wwPDB NMR Structure Validation Summary Report (i)

Aug 22, 2016 – 05:18 PM EDT

PDB ID	:	5JDP
Title	:	E73V mutant of the human voltage-dependent anion channel
Authors	:	Jaremko, M.; Jaremko, L.; Villinger, S.; Schmidt, C.; Giller, K.; Griesinger, C.; Becker, S.; Zweckstetter, M.
Deposited on	:	2016-04-17

This is a wwPDB NMR 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

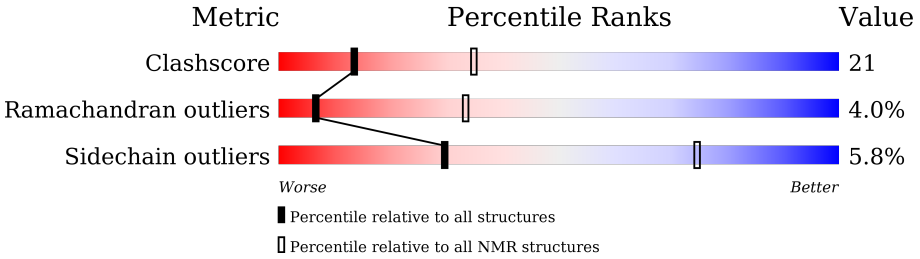
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027939
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	285	58% 28% • 11%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 4 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:2-A:18, A:24-A:89, A:93-A:103, A:109-A:197, A:202-A:211, A:216-A:250, A:254-A:265, A:271-A:283 (253)	0.55	4

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 4 single-model clusters were found.

Cluster number	Models
1	2, 3, 5, 10, 11, 12, 15, 19
2	4, 7, 14, 16, 17, 20
3	13, 18
Single-model clusters	1; 6; 8; 9

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4346 atoms, of which 2162 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Voltage-dependent anion-selective channel protein 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	285	4346	1379	2162	371	430	4	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP P21796
A	73	VAL	GLU	conflict	UNP P21796
A	284	ARG	-	expression tag	UNP P21796
A	285	SER	-	expression tag	UNP P21796

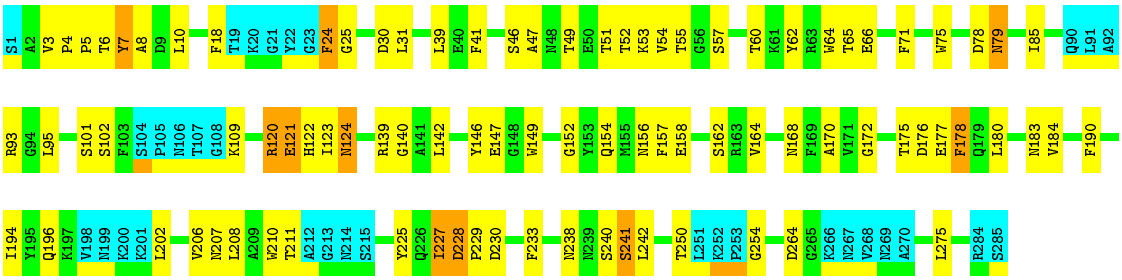
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Voltage-dependent anion-selective channel protein 1

Chain A: 58% 28% • 11%

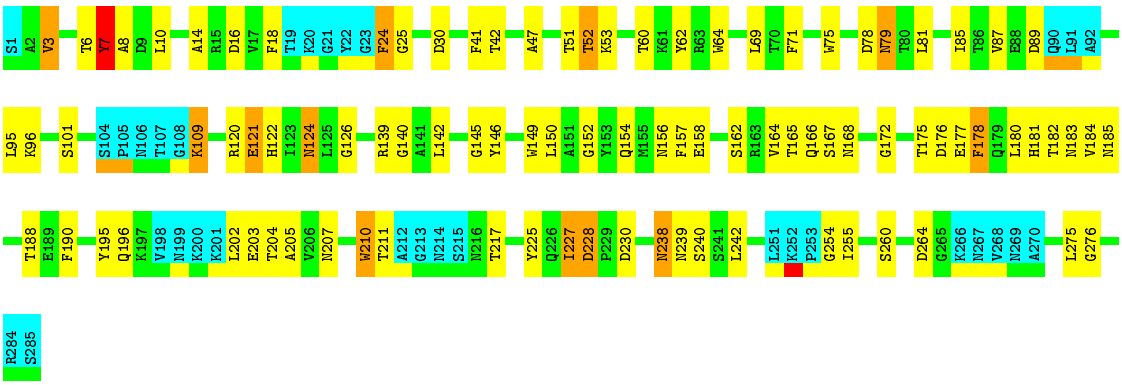


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

- Molecule 1: Voltage-dependent anion-selective channel protein 1

Chain A: 56% 28% • 11%



5 Refinement protocol and experimental data overview (i)

Of the ? calculated structures, 20 were deposited, based on the following criterion: *?*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR NIH	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	5jdp_cs.cif
Number of chemical shift lists	1
Total number of shifts	971
Number of shifts mapped to atoms	0
Number of unparsed shifts	971
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	0%

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality (i)

6.1 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 (average) 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	0.69±0.01	0±0/1992 (0.0±0.0%)	0.88±0.01	1±1/2694 (0.0±0.0%)
All	All	0.69	0/39840 (0.0%)	0.88	23/53880 (0.0%)

There are no bond-length outliers.

5 of 6 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	227	ILE	CA-CB-CG1	7.61	125.45	111.00	20	3
1	A	227	ILE	CA-CB-CG2	6.56	124.02	110.90	1	3
1	A	227	ILE	CB-CA-C	6.56	124.72	111.60	9	10
1	A	227	ILE	CG1-CB-CG2	6.48	125.65	111.40	6	4
1	A	227	ILE	CB-CG1-CD1	-5.49	98.52	113.90	13	2

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1953	1920	1920	81±9
All	All	39060	38400	38400	1627

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 21.

5 of 613 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:238:ASN:ND2	1:A:242:LEU:H	0.92	1.60	3	4
1:A:238:ASN:HD21	1:A:242:LEU:H	0.79	1.20	3	4
1:A:259:LEU:HD22	1:A:259:LEU:N	0.77	1.93	11	1
1:A:63:ARG:O	1:A:64:TRP:CG	0.77	2.38	1	2
1:A:238:ASN:HD21	1:A:242:LEU:HD13	0.77	1.39	10	7

6.3 Torsion angles (i)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	253/285 (89%)	223±3 (88±1%)	20±3 (8±1%)	10±2 (4±1%)	6	33
All	All	5060/5700 (89%)	4458 (88%)	401 (8%)	201 (4%)	6	33

5 of 33 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	121	GLU	19
1	A	122	HIS	19
1	A	79	ASN	18
1	A	120	ARG	14
1	A	24	PHE	13

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	207/232 (89%)	195±2 (94±1%)	12±2 (6±1%)	29	75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	4140/4640 (89%)	3901 (94%)	239 (6%)	29 75

5 of 51 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	7	TYR	18
1	A	228	ASP	18
1	A	102	SER	17
1	A	157	PHE	16
1	A	178	PHE	16

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

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

6.5 Carbohydrates (i)

There are no carbohydrates in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 0% for the well-defined parts and 0% for the entire structure.

7.1 Chemical shift list 1

File name: 5jdp_cs.cif

Chemical shift list name: *assigned_chem_shift_list*

7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	971
Number of shifts mapped to atoms	0
Number of unparsed shifts	971
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following errors were found when reading this chemical shift list.

- Entity instance (chain) must be specified. First 5 (of 971) occurences are reported below.

Shift ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	?	6	THR	H	7.426	0.005	1
2	?	6	THR	N	106.255	0.023	1
3	?	7	TYR	H	7.660	0.005	1
4	?	7	TYR	N	122.565	0.108	1
5	?	166	GLN	H	7.442	0.008	1

7.1.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 0%, i.e. 0 atoms were assigned a chemical

shift out of a possible 2986. 0 out of 37 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	0/1257 (0%)	0/502 (0%)	0/506 (0%)	0/249 (0%)
Sidechain	0/1454 (0%)	0/843 (0%)	0/549 (0%)	0/62 (0%)
Aromatic	0/275 (0%)	0/146 (0%)	0/122 (0%)	0/7 (0%)
Overall	0/2986 (0%)	0/1491 (0%)	0/1177 (0%)	0/318 (0%)

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (assigned_chem_shift_list). RCI is only applicable to proteins.