wwPDB NMR Structure Validation Summary Report

Apr 20, 2024 – 04:05 PM EDT

PDB ID	:	1RXR
Title	:	HIGH RESOLUTION SOLUTION STRUCTURE OF THE RETINOID X RECEPTOR DNA BINDING DOMAIN, NMR, 20 STRUCTURE
Authors	:	Holmbeck, S.M.A.; Foster, M.P.; Casimiro, D.R.; Sem, D.S.; Dyson, H.J.; Wright, P.E.
Deposited on	:	1998-06-12

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:

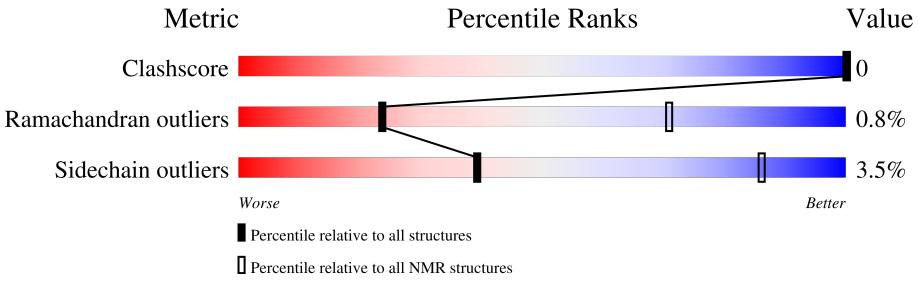
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson 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	83	81% · 17%

2 Ensemble composition and analysis

This entry contains 20 models. Model 19 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:133-A:180, A:187-A:207 (69)	0.82	19

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. No single-model clusters were found.

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1350 atoms, of which 669 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called RETINOIC ACID RECEPTOR-ALPHA.

Mol	Chain	Residues	Atoms						Trace
1	A	83	Total	C	H	N	O	S	0
			1348	413	669	136	120	10	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	195	ALA	CYS	engineered mutation	UNP P19793

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
2	A	2	Total	Zn
			2	2

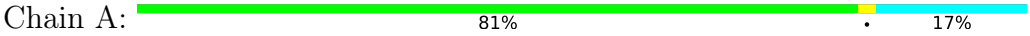
4 Residue-property plots (i)

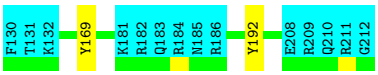
4.1 Average score per residue in the NMR ensemble

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: RETINOIC ACID RECEPTOR-ALPHA

Chain A:



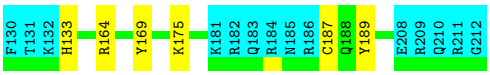


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 19. Colouring as in section 4.1 above.

• Molecule 1: RETINOIC ACID RECEPTOR-ALPHA

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *DISTANCE GEOMETRY AND RESTRAINED MOLECULAR DYNAMICS*.

Of the 83 calculated structures, 20 were deposited, based on the following criterion: *LOWEST RESTRAINT VIOLATIONS AND AMBER ENERGIES*.

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

Software name	Classification	Version
Amber	refinement	
SYBYL TRIAD	structure solution	TRIAD
DIANA	structure solution	
Amber	structure solution	

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

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

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	0.58±0.01	0±0/558 (0.0± 0.0%)	0.87±0.03	0±0/743 (0.0± 0.1%)
All	All	0.58	0/11160 (0.0%)	0.87	4/14860 (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	Planarity
1	A	0.0±0.0	2.4±1.2
All	All	0	48

There are no bond-length outliers.

All 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	164	ARG	NE-CZ-NH2	-6.25	117.17	120.30	14	1
1	A	169	TYR	CB-CG-CD1	-5.33	117.80	121.00	6	1
1	A	141	ARG	NE-CZ-NH2	-5.33	117.64	120.30	9	2

There are no chirality outliers.

5 of 10 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	192	TYR	Sidechain	12
1	A	189	TYR	Sidechain	10
1	A	169	TYR	Sidechain	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	150	TYR	Sidechain	4
1	A	158	PHE	Sidechain	4

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
All	All	11020	10620	10620	-

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 -.

There are no clashes.

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	69/83 (83%)	61±2 (88±3%)	8±2 (11±3%)	1±1 (1±1%)	24	71
All	All	1380/1660 (83%)	1212 (88%)	157 (11%)	11 (1%)	24	71

5 of 7 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	173	ASP	4
1	A	175	LYS	2
1	A	177	CYS	1
1	A	188	GLN	1
1	A	201	LYS	1

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	59/72 (82%)	57±1 (97±2%)	2±1 (3±2%)	39 86
All	All	1180/1440 (82%)	1139 (97%)	41 (3%)	39 86

5 of 17 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	174	ASN	8
1	A	153	GLU	6
1	A	172	ARG	5
1	A	141	ARG	4
1	A	207	GLU	2

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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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