wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:17 AM GMT

PDB ID	:	3HVT
Title	:	STRUCTURAL BASIS OF ASYMMETRY IN THE HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 REVERSE TRANSCRIPTASE HETERODIMER
Authors	:	Steitz, T.A.; Smerdon, S.J.; Jaeger, J.; Wang, J.; Kohlstaedt, L.A.; Chirino, A.J.; Friedman, J.M.; Rice, P.A.
Deposited on	:	1994-07-25
Resolution	:	2.90 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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

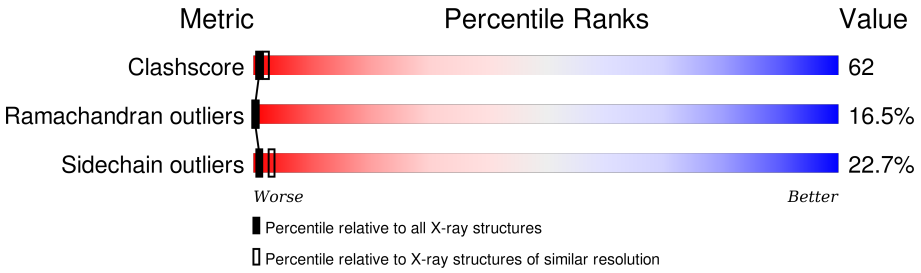
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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))
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)

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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	556	23% 52% 21% •
2	B	428	22% 43% 21% 6% 7%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7426 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.

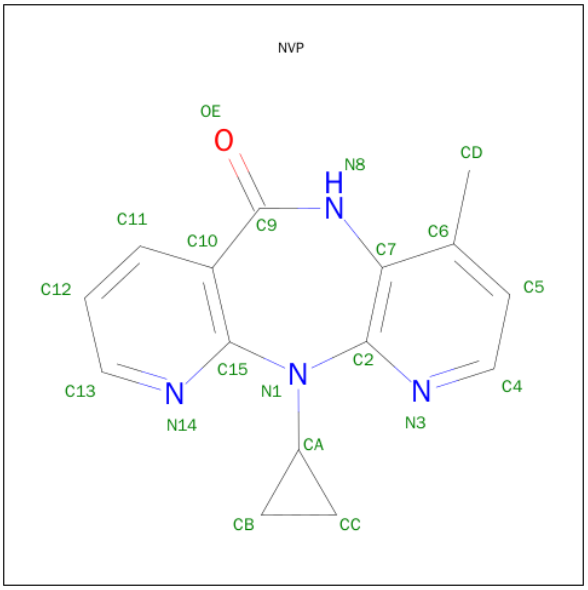
- Molecule 1 is a protein called HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P66).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	555	4248	2744	709	787	8	0	0	0

- Molecule 2 is a protein called HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	398	3158	2050	523	579	6	0	0	0

- 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: $C_{15}H_{14}N_4O$).



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

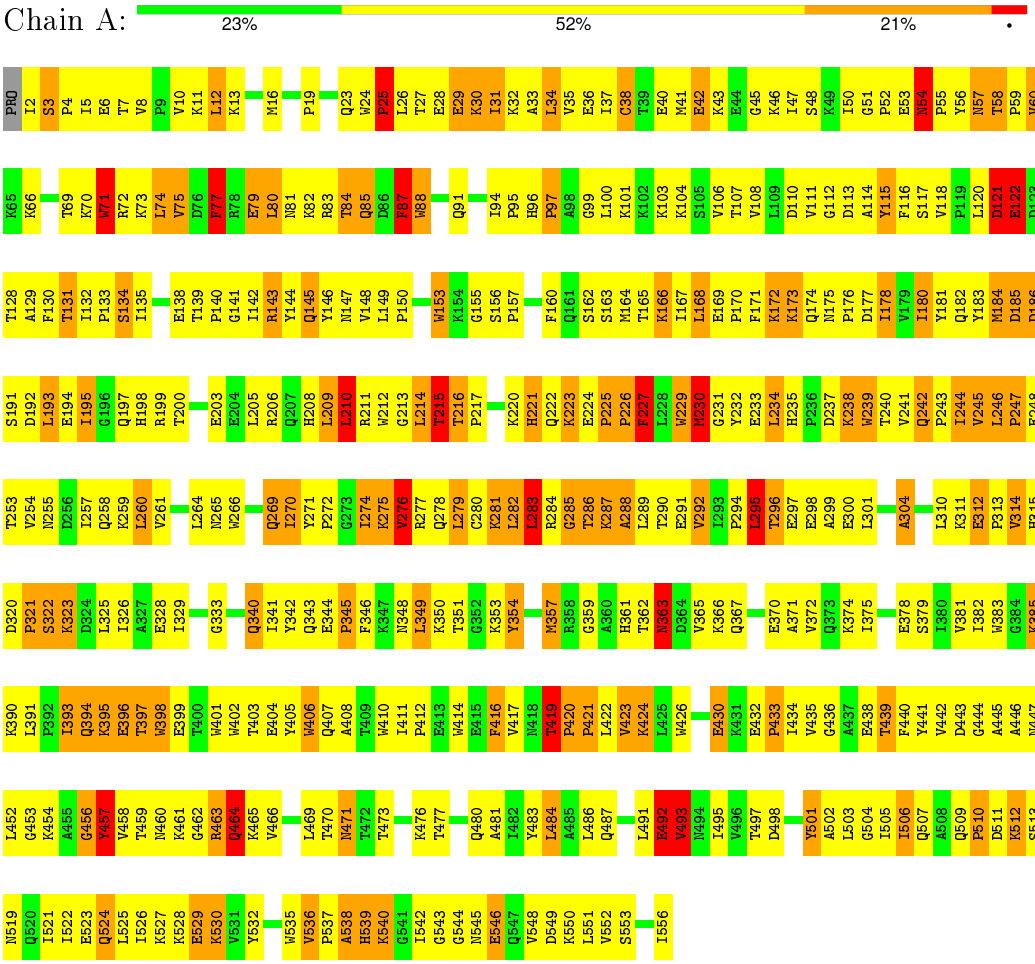
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

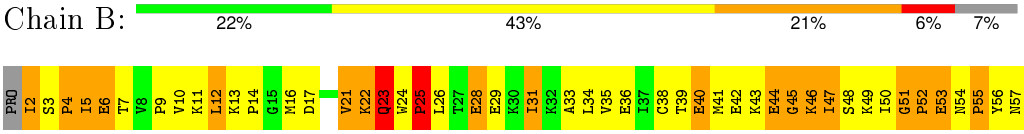
- Molecule 1: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P66)

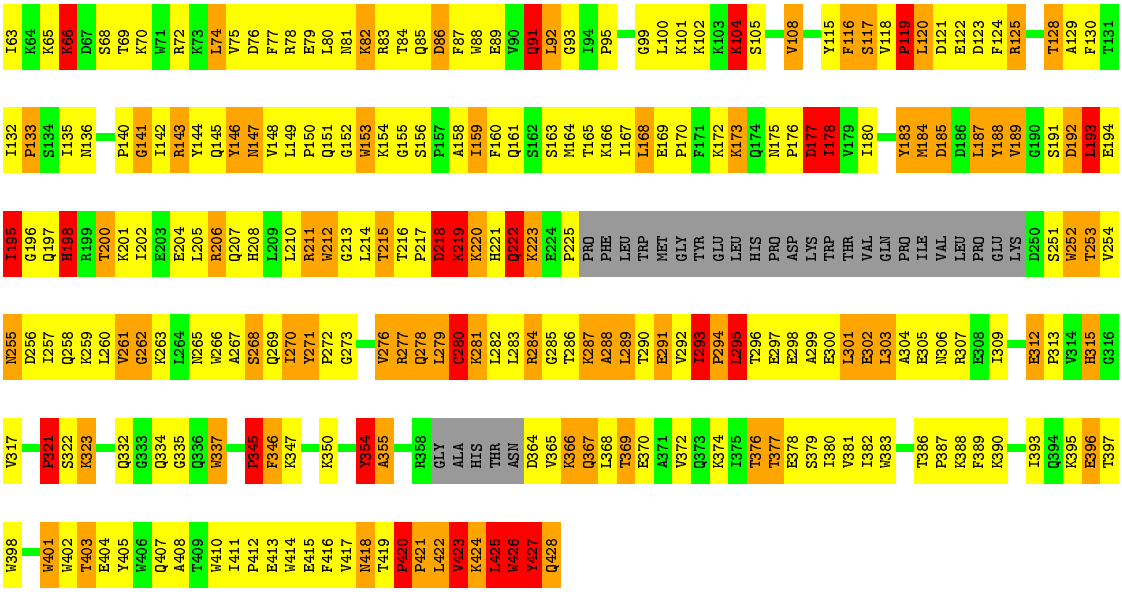
Chain A:



- Molecule 2: HIV-1 REVERSE TRANSCRIPTASE (SUBUNIT P51)

Chain B:




I63 K64 K65 K66 D67 S68 T69 K70 W71 R72 K73 L74 V75 D76 F77 R78 E79 L80 N81 K82 R83 T84 Q85 D86 F87 W88 E89 V90 Q91 L92 G93 I94 P95 G99 L100 K101 K102 K103 K104 S105 V108 Y115 F116 S117 V118 P119 L120 D121 E122 D123 F124 R125 T128 A129 F130 T131
I132 P133 S134 I135 N136 P140 G141 I142 R143 Y144 Q145 Y146 N147 V148 L149 P150 Q151 G152 W153 K154 G155 S156 P157 A158 I159 F160 Q161 S162 S163 M164 T165 K166 I167 L168 E169 P170 F171 K172 K173 Q174 N175 P176 D177 I178 V179 I180 Y183 M184 D185 D186 L187 Y188 V189 G190 S191 D192 L193 E194
I195 G196 Q197 H198 R199 T200 K201 I202 E203 E204 L205 R206 Q207 H208 L209 L210 R211 W212 G213 L214 T215 T216 P217 D218 K219 K220 H221 Q222 K223 E224 P225 PRO PHE LEU TRP MET GLY TYR GLU LEU HIS PRO ASP LYS TRP THR VAL GLN PRO ILE VAL LEU PRO GLU LYS D250 S251 W252 T253 V254
N255 D256 I257 Q258 K259 L260 V261 G262 K263 I264 N265 W266 A267 S268 Q269 I270 Y271 P272 G273 V276 R277 Q278 L279 C280 K281 L282 L283 R284 G285 T286 K287 A288 L289 T290 E291 V292 I293 P294 I295 T296 E297 E298 A299 E300 L301 E302 L303 A304 E305 N306 R307 E308 I309 E312 P313 V314 H315 G316
V317 P321 S322 K323 Q332 G333 Q334 G335 Q336 W337 P345 F346 K347 K350 Y354 A355 R358 GLY ALA HIS THR ASN D364 V365 K366 Q367 L368 T369 E370 A371 V372 Q373 K374 I375 T376 T377 E378 S379 I380 V381 I382 W383 T386 P387 K388 F389 K390 I393 Q394 K395 E396 T397
W398 W401 W402 T403 E404 Y405 W406 Q407 A408 T409 W410 I411 P412 E413 W414 E415 F416 V417 N418 T419 P420 P421 L422 V423 K424 L425 W426 Y427 Q428

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	223.60Å 69.90Å 105.50Å 90.00° 106.40° 90.00°	Depositor
Resolution (Å)	8.00 – 2.90	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.90)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.266 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7426	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.76	1/4358 (0.0%)	1.06	15/5951 (0.3%)
2	B	0.81	0/3242	1.10	21/4411 (0.5%)
All	All	0.78	1/7600 (0.0%)	1.08	36/10362 (0.3%)

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 outliers	#Planarity outliers
2	B	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	250	ASP	CB-CG	5.37	1.63	1.51

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	51	GLY	C-N-CD	-9.27	100.21	120.60
2	B	419	THR	C-N-CD	-8.95	100.92	120.60
2	B	178	ILE	N-CA-C	-8.83	87.15	111.00
1	A	420	PRO	C-N-CD	-8.65	101.56	120.60
2	B	288	ALA	N-CA-C	8.18	133.10	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	183	TYR	Sidechain
2	B	188	TYR	Sidechain
2	B	354	TYR	Sidechain
2	B	56	TYR	Sidechain

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	4248	0	4052	553	4
2	B	3158	0	3082	380	1
3	A	20	0	14	3	0
All	All	7426	0	7148	905	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 62.

The worst 5 of 905 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:441:TYR:HE2	2:B:286:THR:HG23	1.19	1.07
2:B:150:PRO:HD2	2:B:153:TRP:HE1	1.17	1.06
1:A:42:GLU:HA	1:A:46:LYS:HA	1.36	1.05
1:A:79:GLU:HG2	1:A:83:ARG:HH21	1.23	1.01
2:B:180:ILE:HG12	2:B:189:VAL:HG12	1.36	1.00

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:37:ILE:CG1	1:A:449:GLU:OE2[4_647]	1.87	0.33
1:A:37:ILE:CD1	1:A:449:GLU:OE2[4_647]	1.99	0.21
1:A:4:PRO:CB	2:B:211:ARG:NH2[4_646]	2.08	0.12
1:A:70:LYS:NZ	1:A:449:GLU:OE2[4_647]	2.09	0.11

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	553/556 (100%)	338 (61%)	125 (23%)	90 (16%)	0 0
2	B	392/428 (92%)	264 (67%)	62 (16%)	66 (17%)	0 0
All	All	945/984 (96%)	602 (64%)	187 (20%)	156 (16%)	0 0

5 of 156 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	SER
1	A	25	PRO
1	A	54	ASN
1	A	66	LYS
1	A	121	ASP

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	422/496 (85%)	332 (79%)	90 (21%)	1 4
2	B	328/390 (84%)	248 (76%)	80 (24%)	1 2
All	All	750/886 (85%)	580 (77%)	170 (23%)	1 3

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	457	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	23	GLN
2	B	367	GLN
1	A	464	GLN
1	A	523	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	428	GLN
1	A	524	GLN
2	B	407	GLN
1	A	447	ASN
1	A	487	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

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	557	-	18,23,23	1.42	2 (11%)	18,34,34	1.19	3 (16%)

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	557	-	-	0/0/6/6	0/2/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	557	NVP	OE-C9	-3.90	1.15	1.24
3	A	557	NVP	CA-N1	-3.20	1.44	1.49

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	557	NVP	CB-CA-N1	-2.25	115.93	118.25
3	A	557	NVP	CC-CA-N1	-2.20	115.98	118.25
3	A	557	NVP	CD-C6-C7	2.58	122.64	119.84

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	557	NVP	3	0

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.