



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 02:19 am GMT

PDB ID : 1S1U
Title : Crystal structure of L100I mutant HIV-1 reverse transcriptase in complex with nevirapine
Authors : Ren, J.; Nichols, C.E.; Chamberlain, P.P.; Stammers, D.K.
Deposited on : 2004-01-07
Resolution : 3.00 Å(reported)

This is a wwPDB X-ray 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/XrayValidationReportHelp>

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:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk31020
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk31020

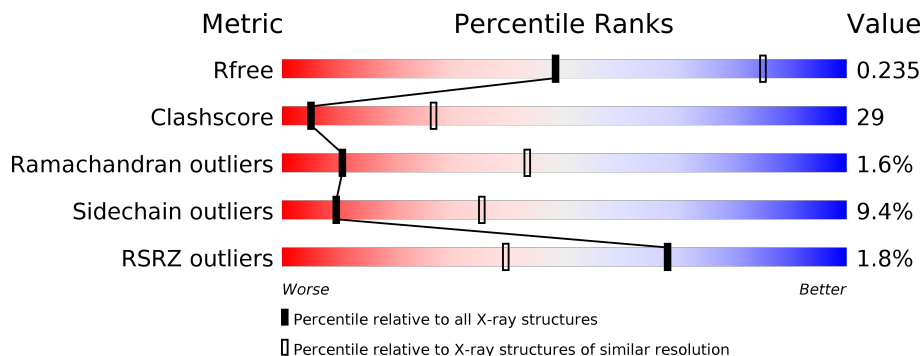
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 Å.

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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	1851 (3.00-3.00)
Clashscore	122126	2167 (3.00-3.00)
Ramachandran outliers	120053	2101 (3.00-3.00)
Sidechain outliers	120020	2104 (3.00-3.00)
RSRZ outliers	108989	1751 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	560	<div> <div>3%</div> <div> <div></div> <div>45%</div> <div>42%</div> <div>7%</div> <div>6%</div> </div> </div>
2	B	440	<div> <div></div> <div> <div>45%</div> <div>43%</div> <div>•</div> <div>9%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7660 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Reverse transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	525	Total	C	N	O	S	0	0	0
			4310	2792	714	796	8			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	ILE	LEU	ENGINEERED	UNP P04585
A	280	CSD	CYS	MODIFIED RESIDUE	UNP P04585

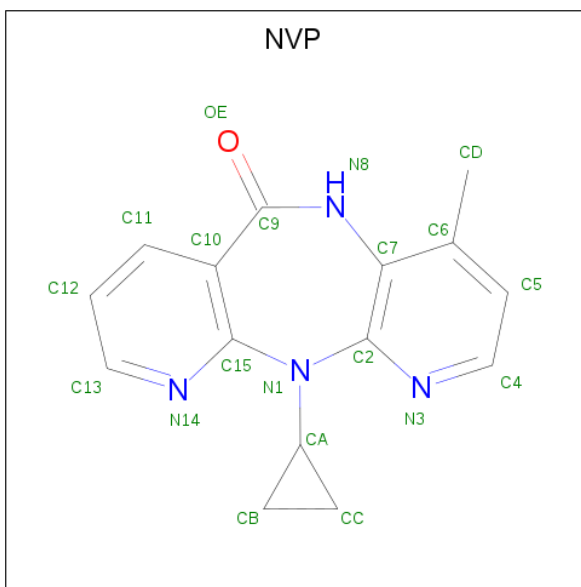
- Molecule 2 is a protein called Reverse transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	401	Total	C	N	O	S	0	0	0
			3330	2173	549	601	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	100	ILE	LEU	ENGINEERED	UNP P04585

- Molecule 3 is 11-CYCLOPROPYL-5,11-DIHYDRO-4-METHYL-6H-DIPYRIDO[3,2-B:2',3'-E][1,4]DIAZEPIN-6-ONE (three-letter code: NVP) (formula: C₁₅H₁₄N₄O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			20	15	4	1		

G384	K385	T386	F387	K388	F389	K390	T393	Q394	K395	K398	W398	W401	W402	T403	E404	Y405	W406	Q407	W414	F415	F416	W417	N418	T419	P420	P421	L422	W423	K424	L425	W426	Y427	Q428	LEU	GLU	LYS	GLU	PRO	ILE	VAL	GLY	ALA	GLU	THR	PHE												
P313	G316	V317	Y318	K323	A327	E328	I329	Q330	T331	Q332	G335	Q336	W337	T338	Y339	Q340	I341	Y342	Q343	E344	P345	F346	K347	N348	L349	K350	T351	G352	K353	Y354	A355	R356	MET	ARG	GLY	ALA	HIS	N362	N363	D364	V365	K366	Q367	L368	T369	E370	A371	V372	I375	S379	I380	K311	W383				
P226	F227	L228	W229	T165	W230	Q231	Y232	E233	L234	W239	T240	Y241	P247	E248	R249	W252	T253	V254	Q258	V261	G262	K263	N265	W266	A267	S268	Q269	I270	V276	R277	C280	L283	T286	L285	T296	E297	E298	I301	E302	L303	A304	E305	N306	I309	L310	K311	E312										
S162	S163	M164	T165	K166	I167	L168	E169	P170	F171	R172	K173	Q174	N175	P176	D177	I178	V179	Y183	M184	D185	D186	L187	Y188	Y189	G190	S191	N192	L193	F194	I195	G196	Q197	K201	I202	E203	E204	L205	R206	Q207	H208	L209	L210	R211	W212	G213	LEU	THR	THR	PRO	ASP	LYS	LYS	HIS	GLN	LYS	GLU	P225

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	139.80Å 115.50Å 66.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.90 – 3.00 29.90 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.5 (29.90-3.00) 98.6 (29.90-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.27 (at 3.00Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.227 , 0.285 0.217 , 0.235	Depositor DCC
R_{free} test set	1060 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å ²)	84.0	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 90.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7660	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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

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 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	0.52	0/4417	0.75	1/6005 (0.0%)
2	B	0.51	0/3426	0.73	1/4653 (0.0%)
All	All	0.52	0/7843	0.74	2/10658 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	90	VAL	N-CA-C	-5.53	96.06	111.00
2	B	229	TRP	CA-CB-CG	-5.09	104.02	113.70

There are no chirality outliers.

There are no planarity outliers.

5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4310	0	4343	243	0
2	B	3330	0	3353	207	0
3	A	20	0	14	0	0
All	All	7660	0	7710	440	0

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 29.

The worst 5 of 440 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:469:LEU:HD21	1:A:480:GLN:HG3	1.17	1.10
2:B:66:LYS:HG2	2:B:230:MET:HA	1.35	1.07
1:A:195:ILE:H	1:A:195:ILE:HD13	1.25	1.01
1:A:295:LEU:HB3	1:A:300:GLU:HG3	1.46	0.96
1:A:517:LEU:HA	1:A:520:GLN:HE21	1.30	0.95

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 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 X-ray entries followed by that with respect to entries of similar resolution.

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	520/560 (93%)	464 (89%)	45 (9%)	11 (2%)	8	36
2	B	393/440 (89%)	346 (88%)	43 (11%)	4 (1%)	17	56
All	All	913/1000 (91%)	810 (89%)	88 (10%)	15 (2%)	11	43

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	114	ALA
1	A	195	ILE
1	A	402	TRP
1	A	170	PRO
1	A	230	MET

5.3.2 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 X-ray entries followed by that with respect to entries of similar resolution.

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	473/499 (95%)	419 (89%)	54 (11%)	6	26
2	B	366/400 (92%)	341 (93%)	25 (7%)	17	52
All	All	839/899 (93%)	760 (91%)	79 (9%)	9	35

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	356	ARG
1	A	470	THR
2	B	297	GLU
1	A	362	THR
1	A	396	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	507	GLN
2	B	147	ASN
2	B	348	ASN
1	A	509	GLN
1	A	520	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
1	CSD	A	280	1	4,7,8	1.02	0	2,8,10	4.89	1 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CSD	A	280	1	-	1/2/6/8	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	280	CSD	OD1-SG-CB	6.79	118.46	105.54

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	280	CSD	CA-CB-SG-OD1

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	NVP	A	999	-	15,23,23	1.42	4 (26%)	13,34,34	1.08	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NVP	A	999	-	-	0/0/6/6	0/2/4/4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	999	NVP	C12-C11	2.16	1.41	1.36
3	A	999	NVP	C13-N14	2.18	1.36	1.32
3	A	999	NVP	C4-N3	2.27	1.37	1.32
3	A	999	NVP	C10-C15	2.67	1.44	1.41

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å ²)	Q<0.9
1	A	524/560 (93%)	-0.36	15 (2%)	51	23	36, 87, 141, 150	0
2	B	401/440 (91%)	-0.34	2 (0%)	90	74	40, 88, 133, 149	0
All	All	925/1000 (92%)	-0.35	17 (1%)	68	39	36, 88, 136, 150	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	67	ASP	4.7
1	A	470	THR	3.8
1	A	471	ASP	3.6
1	A	15	GLY	2.9
1	A	137	ASN	2.8

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
1	CSD	A	280	8/9	0.93	0.16	78,81,85,89	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	NVP	A	999	20/20	0.97	0.23	58,63,77,78	0

6.5 Other polymers

There are no such residues in this entry.