



wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 11:12 PM GMT

PDB ID : 1J5O
Title : CRYSTAL STRUCTURE OF MET184ILE MUTANT OF HIV-1 REVERSE TRANSCRIPTASE IN COMPLEX WITH DOUBLE STRANDED DNA TEMPLATE-PRIMER
Authors : Sarafianos, S.G.; Das, K.; Arnold, E.
Deposited on : 2002-05-24
Resolution : 3.50 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

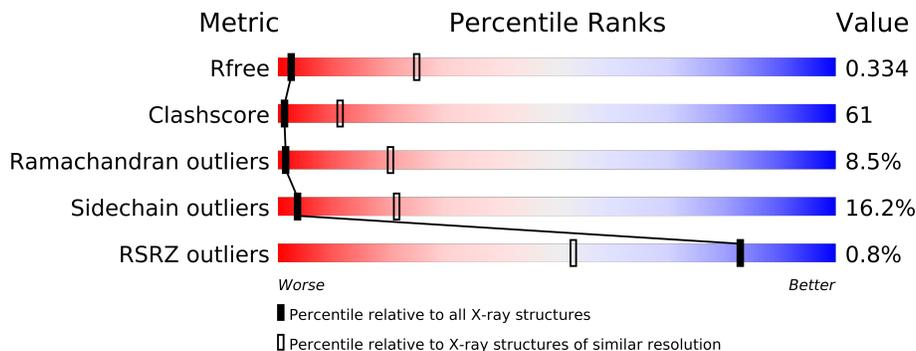
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.50 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}	66092	1243 (3.70-3.30)
Clashscore	79885	1039 (3.66-3.34)
Ramachandran outliers	78287	1000 (3.66-3.34)
Sidechain outliers	78261	1000 (3.66-3.34)
RSRZ outliers	66119	1243 (3.70-3.30)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	T	19	
2	P	18	
3	A	558	
4	B	430	
5	L	214	
6	H	220	

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 11720 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA chain called 5'-D(*AP*TP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	T	19	390	184	80	108	18	0	0	0

- Molecule 2 is a DNA chain called 5'-D(*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*GP*CP*GP*CP*CP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	P	18	363	173	64	109	17	0	0	0

- Molecule 3 is a protein called Reverse transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	558	4292	2778	713	795	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	184	ILE	MET	ENGINEERED	UNP P03366
A	280	SER	CYS	ENGINEERED	UNP P03366

- Molecule 4 is a protein called Reverse transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	430	3411	2217	568	620	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	184	ILE	MET	ENGINEERED	UNP P03366
B	280	SER	CYS	ENGINEERED	UNP P03366

- Molecule 5 is a protein called ANTIBODY (LIGHT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	L	214	Total	C	N	O	S	0	0	0
			1616	1010	256	343	7			

- Molecule 6 is a protein called ANTIBODY (HEAVY CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	220	Total	C	N	O	S	0	0	0
			1648	1037	270	333	8			

3 Residue-property plots

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.

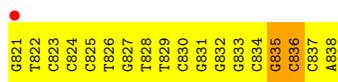
- Molecule 1: 5'-D(*AP*TP*GP*GP*CP*GP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*C)-3'

Chain T: 



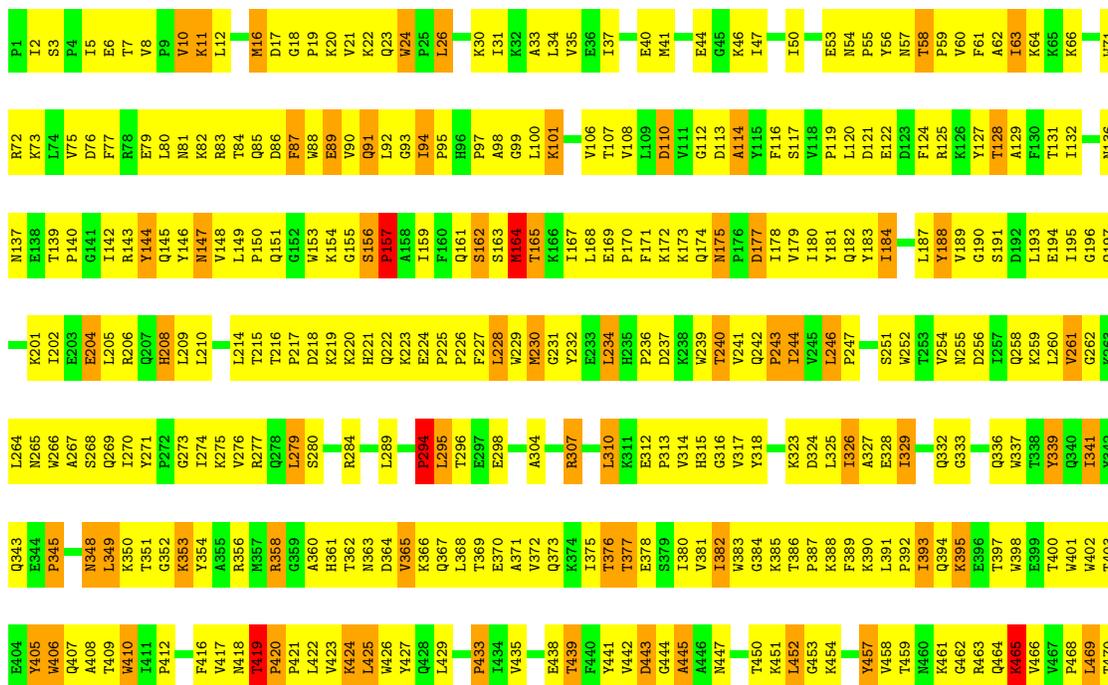
- Molecule 2: 5'-D(*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*GP*CP*GP*CP*CP*A)-3'

Chain P: 



- Molecule 3: Reverse transcriptase

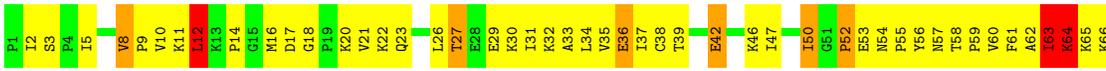
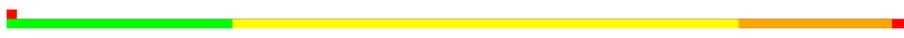
Chain A: 





• Molecule 4: Reverse transcriptase

Chain B:



• Molecule 5: ANTIBODY (LIGHT CHAIN)

Chain L:



• Molecule 6: ANTIBODY (HEAVY CHAIN)

Chain H:



Q1	P63	A125	P194
I2	S64	T126	S196
T3	L65	T127	T197
L4	K66	P128	W198
K5	S67	S129	P199
E6	R68	V130	V203
P9	L69	V131	T204
G10	T70	Y132	C205
I11	V71	P133	N206
V12	S72	L134	V207
V13	K73	Q141	A208
Q13	S76	T142	H209
P14	N77	N143	F210
S15	N78	S144	A211
Q16	Q79	M145	S212
P17	A80	V146	S213
F18	F81	T147	T214
T21	L82	L148	K215
C22	M83	G149	V216
T23	M84	C150	D217
F24	M85	L151	K218
S25	T86	V152	K219
G26	V87	K153	I220
E27	E88	G154	
S28	T89	Y155	
L29	A90	F156	
S30	D91	P157	
T31	E158	E158	
S32	P159	A93	
G33	I94	V160	
T34	Y95	T161	
G35	Y96	V162	
V36	C97	T163	
T37	W164	W164	
W38	N165	N165	
I39	L169	L169	
R40	V173	V173	
Q41	H174	H174	
P42	T175	T175	
G44	F176	F176	
S43	P177	P177	
G45	A178	A178	
G46	V179	V179	
L47	L180	L180	
E48	Q181	Q181	
W49	S182	S182	
L50	D183	D183	
A51	L184	L184	
T52	Y185	Y185	
I53	T186	T186	
M54	L187	L187	
M55	S188	S188	
D56	S189	S189	
D57	S190	S190	
D58	V191	V191	
M59	T192	T192	
R60	S122	S122	
Y61	T123	T123	
N62	A124	A124	

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, α , β , γ	169.18Å 169.18Å 221.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 3.50 39.97 – 3.50	Depositor EDS
% Data completeness (in resolution range)	91.5 (10.00-3.50) 91.4 (39.97-3.50)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.71 (at 3.48Å)	Xtriage
Refinement program	X-PLOR 3.843	Depositor
R, R_{free}	0.262 , 0.338 0.281 , 0.334	Depositor DCC
R_{free} test set	1219 reflections (3.04%)	DCC
Wilson B-factor (Å ²)	98.8	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 72.4	EDS
Estimated twinning fraction	0.049 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 42000 reflections	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	11720	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.19% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality i

5.1 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 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	T	0.55	0/439	0.76	0/676
2	P	0.56	1/405 (0.2%)	0.81	0/623
3	A	0.56	0/4404	0.78	2/6017 (0.0%)
4	B	0.63	0/3510	0.84	4/4784 (0.1%)
5	L	0.54	0/1654	0.85	0/2256
6	H	0.58	0/1691	0.85	0/2320
All	All	0.58	1/12103 (0.0%)	0.82	6/16676 (0.0%)

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	P	0	1
3	A	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	835	DG	O3'-P	-5.99	1.53	1.61

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	247	PRO	N-CA-CB	6.75	111.39	103.30
4	B	243	PRO	N-CA-CB	6.31	110.87	103.30
3	A	294	PRO	N-CA-CB	5.91	110.39	103.30
3	A	247	PRO	N-CA-CB	5.63	110.06	103.30
4	B	313	PRO	N-CA-CB	5.37	109.74	103.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	144	TYR	Sidechain
2	P	836	DC	Sidechain

5.2 Close contacts

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	T	390	0	212	55	0
2	P	363	0	204	103	0
3	A	4292	0	4127	482	0
4	B	3411	0	3300	446	0
5	L	1616	0	1517	174	0
6	H	1648	0	1602	184	0
All	All	11720	0	10962	1380	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 61.

The worst 5 of 1380 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:P:822:DT:H2''	2:P:823:DC:C6	1.75	1.22
1:T:803:DG:H2''	1:T:804:DG:C5'	1.77	1.14
2:P:831:DG:H2''	2:P:832:DG:H5'	1.17	1.11
3:A:20:LYS:HA	3:A:57:ASN:H	1.15	1.10
3:A:333:GLY:H	3:A:336:GLN:HB2	0.99	1.09

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	
3	A	556/558 (100%)	394 (71%)	121 (22%)	41 (7%)	2	24
4	B	428/430 (100%)	284 (66%)	96 (22%)	48 (11%)	1	13
5	L	212/214 (99%)	162 (76%)	35 (16%)	15 (7%)	2	25
6	H	218/220 (99%)	160 (73%)	42 (19%)	16 (7%)	2	25
All	All	1414/1422 (99%)	1000 (71%)	294 (21%)	120 (8%)	1	20

5 of 120 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	63	ILE
3	A	66	LYS
3	A	136	ASN
3	A	195	ILE
3	A	223	LYS

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	
3	A	434/498 (87%)	364 (84%)	70 (16%)	3	22
4	B	350/392 (89%)	290 (83%)	60 (17%)	3	18
5	L	182/182 (100%)	148 (81%)	34 (19%)	2	13
6	H	191/191 (100%)	168 (88%)	23 (12%)	7	37
All	All	1157/1263 (92%)	970 (84%)	187 (16%)	3	21

5 of 187 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	73	LYS
4	B	211	ARG
6	H	81	PHE
4	B	104	LYS
4	B	163	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	207	GLN
4	B	306	ASN
6	H	62	ASN
4	B	222	GLN
4	B	330	GLN

5.3.3 RNA [i](#)

There are no RNA chains in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

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	T	19/19 (100%)	0.69	2 (10%) 7 5	59, 76, 94, 96	0
2	P	18/18 (100%)	0.12	1 (5%) 24 10	61, 74, 86, 87	0
3	A	558/558 (100%)	-0.08	0 100 100	11, 52, 75, 92	0
4	B	430/430 (100%)	-0.17	5 (1%) 75 42	6, 32, 75, 93	0
5	L	214/214 (100%)	-0.08	2 (0%) 81 51	17, 44, 66, 76	0
6	H	220/220 (100%)	-0.16	1 (0%) 88 64	14, 33, 57, 68	0
All	All	1459/1459 (100%)	-0.11	11 (0%) 83 53	6, 43, 75, 96	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	429	LEU	6.0
5	L	214	CYS	5.9
1	T	801	DA	4.7
4	B	427	TYR	3.7
4	B	428	GLN	2.8

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.