



Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 04:39 PM BST

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 Full wwPDB NMR Structure Validation 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

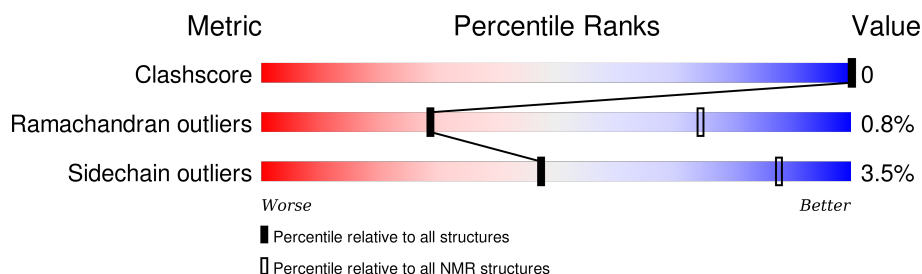
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 79%.

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 $\leq 5\%$.

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)	0.47	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 [i](#)

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	MUTATION	UNP P19793

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

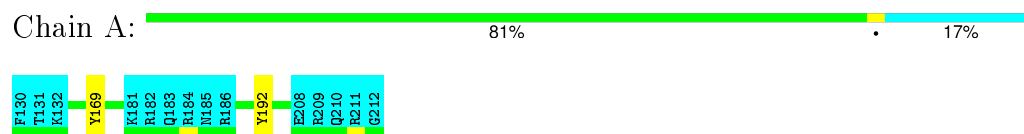
Mol	Chain	Residues	Atoms	
2	A	2	Total	Zn
			2	2

4 Residue-property plots [i](#)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA 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: RETINOIC ACID RECEPTOR-ALPHA

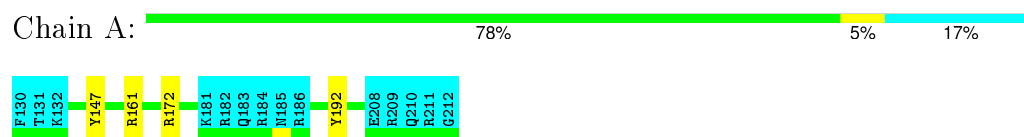


4.2 Scores per residue for each member of the ensemble

Colouring as in section [4.1](#) above.

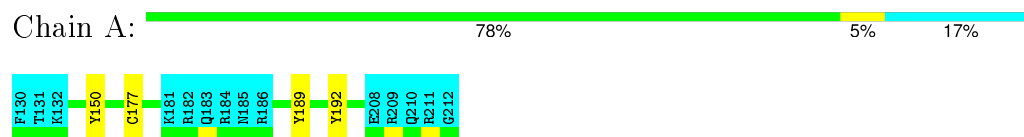
4.2.1 Score per residue for model 1

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



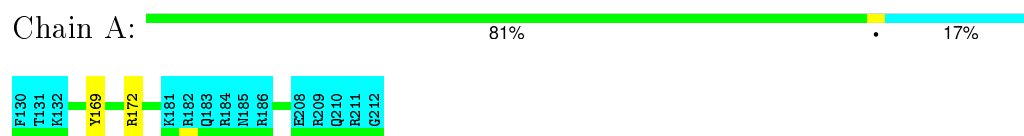
4.2.2 Score per residue for model 2

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



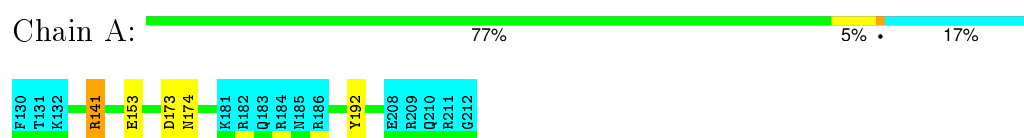
4.2.3 Score per residue for model 3

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



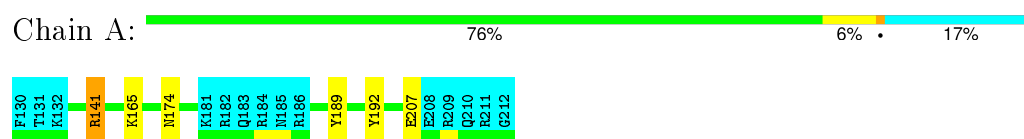
4.2.4 Score per residue for model 4

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



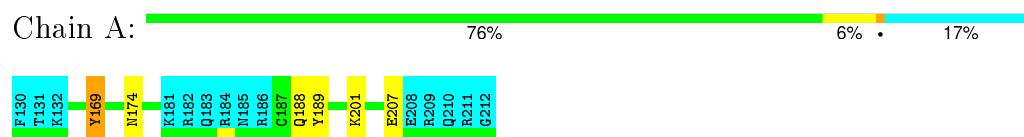
4.2.5 Score per residue for model 5

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



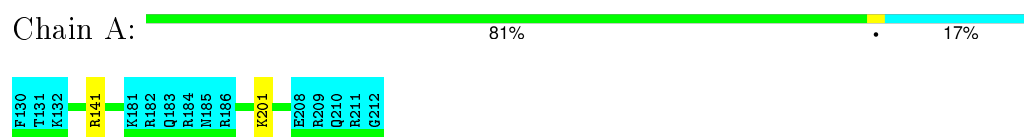
4.2.6 Score per residue for model 6

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



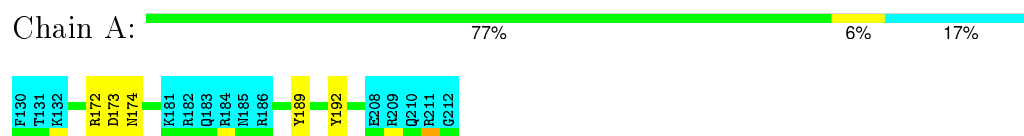
4.2.7 Score per residue for model 7

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



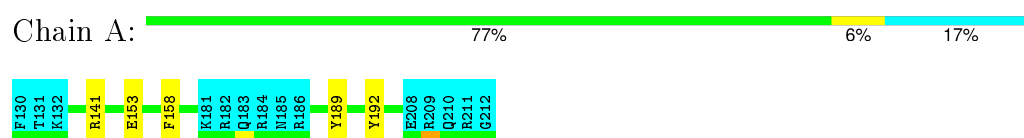
4.2.8 Score per residue for model 8

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



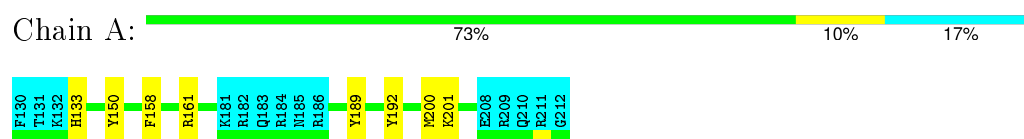
4.2.9 Score per residue for model 9

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



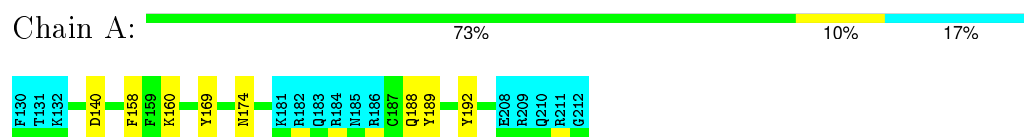
4.2.10 Score per residue for model 10

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



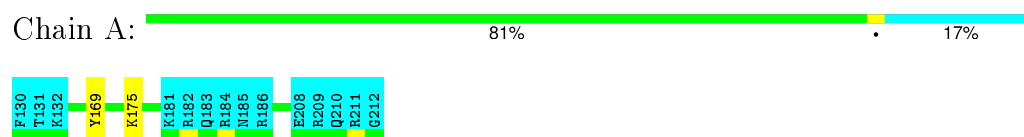
4.2.11 Score per residue for model 11

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



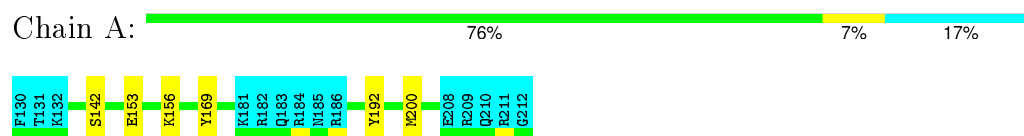
4.2.12 Score per residue for model 12

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



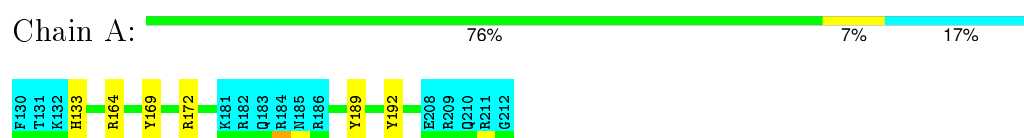
4.2.13 Score per residue for model 13

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



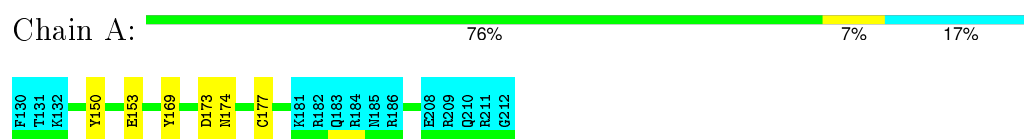
4.2.14 Score per residue for model 14

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



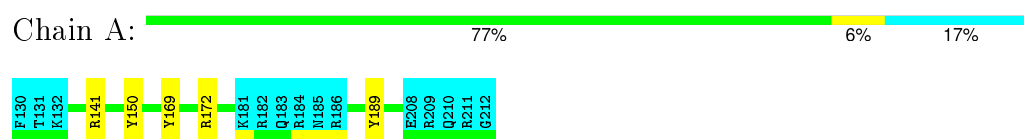
4.2.15 Score per residue for model 15

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



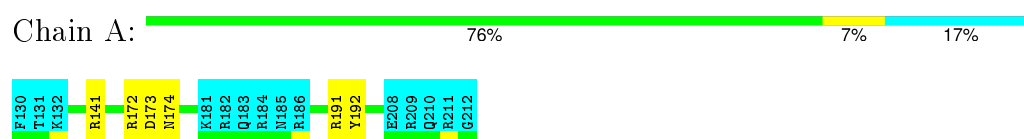
4.2.16 Score per residue for model 16

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



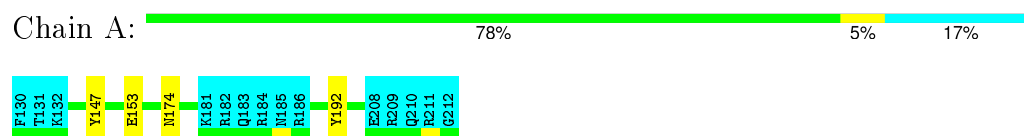
4.2.17 Score per residue for model 17

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



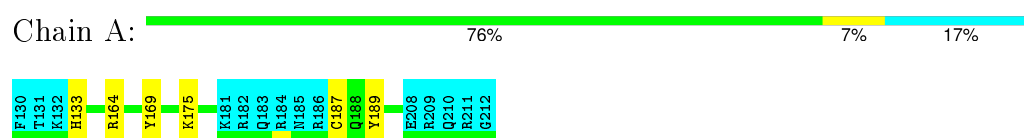
4.2.18 Score per residue for model 18

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



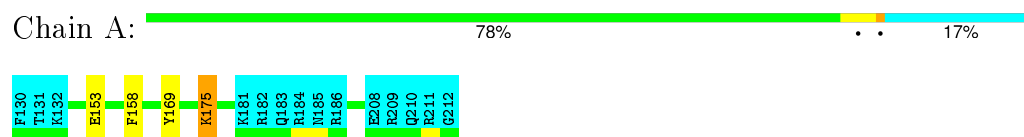
4.2.19 Score per residue for model 19 (medoid)

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



4.2.20 Score per residue for model 20

- Molecule 1: RETINOIC ACID RECEPTOR-ALPHA



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	

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)	BMRB entry 4153
Number of chemical shift lists	1
Total number of shifts	863
Number of shifts mapped to atoms	850
Number of unparsed shifts	0
Number of shifts with mapping errors	13
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	79%

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

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(°)	Ideal(°)	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.

All 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	158	PHE	Sidechain	4
1	A	150	TYR	Sidechain	4
1	A	147	TYR	Sidechain	2
1	A	161	ARG	Sidechain	2
1	A	141	ARG	Sidechain	2
1	A	191	ARG	Sidechain	1
1	A	172	ARG	Sidechain	1

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
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 [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	69/83 (83%)	61±2 (88±3%)	8±2 (11±3%)	1±1 (1±1%)	29	74
All	All	1380/1660 (83%)	1212 (88%)	157 (11%)	11 (1%)	29	74

All 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

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	175	LYS	2
1	A	187	CYS	1
1	A	177	CYS	1
1	A	188	GLN	1
1	A	201	LYS	1
1	A	133	HIS	1

6.3.2 Protein sidechains ⓘ

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%)	47	88
All	All	1180/1440 (82%)	1139 (97%)	41 (3%)	47	88

All 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	201	LYS	2
1	A	175	LYS	2
1	A	133	HIS	2
1	A	200	MET	2
1	A	207	GLU	2
1	A	188	GLN	1
1	A	160	LYS	1
1	A	165	LYS	1
1	A	164	ARG	1
1	A	156	LYS	1
1	A	142	SER	1
1	A	177	CYS	1
1	A	140	ASP	1

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](#)

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

The completeness of assignment taking into account all chemical shift lists is 79% for the well-defined parts and 73% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 4153

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

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	863
Number of shifts mapped to atoms	850
Number of unparsed shifts	0
Number of shifts with mapping errors	13
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. All 13 occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	83	NVA	N	115.86	-1.0	1
A	83	NVA	H	7.97	-1.0	1
A	83	NVA	HG2	1.47	-1.0	2
A	83	NVA	HD1	1.81	-1.0	1
A	83	NVA	CA	36.96	-1.0	1
A	83	NVA	HG3	1.56	-1.0	2
A	83	NVA	CG	55.33	-1.0	1
A	83	NVA	HD2	1.81	-1.0	1
A	83	NVA	HB2	3.07	-1.0	2
A	83	NVA	CB	42.05	-1.0	1
A	83	NVA	HD3	1.81	-1.0	1
A	83	NVA	CD	59.26	-1.0	1
A	83	NVA	HB3	3.28	-1.0	2

7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	81	1.49 ± 0.19	Should be applied
$^{13}\text{C}_\beta$	75	1.98 ± 0.12	Should be applied
$^{13}\text{C}'$	0	—	—
^{15}N	79	-0.82 ± 0.39	Should be applied

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 79%, i.e. 680 atoms were assigned a chemical shift out of a possible 862. 1 out of 6 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	276/345 (80%)	138/138 (100%)	69/138 (50%)	69/69 (100%)
Sidechain	347/445 (78%)	220/265 (83%)	124/151 (82%)	3/29 (10%)
Aromatic	57/72 (79%)	32/38 (84%)	25/32 (78%)	0/2 (0%)
Overall	680/862 (79%)	390/441 (88%)	218/321 (68%)	72/100 (72%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 73%, i.e. 803 atoms were assigned a chemical shift out of a possible 1098. 1 out of 6 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	316/415 (76%)	158/166 (95%)	80/166 (48%)	78/83 (94%)
Sidechain	430/602 (71%)	275/360 (76%)	150/193 (78%)	5/49 (10%)
Aromatic	57/81 (70%)	32/43 (74%)	25/36 (69%)	0/2 (0%)
Overall	803/1098 (73%)	465/569 (82%)	255/395 (65%)	83/134 (62%)

7.1.4 Statistically unusual chemical shifts [i](#)

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	200	MET	CE	48.46	26.97 – 7.37	16.0
1	A	198	MET	CE	48.07	26.97 – 7.37	15.8
1	A	188	GLN	HB2	0.56	3.30 – 0.80	-6.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	148	GLY	N	129.71	129.07 – 90.27	5.2
1	A	156	LYS	HD2	0.43	2.76 – 0.46	-5.1
1	A	157	GLY	N	129.41	129.07 – 90.27	5.1

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:

