wwPDB X-ray Structure Validation Summary Report (i)

Feb 19, 2016 – 11:18 PM GMT

PDB ID	:	5HLF
Title	:	STRUCTURE OF HIV-1 REVERSE TRANSCRIPTASE In COMPLEX WITH A 38-MER HAIRPIN TEMPLATE-PRIMER DNA APTAMER AND AN ALPHA-CARBOXYPHOSPHONATE INHIBITOR
Authors	:	Das, K.; Arnold, E.
Deposited on	:	2016-01-15
Resolution	:	2.95 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 (1)) were used in the production of this report:

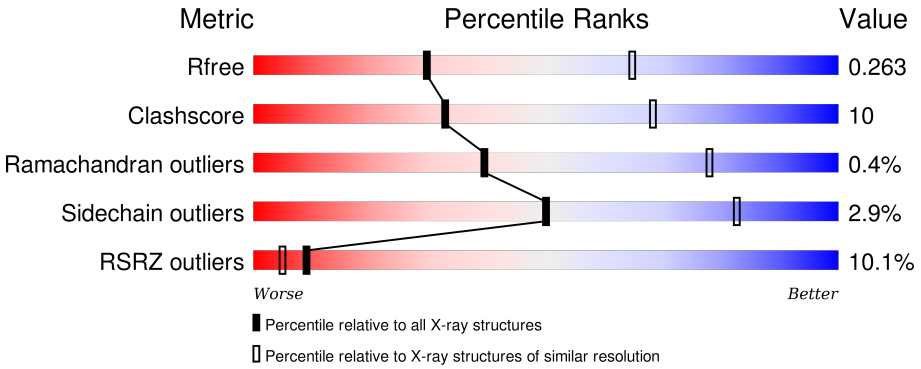
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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 2.95 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.



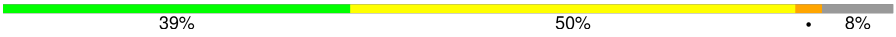
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2184 (3.00-2.92)
Clashscore	102246	2552 (3.00-2.92)
Ramachandran outliers	100387	2468 (3.00-2.92)
Sidechain outliers	100360	2471 (3.00-2.92)
RSRZ outliers	91569	2201 (3.00-2.92)

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.

Mol	Chain	Length	Quality of chain
1	A	555	10%, 72%, 25%, ·
1	C	555	19%, 67%, 31%, ·
2	B	444	4%, 78%, 15%, · 6%
2	D	444	5%, 75%, 17%, · 8%
3	E	38	5%, 42%, 42%, 8%, 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	38	

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
7	GOL	B	2003	-	-	-	X
7	GOL	C	603	-	-	-	X
7	GOL	D	503	-	-	-	X
7	GOL	D	504	-	-	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 17429 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 P66 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	554	Total	C	N	O	S	0	0	0
			4503	2915	749	832	7			
1	C	553	Total	C	N	O	S	0	0	0
			4495	2910	748	830	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	280	SER	CYS	engineered mutation	UNP P03366
C	280	SER	CYS	engineered mutation	UNP P03366

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	416	Total	C	N	O	S	0	1	0
			3436	2238	568	623	7			
2	D	410	Total	C	N	O	S	0	2	0
			3403	2222	559	615	7			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	MET	-	initiating methionine	UNP P03366
B	-14	ALA	-	expression tag	UNP P03366
B	-13	HIS	-	expression tag	UNP P03366
B	-12	HIS	-	expression tag	UNP P03366
B	-11	HIS	-	expression tag	UNP P03366
B	-10	HIS	-	expression tag	UNP P03366
B	-9	HIS	-	expression tag	UNP P03366
B	-8	HIS	-	expression tag	UNP P03366
B	-7	ALA	-	expression tag	UNP P03366
B	-6	LEU	-	expression tag	UNP P03366

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	GLU	-	expression tag	UNP P03366
B	-4	VAL	-	expression tag	UNP P03366
B	-3	LEU	-	expression tag	UNP P03366
B	-2	PHE	-	expression tag	UNP P03366
B	-1	GLN	-	expression tag	UNP P03366
B	0	GLY	-	expression tag	UNP P03366
B	280	SER	CYS	engineered mutation	UNP P03366
D	-15	MET	-	initiating methionine	UNP P03366
D	-14	ALA	-	expression tag	UNP P03366
D	-13	HIS	-	expression tag	UNP P03366
D	-12	HIS	-	expression tag	UNP P03366
D	-11	HIS	-	expression tag	UNP P03366
D	-10	HIS	-	expression tag	UNP P03366
D	-9	HIS	-	expression tag	UNP P03366
D	-8	HIS	-	expression tag	UNP P03366
D	-7	ALA	-	expression tag	UNP P03366
D	-6	LEU	-	expression tag	UNP P03366
D	-5	GLU	-	expression tag	UNP P03366
D	-4	VAL	-	expression tag	UNP P03366
D	-3	LEU	-	expression tag	UNP P03366
D	-2	PHE	-	expression tag	UNP P03366
D	-1	GLN	-	expression tag	UNP P03366
D	0	GLY	-	expression tag	UNP P03366
D	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 3 is a DNA chain called DNA (38-MER).

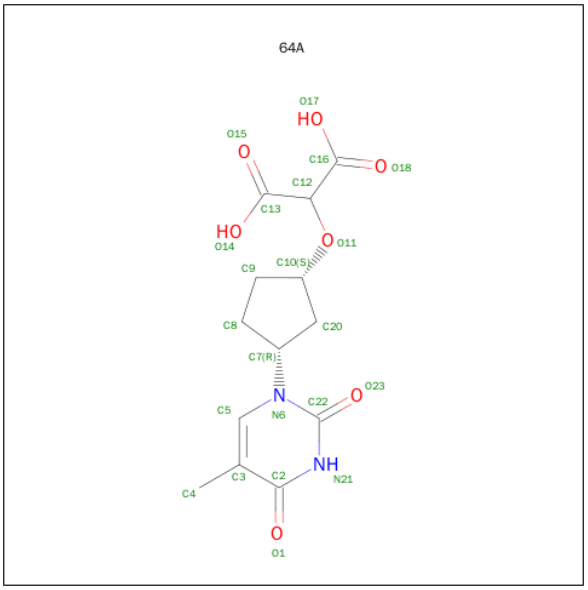
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	35	Total	C	N	O	P	0	0	0
			720	340	130	215	35			
3	F	35	Total	C	N	O	P	0	0	0
			720	340	130	215	35			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Mg	0	0
			2	2		
4	C	1	Total	Mg	0	0
			1	1		

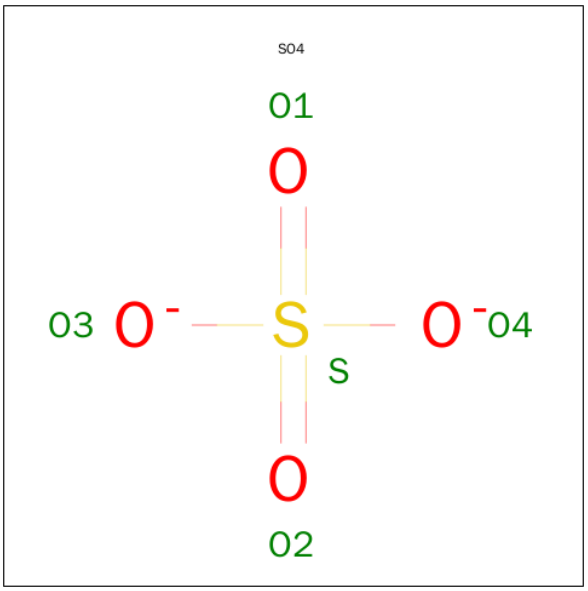
- Molecule 5 is {[(1S,3R)-3-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)cyclopentyl]ox

y}propanedioic acid (three-letter code: 64A) (formula: $C_{13}H_{16}N_2O_7$).



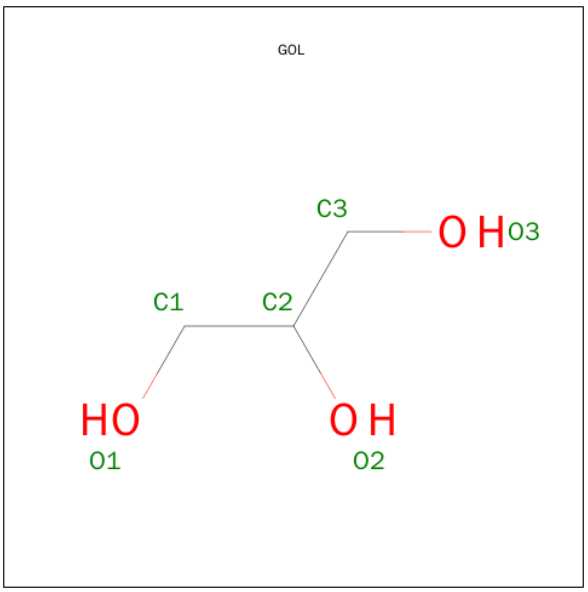
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O 22 13 2 7	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



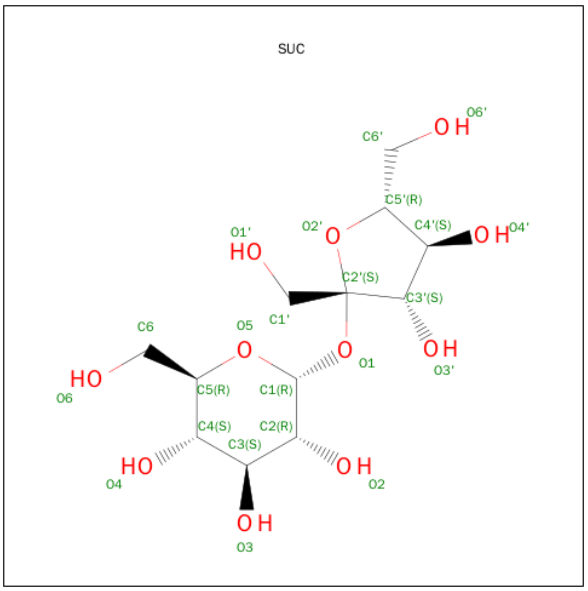
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O S 5 4 1	0	0
6	C	1	Total O S 5 4 1	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	C	1	Total C O 6 3 3	0	0
7	D	1	Total C O 6 3 3	0	0
7	D	1	Total C O 6 3 3	0	0
7	D	1	Total C O 6 3 3	0	0
7	D	1	Total C O 6 3 3	0	0

- Molecule 8 is SUCROSE (three-letter code: SUC) (formula: $C_{12}H_{22}O_{11}$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C O 23 12 11	0	0
8	D	1	Total C O 23 12 11	0	0

- Molecule 9 is water.

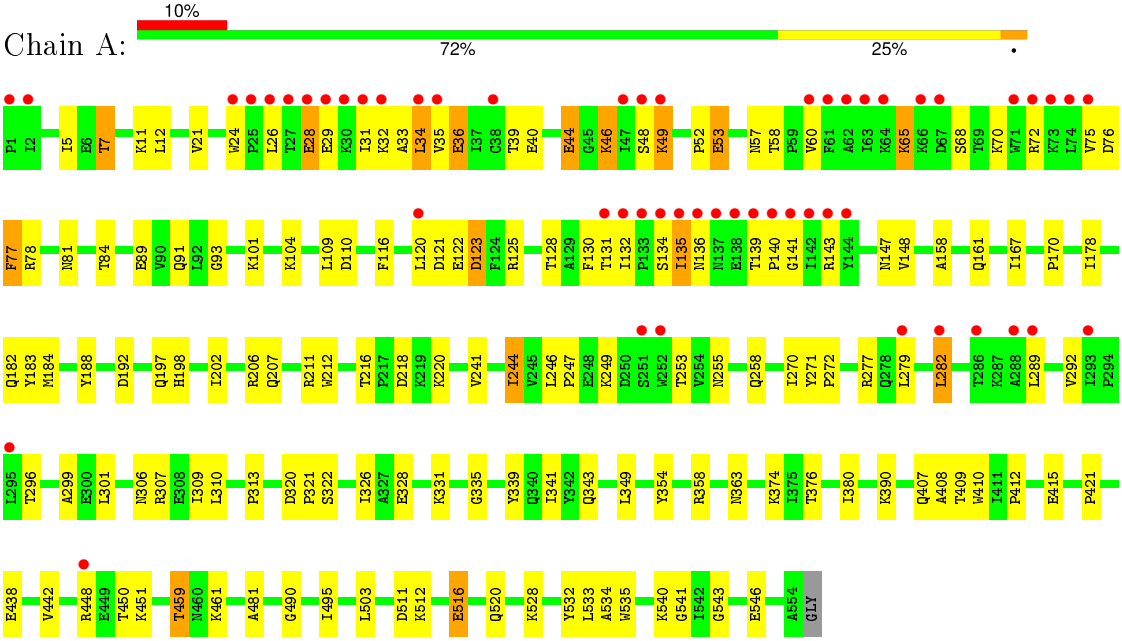
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	2	Total O 2 2	0	0
9	C	3	Total O 3 3	0	0

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.

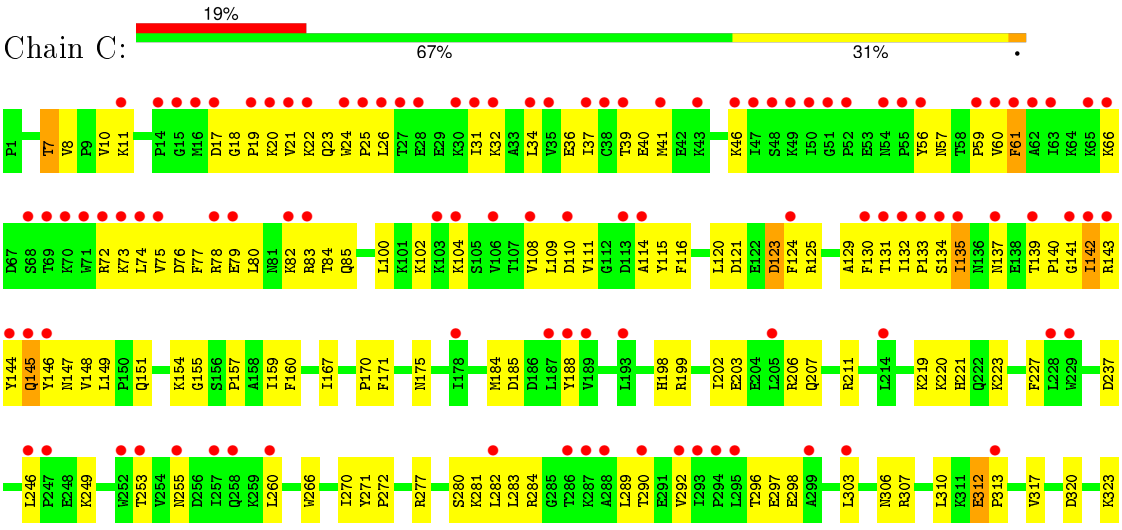
- Molecule 1: HIV-1 REVERSE TRANSCRIPTASE P66 SUBUNIT

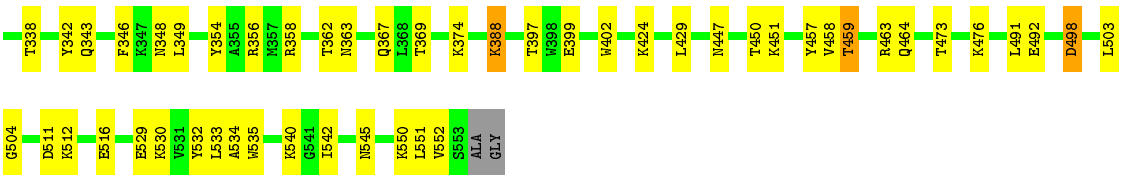
Chain A:



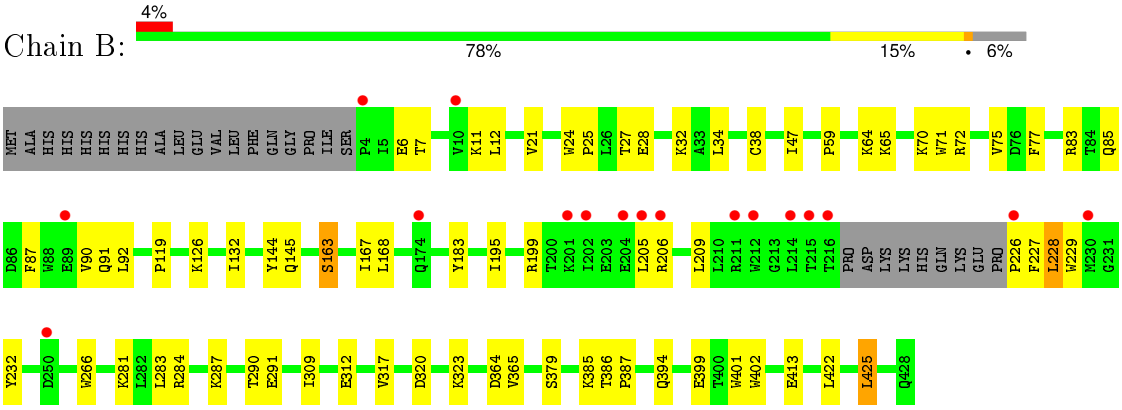
- Molecule 1: HIV-1 REVERSE TRANSCRIPTASE P66 SUBUNIT

Chain C:

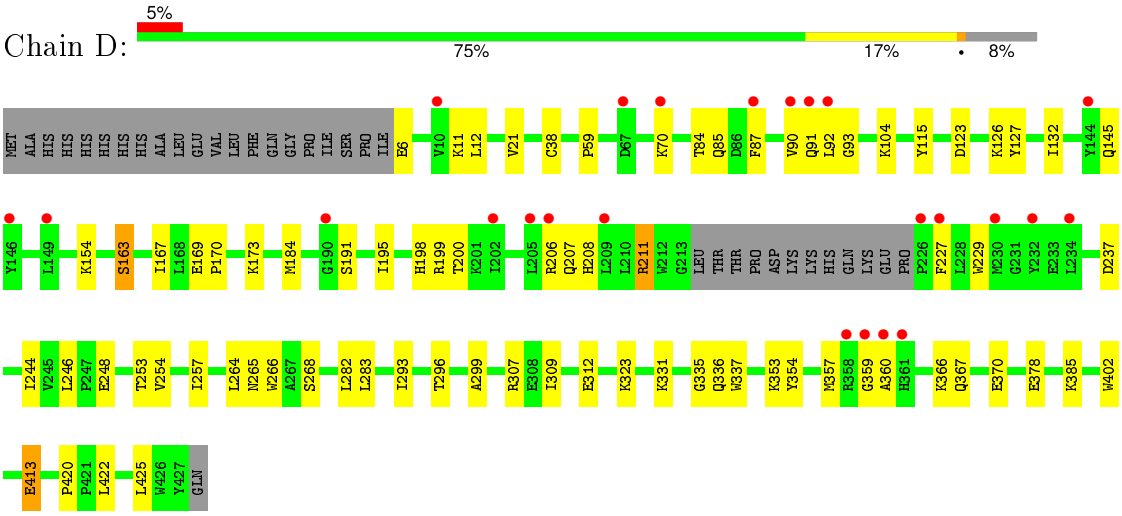




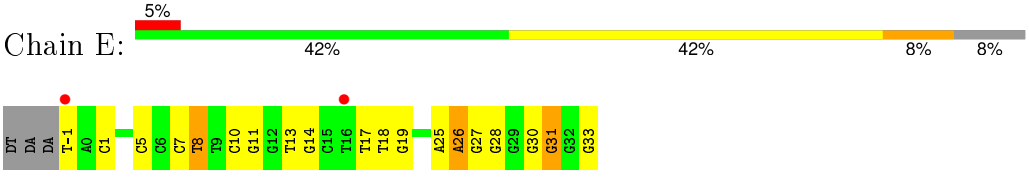
• Molecule 2: HIV-1 REVERSE TRANSCRIPTASE P51 SUBUNIT



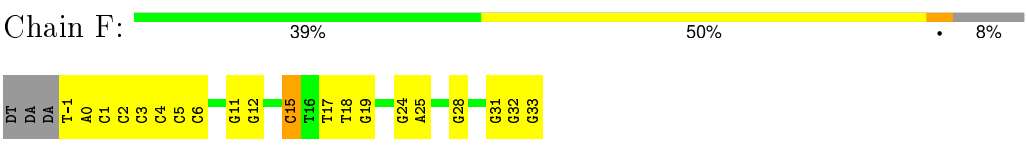
• Molecule 2: HIV-1 REVERSE TRANSCRIPTASE P51 SUBUNIT



• Molecule 3: DNA (38-MER)



• Molecule 3: DNA (38-MER)



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	90.06Å 128.56Å 132.41Å 90.00° 101.46° 90.00°	Depositor
Resolution (Å)	43.00 – 2.95 43.09 – 2.95	Depositor EDS
% Data completeness (in resolution range)	83.6 (43.00-2.95) 83.6 (43.09-2.95)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.55 (at 2.95Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.208 , 0.257 0.215 , 0.263	Depositor DCC
R_{free} test set	2566 reflections (5.18%)	DCC
Wilson B-factor (Å^2)	83.0	Xtriage
Anisotropy	0.295	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 72.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	2 of 52093 reflections (0.004%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	17429	wwPDB-VP
Average B, all atoms (Å^2)	103.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 64A, GOL, MG, SUC, OMC, SO4

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	0.41	0/4621	0.78	5/6279 (0.1%)
1	C	0.43	0/4613	0.85	5/6268 (0.1%)
2	B	0.30	0/3537	0.60	2/4803 (0.0%)
2	D	0.36	0/3509	0.76	3/4768 (0.1%)
3	E	0.92	4/759 (0.5%)	1.28	6/1170 (0.5%)
3	F	1.27	3/759 (0.4%)	1.13	4/1170 (0.3%)
All	All	0.49	7/17798 (0.0%)	0.81	25/24458 (0.1%)

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
1	C	0	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	15	DC	C4'-O4'	27.21	1.72	1.45
3	E	14	DG	C4'-O4'	11.23	1.56	1.45
3	F	11	DG	C4'-O4'	9.42	1.54	1.45
3	F	12	DG	C4'-O4'	6.83	1.51	1.45
3	E	26	DA	C4'-O4'	5.86	1.50	1.45

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	31	DG	O4'-C4'-C3'	-12.83	98.30	106.00
3	E	28	DG	O4'-C4'-C3'	-12.28	98.64	106.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	15	DC	C1'-O4'-C4'	-9.83	100.27	110.10
3	E	14	DG	O4'-C4'-C3'	-8.95	100.63	106.00
1	C	145	GLN	CB-CA-C	-8.33	93.73	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	306	ASN	Mainchain
1	C	542	ILE	Peptide

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	4503	0	4554	117	0
1	C	4495	0	4549	131	0
2	B	3436	0	3468	51	0
2	D	3403	0	3427	50	0
3	E	720	0	397	11	0
3	F	720	0	397	17	0
4	A	2	0	0	0	0
4	C	1	0	0	0	0
5	A	22	0	14	2	0
6	A	5	0	0	0	0
6	C	5	0	0	0	0
7	A	6	0	8	1	0
7	B	30	0	40	6	0
7	C	6	0	8	2	0
7	D	24	0	32	2	0
8	B	23	0	22	2	0
8	D	23	0	22	1	0
9	A	2	0	0	0	0
9	C	3	0	0	0	0
All	All	17429	0	16938	355	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 10.

The worst 5 of 355 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)
3:F:15:DC:C4'	3:F:15:DC:O4'	1.72	1.36
2:D:354:TYR:OH	2:D:378:GLU:OE1	1.88	0.91
1:C:60:VAL:HG21	1:C:130:PHE:CD2	2.12	0.84
1:C:23:GLN:OE1	1:C:131:THR:OG1	1.95	0.83
1:C:72:ARG:HH21	1:C:74:LEU:HD21	1.45	0.81

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	552/555 (100%)	525 (95%)	24 (4%)	3 (0%)	34 74
1	C	551/555 (99%)	511 (93%)	36 (6%)	4 (1%)	26 67
2	B	413/444 (93%)	400 (97%)	13 (3%)	0	100 100
2	D	408/444 (92%)	396 (97%)	12 (3%)	0	100 100
All	All	1924/1998 (96%)	1832 (95%)	85 (4%)	7 (0%)	39 78

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	551	LEU
1	A	282	LEU
1	A	135	ILE
1	A	313	PRO
1	C	249	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
1	A	493/495 (100%)	469 (95%)	24 (5%)	31 68
1	C	493/495 (100%)	472 (96%)	21 (4%)	35 73
2	B	377/403 (94%)	374 (99%)	3 (1%)	86 95
2	D	373/403 (93%)	371 (100%)	2 (0%)	92 97
All	All	1736/1796 (97%)	1686 (97%)	50 (3%)	50 82

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	459	THR
1	C	7	THR
1	C	498	ASP
1	A	516	GLU
2	B	290	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	151	GLN
2	B	361	HIS
1	C	373	GLN
1	A	407	GLN
1	C	255	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	OMC	E	2	3	15,22,23	0.91	1 (6%)	20,31,34	1.66	2 (10%)
3	OMC	E	4	3	15,22,23	0.90	1 (6%)	20,31,34	1.49	1 (5%)
3	OMC	F	2	3	15,22,23	0.91	1 (6%)	20,31,34	1.58	1 (5%)
3	OMC	F	4	3	15,22,23	0.88	1 (6%)	20,31,34	1.44	1 (5%)

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	OMC	E	2	3	-	0/5/27/28	0/2/2/2
3	OMC	E	4	3	-	0/5/27/28	0/2/2/2
3	OMC	F	2	3	-	0/5/27/28	0/2/2/2
3	OMC	F	4	3	-	0/5/27/28	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	2	OMC	C4-N4	2.76	1.43	1.35
3	F	4	OMC	C4-N4	2.78	1.43	1.35
3	E	4	OMC	C4-N4	2.79	1.43	1.35
3	E	2	OMC	C4-N4	2.84	1.43	1.35

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	2	OMC	N4-C4-N3	2.08	120.14	116.50
3	E	4	OMC	C6-C5-C4	5.59	119.63	117.44
3	F	4	OMC	C6-C5-C4	5.65	119.65	117.44
3	F	2	OMC	C6-C5-C4	5.97	119.78	117.44
3	E	2	OMC	C6-C5-C4	6.23	119.88	117.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	2	OMC	1	0
3	F	4	OMC	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 19 ligands modelled in this entry, 3 are monoatomic - leaving 16 for Mogul analysis.

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	$\#\|Z\| > 2$
5	64A	A	602	4	13,23,23	2.66	5 (38%)	11,33,33	3.42	4 (36%)
6	SO4	A	604	-	4,4,4	0.23	0	6,6,6	0.05	0
7	GOL	A	605	-	5,5,5	0.29	0	5,5,5	0.50	0
8	SUC	B	2001	-	24,24,24	0.47	0	36,36,36	0.84	0
7	GOL	B	2002	-	5,5,5	0.33	0	5,5,5	0.62	0
7	GOL	B	2003	-	5,5,5	0.46	0	5,5,5	0.34	0
7	GOL	B	2004	-	5,5,5	0.63	0	5,5,5	0.55	0
7	GOL	B	2005	-	5,5,5	0.35	0	5,5,5	0.51	0
7	GOL	B	2006	-	5,5,5	0.33	0	5,5,5	0.56	0
6	SO4	C	602	-	4,4,4	0.21	0	6,6,6	0.10	0
7	GOL	C	603	-	5,5,5	0.33	0	5,5,5	0.27	0
8	SUC	D	501	-	24,24,24	0.45	0	36,36,36	0.88	0
7	GOL	D	502	-	5,5,5	0.34	0	5,5,5	0.19	0
7	GOL	D	503	-	5,5,5	0.34	0	5,5,5	0.25	0
7	GOL	D	504	-	5,5,5	0.33	0	5,5,5	0.28	0
7	GOL	D	505	-	5,5,5	0.36	0	5,5,5	0.18	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
5	64A	A	602	4	-	0/4/25/25	0/2/2/2
6	SO4	A	604	-	-	0/0/0/0	0/0/0/0
7	GOL	A	605	-	-	0/4/4/4	0/0/0/0
8	SUC	B	2001	-	-	0/12/51/51	0/2/2/2
7	GOL	B	2002	-	-	0/4/4/4	0/0/0/0
7	GOL	B	2003	-	-	0/4/4/4	0/0/0/0
7	GOL	B	2004	-	-	0/4/4/4	0/0/0/0
7	GOL	B	2005	-	-	0/4/4/4	0/0/0/0
7	GOL	B	2006	-	-	0/4/4/4	0/0/0/0
6	SO4	C	602	-	-	0/0/0/0	0/0/0/0
7	GOL	C	603	-	-	0/4/4/4	0/0/0/0
8	SUC	D	501	-	-	0/12/51/51	0/2/2/2
7	GOL	D	502	-	-	0/4/4/4	0/0/0/0
7	GOL	D	503	-	-	0/4/4/4	0/0/0/0
7	GOL	D	504	-	-	0/4/4/4	0/0/0/0
7	GOL	D	505	-	-	0/4/4/4	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	602	64A	O11-C10	-7.64	1.34	1.44
5	A	602	64A	C22-N21	-2.32	1.33	1.38
5	A	602	64A	C20-C7	-2.27	1.50	1.53
5	A	602	64A	O11-C12	-2.08	1.38	1.43
5	A	602	64A	O1-C2	3.85	1.34	1.24

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	602	64A	C3-C2-N21	-7.01	119.47	125.35
5	A	602	64A	C10-C20-C7	2.50	107.57	102.49
5	A	602	64A	C8-C9-C10	2.82	107.76	102.54
5	A	602	64A	C2-N21-C22	7.46	121.38	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

11 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	602	64A	2	0
7	A	605	GOL	1	0
8	B	2001	SUC	2	0
7	B	2003	GOL	3	0
7	B	2004	GOL	1	0
7	B	2005	GOL	1	0
7	B	2006	GOL	1	0
7	C	603	GOL	2	0
8	D	501	SUC	1	0
7	D	502	GOL	1	0
7	D	505	GOL	1	0

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, 95^{th} 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	554/555 (99%)	0.53	54 (9%) 10 5	44, 93, 169, 216	0
1	C	553/555 (99%)	0.96	104 (18%) 2 1	45, 127, 194, 218	0
2	B	416/444 (93%)	0.32	17 (4%) 41 24	41, 80, 136, 170	0
2	D	410/444 (92%)	0.38	24 (5%) 26 14	48, 97, 152, 189	0
3	E	33/38 (86%)	0.44	2 (6%) 25 13	111, 150, 171, 212	0
3	F	33/38 (86%)	-0.24	0 100 100	75, 108, 131, 169	0
All	All	1999/2074 (96%)	0.56	201 (10%) 9 5	41, 95, 179, 218	0

The worst 5 of 201 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	72	ARG	13.9
1	A	133	PRO	10.6
1	C	73	LYS	9.9
1	C	131	THR	9.5
1	C	61	PHE	9.1

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

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, 95^{th} 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	OMC	E	2	21/22	0.93	0.14	-	147,150,154,159	0
3	OMC	E	4	21/22	0.85	0.20	-	127,138,145,157	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	OMC	F	4	21/22	0.97	0.19	-	70,76,81,83	0
3	OMC	F	2	21/22	0.97	0.16	-	86,96,102,111	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, 95^{th} 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
7	GOL	C	603	6/6	0.75	0.42	6.54	112,126,133,133	0
7	GOL	D	503	6/6	0.91	0.30	5.79	66,75,83,83	0
7	GOL	D	504	6/6	0.85	0.37	3.37	79,98,104,106	0
7	GOL	B	2003	6/6	0.95	0.31	2.67	56,59,64,73	0
7	GOL	D	502	6/6	0.83	0.33	0.90	81,87,88,93	0
7	GOL	B	2006	6/6	0.89	0.23	0.32	59,64,74,82	0
5	64A	A	602	22/22	0.90	0.21	-0.03	93,110,118,119	0
7	GOL	D	505	6/6	0.92	0.25	-0.31	75,82,92,92	0
7	GOL	B	2005	6/6	0.91	0.24	-0.38	60,64,66,72	0
8	SUC	D	501	23/23	0.92	0.20	-0.67	75,96,104,105	0
7	GOL	B	2002	6/6	0.92	0.14	-0.79	57,69,71,76	0
6	SO4	C	602	5/5	0.90	0.14	-1.13	104,121,122,131	0
8	SUC	B	2001	23/23	0.94	0.18	-2.18	71,84,95,98	0
7	GOL	A	605	6/6	0.95	0.14	-2.44	61,82,84,89	0
7	GOL	B	2004	6/6	0.96	0.15	-2.51	70,76,80,96	0
4	MG	C	601	1/1	0.90	0.08	-	81,81,81,81	0
4	MG	A	601	1/1	0.96	0.17	-	109,109,109,109	0
6	SO4	A	604	5/5	0.94	0.11	-	100,118,133,135	0
4	MG	A	603	1/1	0.97	0.15	-	82,82,82,82	0

6.5 Other polymers (i)

There are no such residues in this entry.