



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:35 PM GMT

PDB ID : 4IU8
Title : Crystal structure of a membrane transporter (selenomethionine derivative)
Authors : Yan, H.; Huang, W.; Yan, C.; Gong, X.; Jiang, S.; Zhao, Y.; Wang, J.; Shi, Y.
Deposited on : 2013-01-20
Resolution : 3.11 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
<http://wwpdb.org/validation/2016/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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

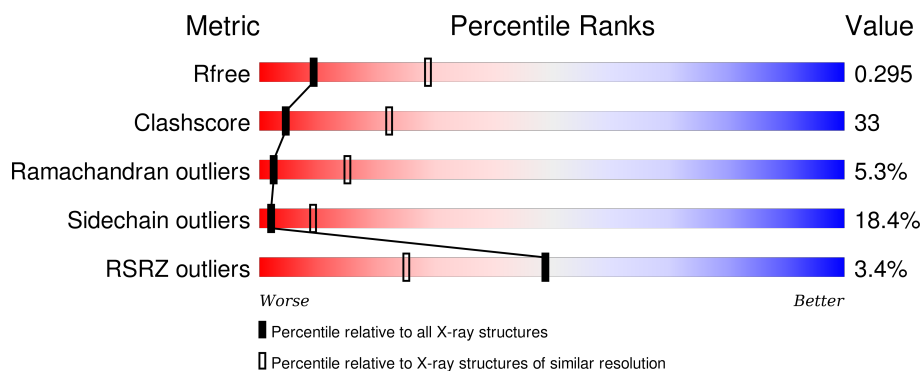
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.11 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}	91344	1112 (3.16-3.08)
Clashscore	102246	1218 (3.16-3.08)
Ramachandran outliers	100387	1175 (3.16-3.08)
Sidechain outliers	100360	1175 (3.16-3.08)
RSRZ outliers	91569	1114 (3.16-3.08)

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	468	<div> <div>2%</div> <div>41%</div> <div>38%</div> <div>11%</div> <div>10%</div> </div>
1	B	468	<div> <div>4%</div> <div>39%</div> <div>43%</div> <div>9%</div> <div>9%</div> </div>

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
2	NO3	A	501	-	-	-	X

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6214 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.

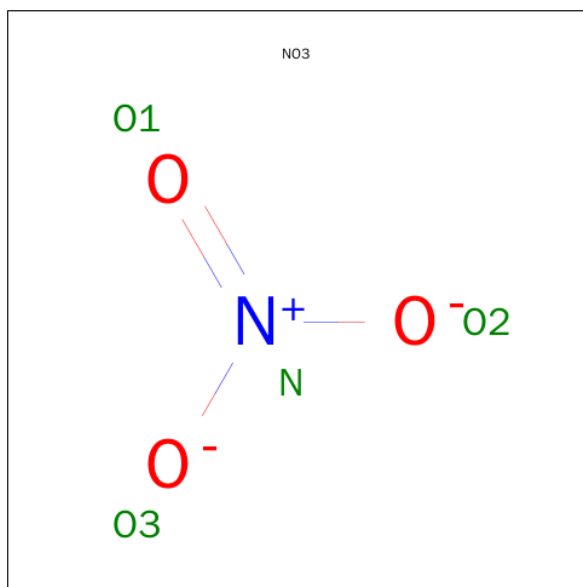
- Molecule 1 is a protein called Nitrite extrusion protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	Se	0	0	0
			3081	2059	490	514	5	13			
1	B	427	Total	C	N	O	S	Se	0	0	0
			3129	2092	500	519	5	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	463	LEU	-	EXPRESSION TAG	UNP P37758
A	464	GLU	-	EXPRESSION TAG	UNP P37758
A	465	VAL	-	EXPRESSION TAG	UNP P37758
A	466	LEU	-	EXPRESSION TAG	UNP P37758
A	467	PHE	-	EXPRESSION TAG	UNP P37758
A	468	GLN	-	EXPRESSION TAG	UNP P37758
B	463	LEU	-	EXPRESSION TAG	UNP P37758
B	464	GLU	-	EXPRESSION TAG	UNP P37758
B	465	VAL	-	EXPRESSION TAG	UNP P37758
B	466	LEU	-	EXPRESSION TAG	UNP P37758
B	467	PHE	-	EXPRESSION TAG	UNP P37758
B	468	GLN	-	EXPRESSION TAG	UNP P37758

- Molecule 2 is NITRATE ION (three-letter code: NO3) (formula: NO₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	N	O	0	0
			4	1	3		

W215	R290	A359	M427
L216	L291	G360	M428
W217	A292	L361	T429
V218	F293	G362	G430
P219	F294	S363	S431
L220	G295	T366	P432
L221	P296	F367	M433
A222	F297	Q368	G434
T225	I298	M369	A435
A228	A300	M369	M436
W229	I301	A371	K437
S230	A302	V372	L440
G231	R303	I373	I441
M232	V304	F374	V445
I235	V305	Q376	C446
ALA	I309	I377	V447
SER	S310	T378	L448
SER	F313	I379	L449
SER	G314	T380	T450
ARG	G315	R381	V453
ALA	SER	VAL	V454
SER	V318	LYS	GLY
A243	T319	MSE	ARG
D244	L320	LYS	ARG
P247	F323	GLY	LYS
L252	I324	GLY	LYS
W255	F325	PHE	PHE
L256	M326	SER	SER
L257	F329	ASP	GLN
S258	S330	GLU	LYS
L259	A331	GLN	LEU
L262	L332	A392	LEU
A263	L333	H393	GLU
T264	F334	K394	VAL
F265	L335	E395	LEU
G266	T336	A401	PHE
S267	L337	A402	GLN
F268	P338	A403	
F269	G339	L404	
G270	T340	G405	
F271	G341	F406	
S272	S342	I407	
A273	G343	I410	
A276	N344	V413	
M277	F345	G414	
L278	F348	G415	
A279	Y349	F416	
Q282	F352	I418	
I288	M353	P419	
L289	F356	Q420	
	I357	A421	
	T358	F422	
		S425	
		L426	

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	106.58Å 117.50Å 127.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.11 40.03 – 3.11	Depositor EDS
% Data completeness (in resolution range)	82.7 (40.00-3.11) 82.9 (40.03-3.11)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.37 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.239 , 0.299 0.240 , 0.295	Depositor DCC
R_{free} test set	1222 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	79.4	Xtriage
Anisotropy	1.095	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 45.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 24412 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	6214	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.92% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NO3

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	0.67	5/3150 (0.2%)	0.79	3/4276 (0.1%)
1	B	0.71	3/3202 (0.1%)	0.83	4/4349 (0.1%)
All	All	0.69	8/6352 (0.1%)	0.81	7/8625 (0.1%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	215	TRP	CD2-CE2	7.37	1.50	1.41
1	B	215	TRP	CD2-CE2	6.32	1.49	1.41
1	A	229	TRP	CD2-CE2	5.53	1.48	1.41
1	A	451	TRP	CD2-CE2	5.34	1.47	1.41
1	A	50	TRP	CD2-CE2	5.20	1.47	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	353	MSE	CG-SE-CE	6.04	112.19	98.90
1	A	52	LEU	CA-CB-CG	6.02	129.14	115.30
1	B	277	MSE	CA-CB-CG	-5.87	103.32	113.30
1	B	247	PRO	N-CA-CB	5.84	110.31	103.30
1	B	33	ARG	NE-CZ-NH1	5.21	122.90	120.30

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	3081	0	3049	207	0
1	B	3129	0	3092	197	0
2	A	4	0	0	1	0
All	All	6214	0	6141	402	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 33.

The worst 5 of 402 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:123:ASN:O	1:A:126:THR:HG22	1.47	1.11
1:B:12:LEU:HD23	1:B:13:LEU:H	1.12	1.08
1:A:63:ILE:HG22	1:A:64:GLY:H	1.23	1.03
1:A:88:VAL:HB	1:A:89:PRO:HD3	1.41	1.01
1:B:93:MSE:HG2	1:B:102:TRP:CH2	1.96	1.00

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.

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	416/468 (89%)	308 (74%)	83 (20%)	25 (6%)	2	11
1	B	421/468 (90%)	327 (78%)	75 (18%)	19 (4%)	3	18
All	All	837/936 (89%)	635 (76%)	158 (19%)	44 (5%)	2	15

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	100	ARG
1	A	231	GLY
1	A	305	VAL
1	A	312	LYS
1	A	432	PRO

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	299/356 (84%)	246 (82%)	53 (18%)	2	10
1	B	304/356 (85%)	246 (81%)	58 (19%)	2	8
All	All	603/712 (85%)	492 (82%)	111 (18%)	2	9

5 of 111 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	436	MSE
1	B	78	LEU
1	B	406	PHE
1	A	448	LEU
1	B	40	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	173	ASN
1	B	420	GLN
1	A	287	ASN
1	A	122	GLN
1	B	173	ASN

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](#)

1 ligand is 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	NO3	A	501	-	3,3,3	2.64	2 (66%)	3,3,3	0.48	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
2	NO3	A	501	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	NO3	O2-N	3.18	1.41	1.25
2	A	501	NO3	O3-N	3.25	1.42	1.25

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	NO3	1	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 [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, 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	408/468 (87%)	-0.25	11 (2%) 58 35	18, 43, 91, 145	0
1	B	414/468 (88%)	-0.16	17 (4%) 41 19	16, 33, 96, 141	0
All	All	822/936 (87%)	-0.21	28 (3%) 49 25	16, 37, 94, 145	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	378	THR	4.9
1	B	393	HIS	4.6
1	B	244	ASP	4.5
1	A	12	LEU	4.5
1	B	12	LEU	4.0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	NO3	A	501	4/4	0.85	0.85	12.90	64,68,68,70	0

6.5 Other polymers ⓘ

There are no such residues in this entry.