1:B:307:GLN:HE22	1.37	0.72
1:A:162:ARG:NH1	1:A:211:ASP:HA	2.08	0.68
1:A:199:GLY:N	4:A:823:HOH:O	2.14	0.68
1:B:228:LYS:HD2	4:B:704:HOH:O	1.94	0.67
1:A:43:SER:O	1:A:223:LYS:HE2	1.96	0.66
1:A:97[A]:VAL:HG11	1:A:99:TYR:CE2	2.32	0.64
1:A:180:GLY:H	1:A:195:GLN:NE2	1.95	0.63
1:A:361:ASN:HD21	1:B:346:GLN:HE22	1.45	0.62
1:A:162:ARG:HH12	1:A:211:ASP:HA	1.65	0.61
1:A:97[A]:VAL:HG11	1:A:99:TYR:CZ	2.37	0.60
1:B:227:PHE:CZ	1:B:228:LYS:HE3	2.38	0.58
1:B:97[B]:VAL:HG11	1:B:99:TYR:CE2	2.38	0.58
1:A:100:TYR:HA	1:A:195:GLN:HE22	1.69	0.57
1:B:127:HIS:HE1	1:B:244:THR:OG1	1.87	0.56
1:B:95:HIS:HD2	4:B:610:HOH:O	1.87	0.55
1:B:97[B]:VAL:HG11	1:B:99:TYR:CZ	2.44	0.53
1:A:361:ASN:HD21	1:B:346:GLN:NE2	2.07	0.52
1:A:347[B]:ILE:HD12	1:B:288[B]:ILE:HD13	1.93	0.51
1:B:118:LEU:HD21	1:B:338:ARG:HA	1.92	0.51
1:A:288:ILE:HD11	1:A:298:LEU:HD21	1.92	0.51
1:A:310:PHE:CE1	1:A:314[A]:ASN:ND2	2.77	0.51
1:A:32:VAL:N	4:A:745:HOH:O	2.43	0.51
1:A:347[B]:ILE:CD1	1:B:288[B]:ILE:HD13	2.42	0.49
1:A:162:ARG:NH2	1:A:205:LEU:O	2.45	0.49
1:B:297:GLN:HE21	1:B:299:ARG:NH2	2.05	0.49
1:A:129:HIS:HD2	4:B:805:HOH:O	1.96	0.48
1:B:32[A]:VAL:HG13	1:B:66:PHE:HA	1.94	0.48
1:B:104:LYS:HE3	1:B:143:HIS:CE1	2.48	0.47
1:B:31:LYS:H	1:B:65:ASN:HD22	1.63	0.46
1:B:127:HIS:HD2	4:B:643:HOH:O	1.99	0.46
1:B:95:HIS:HE1	1:B:150:GLY:O	1.99	0.46
1:A:234:TYR:HA	1:A:299:ARG:O	2.16	0.46
1:A:118:LEU:HD21	1:A:338:ARG:HA	2.00	0.43
1:B:234:TYR:HA	1:B:299:ARG:O	2.19	0.43
1:B:127:HIS:CD2	4:B:643:HOH:O	2.70	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:165:ILE:HG22	1:A:170:ASP:HB3	2.01	0.42
1:A:97[A]:VAL:CG1	1:A:99:TYR:CE2	3.03	0.41

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	336/365 (92%)	325 (97%)	9 (3%)	2 (1%)	27 7
1	B	341/365 (93%)	330 (97%)	9 (3%)	2 (1%)	27 7
All	All	677/730 (93%)	655 (97%)	18 (3%)	4 (1%)	43 7

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	347[A]	ILE
1	A	347[B]	ILE
1	B	347[A]	ILE
1	B	347[B]	ILE

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	268/282 (95%)	265 (99%)	3 (1%)	76	53
1	B	272/282 (96%)	266 (98%)	6 (2%)	55	20
All	All	540/564 (96%)	531 (98%)	9 (2%)	67	33

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	PHE
1	A	163	ARG
1	A	209	THR
1	B	66	PHE
1	B	97[A]	VAL
1	B	97[B]	VAL
1	B	195	GLN
1	B	233	TYR
1	B	289	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	124	ASN
1	A	129	HIS
1	A	195	GLN
1	A	262	GLN
1	A	275	ASN
1	A	359	GLN
1	B	65	ASN
1	B	95	HIS
1	B	127	HIS
1	B	166	ASN
1	B	275	ASN
1	B	297	GLN
1	B	346	GLN

5.3.3 RNA ⓘ

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 2 are 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	PYR	A	601	2	2,5,5	1.71	1 (50%)	2,6,6	0.74	0
3	PYR	B	602	2	2,5,5	1.83	1 (50%)	2,6,6	0.49	0
3	PYR	B	603	-	2,5,5	1.83	1 (50%)	2,6,6	0.36	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	PYR	A	601	2	-	0/0/4/4	0/0/0/0
3	PYR	B	602	2	-	0/0/4/4	0/0/0/0
3	PYR	B	603	-	-	0/0/4/4	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	PYR	O3-C2	2.42	1.30	1.22
3	B	603	PYR	O3-C2	2.58	1.31	1.22
3	B	602	PYR	O3-C2	2.59	1.31	1.22

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95th 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	A	334/365 (91%)	0.25	12 (3%) 42 42	4, 10, 21, 27	0
1	B	337/365 (92%)	0.21	10 (2%) 50 50	4, 9, 18, 25	0
All	All	671/730 (91%)	0.23	22 (3%) 46 46	4, 10, 20, 27	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	207	LYS	4.1
1	A	172	GLN	3.4
1	A	63	ASP	3.2
1	A	189	LEU	3.2
1	A	304	GLU	3.0
1	A	168	VAL	2.7
1	B	195	GLN	2.6
1	B	102	TRP	2.6
1	B	63	ASP	2.6
1	A	303	PRO	2.5
1	A	209	THR	2.5
1	B	304	GLU	2.4
1	B	29	ALA	2.4
1	B	303	PRO	2.4
1	B	315	GLU	2.3
1	A	163	ARG	2.2
1	A	188	ARG	2.1
1	A	174	LEU	2.1
1	B	168	VAL	2.1
1	B	88	GLU	2.1
1	A	198	GLY	2.1
1	B	172	GLN	2.0

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 carbohydrates 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, 95th 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(Å ²)	Q<0.9
3	PYR	B	603	6/6	0.79	0.24	24,25,26,26	0
3	PYR	A	601	6/6	0.97	0.08	8,9,9,10	0
3	PYR	B	602	6/6	0.97	0.09	8,8,9,10	0
2	NA	B	501	1/1	1.00	0.12	4,4,4,4	0
2	NA	A	500	1/1	1.00	0.10	4,4,4,4	0

6.5 Other polymers [i](#)

There are no such residues in this entry.