



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 01:51 PM GMT

PDB ID : 2B5F
Title : Crystal structure of the spinach aquaporin SoPIP2;1 in an open conformation to 3.9 resolution
Authors : Tornroth-Horsefield, S.; Wang, Y.; Hedfalk, K.; Johanson, U.; Karlsson, M.; Tajkhorshid, E.; Neutze, R.; Kjellbom, P.
Deposited on : 2005-09-28
Resolution : 3.90 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

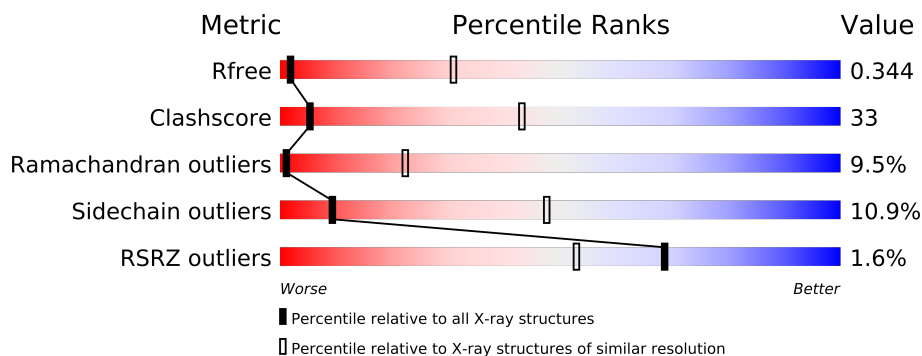
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.90 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}	66092	1022 (4.38-3.42)
Clashscore	79885	1173 (4.30-3.50)
Ramachandran outliers	78287	1118 (4.30-3.50)
Sidechain outliers	78261	1107 (4.30-3.50)
RSRZ outliers	66119	1000 (4.36-3.44)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	303	
1	B	303	
1	C	303	
1	D	303	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 6854 atoms, of which 0 are hydrogen and 0 are deuterium.

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 aquaporin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	236	Total	C	N	O	S	0	0	0
			1767	1175	288	296	8			
1	B	224	Total	C	N	O	S	0	0	0
			1671	1116	267	280	8			
1	C	227	Total	C	N	O	S	0	0	0
			1696	1134	271	283	8			
1	D	230	Total	C	N	O	S	0	0	0
			1720	1150	276	286	8			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	282	LEU	-	EXPRESSION TAG	UNP Q41372
A	283	GLU	-	EXPRESSION TAG	UNP Q41372
A	284	GLN	-	EXPRESSION TAG	UNP Q41372
A	285	LYS	-	EXPRESSION TAG	UNP Q41372
A	286	LEU	-	EXPRESSION TAG	UNP Q41372
A	287	ILE	-	EXPRESSION TAG	UNP Q41372
A	288	SER	-	EXPRESSION TAG	UNP Q41372
A	289	GLU	-	EXPRESSION TAG	UNP Q41372
A	290	GLU	-	EXPRESSION TAG	UNP Q41372
A	291	ASP	-	EXPRESSION TAG	UNP Q41372
A	292	LEU	-	EXPRESSION TAG	UNP Q41372
A	293	ASN	-	EXPRESSION TAG	UNP Q41372
A	294	SER	-	EXPRESSION TAG	UNP Q41372
A	295	ALA	-	EXPRESSION TAG	UNP Q41372
A	296	VAL	-	EXPRESSION TAG	UNP Q41372
A	297	ASP	-	EXPRESSION TAG	UNP Q41372
A	298	HIS	-	EXPRESSION TAG	UNP Q41372
A	299	HIS	-	EXPRESSION TAG	UNP Q41372
A	300	HIS	-	EXPRESSION TAG	UNP Q41372
A	301	HIS	-	EXPRESSION TAG	UNP Q41372
A	302	HIS	-	EXPRESSION TAG	UNP Q41372

Continued on next page...

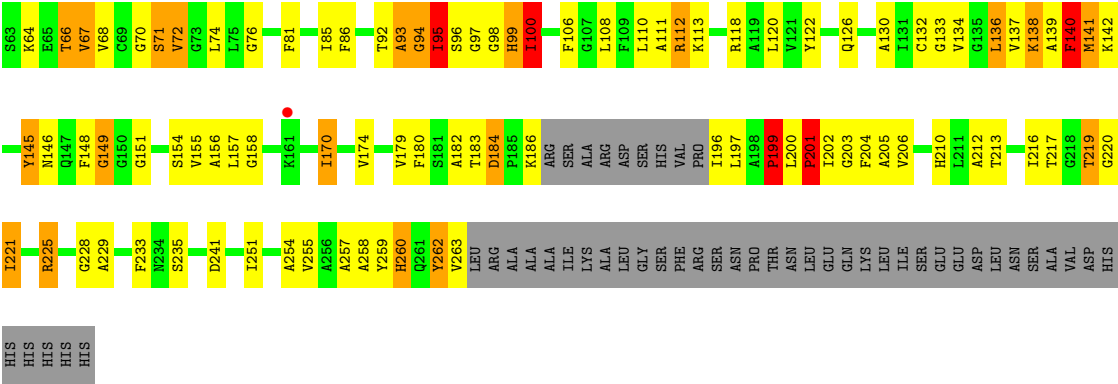
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	303	HIS	-	EXPRESSION TAG	UNP Q41372
B	282	LEU	-	EXPRESSION TAG	UNP Q41372
B	283	GLU	-	EXPRESSION TAG	UNP Q41372
B	284	GLN	-	EXPRESSION TAG	UNP Q41372
B	285	LYS	-	EXPRESSION TAG	UNP Q41372
B	286	LEU	-	EXPRESSION TAG	UNP Q41372
B	287	ILE	-	EXPRESSION TAG	UNP Q41372
B	288	SER	-	EXPRESSION TAG	UNP Q41372
B	289	GLU	-	EXPRESSION TAG	UNP Q41372
B	290	GLU	-	EXPRESSION TAG	UNP Q41372
B	291	ASP	-	EXPRESSION TAG	UNP Q41372
B	292	LEU	-	EXPRESSION TAG	UNP Q41372
B	293	ASN	-	EXPRESSION TAG	UNP Q41372
B	294	SER	-	EXPRESSION TAG	UNP Q41372
B	295	ALA	-	EXPRESSION TAG	UNP Q41372
B	296	VAL	-	EXPRESSION TAG	UNP Q41372
B	297	ASP	-	EXPRESSION TAG	UNP Q41372
B	298	HIS	-	EXPRESSION TAG	UNP Q41372
B	299	HIS	-	EXPRESSION TAG	UNP Q41372
B	300	HIS	-	EXPRESSION TAG	UNP Q41372
B	301	HIS	-	EXPRESSION TAG	UNP Q41372
B	302	HIS	-	EXPRESSION TAG	UNP Q41372
B	303	HIS	-	EXPRESSION TAG	UNP Q41372
C	282	LEU	-	EXPRESSION TAG	UNP Q41372
C	283	GLU	-	EXPRESSION TAG	UNP Q41372
C	284	GLN	-	EXPRESSION TAG	UNP Q41372
C	285	LYS	-	EXPRESSION TAG	UNP Q41372
C	286	LEU	-	EXPRESSION TAG	UNP Q41372
C	287	ILE	-	EXPRESSION TAG	UNP Q41372
C	288	SER	-	EXPRESSION TAG	UNP Q41372
C	289	GLU	-	EXPRESSION TAG	UNP Q41372
C	290	GLU	-	EXPRESSION TAG	UNP Q41372
C	291	ASP	-	EXPRESSION TAG	UNP Q41372
C	292	LEU	-	EXPRESSION TAG	UNP Q41372
C	293	ASN	-	EXPRESSION TAG	UNP Q41372
C	294	SER	-	EXPRESSION TAG	UNP Q41372
C	295	ALA	-	EXPRESSION TAG	UNP Q41372
C	296	VAL	-	EXPRESSION TAG	UNP Q41372
C	297	ASP	-	EXPRESSION TAG	UNP Q41372
C	298	HIS	-	EXPRESSION TAG	UNP Q41372
C	299	HIS	-	EXPRESSION TAG	UNP Q41372
C	300	HIS	-	EXPRESSION TAG	UNP Q41372

Continued on next page...

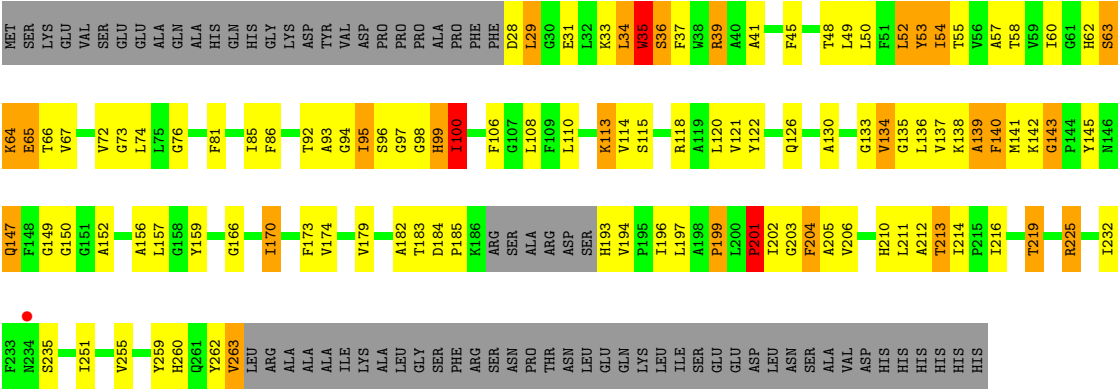
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	301	HIS	-	EXPRESSION TAG	UNP Q41372
C	302	HIS	-	EXPRESSION TAG	UNP Q41372
C	303	HIS	-	EXPRESSION TAG	UNP Q41372
D	282	LEU	-	EXPRESSION TAG	UNP Q41372
D	283	GLU	-	EXPRESSION TAG	UNP Q41372
D	284	GLN	-	EXPRESSION TAG	UNP Q41372
D	285	LYS	-	EXPRESSION TAG	UNP Q41372
D	286	LEU	-	EXPRESSION TAG	UNP Q41372
D	287	ILE	-	EXPRESSION TAG	UNP Q41372
D	288	SER	-	EXPRESSION TAG	UNP Q41372
D	289	GLU	-	EXPRESSION TAG	UNP Q41372
D	290	GLU	-	EXPRESSION TAG	UNP Q41372
D	291	ASP	-	EXPRESSION TAG	UNP Q41372
D	292	LEU	-	EXPRESSION TAG	UNP Q41372
D	293	ASN	-	EXPRESSION TAG	UNP Q41372
D	294	SER	-	EXPRESSION TAG	UNP Q41372
D	295	ALA	-	EXPRESSION TAG	UNP Q41372
D	296	VAL	-	EXPRESSION TAG	UNP Q41372
D	297	ASP	-	EXPRESSION TAG	UNP Q41372
D	298	HIS	-	EXPRESSION TAG	UNP Q41372
D	299	HIS	-	EXPRESSION TAG	UNP Q41372
D	300	HIS	-	EXPRESSION TAG	UNP Q41372
D	301	HIS	-	EXPRESSION TAG	UNP Q41372
D	302	HIS	-	EXPRESSION TAG	UNP Q41372
D	303	HIS	-	EXPRESSION TAG	UNP Q41372



• Molecule 1: aquaporin

Chain D:



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	181.30Å 103.97Å 67.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.90 32.37 – 3.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.90) 89.9 (32.37-3.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.50 (at 3.75Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.290 , 0.332 0.290 , 0.344	Depositor DCC
R_{free} test set	519 reflections (4.78%)	DCC
Wilson B-factor (Å ²)	103.8	Xtriage
Anisotropy	0.488	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.16 , 43.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	0 of 11748 reflections	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	6854	wwPDB-VP
Average B, all atoms (Å ²)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.73% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality i

5.1 Standard geometry i

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	1.11	4/1817 (0.2%)	0.97	6/2479 (0.2%)
1	B	1.12	4/1718 (0.2%)	1.03	7/2344 (0.3%)
1	C	1.12	4/1743 (0.2%)	1.01	6/2377 (0.3%)
1	D	1.12	5/1769 (0.3%)	1.00	7/2414 (0.3%)
All	All	1.12	17/7047 (0.2%)	1.00	26/9614 (0.3%)

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	1
1	C	0	1
1	D	0	1
All	All	0	4

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	122	TYR	CE2-CZ	-12.34	1.22	1.38
1	C	122	TYR	CE2-CZ	-12.30	1.22	1.38
1	A	122	TYR	CE2-CZ	-12.30	1.22	1.38
1	B	122	TYR	CE2-CZ	-12.27	1.22	1.38
1	D	122	TYR	CG-CD1	-10.79	1.25	1.39

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	201	PRO	CA-N-CD	-15.52	89.77	111.50
1	D	201	PRO	CA-N-CD	-15.52	89.77	111.50
1	C	201	PRO	CA-N-CD	-15.48	89.82	111.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	201	PRO	CA-N-CD	-15.48	89.83	111.50
1	A	122	TYR	CB-CG-CD1	6.54	124.93	121.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	99	HIS	Sidechain
1	B	99	HIS	Sidechain
1	C	99	HIS	Sidechain
1	D	99	HIS	Sidechain

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1767	0	1796	126	0
1	B	1671	0	1692	106	0
1	C	1696	0	1727	144	0
1	D	1720	0	1750	121	0
All	All	6854	0	6965	452	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 33.

The worst 5 of 452 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:39:ARG:HD3	1:D:263:VAL:HG13	1.42	1.01
1:C:66:THR:HG23	1:C:67:VAL:H	1.28	0.98
1:C:134:VAL:HG11	1:C:229:ALA:HA	1.48	0.95
1:C:202:ILE:O	1:C:203:GLY:C	2.02	0.95
1:A:202:ILE:O	1:A:203:GLY:C	2.02	0.95

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	234/303 (77%)	170 (73%)	42 (18%)	22 (9%)	1	23
1	B	220/303 (73%)	173 (79%)	29 (13%)	18 (8%)	1	27
1	C	223/303 (74%)	175 (78%)	21 (9%)	27 (12%)	1	14
1	D	226/303 (75%)	180 (80%)	27 (12%)	19 (8%)	1	26
All	All	903/1212 (74%)	698 (77%)	119 (13%)	86 (10%)	1	22

5 of 86 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	36	SER
1	A	96	SER
1	A	145	TYR
1	A	187	ARG
1	A	195	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	176/234 (75%)	157 (89%)	19 (11%)	9	48
1	B	165/234 (70%)	147 (89%)	18 (11%)	9	47
1	C	168/234 (72%)	151 (90%)	17 (10%)	11	51
1	D	171/234 (73%)	151 (88%)	20 (12%)	8	43
All	All	680/936 (73%)	606 (89%)	74 (11%)	9	47

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	201	PRO
1	C	100	ILE
1	D	201	PRO
1	B	225	ARG
1	C	52	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	147	GLN
1	B	153	ASN
1	C	210	HIS
1	B	99	HIS
1	C	146	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates 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(Å ²)	Q<0.9
1	A	236/303 (77%)	0.30	7 (2%)	48	36	34, 60, 87, 97	236 (100%)
1	B	224/303 (73%)	0.13	3 (1%)	74	58	41, 66, 89, 103	224 (100%)
1	C	227/303 (74%)	0.13	4 (1%)	65	50	44, 68, 94, 103	227 (100%)
1	D	230/303 (75%)	0.17	1 (0%)	90	81	52, 77, 93, 107	230 (100%)
All	All	917/1212 (75%)	0.18	15 (1%)	68	53	34, 68, 92, 107	917 (100%)

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	31	GLU	3.0
1	A	190	ARG	2.9
1	B	28	ASP	2.9
1	B	33	LYS	2.8
1	A	195	PRO	2.8

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 carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.