WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 2, 2023 – 04:56 AM EDT

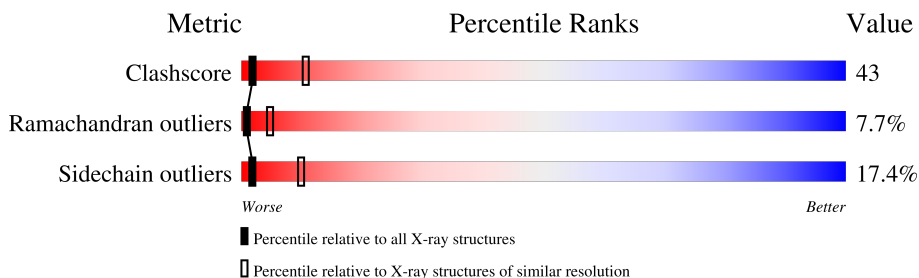|		
PDB ID	:	1C6V
Title	:	SIV INTEGRASE (CATALYTIC DOMAIN + DNA BIDING DOMAIN COMPRISING RESIDUES 50-293) MUTANT WITH PHE 185 REPLACED BY HIS (F185H)
Authors	:	Chen, Z.; Yan, Y.; Munshi, S.; Li, Y.; Zruygay-Murphy, J.; Xu, B.; Witmer, M.; Felock, P.; Wolfe, A.; Sardana, V.; Emini, E.A.; Hazuda, D.; Kuo, L.C.
Deposited on	:	1999-12-21
Resolution	:	3.00 Å(reported)

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 3.00 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	Percentile Ranks	Value
Clashscore		43
Ramachandran outliers		7.7%
Sidechain outliers		17.4%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	164	30% · 48% · 11% · 10%
1	B	164	28% · 48% · 7% · 13%
1	C	164	25% · 46% · 16% · 13%
1	D	164	40% · 37% · 9% · 13%
2	X	81	22% · 33% · 12% · 32%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5030 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 PROTEIN (SIV INTEGRASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	148	Total	C	N	O	S	0	0	0
			1169	733	213	214	9			
1	B	142	Total	C	N	O	S	0	0	0
			1111	697	203	202	9			
1	C	143	Total	C	N	O	S	0	0	0
			1118	702	204	203	9			
1	D	142	Total	C	N	O	S	0	0	0
			1111	697	203	202	9			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	185	HIS	PHE	engineered mutation	UNP Q88016
B	185	HIS	PHE	engineered mutation	UNP Q88016
C	185	HIS	PHE	engineered mutation	UNP Q88016
D	185	HIS	PHE	engineered mutation	UNP Q88016

- Molecule 2 is a protein called PROTEIN (SIU89134).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	X	55	Total	C	N	O	0	0	0
			455	296	85	74			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	250	LEU	ILE	conflict	UNP Q87706

- Molecule 3 is water.

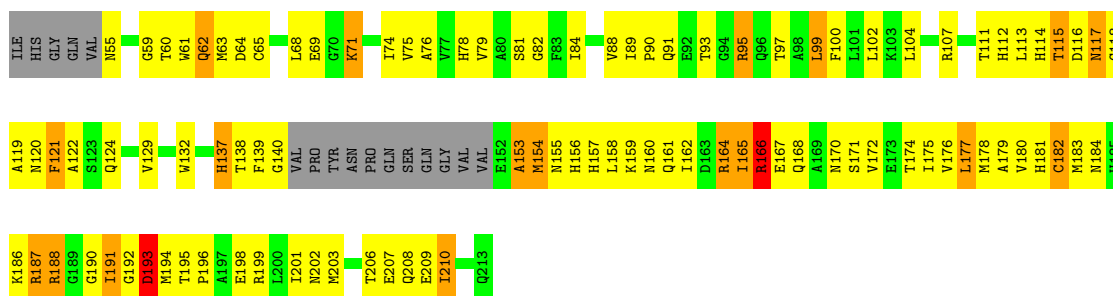| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	16	Total 16	O 16	0	0
3	B	10	Total 10	O 10	0	0
3	C	24	Total 24	O 24	0	0
3	D	12	Total 12	O 12	0	0
3	X	4	Total 4	O 4	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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. 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.
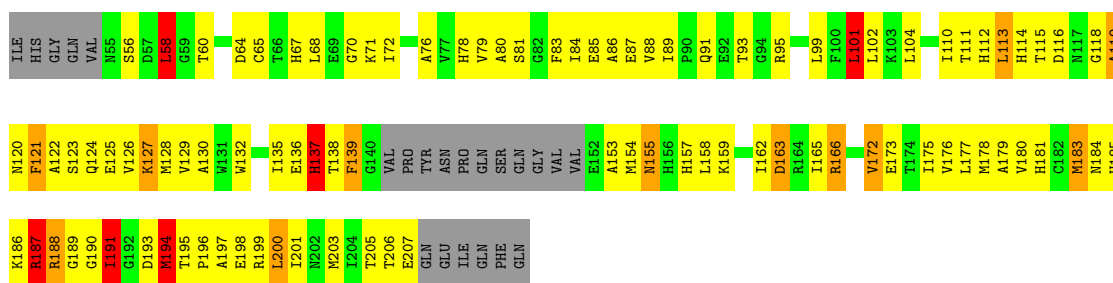
Note EDS was not executed.

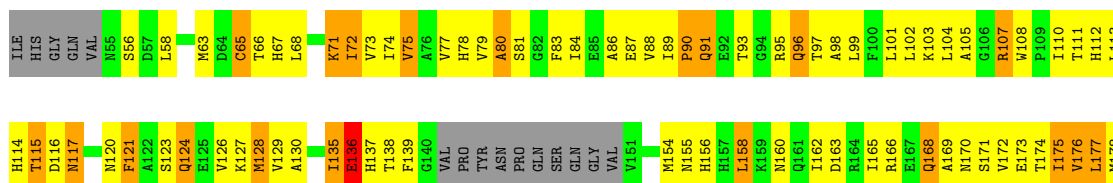- Molecule 1: PROTEIN (SIV INTEGRASE)

Chain A:
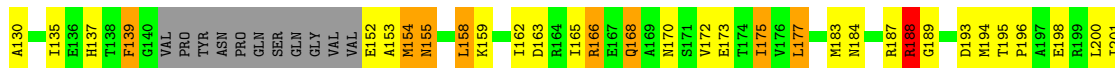
- Molecule 1: PROTEIN (SIV INTEGRASE)

Chain B:

- Molecule 1: PROTEIN (SIV INTEGRASE)

Chain C:

A179 V180 H181 C182 M183 N184 R187 R188 G189 G190 I191 G192 D193 M194 T195 P196 A197 E198 R199 L200 I201 N202 M203 I204 T205 E207 GLN ILE GLN PHE GLN

- Molecule 1: PROTEIN (SIV INTEGRASE)

Chain D:

40%   37%   9%   •   13%

ILE HIS GLY GLN VAL N55 T60 W61 Q62 M63 D64 C65 T66 H67 L68 I72 V73 I74 V75 A76 V77 H78 V79 A80 I84 E85 A86 E87 V88 I89 R95 Q96 L99 F100 L101 L102 K103 R107 I110 T111 H112 L113 H114 T115 D116 N117 N120 F121 A122 S123 V129

A130 I135 E136 H137 T138 F139 G140 VAL PRO TYR ASN PRO GLN SER GLN GLY VAL VAL E152 A153 M154 N155 L158 K159 I162 D163 R164 I165 R166 E167 Q168 A169 N170 S171 V172 E173 T174 I175 V176 L177 M183 N184 R187 H188 G189 D193 M194 T195 P196 A197 E198 R199 L200 I201

N202 M203 I204 T205 T206 E207 GLN GLU ILE GLN PHE GLN

- Molecule 2: PROTEIN (SIU89134)

Chain X:

22%   33%   12%   32%

GLN GLN SER K216 N217 S218 K221 R224 V225 Y226 Y227 R228 E229 G230 R231 L234 W235 K236 G237 P238 G239 E240 L241 L242 W243 K244 G245 E246 G247 A248 V249 L250 L251 K252 K258 V259 V260 P261 R262 R263 K264 I267 I268 K269 D270 TYR GLY GLY LYS GLU VAL ASP SER SER

SER HIS MET GLU ASP THR GLY GLU ALA ARG GLU VAL ALA
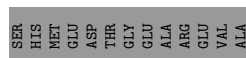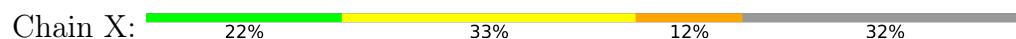
4 Data and refinement statistics ⓘ

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	79.57Å 100.00Å 150.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) (6.00-3.00)	Depositor
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.203 , 0.362	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5030	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.52	0/1192	0.81	0/1612
1	B	0.48	0/1133	0.84	2/1533 (0.1%)
1	C	0.51	0/1140	0.85	4/1543 (0.3%)
1	D	0.46	0/1133	0.74	0/1533
2	X	0.49	0/464	0.77	0/618
All	All	0.49	0/5062	0.80	6/6839 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	190	GLY	N-CA-C	-6.53	96.76	113.10
1	C	121	PHE	N-CA-C	5.49	125.81	111.00
1	B	101	LEU	CA-CB-CG	5.47	127.88	115.30
1	B	187	ARG	N-CA-C	5.41	125.59	111.00
1	C	104	LEU	CA-CB-CG	5.24	127.36	115.30

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	1169	0	1146	118	0
1	B	1111	0	1092	104	0
1	C	1118	0	1101	115	0

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1111	0	1092	78	0
2	X	455	0	487	44	0
3	A	16	0	0	2	0
3	B	10	0	0	0	0
3	C	24	0	0	1	0
3	D	12	0	0	0	0
3	X	4	0	0	0	0
All	All	5030	0	4918	423	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 43.

The worst 5 of 423 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:B:119:ALA:HA	1:B:122:ALA:HB3	1.22	1.11
1:B:71:LYS:HG3	1:B:172:VAL:HG21	1.44	1.00
1:A:165:ILE:HD11	1:A:182:CYS:SG	2.03	0.98
1:B:158:LEU:HA	1:B:183:MET:HE1	1.46	0.95
1:B:81:SER:HA	1:B:199:ARG:HD2	1.46	0.94

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	
1	A	144/164 (88%)	107 (74%)	25 (17%)	12 (8%)	1	4
1	B	138/164 (84%)	109 (79%)	18 (13%)	11 (8%)	1	4
1	C	139/164 (85%)	105 (76%)	25 (18%)	9 (6%)	1	7
1	D	138/164 (84%)	115 (83%)	15 (11%)	8 (6%)	1	10

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	X	53/81 (65%)	33 (62%)	13 (24%)	7 (13%)	0	1
All	All	612/737 (83%)	469 (77%)	96 (16%)	47 (8%)	1	5

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	79	VAL
1	A	117	ASN
1	A	153	ALA
1	A	166	ARG
1	B	139	PHE

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	124/138 (90%)	104 (84%)	20 (16%)	2	12
1	B	117/138 (85%)	98 (84%)	19 (16%)	2	12
1	C	118/138 (86%)	96 (81%)	22 (19%)	1	8
1	D	117/138 (85%)	96 (82%)	21 (18%)	2	9
2	X	47/67 (70%)	38 (81%)	9 (19%)	1	8
All	All	523/619 (84%)	432 (83%)	91 (17%)	2	10

5 of 91 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	177	LEU
1	D	152	GLU
1	C	203	MET
1	D	102	LEU
1	D	166	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	120	ASN
1	C	155	ASN
1	D	160	ASN
1	C	137	HIS
1	C	156	HIS

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 ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

There are no ligands in this entry.

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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.