



Full wwPDB NMR Structure Validation Report ⓘ

Dec 13, 2023 – 10:15 PM EST

PDB ID : 2JNR

Title : Discovery and optimization of a natural HIV-1 entry inhibitor targeting the gp41 fusion peptide

Authors : Munch, J.; Standker, L.; Adermann, K.; Schulz, A.; Pohlmann, S.; Chaipan, C.; Biet, T.; Peters, T.; Meyer, B.; Wilhelm, D.; Lu, H.; Jing, W.; Jiang, S.; Forssmann, W.; Kirchhoff, F.

Deposited on : 2007-02-01

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

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

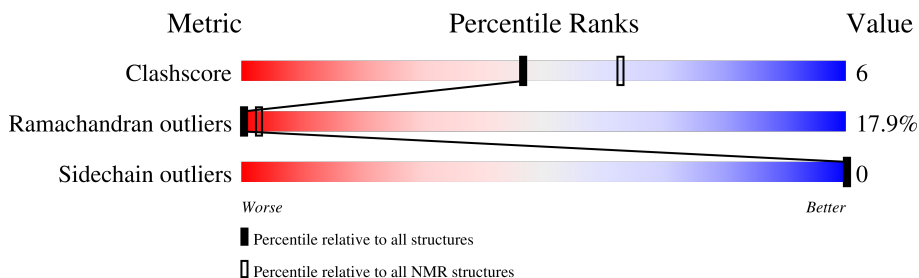
1 Overall quality at a glance

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	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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	20	50% 35% 10% 5%
2	B	23	65% 22% 9% .

2 Ensemble composition and analysis

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 619 atoms, of which 312 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called VIR165.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	20	315	109	157	22	25	2	0

- Molecule 2 is a protein called ENV polyprotein.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
2	B	23	304	95	155	26	27	1	0

There is a discrepancy between the modelled and reference sequences:

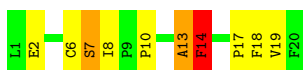
Chain	Residue	Modelled	Actual	Comment	Reference
B	122	ARG	ALA	engineered mutation	UNP Q72502

4 Residue-property plots [i](#)

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.

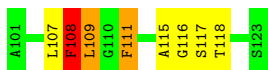
- Molecule 1: VIR165

Chain A:  50% 35% 10% 5%



- Molecule 2: ENV polyprotein

Chain B:  65% 22% 9%



5 Refinement protocol and experimental data overview

The models were refined using the following method: *molecular dynamics*.

Of the 1 calculated structures, 1 were deposited, based on the following criterion: *all calculated structures submitted*.

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

Software name	Classification	Version
SYBYL	refinement	

No chemical shift data was provided.

6 Model quality i

6.1 Standard geometry i

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	1.79	1/165 (0.6%)	2.20	6/224 (2.7%)
2	B	1.59	0/150 (0.0%)	2.22	4/198 (2.0%)
All	All	1.70	1/315 (0.3%)	2.21	10/422 (2.4%)

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	2
2	B	0	2
All	All	0	4

All bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2	GLU	CD-OE1	11.29	1.38	1.25

All angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	108	PHE	CG-CD2-CE2	7.64	129.21	120.80
2	B	108	PHE	CZ-CE2-CD2	-7.37	111.26	120.10
1	A	14	PHE	CB-CG-CD1	-7.15	115.80	120.80
1	A	19	VAL	CB-CA-C	-6.85	98.39	111.40
2	B	111	PHE	CB-CG-CD2	-6.34	116.36	120.80
2	B	109	LEU	CA-C-N	6.12	128.45	116.20
1	A	14	PHE	CB-CG-CD2	5.48	124.63	120.80
1	A	7	SER	CB-CA-C	5.28	120.14	110.10
1	A	17	PRO	N-CA-CB	-5.20	96.89	102.60
1	A	10	PRO	CA-N-CD	5.09	118.83	111.70

There are no chirality outliers.

All planar outliers are listed below.

Mol	Chain	Res	Type	Group
1	A	14	PHE	Sidechain
1	A	18	PHE	Sidechain
2	B	108	PHE	Sidechain
2	B	111	PHE	Sidechain

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	158	157	156	1
2	B	149	155	152	3
All	All	307	312	308	4

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 7.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)
2:B:107:LEU:O	2:B:108:PHE:C	0.53	2.45
1:A:13:ALA:O	1:A:14:PHE:CD1	0.48	2.66
2:B:108:PHE:O	2:B:109:LEU:HB2	0.44	2.11
2:B:117:SER:O	2:B:118:THR:C	0.43	2.55

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	18/20 (90%)	6 (33%)	8 (44%)	4 (22%)	0 2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	21/23 (91%)	11 (52%)	7 (33%)	3 (14%)	1	5
All	All	39/43 (91%)	17 (44%)	15 (38%)	7 (18%)	0	3

All 7 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	6	CYS
1	A	7	SER
1	A	8	ILE
1	A	13	ALA
2	B	108	PHE
2	B	115	ALA
2	B	116	GLY

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	18/18 (100%)	18 (100%)	0 (0%)	100	100
2	B	12/12 (100%)	12 (100%)	0 (0%)	100	100
All	All	30/30 (100%)	30 (100%)	0 (0%)	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

No chemical shift data were provided