wwPDB NMR Structure Validation Summary Report (i)

Apr 28, 2016 – 09:18 PM EDT

PDB ID	:	4BH9
Title	:	A structural model of CAP mutant (T127L and S128I) in the apo state
Authors	:	Tzeng, S.R.; Kalodimos, C.G.
Deposited on	:	2013-03-30

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:

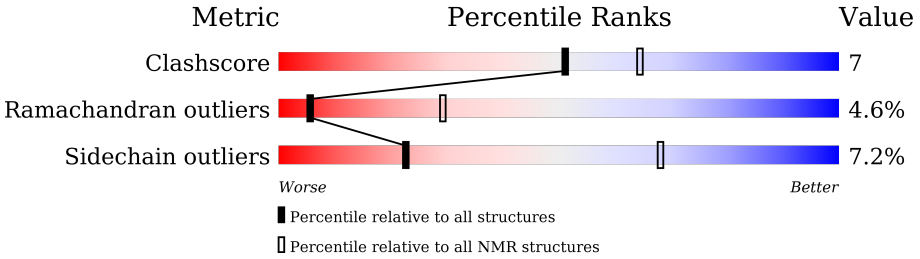
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027457

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 is 12%.

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	209	77% 17% 6%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:9-A:128, A:134-A:209 (196)	0.77	3

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 4 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3365 atoms, of which 1712 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CAMP RECEPTOR PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	209	3365	1049	1712	290	305	9	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	127	LEU	THR	ENGINEERED MUTATION	UNP P0ACJ8
A	128	ILE	SER	ENGINEERED MUTATION	UNP P0ACJ8

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: CAMP RECEPTOR PROTEIN

Chain A: 77% 17% 6%

V1 L2 G3 K4 P5 Q6 T7 D8 L11 L15 C18 K22 L29 I30 H31 K35 A36 V47 L50 I60 L61 L64 I70 G71 E72 E81 R82 S83 A84 W85 V86 C92 E93 V94 A95 E96 I97 I106 Q107 R123 V126 E129 K130

V131 G132 N133 V139 I143 D161 G173 R180 V183 L190 R209

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

The representative model is number 3. Colouring as in section 4.1 above.

- Molecule 1: CAMP RECEPTOR PROTEIN

Chain A: 73% 19% • 6%

V1 L2 G3 K4 P5 Q6 T7 D8 L11 L15 C18 H19 I20 H21 K22 K26 L29 I30 H31 Q32 G33 E34 K35 L50 E58 M59 I60 L61 L75 F76 E77 E81 R82 S83 A84 W85 V86 V94 A95 E96 I97 K101 I106 L113 A118

R122 L127 I128 E129 K130 V131 G132 N133 T140 Q145 L150 I165 K166 I167 I172 G173 Q174 V183 I186 L190 R209

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *TORSION ANGLE DYNAMICS; SIMULATED ANNEALING.*

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *LOWEST ENERGY.*

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

Software name	Classification	Version
CNS	refinement	
CYANA	structure solution	
CNS	structure solution	

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)	4bh9_cs.cif
Number of chemical shift lists	1
Total number of shifts	339
Number of shifts mapped to atoms	339
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	12%

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)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

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	1557	1611	1605	22±4
All	All	31140	32220	32100	448

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 7.

5 of 199 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:15:LEU:HA	1:A:18:CYS:SG	0.73	2.24	1	18
1:A:18:CYS:SG	1:A:95:ALA:HB1	0.67	2.29	20	16
1:A:59:MET:HA	1:A:174:GLN:HB3	0.65	1.69	17	3
1:A:150:LEU:HB3	1:A:165:ILE:HG21	0.63	1.70	16	2
1:A:173:GLY:HA2	1:A:183:VAL:HG21	0.63	1.70	12	15

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	195/209 (93%)	162±3 (83±1%)	25±3 (13±2%)	9±2 (5±1%)	5 29
All	All	3900/4180 (93%)	3231 (83%)	491 (13%)	178 (5%)	5 29

5 of 43 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	85	TRP	19
1	A	31	HIS	18
1	A	82	ARG	12
1	A	84	ALA	11
1	A	161	ASP	10

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	169/180 (94%)	157±4 (93±2%)	12±4 (7±2%)	23 68
All	All	3380/3600 (94%)	3138 (93%)	242 (7%)	23 68

5 of 79 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	61	LEU	18
1	A	29	LEU	13
1	A	180	ARG	10
1	A	50	LEU	9
1	A	115	ARG	8

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 12% for the well-defined parts and 13% for the entire structure.

7.1 Chemical shift list 1

File name: 4bh9_cs.cif

Chemical shift list name: *CAP_127128_APO_1.str.csh*

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	339
Number of shifts mapped to atoms	339
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	112	-0.07 ± 0.10	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	0	—	—
$^{13}C'$	105	-0.43 ± 0.13	None needed (< 0.5 ppm)
^{15}N	122	-0.03 ± 0.37	None needed (< 0.5 ppm)

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 12%, i.e. 309 atoms were assigned a chemical shift out of a possible 2484. 0 out of 34 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	309/970 (32%)	0/387 (0%)	198/392 (51%)	111/191 (58%)
Sidechain	0/1349 (0%)	0/788 (0%)	0/498 (0%)	0/63 (0%)

Continued on next page...

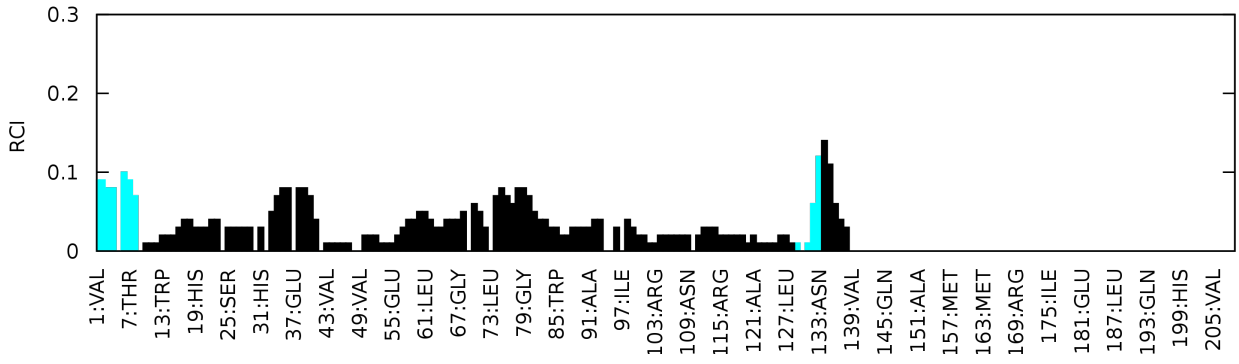
Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/165 (0%)	0/85 (0%)	0/66 (0%)	0/14 (0%)
Overall	309/2484 (12%)	0/1260 (0%)	198/956 (21%)	111/268 (41%)

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A: