wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 03:54 PM BST

PDB ID	:	1LO1
Title	:	ESTROGEN RELATED RECEPTOR 2 DNA BINDING DOMAIN IN COMPLEX WITH DNA
Authors	:	Gearhart, M.D.; Holmbeck, S.M.A.; Evans, R.M.; Dyson, H.J.; Wright, P.E.
Deposited on	:	2002-05-05

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:

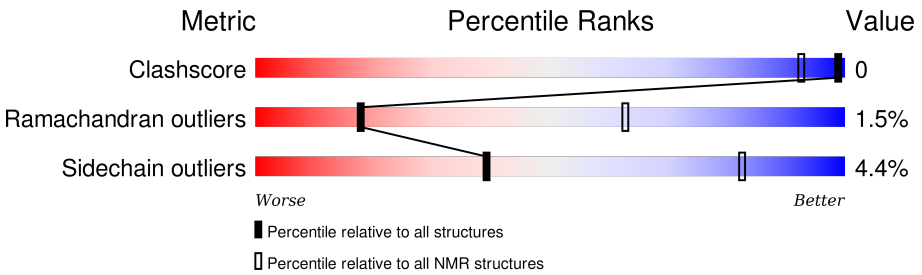
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 66%.

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	B	13	15% 77% 8%
2	C	13	23% 77%
3	A	98	86% • • 8%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 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:101-A:186 (86)	0.62	1

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 5 clusters and 3 single-model clusters were found.

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

3 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2254 atoms, of which 1020 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called 5'-D(*GP*CP*TP*CP*AP*AP*GP*GP*TP*CP*AP*CP*G)-3'.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	B	13	409	126	145	51	75	12	0

- Molecule 2 is a DNA chain called 5'-D(*CP*GP*TP*GP*AP*CP*CP*TP*TP*GP*AP*GP*C)-3'.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	C	13	409	126	146	48	77	12	0

- Molecule 3 is a protein called Steroid hormone receptor ERR2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
3	A	90	1434	438	729	137	120	10	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	163	ALA	CYS	ENGINEERED	UNP O95718
A	193	ASN	SER	SEE REMARK 999	UNP O95718

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

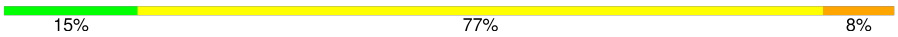
Mol	Chain	Residues	Atoms	
			Total	Zn
4	A	2	2	2

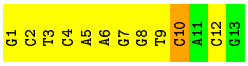
4 Residue-property plots (i)

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.

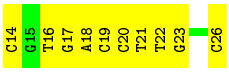
- Molecule 1: 5'-D(*GP*CP*TP*CP*AP*AP*GP*GP*TP*CP*AP*CP*G)-3'

Chain B: 



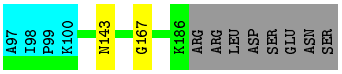
- Molecule 2: 5'-D(*CP*GP*TP*GP*AP*CP*CP*TP*TP*GP*AP*GP*C)-3'

Chain C:



- Molecule 3: Steroid hormone receptor ERR2

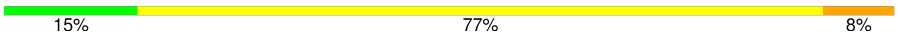
Chain A:



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

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: 5'-D(*GP*CP*TP*CP*AP*AP*GP*GP*TP*CP*AP*CP*G)-3'

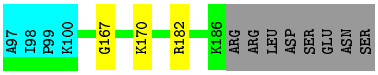
Chain B: 

- Molecule 2: 5'-D(*CP*GP*TP*GP*AP*CP*CP*TP*TP*GP*AP*GP*C)-3'

Chain C:

- Molecule 3: Steroid hormone receptor ERR2

Chain A:



5 Refinement protocol and experimental data overview (i)

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

Of the 32 calculated structures, 20 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
Amber	refinement	6.0
DYANA	refinement	1.5
Amber	structure solution	6.0

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 5363
Number of chemical shift lists	3
Total number of shifts	1276
Number of shifts mapped to atoms	1276
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	66%

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)

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	B	1.25±0.02	0±0/296 (0.0±0.0%)	2.00±0.05	14±2/455 (3.2±0.3%)
2	C	1.28±0.02	0±0/294 (0.0±0.0%)	1.96±0.06	13±2/452 (2.9±0.4%)
3	A	0.63±0.00	0±0/684 (0.0±0.0%)	0.92±0.02	0±1/907 (0.0±0.1%)
All	All	0.98	0/25480 (0.0%)	1.55	555/36280 (1.5%)

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	B	0.0±0.0	3.7±1.1
2	C	0.0±0.0	4.0±1.5
3	A	0.0±0.0	0.1±0.4
All	All	0	156

There are no bond-length outliers.

5 of 58 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	B	7	DG	O4'-C1'-N9	12.68	116.88	108.00	1	15
1	B	1	DG	O4'-C1'-N9	10.20	115.14	108.00	18	10
2	C	23	DG	O4'-C1'-N9	9.64	114.75	108.00	2	19
1	B	5	DA	O4'-C1'-N9	9.42	114.59	108.00	12	14
2	C	14	DC	O4'-C1'-N1	9.30	114.51	108.00	9	8

There are no chirality outliers.

5 of 28 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	B	3	DT	Sidechain	15
2	C	22	DT	Sidechain	14
2	C	16	DT	Sidechain	14
1	B	10	DC	Sidechain	12
1	B	8	DG	Sidechain	11

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
3	A	676	693	692	0±0
All	All	24100	19680	19740	3

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 0.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
3:A:179:ARG:NH1	3:A:180:GLY:H	0.49	2.05	19	1
3:A:114:HIS:CE1	3:A:124:LYS:HE2	0.42	2.50	12	1
3:A:155:CYS:SG	3:A:158:CYS:HB2	0.40	2.56	7	1

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	
3	A	85/98 (87%)	74±2 (87±2%)	10±2 (11±3%)	1±1 (1±1%)	18	63
All	All	1700/1960 (87%)	1485 (87%)	190 (11%)	25 (1%)	18	63

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

Mol	Chain	Res	Type	Models (Total)
3	A	167	GLY	13
3	A	107	GLY	5
3	A	105	VAL	4
3	A	141	ALA	2
3	A	178	VAL	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
3	A	71/82 (87%)	68±2 (96±3%)	3±2 (4±3%)	39 82
All	All	1420/1640 (87%)	1357 (96%)	63 (4%)	39 82

5 of 20 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)
3	A	143	ASN	11
3	A	170	LYS	10
3	A	134	ASN	5
3	A	124	LYS	5
3	A	159	ARG	5

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 (i)

The completeness of assignment taking into account all chemical shift lists is 66% for the well-defined parts and 67% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 5363

Chemical shift list name: *assigned_ chem_ shift_ list_ 1*

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	1090
Number of shifts mapped to atoms	1090
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	16

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	97	1.37 ± 0.12	Should be applied
$^{13}C_{\beta}$	87	2.31 ± 0.19	Should be applied
$^{13}C'$	0	—	—
^{15}N	91	0.15 ± 0.54	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 54%, i.e. 872 atoms were assigned a chemical shift out of a possible 1619. 0 out of 10 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	338/428 (79%)	169/171 (99%)	86/172 (50%)	83/85 (98%)
Sidechain	476/606 (79%)	300/361 (83%)	170/201 (85%)	6/44 (14%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	58/67 (87%)	31/35 (89%)	27/30 (90%)	0/2 (0%)
Overall	872/1619 (54%)	500/877 (57%)	283/577 (49%)	89/165 (54%)

7.1.4 Statistically unusual chemical shifts (i)

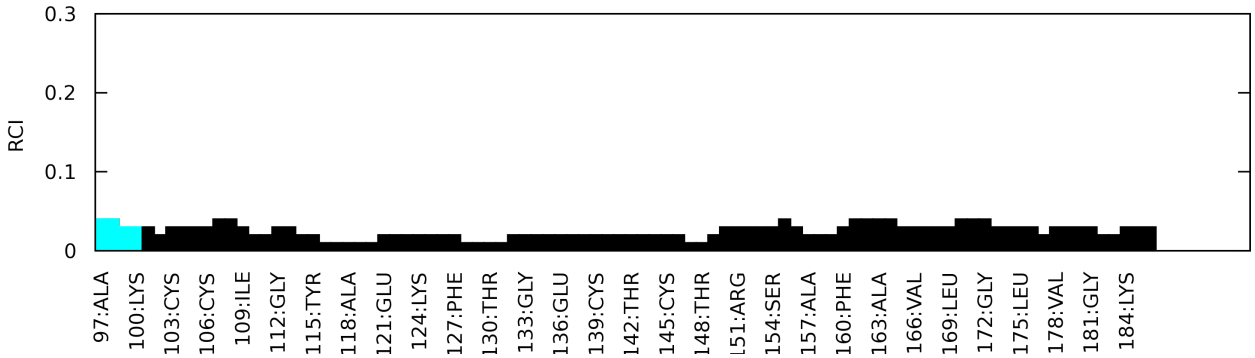
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	126	PHE	CD2	62.30	137.34 – 125.84	-60.3
1	A	127	PHE	CD2	64.70	137.34 – 125.84	-58.2
1	A	160	PHE	CD2	64.74	137.34 – 125.84	-58.1
1	A	126	PHE	CE2	65.20	136.81 – 124.71	-54.2
1	A	160	PHE	CE2	65.70	136.81 – 124.71	-53.8
1	A	113	TYR	CE1	52.10	124.14 – 111.74	-53.1
1	A	185	TYR	CE2	51.50	124.68 – 111.18	-49.2
1	A	115	TYR	CE2	52.53	124.68 – 111.18	-48.4
1	A	115	TYR	CD2	66.23	140.11 – 125.31	-44.9
1	A	113	TYR	CD2	66.36	140.11 – 125.31	-44.8
1	A	137	TYR	CD2	66.40	140.11 – 125.31	-44.8
1	A	185	TYR	CD2	67.20	140.11 – 125.31	-44.3
1	A	114	HIS	CE1	56.50	149.70 – 125.30	-33.2
1	A	124	LYS	NZ	81.06	49.86 – 18.16	14.8
1	A	152	ARG	CD	49.83	47.57 – 38.77	7.6
1	A	156	GLN	HB3	0.28	3.37 – 0.67	-6.4

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:



7.2 Chemical shift list 2

File name: BMRB entry 5363

Chemical shift list name: *assigned_chem_shift_list_2*

7.2.1 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	92
Number of shifts mapped to atoms	92
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

7.2.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

7.2.3 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 6%, i.e. 100 atoms were assigned a chemical shift out of a possible 1619. 0 out of 10 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}**H**	^{13}**C**	^{15}**N**
Backbone	0/428 (0%)	0/171 (0%)	0/172 (0%)	0/85 (0%)
Sidechain	0/606 (0%)	0/361 (0%)	0/201 (0%)	0/44 (0%)
Aromatic	0/67 (0%)	0/35 (0%)	0/30 (0%)	0/2 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Overall	100/1619 (6%)	100/877 (11%)	0/577 (0%)	0/165 (0%)

7.2.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.2.5 Random Coil Index (RCI) plots (i)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (assigned_chem_shift_list_2). RCI is only applicable to proteins.

7.3 Chemical shift list 3

File name: BMRB entry 5363

Chemical shift list name: *assigned_chem_shift_list_3*

7.3.1 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	94
Number of shifts mapped to atoms	94
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

7.3.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

7.3.3 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 6%, i.e. 103 atoms were assigned a chemical shift out of a possible 1619. 0 out of 10 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	0/428 (0%)	0/171 (0%)	0/172 (0%)	0/85 (0%)
Sidechain	0/606 (0%)	0/361 (0%)	0/201 (0%)	0/44 (0%)
Aromatic	0/67 (0%)	0/35 (0%)	0/30 (0%)	0/2 (0%)
Overall	103/1619 (6%)	103/877 (12%)	0/577 (0%)	0/165 (0%)

7.3.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.3.5 Random Coil Index (RCI) plots (i)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (assigned_chem_shift_list_3). RCI is only applicable to proteins.