



Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 07:45 am BST

PDB ID : 5MDQ
Title : Crystal structure of outer membrane expressed Chitoporin VhChip from *Vibrio harveyi*
Authors : Zahn, M.; van den Berg, B.
Deposited on : 2016-11-13
Resolution : 2.50 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

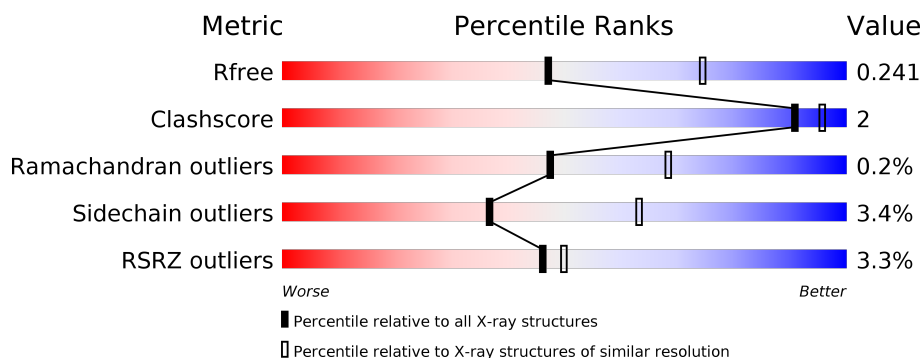
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 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 $\leq 5\%$. 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	350	<div> <div>90%</div> <div>10%</div> </div>
1	B	350	<div> <div>92%</div> <div>7%</div> </div>
1	C	350	<div> <div>9%</div> <div>89%</div> <div>10%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8210 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 Chitoporin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	350	Total	C	N	O	S	0	0	0
			2723	1712	449	553	9			
1	B	350	Total	C	N	O	S	0	0	0
			2723	1712	449	553	9			
1	C	350	Total	C	N	O	S	0	0	0
			2723	1712	449	553	9			

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

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

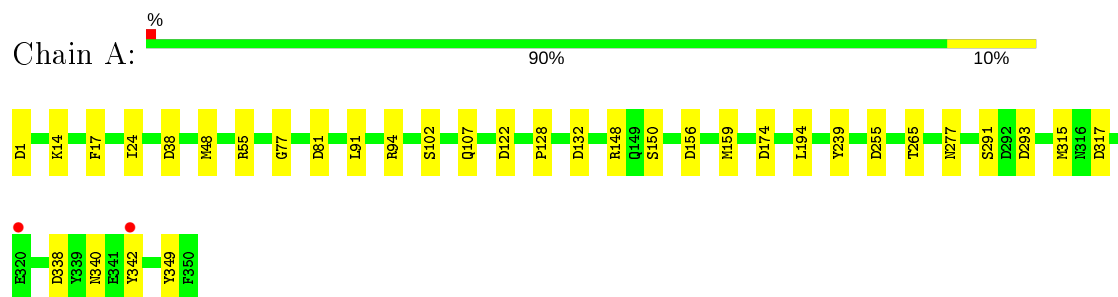
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	19	Total	O	0	0
			19	19		
3	B	15	Total	O	0	0
			15	15		
3	C	4	Total	O	0	0
			4	4		

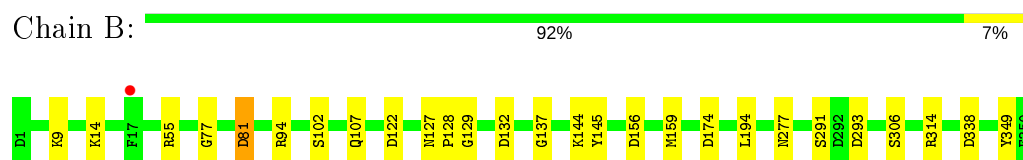
3 Residue-property plots [i](#)

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.

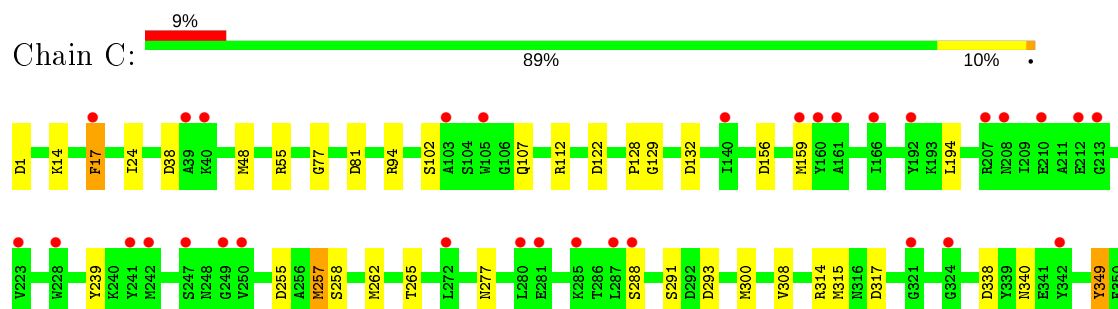
• Molecule 1: Chitoporin



• Molecule 1: Chitoporin



• Molecule 1: Chitoporin



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	255.73 Å 148.64 Å 54.36 Å 90.00° 94.94° 90.00°	Depositor
Resolution (Å)	128.39 – 2.50 49.15 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.4 (128.39-2.50) 95.4 (49.15-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.27 (at 2.51 Å)	Xtriage
Refinement program	REFMAC 5.8.0124	Depositor
R, R_{free}	0.230 , 0.241 0.233 , 0.241	Depositor DCC
R_{free} test set	2089 reflections (3.13%)	wwPDB-VP
Wilson B-factor (Å ²)	39.4	Xtriage
Anisotropy	0.646	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 30.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	8210	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.59% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	1.08	1/2792 (0.0%)	1.05	12/3782 (0.3%)
1	B	1.10	2/2792 (0.1%)	1.04	8/3782 (0.2%)
1	C	1.01	2/2792 (0.1%)	1.04	12/3782 (0.3%)
All	All	1.07	5/8376 (0.1%)	1.05	32/11346 (0.3%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	ASP	CB-CG	-5.59	1.40	1.51
1	B	306	SER	CB-OG	5.54	1.49	1.42
1	C	349	TYR	CG-CD2	5.46	1.46	1.39
1	C	17	PHE	CG-CD2	5.32	1.46	1.38
1	B	137	GLY	N-CA	5.28	1.53	1.46

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	338	ASP	CB-CG-OD2	9.01	126.41	118.30
1	B	81	ASP	CB-CG-OD1	8.97	126.37	118.30
1	A	81	ASP	CB-CG-OD1	8.72	126.14	118.30
1	A	293	ASP	CB-CG-OD1	8.64	126.07	118.30
1	C	1	ASP	CB-CG-OD1	8.61	126.05	118.30
1	A	122	ASP	CB-CG-OD2	-8.18	110.94	118.30
1	B	81	ASP	CB-CG-OD2	-7.93	111.16	118.30
1	C	338	ASP	CB-CG-OD2	7.85	125.37	118.30
1	B	338	ASP	CB-CG-OD2	7.71	125.24	118.30
1	B	174	ASP	CB-CG-OD1	7.60	125.14	118.30
1	C	81	ASP	CB-CG-OD1	7.58	125.12	118.30
1	C	122	ASP	CB-CG-OD2	-7.33	111.71	118.30
1	C	293	ASP	CB-CG-OD1	7.06	124.65	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1	ASP	CB-CG-OD2	-7.03	111.97	118.30
1	A	1	ASP	N-CA-CB	-6.82	98.33	110.60
1	A	194	LEU	CA-CB-CG	6.79	130.92	115.30
1	A	174	ASP	CB-CG-OD1	6.68	124.31	118.30
1	C	81	ASP	CB-CG-OD2	-6.57	112.39	118.30
1	C	1	ASP	CB-CG-OD2	-6.53	112.43	118.30
1	B	293	ASP	CB-CG-OD1	6.52	124.17	118.30
1	C	194	LEU	CA-CB-CG	6.42	130.05	115.30
1	A	1	ASP	CB-CG-OD1	6.18	123.86	118.30
1	C	112	ARG	NE-CZ-NH1	6.15	123.37	120.30
1	C	317	ASP	CB-CG-OD2	-6.14	112.77	118.30
1	B	194	LEU	CA-CB-CG	5.98	129.06	115.30
1	C	262	MET	CG-SD-CE	-5.92	90.72	100.20
1	C	38	ASP	CB-CG-OD1	5.86	123.57	118.30
1	B	174	ASP	OD1-CG-OD2	-5.83	112.22	123.30
1	B	122	ASP	CB-CG-OD2	-5.26	113.57	118.30
1	A	38	ASP	CB-CG-OD1	5.18	122.96	118.30
1	A	148	ARG	NE-CZ-NH2	-5.04	117.78	120.30
1	A	317	ASP	CB-CG-OD1	5.04	122.83	118.30

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	2723	0	2473	12	0
1	B	2723	0	2473	9	0
1	C	2723	0	2473	12	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	19	0	0	0	0
3	B	15	0	0	0	0
3	C	4	0	0	0	0
All	All	8210	0	7419	27	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 2.

All (27) 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:14:LYS:HG3	1:C:349:TYR:CZ	2.36	0.61
1:C:257:MET:HG2	1:C:258:SER:N	2.15	0.60
1:A:349:TYR:CZ	1:C:14:LYS:HG3	2.39	0.57
1:A:14:LYS:HG3	1:B:349:TYR:CZ	2.40	0.57
1:C:300:MET:HE3	1:C:308:VAL:HG11	1.87	0.56
1:C:315:MET:HG2	1:C:340:ASN:OD1	2.06	0.55
1:A:315:MET:HG3	1:A:340:ASN:OD1	2.06	0.53
1:A:239:TYR:OH	1:A:255:ASP:OD2	2.25	0.53
1:B:77:GLY:O	1:B:94:ARG:HD3	2.13	0.49
1:C:77:GLY:O	1:C:94:ARG:HD3	2.14	0.47
1:C:300:MET:CE	1:C:308:VAL:HG11	2.45	0.46
1:B:129:GLY:HA3	1:B:314:ARG:HD3	1.97	0.46
1:A:349:TYR:CE2	1:C:14:LYS:HG3	2.49	0.46
1:B:277:ASN:O	1:B:291:SER:HB2	2.17	0.45
1:A:77:GLY:O	1:A:94:ARG:HD3	2.16	0.45
1:A:24:ILE:HD12	1:A:24:ILE:N	2.32	0.45
1:C:277:ASN:O	1:C:291:SER:HB2	2.18	0.44
1:A:91:LEU:HA	1:A:91:LEU:HD23	1.90	0.43
1:A:277:ASN:O	1:A:291:SER:HB2	2.17	0.43
1:A:342:TYR:CD1	1:A:342:TYR:N	2.87	0.43
1:B:144:LYS:HG3	1:B:145:TYR:CD2	2.54	0.43
1:A:14:LYS:HG3	1:B:349:TYR:CE2	2.54	0.42
1:C:129:GLY:HA3	1:C:314:ARG:HD3	2.00	0.42
1:C:239:TYR:OH	1:C:255:ASP:OD2	2.30	0.42
1:C:24:ILE:HD12	1:C:24:ILE:N	2.35	0.41
1:A:150:SER:OG	1:B:81:ASP:OD2	2.30	0.41
1:B:127:ASN:HA	1:B:128:PRO:HA	1.84	0.40

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	348/350 (99%)	330 (95%)	17 (5%)	1 (0%)	41	61
1	B	348/350 (99%)	332 (95%)	16 (5%)	0	100	100
1	C	348/350 (99%)	333 (96%)	14 (4%)	1 (0%)	41	61
All	All	1044/1050 (99%)	995 (95%)	47 (4%)	2 (0%)	47	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	128	PRO
1	C	128	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	267/267 (100%)	258 (97%)	9 (3%)	37	63
1	B	267/267 (100%)	260 (97%)	7 (3%)	46	72
1	C	267/267 (100%)	256 (96%)	11 (4%)	30	55
All	All	801/801 (100%)	774 (97%)	27 (3%)	37	63

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	PHE
1	A	48	MET
1	A	55	ARG
1	A	102	SER
1	A	107	GLN
1	A	132	ASP
1	A	156	ASP
1	A	159	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	265	THR
1	B	9	LYS
1	B	55	ARG
1	B	102	SER
1	B	107	GLN
1	B	132	ASP
1	B	156	ASP
1	B	159	MET
1	C	17	PHE
1	C	48	MET
1	C	55	ARG
1	C	102	SER
1	C	107	GLN
1	C	132	ASP
1	C	156	ASP
1	C	159	MET
1	C	257	MET
1	C	265	THR
1	C	288	SER

Some 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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

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 [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

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	350/350 (100%)	-0.33	2 (0%) 89 90	26, 47, 73, 93	0
1	B	350/350 (100%)	-0.38	1 (0%) 94 94	27, 43, 68, 85	0
1	C	350/350 (100%)	0.30	32 (9%) 9 9	35, 72, 113, 132	0
All	All	1050/1050 (100%)	-0.13	35 (3%) 46 50	26, 52, 96, 132	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	280	LEU	5.7
1	C	321	GLY	4.9
1	C	287	LEU	3.8
1	C	324	GLY	3.7
1	B	17	PHE	3.5
1	C	39	ALA	3.5
1	C	247	SER	3.2
1	C	241	TYR	3.1
1	C	250	VAL	3.0
1	C	208	ASN	2.7
1	C	342	TYR	2.7
1	C	160	TYR	2.6
1	C	213	GLY	2.6
1	A	342	TYR	2.6
1	C	17	PHE	2.4
1	C	242	MET	2.4
1	A	320	GLU	2.4
1	C	285	LYS	2.4
1	C	281	GLU	2.4
1	C	159	MET	2.3
1	C	212	GLU	2.3
1	C	140	ILE	2.3
1	C	103	ALA	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	105	TRP	2.2
1	C	288	SER	2.2
1	C	166	ILE	2.2
1	C	192	TYR	2.2
1	C	228	TRP	2.2
1	C	272	LEU	2.1
1	C	161	ALA	2.1
1	C	210	GLU	2.1
1	C	223	VAL	2.0
1	C	249	GLY	2.0
1	C	207	ARG	2.0
1	C	40	LYS	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
2	NA	B	401	1/1	0.93	0.20	26,26,26,26	0
2	NA	C	401	1/1	0.96	0.17	32,32,32,32	0
2	NA	A	401	1/1	0.97	0.15	24,24,24,24	0

6.5 Other polymers [i](#)

There are no such residues in this entry.