wwPDB X-ray Structure Validation Summary Report (i)

Sep 19, 2023 – 11:02 AM EDT

PDB ID	:	5HJ8
Title	:	Bacterial sodium channel neck 3G mutant
Authors	:	Rohaim, A.; Minor, D.L.
Deposited on	:	2016-01-13
Resolution	:	3.70 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

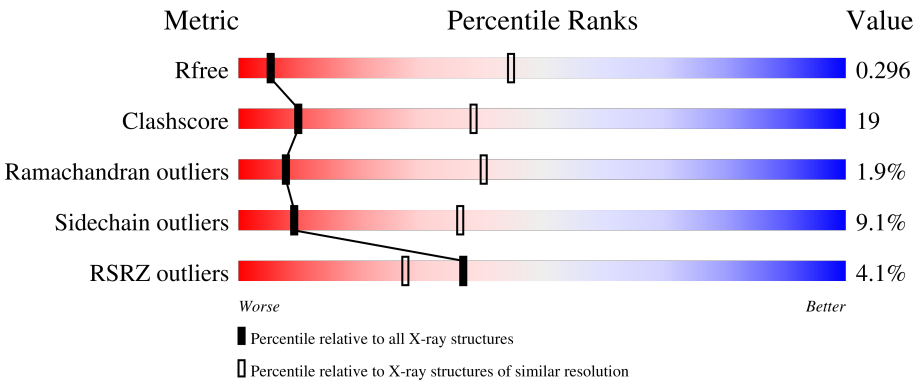
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)
RSRZ outliers	127900	1578 (3.90-3.50)

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 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	152	3% 49% 30% • • 18%
1	B	152	2% 46% 27% • • 22%
1	C	152	5% 53% 26% • • 18%
1	D	152	3% 48% 30% • 18%

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	CA	A	301	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3743 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 Ion transport protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	124	948	636	141	165	6	0	0	0
1	B	118	925	622	139	158	6	0	0	0
1	C	125	933	626	140	161	6	0	0	0
1	D	125	933	629	141	157	6	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	137	GLY	-	expression tag	UNP Q0ABW0
A	138	PRO	-	expression tag	UNP Q0ABW0
A	139	SER	-	expression tag	UNP Q0ABW0
A	140	SER	-	expression tag	UNP Q0ABW0
A	141	PRO	-	expression tag	UNP Q0ABW0
A	142	SER	-	expression tag	UNP Q0ABW0
A	248	GLY	ALA	engineered mutation	UNP Q0ABW0
A	249	GLY	GLU	engineered mutation	UNP Q0ABW0
A	250	GLY	ASP	engineered mutation	UNP Q0ABW0
B	137	GLY	-	expression tag	UNP Q0ABW0
B	138	PRO	-	expression tag	UNP Q0ABW0
B	139	SER	-	expression tag	UNP Q0ABW0
B	140	SER	-	expression tag	UNP Q0ABW0
B	141	PRO	-	expression tag	UNP Q0ABW0
B	142	SER	-	expression tag	UNP Q0ABW0
B	248	GLY	ALA	engineered mutation	UNP Q0ABW0
B	249	GLY	GLU	engineered mutation	UNP Q0ABW0
B	250	GLY	ASP	engineered mutation	UNP Q0ABW0
C	137	GLY	-	expression tag	UNP Q0ABW0
C	138	PRO	-	expression tag	UNP Q0ABW0
C	139	SER	-	expression tag	UNP Q0ABW0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	140	SER	-	expression tag	UNP Q0ABW0
C	141	PRO	-	expression tag	UNP Q0ABW0
C	142	SER	-	expression tag	UNP Q0ABW0
C	248	GLY	ALA	engineered mutation	UNP Q0ABW0
C	249	GLY	GLU	engineered mutation	UNP Q0ABW0
C	250	GLY	ASP	engineered mutation	UNP Q0ABW0
D	137	GLY	-	expression tag	UNP Q0ABW0
D	138	PRO	-	expression tag	UNP Q0ABW0
D	139	SER	-	expression tag	UNP Q0ABW0
D	140	SER	-	expression tag	UNP Q0ABW0
D	141	PRO	-	expression tag	UNP Q0ABW0
D	142	SER	-	expression tag	UNP Q0ABW0
D	248	GLY	ALA	engineered mutation	UNP Q0ABW0
D	249	GLY	GLU	engineered mutation	UNP Q0ABW0
D	250	GLY	ASP	engineered mutation	UNP Q0ABW0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	A	1	1	1	0	0

- Molecule 3 is water.

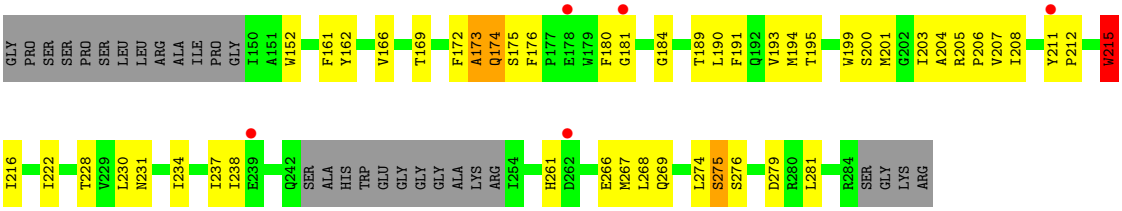
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	B	1	1	1	0	0
			Total	O		
3	C	1	1	1	0	0
			Total	O		
3	D	1	1	1	0	0

3 Residue-property plots (i)

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.

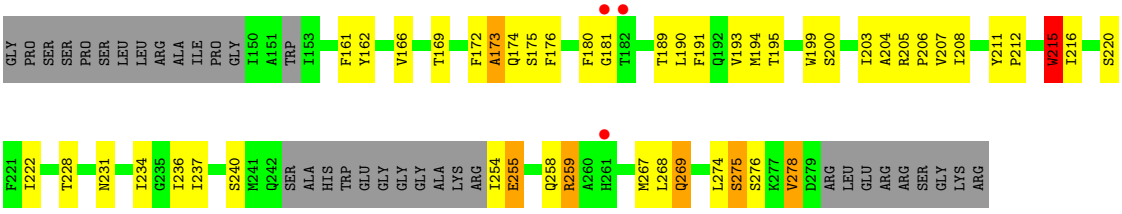
• Molecule 1: Ion transport protein

Chain A:



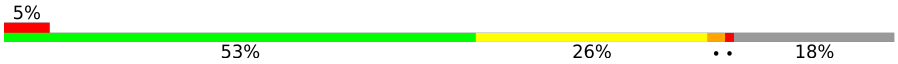
• Molecule 1: Ion transport protein

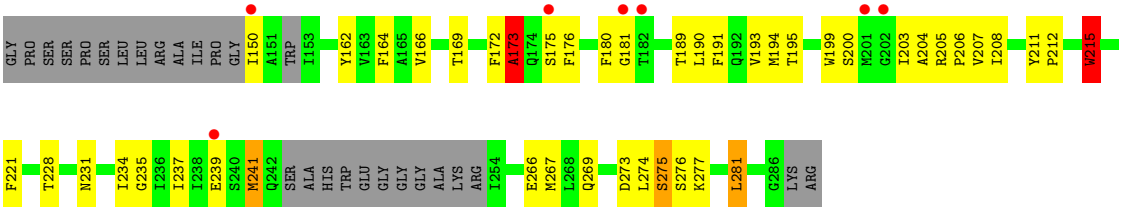
Chain B:



• Molecule 1: Ion transport protein

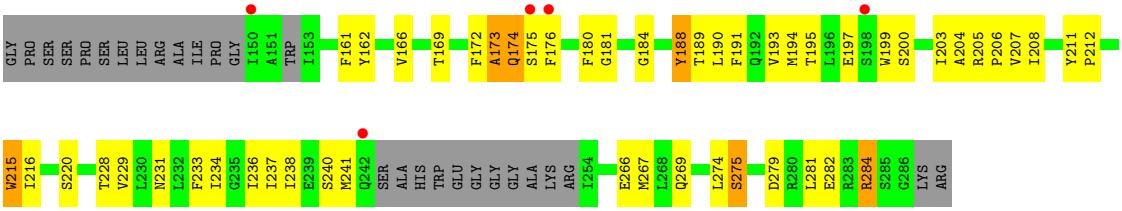
Chain C:





• Molecule 1: Ion transport protein

Chain D:


GLY
PRO
SER
SER
PRO
SER
LEU
LEU
ARG
ALA
ILE
PRO
GLY
I150
A151
TRP
I153
F161
Y162
V166
T169
F172
A173
Q174
S175
F176
F180
G181
G184
Y188
T189
L190
F191
Q192
V193
M194
T195
L196
E197
S198
W199
S200
I203
A204
R205
P206
V207
I208
Y211
P212
W215
I216
S220
T228
V229
L230
N231
L232
F233
I234
G235
I236
I237
I238
E239
S240
M241
Q242
SER
ALA
HIS
TRP
GLU
GLY
GLY
GLY
ALA
LYS
ARG
I254
E266
M267
L268
Q269
L274
S275
D279
R280
L281
E282
R283
R284
S285
G286
LYS
ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	138.31Å 153.35Å 165.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.70 15.00 – 3.70	Depositor EDS
% Data completeness (in resolution range)	98.2 (15.00-3.70) 99.9 (15.00-3.70)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 3.67Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, R_{free}	0.278 , 0.304 0.272 , 0.296	Depositor DCC
R_{free} test set	958 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	164.1	Xtriage
Anisotropy	0.659	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 269.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.34$, $< L^2 > = 0.18$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3743	wwPDB-VP
Average B, all atoms (Å^2)	316.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.84% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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

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	0.55	0/972	0.75	0/1327
1	B	0.60	0/948	0.75	0/1290
1	C	0.58	1/954 (0.1%)	0.76	1/1300 (0.1%)
1	D	0.57	0/956	0.73	0/1305
All	All	0.57	1/3830 (0.0%)	0.75	1/5222 (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	B	0	2
1	C	0	2
1	D	0	1
All	All	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	173	ALA	C-N	-5.67	1.21	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	173	ALA	O-C-N	-7.57	110.60	122.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	276	SER	Peptide
1	B	278	VAL	Peptide
1	C	173	ALA	Mainchain
1	C	281	LEU	Peptide
1	D	284	ARG	Peptide

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	948	0	901	47	0
1	B	925	0	903	37	0
1	C	933	0	884	37	0
1	D	933	0	883	46	0
2	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	4	0
3	D	1	0	0	0	0
All	All	3743	0	3571	140	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 19.

The worst 5 of 140 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:C:273:ASP:O	3:C:301:HOH:O	1.54	1.23
1:D:281:LEU:HA	1:D:284:ARG:CB	1.95	0.97
1:C:277:LYS:N	3:C:301:HOH:O	2.02	0.90
1:C:276:SER:N	3:C:301:HOH:O	2.05	0.88
1:A:234:ILE:HD11	1:D:237:ILE:HA	1.59	0.83

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	120/152 (79%)	101 (84%)	17 (14%)	2 (2%)	9	42
1	B	112/152 (74%)	95 (85%)	14 (12%)	3 (3%)	5	33
1	C	119/152 (78%)	98 (82%)	19 (16%)	2 (2%)	9	42
1	D	119/152 (78%)	101 (85%)	16 (13%)	2 (2%)	9	42
All	All	470/608 (77%)	395 (84%)	66 (14%)	9 (2%)	8	40

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	173	ALA
1	C	173	ALA
1	D	173	ALA
1	D	215	TRP
1	B	173	ALA

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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	93/129 (72%)	83 (89%)	10 (11%)	6	29
1	B	94/129 (73%)	86 (92%)	8 (8%)	10	40
1	C	89/129 (69%)	83 (93%)	6 (7%)	16	47
1	D	88/129 (68%)	79 (90%)	9 (10%)	7	31
All	All	364/516 (70%)	331 (91%)	33 (9%)	9	36

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	228	THR
1	D	241	MET
1	D	275	SER
1	B	228	THR
1	B	215	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	269	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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 (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	124/152 (81%)	-0.07	5 (4%)	38	28	237, 315, 407, 434	0
1	B	118/152 (77%)	-0.20	3 (2%)	57	45	232, 306, 398, 442	0
1	C	125/152 (82%)	0.18	7 (5%)	24	17	243, 325, 404, 452	0
1	D	125/152 (82%)	-0.03	5 (4%)	38	28	226, 298, 392, 446	0
All	All	492/608 (80%)	-0.03	20 (4%)	37	27	226, 310, 404, 452	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	181	GLY	6.4
1	C	182	THR	5.7
1	C	239	GLU	5.2
1	C	202	GLY	4.2
1	C	201	MET	3.9

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 monosaccharides 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	CA	A	301	1/1	-0.20	0.70	306,306,306,306	0

6.5 Other polymers (i)

There are no such residues in this entry.