W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 20, 2024 – 06:55 AM EST

|		
PDB ID	:	4M64
Title	:	3D crystal structure of Na+/melibiose symporter of Salmonella typhimurium
Authors	:	Ethayathulla, A.S.; Guan, L.
Deposited on	:	2013-08-08
Resolution	:	3.35 Å(reported)

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.
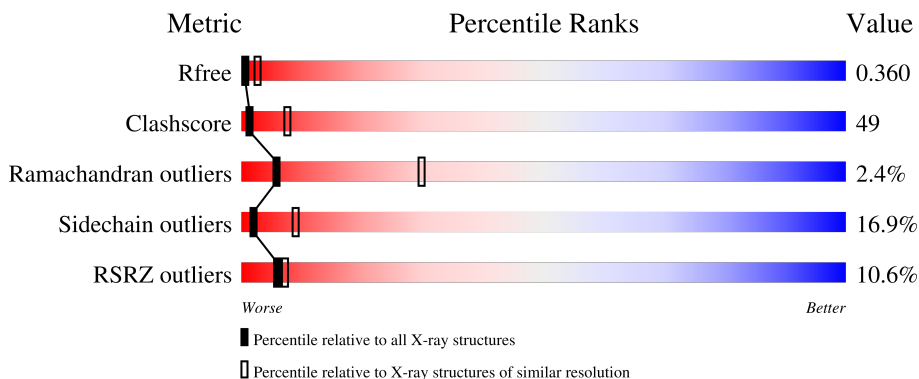
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (3.42-3.30)

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 <=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	486	
1	B	486	
1	C	486	
1	D	486	

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 12982 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 Melibiose carrier protein.

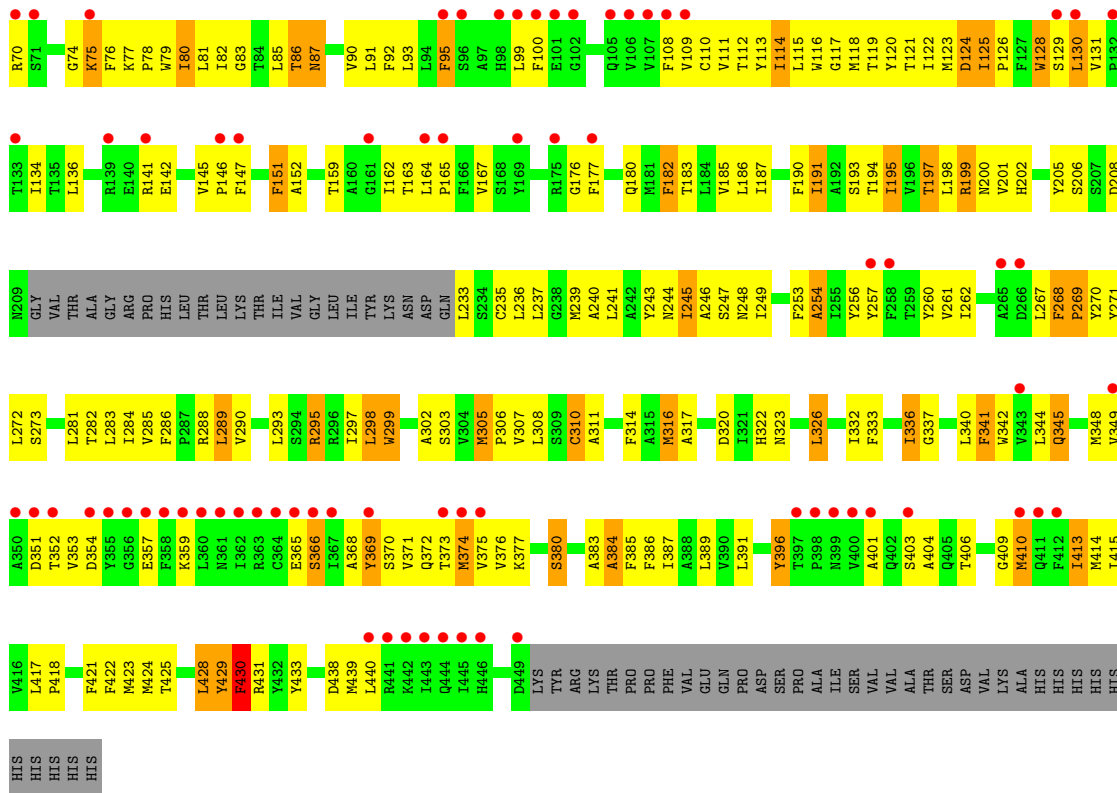| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	448	Total	C	N	O	S	0	0	0
			3442	2295	540	585	22			
1	B	426	Total	C	N	O	S	0	0	0
			3258	2176	503	558	21			
1	C	431	Total	C	N	O	S	0	0	0
			3352	2242	520	568	22			
1	D	382	Total	C	N	O	S	0	0	0
			2930	1965	449	497	19			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	MET	LEU	engineered mutation	UNP P30878
A	477	HIS	-	expression tag	UNP P30878
A	478	HIS	-	expression tag	UNP P30878
A	479	HIS	-	expression tag	UNP P30878
A	480	HIS	-	expression tag	UNP P30878
A	481	HIS	-	expression tag	UNP P30878
A	482	HIS	-	expression tag	UNP P30878
A	483	HIS	-	expression tag	UNP P30878
A	484	HIS	-	expression tag	UNP P30878
A	485	HIS	-	expression tag	UNP P30878
A	486	HIS	-	expression tag	UNP P30878
B	5	MET	LEU	engineered mutation	UNP P30878
B	477	HIS	-	expression tag	UNP P30878
B	478	HIS	-	expression tag	UNP P30878
B	479	HIS	-	expression tag	UNP P30878
B	480	HIS	-	expression tag	UNP P30878
B	481	HIS	-	expression tag	UNP P30878
B	482	HIS	-	expression tag	UNP P30878
B	483	HIS	-	expression tag	UNP P30878
B	484	HIS	-	expression tag	UNP P30878
B	485	HIS	-	expression tag	UNP P30878

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	486	HIS	-	expression tag	UNP P30878
C	5	MET	LEU	engineered mutation	UNP P30878
C	477	HIS	-	expression tag	UNP P30878
C	478	HIS	-	expression tag	UNP P30878
C	479	HIS	-	expression tag	UNP P30878
C	480	HIS	-	expression tag	UNP P30878
C	481	HIS	-	expression tag	UNP P30878
C	482	HIS	-	expression tag	UNP P30878
C	483	HIS	-	expression tag	UNP P30878
C	484	HIS	-	expression tag	UNP P30878
C	485	HIS	-	expression tag	UNP P30878
C	486	HIS	-	expression tag	UNP P30878
D	5	MET	LEU	engineered mutation	UNP P30878
D	477	HIS	-	expression tag	UNP P30878
D	478	HIS	-	expression tag	UNP P30878
D	479	HIS	-	expression tag	UNP P30878
D	480	HIS	-	expression tag	UNP P30878
D	481	HIS	-	expression tag	UNP P30878
D	482	HIS	-	expression tag	UNP P30878
D	483	HIS	-	expression tag	UNP P30878
D	484	HIS	-	expression tag	UNP P30878
D	485	HIS	-	expression tag	UNP P30878
D	486	HIS	-	expression tag	UNP P30878

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: Melibiose carrier protein

• Molecule 1: Melibiose carrier protein

• Molecule 1: Melibiose carrier protein

Chain C:

ALA HIS HIS HIS HIS HIS HIS HIS HIS

• Molecule 1: Melibiose carrier protein

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	127.21Å 127.21Å 206.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.64 – 3.35 38.64 – 3.35	Depositor EDS
% Data completeness (in resolution range)	99.4 (38.64-3.35) 99.5 (38.64-3.35)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	47.41 (at 3.32Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.312 , 0.359 0.311 , 0.360	Depositor DCC
R_{free} test set	2712 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	94.7	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.16 , 85.2	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.41, < L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.011 for -h,-k,l 0.039 for h,-h-k,-l 0.012 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	12982	wwPDB-VP
Average B, all atoms (Å^2)	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 37.36 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.3433e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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 ⓘ

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	0.63	0/3530	0.87	2/4813 (0.0%)
1	B	0.60	0/3342	0.82	3/4561 (0.1%)
1	C	0.66	0/3440	0.91	1/4689 (0.0%)
1	D	0.59	0/3004	0.81	1/4097 (0.0%)
All	All	0.62	0/13316	0.86	7/18160 (0.0%)

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
1	A	0	1
1	B	0	4
1	C	0	4
1	D	0	1
All	All	0	10

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	199	ARG	NE-CZ-NH1	5.77	123.19	120.30
1	C	55	ASP	CB-CG-OD2	-5.76	113.11	118.30
1	A	55	ASP	CB-CG-OD2	5.43	123.18	118.30
1	B	308	LEU	CA-CB-CG	5.24	127.36	115.30
1	A	340	LEU	CB-CG-CD2	-5.21	102.14	111.00

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Res	Type	Group
1	A	65	ILE	Peptide
1	B	254	ALA	Peptide
1	B	38	GLY	Peptide
1	B	429	TYR	Peptide
1	B	74	GLY	Peptide

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	3442	0	3479	346	0
1	B	3258	0	3249	299	0
1	C	3352	0	3433	352	0
1	D	2930	0	2965	285	0
All	All	12982	0	13126	1273	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 49.

The worst 5 of 1273 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:D:299:TRP:CB	1:D:344:LEU:HD13	1.63	1.26
1:D:86:THR:CG2	1:D:115:LEU:HG	1.65	1.26
1:C:75:LYS:HG3	1:C:205:TYR:CB	1.77	1.15
1:D:299:TRP:HB2	1:D:344:LEU:CD1	1.77	1.15
1:C:128:TRP:CH2	1:C:373:THR:HG21	1.83	1.13

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 X-ray entries followed by that with respect to entries of similar resolution.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

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	446/486 (92%)	365 (82%)	65 (15%)	16 (4%)	3	22
1	B	422/486 (87%)	360 (85%)	54 (13%)	8 (2%)	8	34
1	C	429/486 (88%)	358 (83%)	63 (15%)	8 (2%)	8	34
1	D	374/486 (77%)	314 (84%)	52 (14%)	8 (2%)	7	32
All	All	1671/1944 (86%)	1397 (84%)	234 (14%)	40 (2%)	6	30

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	134	ILE
1	A	269	PRO
1	A	271	TYR
1	B	430	PHE
1	C	134	ILE

5.3.2 Protein sidechains ⓘ

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	354/406 (87%)	303 (86%)	51 (14%)	3	14
1	B	331/406 (82%)	272 (82%)	59 (18%)	2	7
1	C	353/406 (87%)	281 (80%)	72 (20%)	1	4
1	D	302/406 (74%)	258 (85%)	44 (15%)	3	13
All	All	1340/1624 (82%)	1114 (83%)	226 (17%)	2	9

5 of 226 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	33	TYR
1	D	386	PHE
1	C	181	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	344	LEU
1	D	138	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	105	GLN
1	C	411	GLN
1	D	244	ASN
1	C	402	GLN
1	B	67	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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(Å²)	Q<0.9
1	A	448/486 (92%)	0.03	30 (6%)	17	20	27, 80, 183, 265	25 (5%)
1	B	426/486 (87%)	0.76	78 (18%)	1	1	28, 89, 247, 360	9 (2%)
1	C	431/486 (88%)	0.03	24 (5%)	24	26	24, 76, 168, 340	0
1	D	382/486 (78%)	0.51	46 (12%)	4	5	31, 85, 220, 331	1 (0%)
All	All	1687/1944 (86%)	0.32	178 (10%)	6	7	24, 82, 213, 360	35 (2%)

The worst 5 of 178 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	362	ILE	21.4
1	B	350	ALA	19.3
1	C	321	ILE	16.4
1	B	351	ASP	15.7
1	D	412	PHE	15.5

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.