



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:05 PM GMT

PDB ID : 1S1X
Title : Crystal structure of V108I 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 : 2.80 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
<http://wwpdb.org/validation/2016/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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

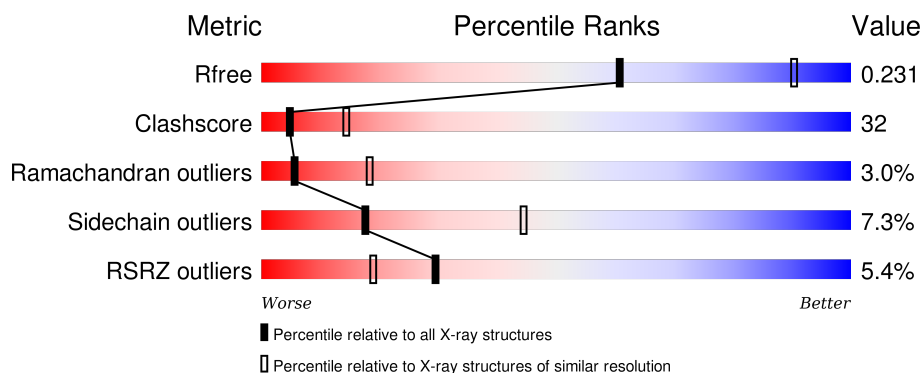
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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>5%</div> <div>43%</div> <div>48%</div> <div>7%</div> <div>.</div> </div>
2	B	440	<div> <div>5%</div> <div>45%</div> <div>41%</div> <div>5%</div> <div>10%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7790 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	546	Total	C	N	O	S	0	0	0
			4459	2883	743	825	8			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	108	ILE	VAL	ENGINEERED	UNP P04585
A	280	CSD	CYS	MODIFIED RESIDUE	UNP P04585

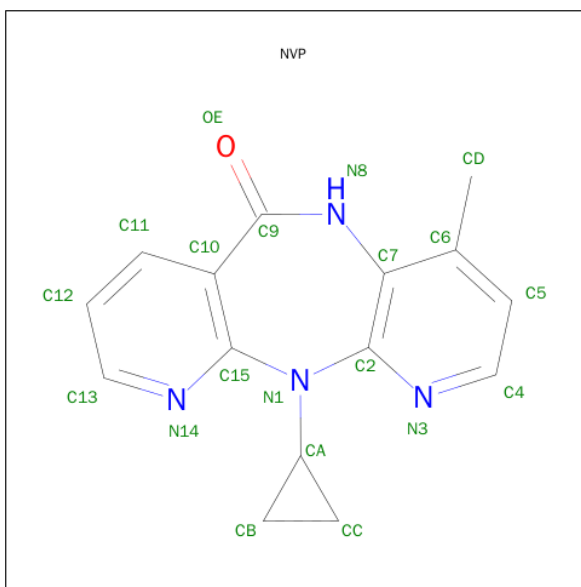
- Molecule 2 is a protein called Reverse transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	397	Total	C	N	O	S	0	0	0
			3290	2139	547	597	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	108	ILE	VAL	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		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	14	Total	O	0	0
			14	14		
4	B	7	Total	O	0	0
			7	7		

I380	I381	I382	K383	G384	K385	T386	F387	K388	F389	K390	L391	P392	I393	K394	K395	W401	W402	T403	W406	Q407	A408	T409	F416	V417	W418	P421	L422	W423	K424	L425	W426	W427	Q428	LEU	GLU	LYS	GLU	PRO	ILE	VAL	GLY	ALA	ALA	GLU	THR	PHE															
E291	L214	T215	THR	PRO	ASP	LYS	LYS	HIS	GLN	LYS	GLU	PRO	PRO	PHE	LEU	TRP	MET	GLY	TYR	E233	L234	H235	P236	D237	K238	W239	V241	I244	E248	K249	W252	T253	V254	Q258	K259	L260	W266	A267	S268	Q269	I274	R277	Q278	K281	L282	L283	R284	G285	T286	K287											
K154	G155	S156	P157	A158	I159	F160	Q161	S162	S163	M164	T165	K166	I167	L168	E169	P170	F171	R172	K173	Q174	M175	P176	D177	I178	V179	I180	Y181	Q182	Y183	M184	D185	L186	L187	K126	Y188	V189	G190	S191	D192	L193	E194	I195	G196	Q197	H198	R199	T200	K201	I202	E203	E204	L205	R206	Q207	H147	V148	L149	F150	Q151	G152	W153

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	138.40 Å 109.00 Å 72.80 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.57 – 2.80 29.57 – 2.82	Depositor EDS
% Data completeness (in resolution range)	98.6 (29.57-2.80) 98.7 (29.57-2.82)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 2.80 Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.235 , 0.295 0.231 , 0.231	Depositor DCC
R_{free} test set	1313 reflections (5.11%)	DCC
Wilson B-factor (Å ²)	74.6	Xtriage
Anisotropy	0.542	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 78.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 27001 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7790	wwPDB-VP
Average B, all atoms (Å ²)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.51	0/4564	0.73	2/6199 (0.0%)
2	B	0.49	0/3380	0.69	2/4587 (0.0%)
All	All	0.50	0/7944	0.71	4/10786 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	356	ARG	N-CA-C	5.74	126.50	111.00
1	A	19	PRO	N-CA-C	5.59	126.65	112.10
1	A	93	GLY	N-CA-C	-5.36	99.70	113.10
2	B	401	TRP	N-CA-C	5.34	125.42	111.00

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	4459	0	4501	283	0
2	B	3290	0	3321	217	0
3	A	20	0	14	0	0
4	A	14	0	0	1	0
4	B	7	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7790	0	7836	492	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 32.

The worst 5 of 492 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)
2:B:358:ARG:HD2	2:B:358:ARG:H	1.16	1.07
2:B:191:SER:HB2	2:B:193:LEU:HD13	1.38	1.04
2:B:180:ILE:HG12	2:B:189:VAL:HG13	1.45	0.96
1:A:50:ILE:HG13	1:A:51:GLY:H	1.30	0.96
1:A:57:ASN:HB2	1:A:143:ARG:HH22	1.32	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	539/560 (96%)	446 (83%)	74 (14%)	19 (4%)	4	15
2	B	389/440 (88%)	335 (86%)	45 (12%)	9 (2%)	8	26
All	All	928/1000 (93%)	781 (84%)	119 (13%)	28 (3%)	5	18

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	GLU
1	A	135	ILE
1	A	538	ALA
2	B	65	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	423	VAL

5.3.2 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 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	487/499 (98%)	445 (91%)	42 (9%)	13	36
2	B	362/400 (90%)	342 (94%)	20 (6%)	27	59
All	All	849/899 (94%)	787 (93%)	62 (7%)	17	44

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	324	ASP
1	A	402	TRP
2	B	325	LEU
1	A	362	THR
1	A	413	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	278	GLN
1	A	474	ASN
2	B	332	GLN
1	A	336	GLN
1	A	367	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	3,7,8	0.92	0	3,8,10	4.84	1 (33%)

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	8.15	118.99	105.40

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	280	CSD	1	0

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

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	-	18,23,23	1.34	4 (22%)	18,34,34	1.12	1 (5%)

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	C4-N3	2.11	1.36	1.32
3	A	999	NVP	C12-C11	2.16	1.41	1.36
3	A	999	NVP	C13-N14	2.31	1.37	1.32
3	A	999	NVP	C10-C15	2.67	1.43	1.40

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	999	NVP	CB-CA-N1	-2.07	116.12	118.25

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

There are no such residues in this entry.

5.8 Polymer linkage issues

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	545/560 (97%)	0.02	28 (5%) 32 21	31, 79, 133, 150	0
2	B	397/440 (90%)	0.05	23 (5%) 26 16	39, 81, 130, 150	0
All	All	942/1000 (94%)	0.03	51 (5%) 29 19	31, 80, 132, 150	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	54	ASN	4.8
1	A	3	SER	4.7
1	A	245	VAL	4.6
1	A	19	PRO	4.4
1	A	448	ARG	4.4

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
1	CSD	A	280	8/9	0.87	0.15	-	71,79,95,99	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
3	NVP	A	999	20/20	0.97	0.16	-0.03	47,55,61,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.