wwPDB NMR Structure Validation Summary Report

Jun 12, 2024 – 02:34 AM EDT

PDB ID	:	2QMV
Title	:	High Resolution Structure of Peroxisone Proliferation-Activated Receptor gamma and Characterisation of its Interaction with the Co-activator Transcriptional Intermediary Factor 2
Authors	:	Hartl, R.; Riepl, H.; Kauschke, S.; Kalbitzer, H.R.; Maurer, T.
Deposited on	:	2007-07-17

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
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:

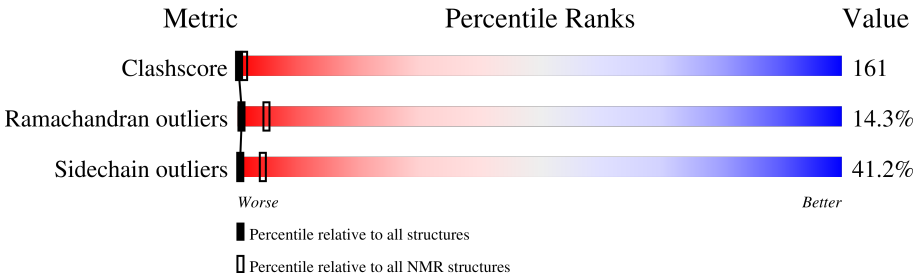
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	270	. 47% 33% . 16%

2 Ensemble composition and analysis

This entry contains 9 models. Model 7 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:210-A:239, A:246-A:249, A:276-A:341, A:346-A:473 (228)	1.04	7

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 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4441 atoms, of which 2276 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Peroxisome proliferator-activated receptor gamma.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	270	4441	1397	2276	354	404	10	0

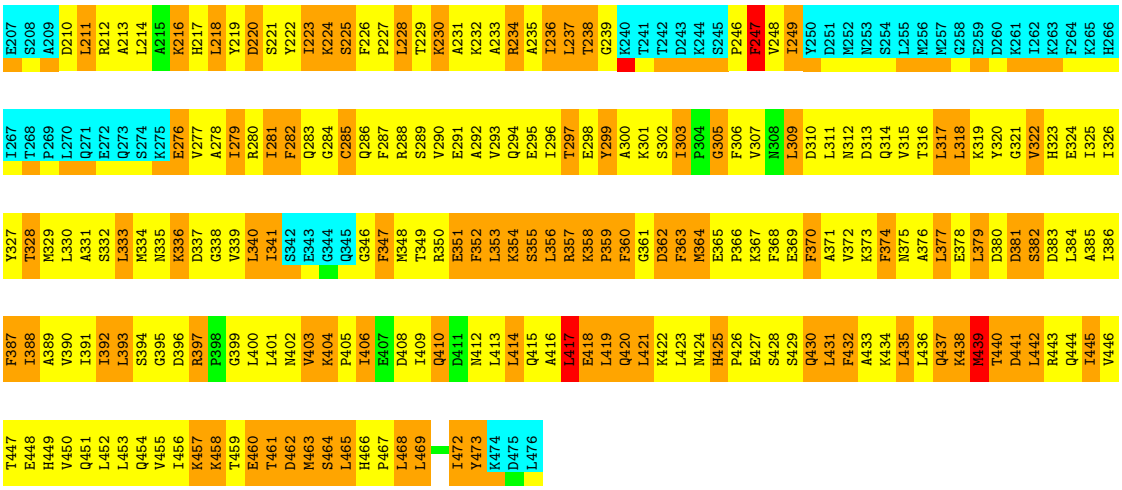
4 Residue-property plots

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: Peroxisome proliferator-activated receptor gamma

Chain A: 

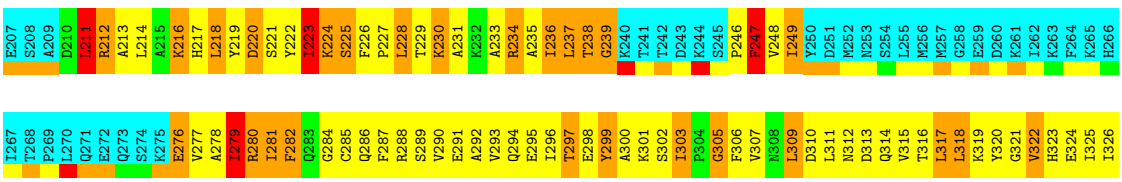


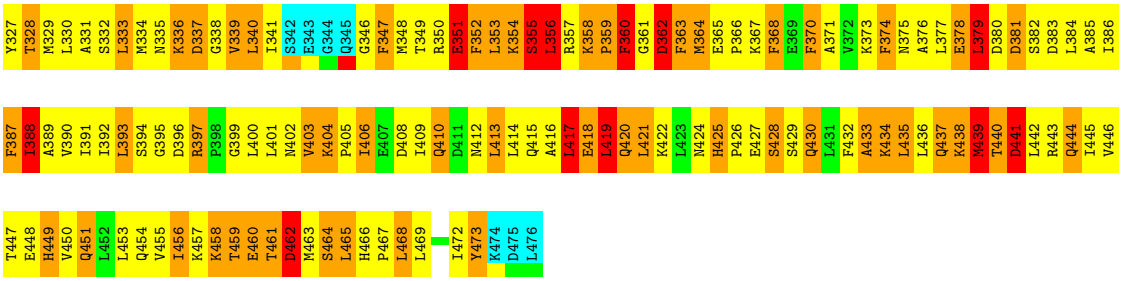
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 7. Colouring as in section 4.1 above.

• Molecule 1: Peroxisome proliferator-activated receptor gamma

Chain A: 




Y327 T328 M329 L330 A331 S332 L333 M334 N335 K336 D337 G338 V339 L340 T341 S342 E343 G344 Q345 G346 F347 M348 T349 R350 E351 F352 L353 K354 S355 L356 R357 K358 P359 F360 G361 D362 F363 M364 E365 P366 K367 F368 E369 F370 A371 V372 K373 F374 N375 A376 L377 E378 L379 D380 D381 S382 D383 L384 A385 I386
F387 T388 A389 V390 I391 I392 L393 S394 G395 D396 R397 P398 G399 L400 L401 N402 V403 K404 P405 I406 E407 D408 I409 Q410 D411 N412 L413 L414 Q415 A416 L417 E418 L419 Q420 L421 K422 L423 N424 H425 P426 E427 S428 S429 Q430 L431 F432 A433 K434 L435 L436 Q437 K438 M439 T440 D441 L442 R443 Q444 I445 V446
T447 E448 H449 V450 Q451 L452 L453 Q454 V455 I456 K457 K458 T459 E460 T461 D462 M463 S464 L465 H466 P467 L468 L469 I472 Y473 K474 D475 L476

5 Refinement protocol and experimental data overview

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

Of the 1024 calculated structures, 9 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
DYANA	structure solution	1.5
DYANA	refinement	1.5

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

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

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	1831	1935	1901	603±40
All	All	16479	17415	17109	5423

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 161.

5 of 2420 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:402:ASN:O	1:A:406:ILE:CD1	1.28	1.80	4	8
1:A:221:SER:O	1:A:299:TYR:CD2	1.26	1.88	5	1
1:A:402:ASN:O	1:A:405:PRO:HD2	1.26	1.24	4	8
1:A:402:ASN:O	1:A:406:ILE:HD13	1.22	1.34	2	3
1:A:359:PRO:O	1:A:361:GLY:N	1.16	1.78	2	9

6.3 Torsion angles

6.3.1 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 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	228/270 (84%)	137±2 (60±1%)	59±3 (26±1%)	33±3 (14±1%)	1 5
All	All	2052/2430 (84%)	1230 (60%)	528 (26%)	294 (14%)	1 5

5 of 73 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	301	LYS	9
1	A	360	PHE	9
1	A	417	LEU	9
1	A	439	MET	9
1	A	440	THR	9

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	204/243 (84%)	120±7 (59±4%)	84±7 (41±4%)	0 4
All	All	1836/2187 (84%)	1080 (59%)	756 (41%)	0 4

5 of 157 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	211	LEU	9
1	A	216	LYS	9
1	A	228	LEU	9
1	A	234	ARG	9
1	A	297	THR	9

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