



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:42 PM GMT

PDB ID : 1UWB
Title : TYR 181 CYS HIV-1 RT/8-CL TIBO
Authors : Das, K.; Ding, J.; Hsiou, Y.; Arnold, E.
Deposited on : 1996-11-21
Resolution : 3.20 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 (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

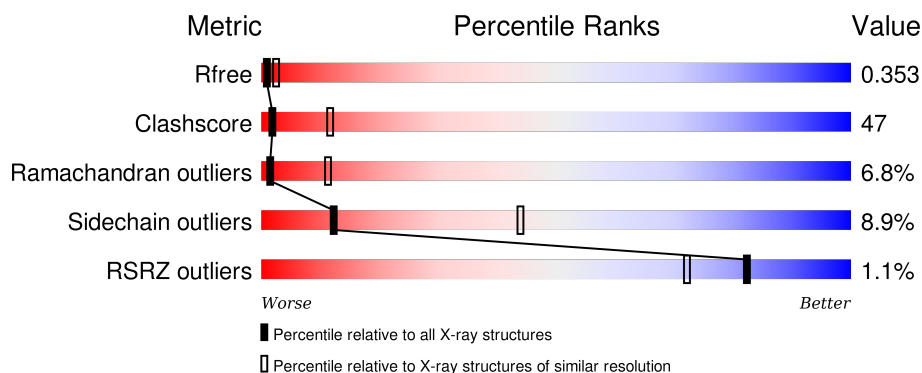
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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}	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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	558	<div> <div></div> <div> <div></div> <div>36%</div> <div>56%</div> <div>8%</div> </div> </div>
2	B	427	<div> <div></div> <div> <div></div> <div>29%</div> <div>56%</div> <div>14%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7822 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	558	Total	C	N	O	S	0	0	0
			4365	2822	725	812	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	181	CYS	TYR	ENGINEERED	UNP P03366
A	280	SER	CYS	ENGINEERED	UNP P03366

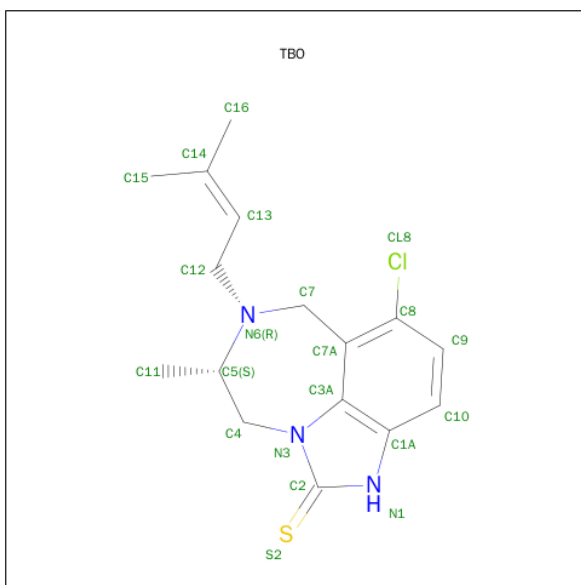
- Molecule 2 is a protein called REVERSE TRANSCRIPTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	427	Total	C	N	O	S	0	0	0
			3436	2234	567	629	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	181	CYS	TYR	ENGINEERED	UNP P03366
B	280	SER	CYS	ENGINEERED	UNP P03366

- Molecule 3 is 5-CHLORO-8-METHYL-7-(3-METHYL-BUT-2-ENYL)-6,7,8,9-TETRAHYDRO-2H-2,7,9A-TRIAZA-BENZO[CD]AZULENE-1-THIONE (three-letter code: TBO) (formula: C₁₆H₂₀ClN₃S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	Cl	N	S	0	0
			21	16	1	3	1		

T400	T338	G273	E203	T139	L74
W401	Y339	I274	E204	P140	V75
W402	Q340	L279	L205	G141	D76
T403	I341	L280	Q207	I142	F77
E404	Y342	K281	E208	R143	R78
Y405	Q343	L282	L209	Y144	E79
W406	E344	L283	L210	Q145	L80
Q407	P345	L284	G213	Y146	N81
W410	F346	K285	L214	N147	K82
L411	K347	G286	T215	V148	R83
P412	N348	T287	T216	L149	T84
W413	L349	K288	P217	G152	Q85
W414	K350	L289	D218	W153	D86
F415	G352	T290	G155	K154	F87
F416	K353	E291	E224	E88	W88
V417	Y354	V292	P157	S156	E89
F421	A355	I293	P225	S157	V90
K424	R356	P294	P226	Q161	Q91
L425	M357	L295	W229	S162	L92
W426	K358	T296	G230	A158	G93
Y427	G359	E297	G231	I159	I94
	N363	A299	Y232	E158	P95
	D364	E300	E233	M164	H96
	W365	L301	L234	T165	P97
	K366	E302	H235	K166	A98
	Q367	L303	W239	I167	G99
	L368	A304	E240	L168	L100
	T369	E305	T241	E169	K101
	E370	N306	Q242	K102	K103
	A371	R307	P243	P170	K104
	V372	E308	I244	F171	T107
	Q373	L309	V245	K172	V108
	K374	L310	L246	K173	L109
	I375	K311	D250	P176	D110
	T376	V314	S251	D177	V111
	T377	R315	W252	I178	Y115
	E378	G316	T253	V179	F116
	S379	V317	V254	C181	F117
	I380	V318	N255	D185	V118
	V381	V319	D256	D186	E122
	I382	D320	I257	L187	D123
	W383	P321	Q258	Y188	F124
	G384	K322	K259	V189	R125
	K385	K323	L260	Q190	K126
	T386	D324	V261	S191	Y127
	P387	I325	G262	D192	T128
	K388	L326	K263	L193	A129
	F389	K327	L264	E194	F130
	K390	R328	N265	T195	I131
	L391	G329	W266	G196	I132
	P392	L329	A267	Q197	P133
	I393	Q330	S268	H198	S134
	Q394	K331	Q269	A199	I135
	K395	G332	I270	K201	N136
	E396	G333	Y271	I202	I137
	T397	Q336	P272		E138
	W398	W337			
	E399				

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	225.70Å 69.20Å 104.90Å 90.00° 106.60° 90.00°	Depositor
Resolution (Å)	(Not available) – 3.20 14.98 – 3.20	Depositor EDS
% Data completeness (in resolution range)	85.0 ((Not available)-3.20) 85.0 (14.98-3.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.67 (at 3.19Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.274 , 0.360 0.275 , 0.353	Depositor DCC
R_{free} test set	1059 reflections (5.10%)	DCC
Wilson B-factor (Å ²)	68.5	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.14 , 14.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.37$, $\langle L^2 \rangle = 0.20$	Xtriage
Outliers	0 of 21825 reflections	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	7822	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% 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: TBO

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	0.55	0/4479	0.78	3/6108 (0.0%)
2	B	0.60	0/3534	0.85	5/4812 (0.1%)
All	All	0.57	0/8013	0.81	8/10920 (0.1%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	287	LYS	N-CA-C	6.92	129.68	111.00
1	A	517	LEU	CA-CB-CG	5.82	128.69	115.30
2	B	226	PRO	N-CA-CB	5.54	109.95	103.30
2	B	225	PRO	N-CA-CB	5.50	109.89	103.30
2	B	260	LEU	CA-CB-CG	5.50	127.94	115.30

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	4365	0	4260	397	0
2	B	3436	0	3401	363	0
3	A	21	0	20	8	0
All	All	7822	0	7681	736	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 47.

The worst 5 of 736 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:60:VAL:HA	1:A:75:VAL:HG22	1.39	1.03
2:B:180:ILE:HG23	2:B:189:VAL:HG22	1.41	1.00
1:A:272:PRO:HA	1:A:351:THR:HG21	1.43	0.99
1:A:318:TYR:CE2	3:A:559:TBO:H10	2.03	0.93
2:B:12:LEU:HA	2:B:84:THR:HG22	1.48	0.92

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	556/558 (100%)	415 (75%)	114 (20%)	27 (5%)	3	22
2	B	425/427 (100%)	309 (73%)	76 (18%)	40 (9%)	1	5
All	All	981/985 (100%)	724 (74%)	190 (19%)	67 (7%)	1	11

5 of 67 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	195	ILE
1	A	287	LYS
1	A	420	PRO
1	A	556	ILE
2	B	71	TRP

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	453/498 (91%)	424 (94%)	29 (6%)	22	62
2	B	367/389 (94%)	323 (88%)	44 (12%)	6	28
All	All	820/887 (92%)	747 (91%)	73 (9%)	12	44

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	49	LYS
2	B	101	LYS
2	B	405	TYR
2	B	70	LYS
2	B	139	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	480	GLN
2	B	96	HIS
2	B	207	GLN
1	A	474	ASN
2	B	255	ASN

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	TBO	A	559	-	19,23,23	2.40	7 (36%)	15,34,34	2.44	2 (13%)

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	TBO	A	559	-	-	0/4/17/17	0/2/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	559	TBO	C4-N3	-6.46	1.44	1.49
3	A	559	TBO	C7-N6	-2.81	1.43	1.47
3	A	559	TBO	C7A-C3A	-2.59	1.40	1.43
3	A	559	TBO	C3A-C1A	-2.59	1.34	1.43
3	A	559	TBO	C8-C7A	2.12	1.43	1.39

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	559	TBO	C15-C14-C13	-5.11	106.17	122.61
3	A	559	TBO	C16-C14-C15	6.71	131.14	114.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	559	TBO	8	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	558/558 (100%)	-0.45	5 (0%) 85 78	22, 71, 97, 100	0
2	B	427/427 (100%)	-0.46	6 (1%) 78 65	18, 59, 97, 100	0
All	All	985/985 (100%)	-0.45	11 (1%) 82 72	18, 68, 97, 100	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	95	PRO	3.0
2	B	225	PRO	2.7
1	A	557	ARG	2.7
1	A	24	TRP	2.6
1	A	66	LYS	2.6

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](#)

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	TBO	A	559	21/21	0.96	0.17	-0.40	59,67,70,73	0

6.5 Other polymers [i](#)

There are no such residues in this entry.