



# wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 03:23 PM BST

PDB ID : 1JUA  
Title : Solution Structure of the Deoxyribose HIV-1Lai Initiation Sequence Stable Dimer  
Authors : Barbault, F.; Huynh-Dinh, T.; Paoletti, J.; Lancelot, G.  
Deposited on : 2001-08-24

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto: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

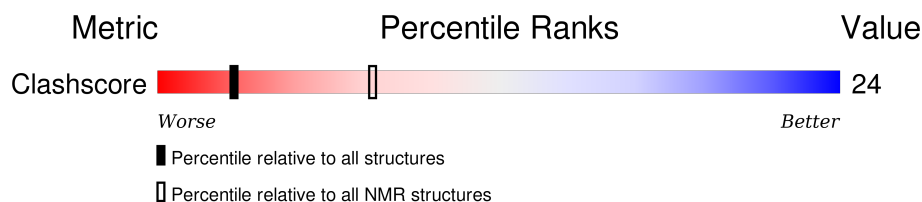
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

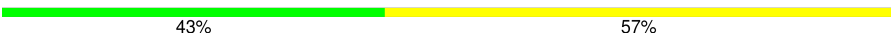

The overall completeness of chemical shifts assignment was not calculated.

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	114402	11133

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  $\geq 3$ , 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	23		
1	B	23		

## 2 Ensemble composition and analysis ⓘ

This entry contains 13 models. This entry does not contain protein, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1458 atoms, of which 516 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
1	A	23	Total	C	H	N	O	P	0
			729	223	258	92	134	22	
1	B	23	Total	C	H	N	O	P	0
			729	223	258	92	134	22	

## 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(\*CP\*TP\*TP\*GP\*CP\*TP\*GP\*AP\*AP\*GP\*CP\*GP\*CP\*GP\*CP\*AP\*CP\*GP\*GP\*CP\*AP\*AP\*G)-3'

Chain A: 



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

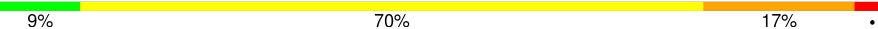
Chain B: 



### 4.2 Residue scores for the representative (author defined) model from the NMR ensemble

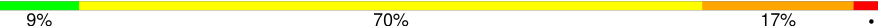
The representative model is number 9. Colouring as in section 4.1 above.

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

Chain A: 



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

Chain B: 

C1	T2	T3	G4	C5	T6	G7	A8	A9	G10	C11	G12	G13	G14	C15	A16	C17	G18	G19	C20	A21	A22	G23
----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *Simulated annealing Relaxation matrix refinement Chemical shifts back-calculation refinement*.

Of the 13 calculated structures, 13 were deposited, based on the following criterion: *back calculated data agree with experimental NOESY spectrum, structures with acceptable covalent geometry, structures with the least restraint violations, structures with the lowest energy*.

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

Software name	Classification	Version
CNS	structure solution	1.0
Nuchemics	refinement	

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

## 6 Model quality

### 6.1 Standard geometry

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	0.35±0.11	0±0/529 (0.0±0.0%)	0.86±0.11	1±1/815 (0.1±0.1%)
1	B	0.35±0.11	0±0/529 (0.0±0.0%)	0.87±0.11	1±1/815 (0.1±0.1%)
All	All	0.37	0/13754 (0.0%)	0.87	20/21190 (0.1%)

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	0.2±0.8
1	B	0.0±0.0	0.2±0.8
All	All	0	6

There are no bond-length outliers.

5 of 8 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	B	10	DG	C5'-C4'-O4'	-5.77	98.33	109.30	9	1
1	A	10	DG	C5'-C4'-O4'	-5.74	98.40	109.30	9	1
1	B	15	DC	C1'-O4'-C4'	-5.73	104.37	110.10	9	1
1	A	15	DC	C1'-O4'-C4'	-5.71	104.39	110.10	9	1
1	B	13	DC	C1'-O4'-C4'	-5.50	104.59	110.10	9	1

There are no chirality outliers.

5 of 6 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	13	DC	Sidechain	1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group	Models (Total)
1	B	13	DC	Sidechain	1
1	B	12	DG	Sidechain	1
1	A	12	DG	Sidechain	1
1	B	11	DC	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
1	A	471	258	258	23±22
1	B	471	258	258	23±22
All	All	12246	6708	6708	448

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 24.

5 of 168 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:DG:N2	1:B:15:DC:N3	0.91	2.18	9	1
1:A:7:DG:H1	1:A:17:DC:H42	0.90	1.10	9	1
1:A:15:DC:N3	1:B:10:DG:N2	0.89	2.19	9	1
1:B:15:DC:H2'	1:B:16:DA:N7	0.88	1.83	9	10
1:A:14:DG:O6	1:B:8:DA:N1	0.88	2.05	9	1

## 6.3 Torsion angles [i](#)

### 6.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

### 6.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

### 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](#)

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

No chemical shift data were provided