wwPDB NMR Structure Validation Summary Report (i)

Dec 13, 2023 – 09:01 PM EST

PDB ID	:	2JOG
Title	:	Structure of the calcineurin-NFAT complex
Authors	:	Takeuchi, K.; Roehrl, M.H.; Sun, Z.Y.; Wagner, G.
Deposited on	:	2007-03-10

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
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
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 (i)) were used in the production of this report:

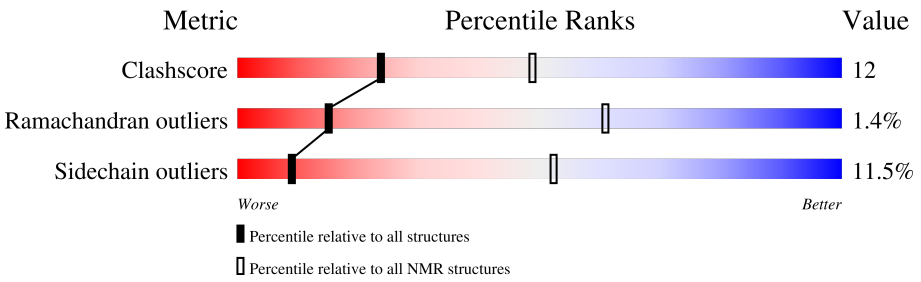
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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	327	60% 26% • 11%
2	B	16	100%

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 5 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:21-A:79, A:85-A:99, A:111-A:289, A:300-A:324, A:333-A:346 (292)	0.12	5

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 2 clusters and 2 single-model clusters were found.

Cluster number	Models
1	2, 3, 4, 5, 10
2	1, 6, 7
Single-model clusters	8; 9

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5548 atoms, of which 2764 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Calmodulin-dependent calcineurin A subunit alpha isoform.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	327	5306	1714	2643	448	482	19	0

- Molecule 2 is a protein called NFAT.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	16	242	78	121	20	23	0

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: Calmodulin-dependent calcineurin A subunit alpha isoform

Chain A: 60% 26% • 11%

V21 R28 K32 K52 E53 G54 R55 I65 E75 D80 I81 D82 A83 P84 V85 T86 V87 L98 M99 K100 L101 F102 E103 V104 G105 G106 S107 P108 A109 N110 T111 R112 Y113 L114 F115 L116 G117 D118 Y119 V120 D121 R122 V130 L131 W134 Y140 T143 L144

L147 R148 H151 E152 C153 R154 H155 L156 T157 K167 I168 K169 D183 P186 L187 L190 M191 N192 Q193 Q194 F195 L196 C197 V198 L202 I206 N207 T208 L209 D210 R213 R217 P222 A223 Y224 D229 I230 S233 D234 P235 L236 E237 D238 F239 G240 N241 E242 K243

T244 Q245 E246 H247 F248 T249 H250 N251 R254 G255 C256 S257 Y258 F259 Y260 S261 Y262 P263 C266 E267 F268 L269 N273 L274 L275 R279 A280 Y288 R289 M290 Y291 R292 K293 S294 Q295 T296 T297 G298 F299 F306 S307 A308 P309 N310 D313 V314 Y315 L322 K323 Y324 E325 N326

N327 V328 M329 N330 I331 R332 Q333 F334 M347

• Molecule 2: NFAT

Chain B: 100%

G376 P377 H378 P379 V380 I381 V382 I383 T384 G385 P386 H387 E388 E389 L390 E391

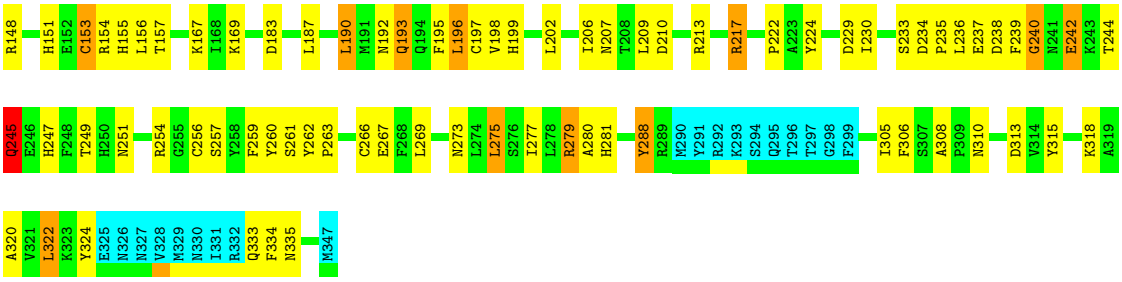
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 5. Colouring as in section 4.1 above.

• Molecule 1: Calmodulin-dependent calcineurin A subunit alpha isoform

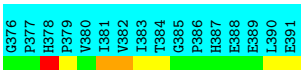
Chain A: 58% 27% • 11%

V21 R28 K32 K52 E53 G54 R55 I65 E75 D80 I81 D82 A83 P84 V85 T86 V87 L98 M99 K100 L101 F102 E103 V104 G105 G106 S107 P108 A109 N110 T111 R112 Y113 L114 F115 L116 G117 D118 Y119 V120 D121 R122 V130 L131 W134 Y140 T143 L144



- Molecule 2: NFAT

Chain B: 100%



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 50 calculated structures, 10 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
CYANA	structure solution	
CYANA	refinement	

No chemical shift data was provided.

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	2390	2373	2327	57±4
2	B	0	0	0	0±0
All	All	23900	23730	23270	569

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 12.

5 of 98 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:229:ASP:O	1:A:233:SER:N	0.88	2.05	8	10
1:A:75:GLU:O	1:A:217:ARG:NH2	0.81	2.14	7	10
1:A:262:TYR:N	1:A:263:PRO:HD2	0.72	2.00	1	10
1:A:115:PHE:HB3	1:A:119:TYR:CE1	0.64	2.28	10	10
1:A:251:ASN:HB3	1:A:256:CYS:O	0.64	1.91	2	10

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	291/327 (89%)	267±0 (92±0%)	20±0 (7±0%)	4±0 (1±0%)	15 61
2	B	0	-	-	-	-
All	All	2910/3430 (85%)	2671 (92%)	199 (7%)	40 (1%)	15 61

All 4 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	239	PHE	10
1	A	240	GLY	10
1	A	245	GLN	10
1	A	257	SER	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	261/291 (90%)	231±1 (89±0%)	30±1 (11±0%)	9 52
2	B	0	-	-	-
All	All	2610/3050 (86%)	2311 (89%)	299 (11%)	9 52

5 of 39 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	21	VAL	10
1	A	28	ARG	10
1	A	85	VAL	10
1	A	98	LEU	10
1	A	114	LEU	10

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 monosaccharides 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)

No chemical shift data were provided