



wwPDB X-ray Structure Validation Summary Report i

Mar 1, 2014 – 12:49 AM GMT

PDB ID : 3DRS
Title : HIV reverse transcriptase K103N mutant in complex with inhibitor R8D
Authors : Yan, Y.; Prasad, S.
Deposited on : 2008-07-11
Resolution : 3.15 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>

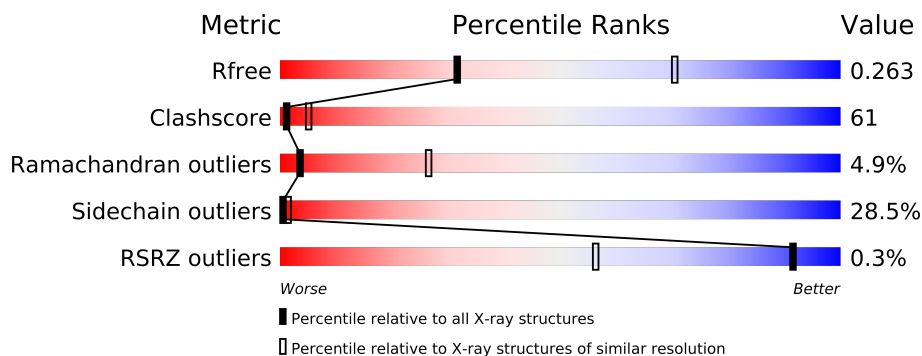
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.15 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	1360 (3.22-3.10)
Clashscore	79885	1681 (3.22-3.10)
Ramachandran outliers	78287	1639 (3.22-3.10)
Sidechain outliers	78261	1638 (3.22-3.10)
RSRZ outliers	66119	1361 (3.22-3.10)

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	A	563	
2	B	443	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	R8D	A	601	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7854 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 protein called Reverse transcriptase/ribonucleaseH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	549	Total	C	N	O	S	0	0	0
			4475	2891	747	829	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	EXPRESSION TAG	UNP P04585
A	-1	ASN	-	EXPRESSION TAG	UNP P04585
A	0	SER	-	EXPRESSION TAG	UNP P04585
A	103	ASN	LYS	ENGINEERED	UNP P04585

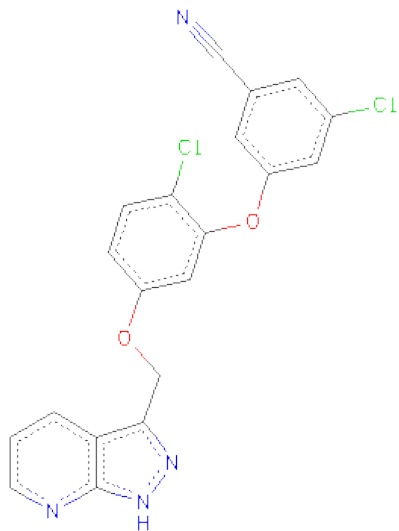
- Molecule 2 is a protein called p66 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	405	Total	C	N	O	S	0	0	0
			3351	2180	555	610	6			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	MET	-	EXPRESSION TAG	UNP P04585
B	-1	ASN	-	EXPRESSION TAG	UNP P04585
B	0	SER	-	EXPRESSION TAG	UNP P04585
B	103	ASN	LYS	ENGINEERED	UNP P04585

- Molecule 3 is 3-CHLORO-5-[2-CHLORO-5-(1H-PYRAZOLO[3,4-B]PYRIDIN-3-YLMETH OXY)PHENOXY]BENZONITRILE (three-letter code: R8D) (formula: C₂₀H₁₂Cl₂N₄O₂).



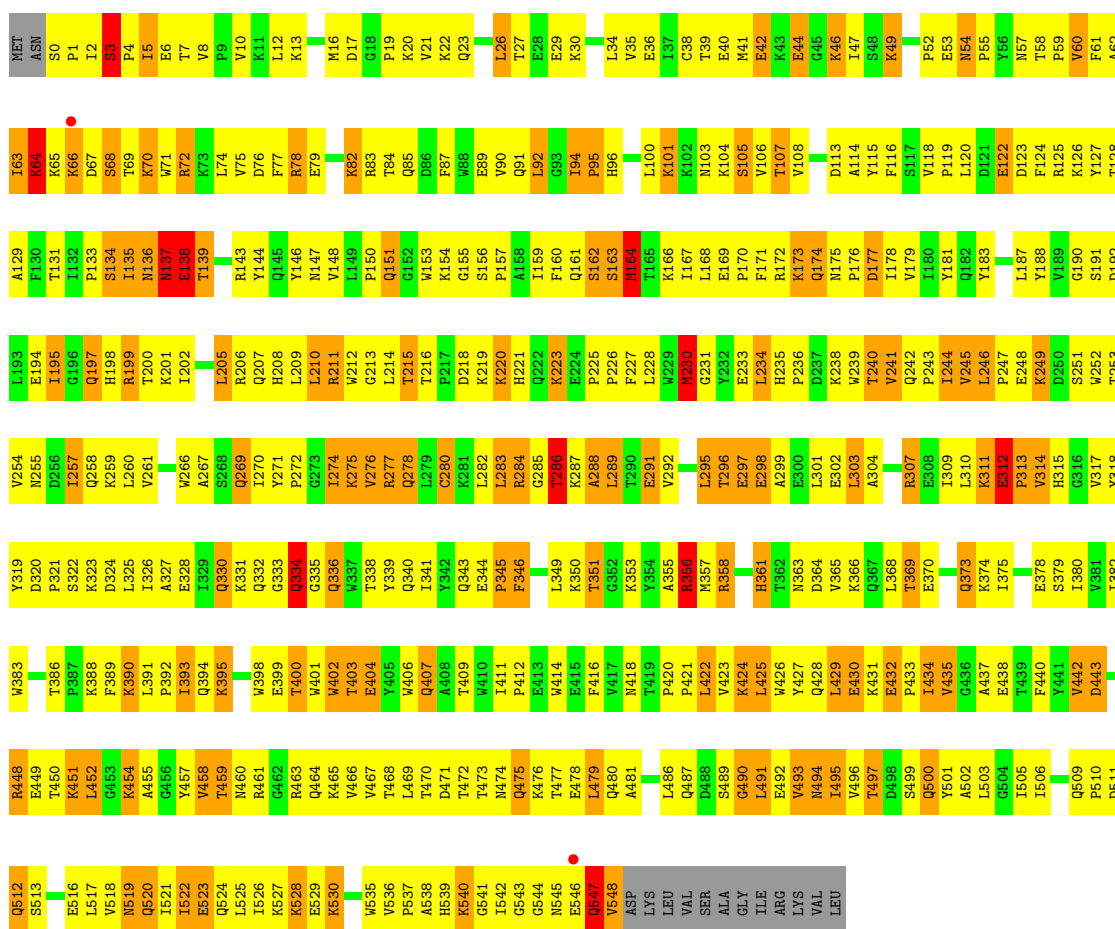
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Cl	N	O		
3	A	1	28	20	2	4	2	0	0

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.

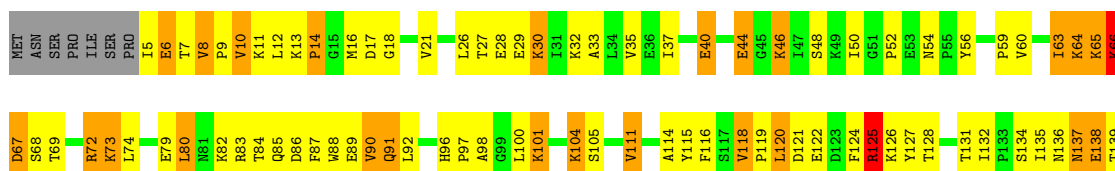
• Molecule 1: Reverse transcriptase/ribonucleaseH

Chain A:



• Molecule 2: p66 RT

Chain B:



I393	Q394	E396	T397	W398	E399	T400	W401	W402	T403	E404	Y405	W406	Q407	A408	T409	W410	I411	F412	E413	W414	E415	F416	W417	W418	T419	P420	P421	L422	W423	K424	L425	W426	Y427	Q428	LEU	GLU	GLU	GLU	PRO	THR	PHE																		
I329	Q330	K331	Q332	G333	Q334	G335	Q336	W337	T338	Y339	K340	T341	Y342	Q343	E344	P345	F346	K347	N348	L349	G352	K353	Y354	A355	R356	MET	ARG	GLY	ALA	H361	T362	G366	Q367	L368	T369	E370	Q373	K374	I375	T376	T377	E378	S379	I380	V381	I382	W383	G384	K385	T386	P387	K388	F389	K390	L391	P392			
N265	W266	A267	S268	Q269	I270	Y271	P272	G273	I274	K275	W276	R277	Q278	L279	C280	K281	L282	L283	K284	G285	T286	K287	A288	L289	T290	E291	W292	I293	P294	L295	T296	E297	E298	A299	E300	L301	E302	L303	A304	E305	N306	R307	E308	I309	L310	K311	E312	P313	V314	V317	Y318	Y319	D320	P321	S322	K323	E328		
E203	E204	L205	R206	Q207	H208	L209	L210	R211	L214	T215	THR	PRO	ASP	LYS	LYS	HIS	GLN	LYS	GLU	S162	S163	PRO	M164	PHE	K166	T167	L168	E169	P170	F171	R172	K173	Q174	N175	P176	D177	I178	Y179	I180	Y181	Q182	Y183	M184	L187	Y188	V189	G190	S191	D192	L193	E194	I195	G196	Q197	H198	R199	T200	K201	I202

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	118.55Å 154.63Å 154.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.50 – 3.15 37.50 – 3.15	Depositor EDS
% Data completeness (in resolution range)	99.4 (37.50-3.15) 99.5 (37.50-3.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.98 (at 3.18Å)	Xtriage
Refinement program	BUSTER-TNT 2.1.1	Depositor
R, R_{free}	0.184 , 0.250 0.193 , 0.263	Depositor DCC
R_{free} test set	1255 reflections (5.36%)	DCC
Wilson B-factor (Å ²)	68.6	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 65.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 24682 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7854	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.80% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: R8D

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	0.67	0/4592	0.89	6/6242 (0.1%)
2	B	0.67	0/3445	0.95	7/4682 (0.1%)
All	All	0.67	0/8037	0.92	13/10924 (0.1%)

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	344	GLU	C-N-CD	-19.16	78.46	120.60
2	B	312	GLU	C-N-CD	-11.88	94.47	120.60
2	B	420	PRO	C-N-CD	-10.86	96.70	120.60
1	A	312	GLU	C-N-CD	-10.29	97.96	120.60
1	A	94	ILE	C-N-CD	-8.80	101.23	120.60

There are no chirality outliers.

There are no planarity outliers.

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	A	4475	0	4514	602	1
2	B	3351	0	3373	391	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	28	0	12	16	0
All	All	7854	0	7899	967	1

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 967 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:288:ALA:HB3	1:A:291:GLU:HB3	1.17	1.17
1:A:318:TYR:CE2	3:A:601:R8D:H15	1.82	1.14
1:A:469:LEU:HD21	1:A:480:GLN:HG2	1.32	1.12
2:B:241:VAL:HG23	2:B:243:PRO:HD3	1.16	1.08
1:A:122:GLU:HA	1:A:125:ARG:HD2	1.41	1.02

All (1) 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	Distance(Å)	Clash(Å)
1:A:448:ARG:NH2	1:A:448:ARG:NH2[3_555]	1.54	0.66

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	
1	A	547/563 (97%)	427 (78%)	91 (17%)	29 (5%)	3	24
2	B	399/443 (90%)	338 (85%)	44 (11%)	17 (4%)	4	30
All	All	946/1006 (94%)	765 (81%)	135 (14%)	46 (5%)	3	26

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	VAL
1	A	138	GLU
1	A	195	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	230	MET
1	A	286	THR

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	491/503 (98%)	344 (70%)	147 (30%)	0	1
2	B	369/403 (92%)	271 (73%)	98 (27%)	1	2
All	All	860/906 (95%)	615 (72%)	245 (28%)	0	1

5 of 245 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	422	LEU
1	A	503	LEU
2	B	310	LEU
1	A	430	GLU
1	A	454	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	474	ASN
2	B	96	HIS
2	B	373	GLN
1	A	524	GLN
2	B	103	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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	R8D	A	601	-	31,31,31	2.82	16 (51%)	40,43,43	4.08	22 (55%)

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	R8D	A	601	-	-	0/9/11/11	0/2/4/4

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	R8D	C16-C17	7.06	1.53	1.39
3	A	601	R8D	C18-N24	6.28	1.42	1.34
3	A	601	R8D	C2-C25	-5.03	1.30	1.44
3	A	601	R8D	C6-CL27	-3.91	1.65	1.74
3	A	601	R8D	C5-C4	-3.85	1.31	1.38

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	R8D	C17-C16-N20	-15.90	105.48	110.46
3	A	601	R8D	C2-C1-C6	-8.04	114.78	119.51
3	A	601	R8D	C1-C6-C5	7.30	131.76	121.65
3	A	601	R8D	C5-C6-CL27	-6.84	110.69	119.14
3	A	601	R8D	O7-C8-C13	-6.30	106.05	119.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	549/563 (97%)	-0.15	2 (0%)	90 49	26, 58, 95, 145	0
2	B	405/443 (91%)	-0.13	1 (0%)	93 64	23, 51, 107, 124	0
All	All	954/1006 (94%)	-0.14	3 (0%)	91 56	23, 56, 105, 145	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	546	GLU	2.5
2	B	277	ARG	2.3
1	A	66	LYS	2.2

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

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

6.3 Carbohydrates ⓘ

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. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	R8D	A	601	28/28	0.31	1.49	64,69,79,79	0

6.5 Other polymers ⓘ

There are no such residues in this entry.