



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 12:03 pm GMT

PDB ID : 3DLG
Title : Crystal structure of hiv-1 reverse transcriptase in complex with GW564511.
Authors : Ren, J.; Chamberlain, P.P.; Stammers, D.K.
Deposited on : 2008-06-27
Resolution : 2.20 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

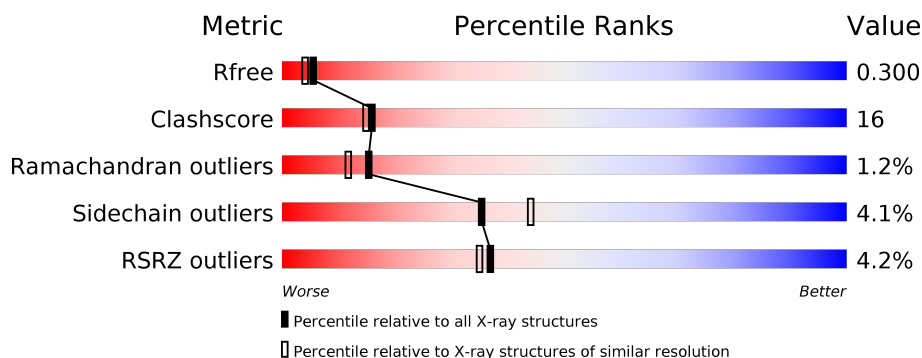
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	100719	4002 (2.20-2.20)
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)
RSRZ outliers	101464	4033 (2.20-2.20)

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	
2	B	440	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PO4	A	1301	-	-	-	X

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7928 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.

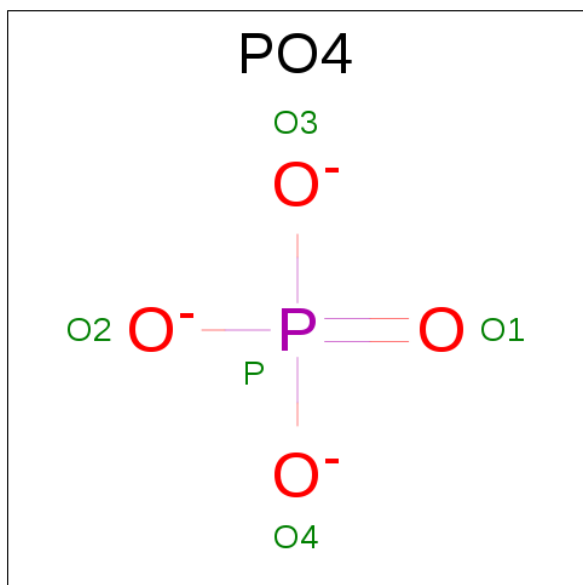
- Molecule 1 is a protein called REVERSE TRANSCRIPTASE/RIBONUCLEASE H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	532	Total	C	N	O	S	0	0	0
			4355	2822	723	802	8			

- Molecule 2 is a protein called P51 RT.

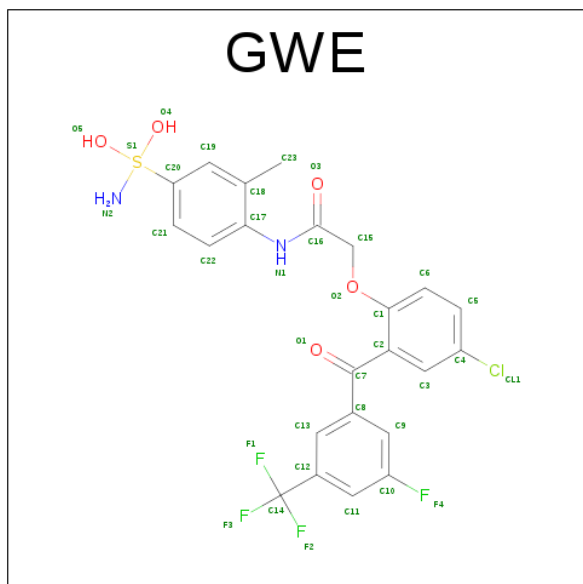
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	400	Total	C	N	O	S	0	0	0
			3303	2148	550	598	7			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is N-{4-[AMINO(DIHYDROXY)-LAMBDA 4 -SULFANYL]-2-METHYLPHENYL}-2-(4-CHLORO-2-{[3-FLUORO-5-(TRIFLUOROMETHYL)PHENYL]CARBONYL}PHENOXY)ACETAMIDE (three-letter code: GWE) (formula: $C_{23}H_{19}ClF_4N_2O_5S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Cl	F	N	O		
4	A	1	36	23	1	4	2	5	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	140	Total	O	0	0
			140	140		
5	B	84	Total	O	0	0
			84	84		

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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 ($\text{RSRZ} > 2$). 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.

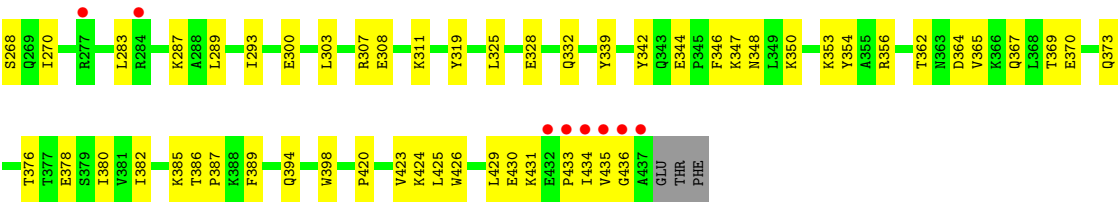
[illegible]

Chain B:

5% 60% 30% 9%

PRO ILE ILE SER PRO PRO ILE GLU T7 V8 P9 K13 M16 V21 K22 Q23 W24 E28 K30 K39 I31 L34 V35 E40 E44 G45 K46 E53 N54 P55 T56 N57 T58 P59 V60 F61 A62 I63 K64 LYS ASP SER T69 K70 K73 L74 E79 R83 E87

TRP GLU VAL LEU LEU GLY ILE P95 K101 K102 K103 K104 Y115 F116 P119 L120 D121 E122 D123 P133 S134 L135 N136 N137 T139 P140 G141 I142 R143 L149 P150 Q151 F160 Q161 S162 S163 K166 I167 L168 E169 P170 Q174 N175 I180 V189 E194 I195 C196



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	137.22Å 109.52Å 72.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.43 – 2.20 29.43 – 2.20	Depositor EDS
% Data completeness (in resolution range)	97.8 (29.43-2.20) 97.9 (29.43-2.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.13 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.220 , 0.299 0.221 , 0.300	Depositor DCC
R_{free} test set	2769 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	46.8	Xtriage
Anisotropy	0.263	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7928	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GWE, CSD, PO4

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	0.38	0/4462	0.62	0/6064
2	B	0.37	0/3394	0.62	0/4608
All	All	0.38	0/7856	0.62	0/10672

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	4355	0	4398	149	0
2	B	3303	0	3338	113	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	36	0	17	3	0
5	A	140	0	0	3	0
5	B	84	0	0	2	0
All	All	7928	0	7753	252	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 16.

The worst 5 of 252 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:353:LYS:HE3	2:B:430:GLU:HB3	1.46	0.97
2:B:242:GLN:HE21	2:B:243:PRO:HD2	1.28	0.96
1:A:270:ILE:HG22	1:A:314:VAL:HG11	1.48	0.93
1:A:216:THR:OG1	1:A:217:PRO:HD2	1.73	0.89
2:B:115:TYR:HB3	2:B:149:LEU:HB2	1.56	0.88

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	525/560 (94%)	490 (93%)	26 (5%)	9 (2%)	11	7
2	B	392/440 (89%)	370 (94%)	20 (5%)	2 (0%)	32	34
All	All	917/1000 (92%)	860 (94%)	46 (5%)	11 (1%)	15	12

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	VAL
1	A	217	PRO
1	A	91	GLN
2	B	239	TRP
1	A	14	PRO

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	476/499 (95%)	451 (95%)	25 (5%)	26	31
2	B	363/400 (91%)	354 (98%)	9 (2%)	53	65
All	All	839/899 (93%)	805 (96%)	34 (4%)	35	44

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	241	VAL
1	A	368	LEU
2	B	283	LEU
1	A	255	ASN
1	A	116	PHE

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	336	GLN
1	A	519	ASN
2	B	394	GLN
1	A	487	GLN
1	A	500	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	1.49	1 (25%)	2,8,10	1.09	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
1	CSD	A	280	1	-	0/2/6/8	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	280	CSD	CA-C	2.65	1.53	1.50

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.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

3 ligands are 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	PO4	A	1301	-	4,4,4	1.36	0	6,6,6	0.39	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	GWE	A	999	-	34,38,38	2.78	12 (35%)	48,57,57	1.20	6 (12%)
3	PO4	B	1300	-	4,4,4	1.39	0	6,6,6	0.38	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	PO4	A	1301	-	-	0/0/0/0	0/0/0/0
4	GWE	A	999	-	-	0/25/29/29	0/3/3/3
3	PO4	B	1300	-	-	0/0/0/0	0/0/0/0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	999	GWE	C19-C18	-2.77	1.35	1.39
4	A	999	GWE	C13-C8	2.14	1.42	1.39
4	A	999	GWE	C8-C7	2.26	1.53	1.49
4	A	999	GWE	C9-C8	2.32	1.42	1.39
4	A	999	GWE	C11-C10	2.56	1.42	1.37

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	999	GWE	C11-C10-C9	-2.58	120.25	123.53
4	A	999	GWE	C22-C17-C18	-2.10	118.35	120.75
4	A	999	GWE	C1-C2-C7	-2.08	119.48	122.76
4	A	999	GWE	O2-C15-C16	-2.03	105.05	110.83
4	A	999	GWE	C15-O2-C1	2.07	122.00	117.60

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
4	A	999	GWE	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 [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	531/560 (94%)	-0.06	18 (3%) 46 43	28, 53, 95, 132	0
2	B	400/440 (90%)	0.07	21 (5%) 27 26	29, 53, 101, 127	0
All	All	931/1000 (93%)	-0.01	39 (4%) 37 35	28, 53, 98, 132	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	195	ILE	9.7
2	B	435	VAL	7.4
1	A	52	PRO	5.2
1	A	245	VAL	5.0
1	A	92	LEU	5.0

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	0.91	0.14	-	51,55,69,77	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(Å ²)	Q<0.9
3	PO4	A	1301	5/5	0.86	0.18	2.25	97,112,128,130	0
3	PO4	B	1300	5/5	0.79	0.21	0.59	117,123,129,132	0
4	GWE	A	999	36/36	0.97	0.11	-0.73	25,38,60,66	0

6.5 Other polymers [i](#)

There are no such residues in this entry.