Full wwPDB X-ray Structure Validation Report (i)

Aug 29, 2023 – 09:18 AM EDT

PDB ID	:	3P0A
Title	:	X-ray structure of pentameric HIV-1 CA
Authors	:	Pornillos, O.
Deposited on	:	2010-09-27
Resolution	:	5.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

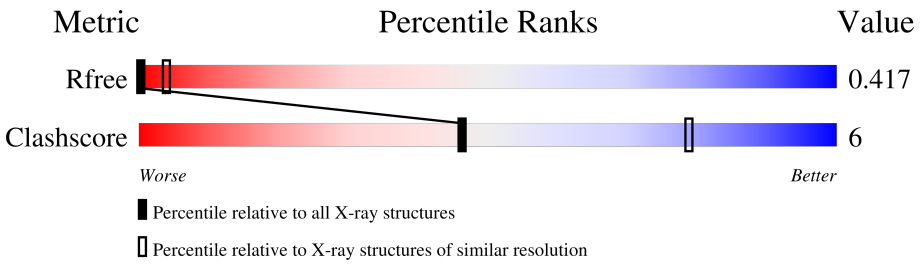
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 5.95 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}	130704	1000 (8.00-3.88)
Clashscore	141614	1049 (8.00-3.90)

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%

Mol	Chain	Length	Quality of chain
1	A	231	84% • 15%
1	B	231	85% • 14%
1	C	231	84% • 14%
1	D	231	85% • 14%
1	E	231	84% • 15%
1	F	231	84% • 15%
1	G	231	84% • 14%
1	H	231	85% • 14%
1	I	231	84% • 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	231	84% • 15%
1	K	231	84% • 15%
1	L	231	84% • 14%
1	M	231	85% • 14%
1	N	231	85% • 14%
1	O	231	84% • 15%
1	P	231	84% • 15%
1	Q	231	85% • 14%
1	R	231	84% • 14%
1	S	231	85% • 14%
1	T	231	84% • 15%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3952 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 CA.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
			Total	C			
1	A	197	197	197	0	0	197
1	B	198	198	198	0	0	198
1	C	198	198	198	0	0	198
1	D	198	198	198	0	0	198
1	E	197	197	197	0	0	197
1	F	197	197	197	0	0	197
1	G	198	198	198	0	0	198
1	H	198	198	198	0	0	198
1	I	198	198	198	0	0	198
1	J	197	197	197	0	0	197
1	K	197	197	197	0	0	197
1	L	198	198	198	0	0	198
1	M	198	198	198	0	0	198
1	N	198	198	198	0	0	198
1	O	197	197	197	0	0	197
1	P	197	197	197	0	0	197

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
			Total	C			
1	Q	198	198	198	0	0	198
			Total	C			
1	R	198	198	198	0	0	198
			Total	C			
1	S	198	198	198	0	0	198
			Total	C			
1	T	197	197	197	0	0	197

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	CYS	PRO	engineered mutation	UNP Q72497
A	18	LEU	ARG	engineered mutation	UNP Q72497
A	19	CYS	THR	engineered mutation	UNP Q72497
A	184	ALA	TRP	engineered mutation	UNP Q72497
A	185	ALA	MET	engineered mutation	UNP Q72497
B	17	CYS	PRO	engineered mutation	UNP Q72497
B	18	LEU	ARG	engineered mutation	UNP Q72497
B	19	CYS	THR	engineered mutation	UNP Q72497
B	184	ALA	TRP	engineered mutation	UNP Q72497
B	185	ALA	MET	engineered mutation	UNP Q72497
C	17	CYS	PRO	engineered mutation	UNP Q72497
C	18	LEU	ARG	engineered mutation	UNP Q72497
C	19	CYS	THR	engineered mutation	UNP Q72497
C	184	ALA	TRP	engineered mutation	UNP Q72497
C	185	ALA	MET	engineered mutation	UNP Q72497
D	17	CYS	PRO	engineered mutation	UNP Q72497
D	18	LEU	ARG	engineered mutation	UNP Q72497
D	19	CYS	THR	engineered mutation	UNP Q72497
D	184	ALA	TRP	engineered mutation	UNP Q72497
D	185	ALA	MET	engineered mutation	UNP Q72497
E	17	CYS	PRO	engineered mutation	UNP Q72497
E	18	LEU	ARG	engineered mutation	UNP Q72497
E	19	CYS	THR	engineered mutation	UNP Q72497
E	184	ALA	TRP	engineered mutation	UNP Q72497
E	185	ALA	MET	engineered mutation	UNP Q72497
F	17	CYS	PRO	engineered mutation	UNP Q72497
F	18	LEU	ARG	engineered mutation	UNP Q72497
F	19	CYS	THR	engineered mutation	UNP Q72497
F	184	ALA	TRP	engineered mutation	UNP Q72497
F	185	ALA	MET	engineered mutation	UNP Q72497
G	17	CYS	PRO	engineered mutation	UNP Q72497

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	18	LEU	ARG	engineered mutation	UNP Q72497
G	19	CYS	THR	engineered mutation	UNP Q72497
G	184	ALA	TRP	engineered mutation	UNP Q72497
G	185	ALA	MET	engineered mutation	UNP Q72497
H	17	CYS	PRO	engineered mutation	UNP Q72497
H	18	LEU	ARG	engineered mutation	UNP Q72497
H	19	CYS	THR	engineered mutation	UNP Q72497
H	184	ALA	TRP	engineered mutation	UNP Q72497
H	185	ALA	MET	engineered mutation	UNP Q72497
I	17	CYS	PRO	engineered mutation	UNP Q72497
I	18	LEU	ARG	engineered mutation	UNP Q72497
I	19	CYS	THR	engineered mutation	UNP Q72497
I	184	ALA	TRP	engineered mutation	UNP Q72497
I	185	ALA	MET	engineered mutation	UNP Q72497
J	17	CYS	PRO	engineered mutation	UNP Q72497
J	18	LEU	ARG	engineered mutation	UNP Q72497
J	19	CYS	THR	engineered mutation	UNP Q72497
J	184	ALA	TRP	engineered mutation	UNP Q72497
J	185	ALA	MET	engineered mutation	UNP Q72497
K	17	CYS	PRO	engineered mutation	UNP Q72497
K	18	LEU	ARG	engineered mutation	UNP Q72497
K	19	CYS	THR	engineered mutation	UNP Q72497
K	184	ALA	TRP	engineered mutation	UNP Q72497
K	185	ALA	MET	engineered mutation	UNP Q72497
L	17	CYS	PRO	engineered mutation	UNP Q72497
L	18	LEU	ARG	engineered mutation	UNP Q72497
L	19	CYS	THR	engineered mutation	UNP Q72497
L	184	ALA	TRP	engineered mutation	UNP Q72497
L	185	ALA	MET	engineered mutation	UNP Q72497
M	17	CYS	PRO	engineered mutation	UNP Q72497
M	18	LEU	ARG	engineered mutation	UNP Q72497
M	19	CYS	THR	engineered mutation	UNP Q72497
M	184	ALA	TRP	engineered mutation	UNP Q72497
M	185	ALA	MET	engineered mutation	UNP Q72497
N	17	CYS	PRO	engineered mutation	UNP Q72497
N	18	LEU	ARG	engineered mutation	UNP Q72497
N	19	CYS	THR	engineered mutation	UNP Q72497
N	184	ALA	TRP	engineered mutation	UNP Q72497
N	185	ALA	MET	engineered mutation	UNP Q72497
O	17	CYS	PRO	engineered mutation	UNP Q72497
O	18	LEU	ARG	engineered mutation	UNP Q72497
O	19	CYS	THR	engineered mutation	UNP Q72497

Continued on next page...

Continued from previous page...

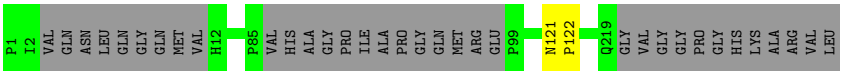
Chain	Residue	Modelled	Actual	Comment	Reference
O	184	ALA	TRP	engineered mutation	UNP Q72497
O	185	ALA	MET	engineered mutation	UNP Q72497
P	17	CYS	PRO	engineered mutation	UNP Q72497
P	18	LEU	ARG	engineered mutation	UNP Q72497
P	19	CYS	THR	engineered mutation	UNP Q72497
P	184	ALA	TRP	engineered mutation	UNP Q72497
P	185	ALA	MET	engineered mutation	UNP Q72497
Q	17	CYS	PRO	engineered mutation	UNP Q72497
Q	18	LEU	ARG	engineered mutation	UNP Q72497
Q	19	CYS	THR	engineered mutation	UNP Q72497
Q	184	ALA	TRP	engineered mutation	UNP Q72497
Q	185	ALA	MET	engineered mutation	UNP Q72497
R	17	CYS	PRO	engineered mutation	UNP Q72497
R	18	LEU	ARG	engineered mutation	UNP Q72497
R	19	CYS	THR	engineered mutation	UNP Q72497
R	184	ALA	TRP	engineered mutation	UNP Q72497
R	185	ALA	MET	engineered mutation	UNP Q72497
S	17	CYS	PRO	engineered mutation	UNP Q72497
S	18	LEU	ARG	engineered mutation	UNP Q72497
S	19	CYS	THR	engineered mutation	UNP Q72497
S	184	ALA	TRP	engineered mutation	UNP Q72497
S	185	ALA	MET	engineered mutation	UNP Q72497
T	17	CYS	PRO	engineered mutation	UNP Q72497
T	18	LEU	ARG	engineered mutation	UNP Q72497
T	19	CYS	THR	engineered mutation	UNP Q72497
T	184	ALA	TRP	engineered mutation	UNP Q72497
T	185	ALA	MET	engineered mutation	UNP Q72497

3 Residue-property plots (i)

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.

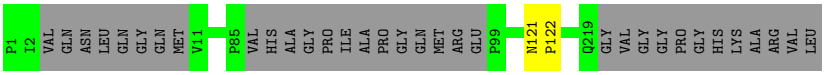
- Molecule 1: HIV-1 CA

Chain A:



- Molecule 1: HIV-1 CA

Chain B:



- Molecule 1: HIV-1 CA

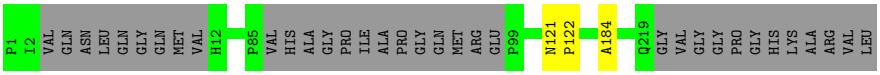
Chain C:

- Molecule 1: HIV-1 CA

Chain D:

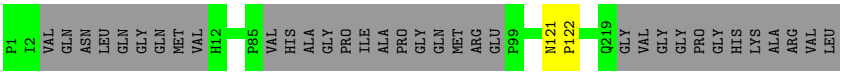
- Molecule 1: HIV-1 CA

Chain E:



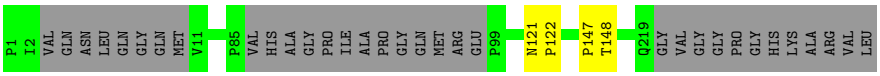
- Molecule 1: HIV-1 CA

Chain F:



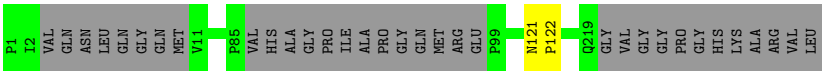
• Molecule 1: HIV-1 CA

Chain G: 84% • 14%



• Molecule 1: HIV-1 CA

Chain H: 85% • 14%

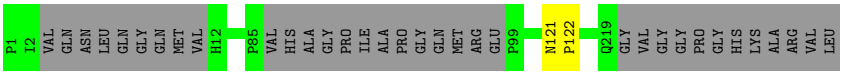


• Molecule 1: HIV-1 CA

Chain I: 84% • 14%

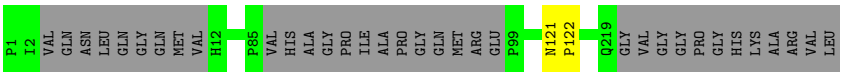
• Molecule 1: HIV-1 CA

Chain J: 84% • 15%



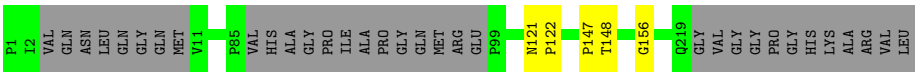
• Molecule 1: HIV-1 CA

Chain K: 84% • 15%



• Molecule 1: HIV-1 CA

Chain L: 84% • 14%



• Molecule 1: HIV-1 CA

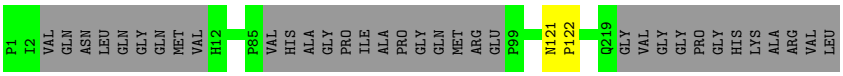
Chain M: 85% • 14%

• Molecule 1: HIV-1 CA

Chain N: 85% 14%

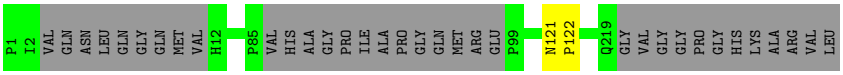
• Molecule 1: HIV-1 CA

Chain O: 84% 15%



• Molecule 1: HIV-1 CA

Chain P: 84% 15%



• Molecule 1: HIV-1 CA

Chain Q: 85% 14%

• Molecule 1: HIV-1 CA

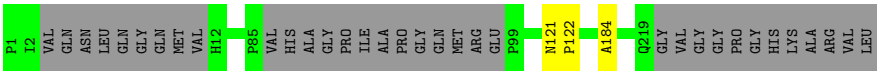
Chain R: 84% 14%

• Molecule 1: HIV-1 CA

Chain S: 85% 14%

• Molecule 1: HIV-1 CA

Chain T: 84% 15%


P1
I2
VAL
GLN
ASN
LEU
GLN
GLY
GLN
MET
VAL
H12
P85
VAL
HIS
ALA
GLY
PRO
ILE
ALA
PRO
GLY
GLN
MET
ARG
GLU
P99
N121
P122
A184
Q219
GLY
VAL
GLY
GLY
PRO
GLY
HIS
LYS
ALA
ARG
VAL
LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	86.79Å 122.13Å 149.30Å 74.41° 74.36° 81.48°	Depositor
Resolution (Å)	45.56 – 5.95 49.59 – 5.95	Depositor EDS
% Data completeness (in resolution range)	97.9 (45.56-5.95) 98.1 (49.59-5.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.16 (at 6.15Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, R_{free}	0.297 , 0.316 0.414 , 0.417	Depositor DCC
R_{free} test set	1420 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å^2)	287.8	Xtriage
Anisotropy	0.420	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.46 , -10.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.79	EDS
Total number of atoms	3952	wwPDB-VP
Average B, all atoms (Å^2)	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.52% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

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 (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	197	0	0	1	0
1	B	198	0	0	1	0
1	C	198	0	0	3	0
1	D	198	0	0	1	0
1	E	197	0	0	1	1
1	F	197	0	0	1	0
1	G	198	0	0	2	0
1	H	198	0	0	1	0
1	I	198	0	0	2	0
1	J	197	0	0	1	0
1	K	197	0	0	1	0
1	L	198	0	0	3	0
1	M	198	0	0	1	0
1	N	198	0	0	1	0
1	O	197	0	0	1	0
1	P	197	0	0	1	0
1	Q	198	0	0	1	0
1	R	198	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	S	198	0	0	1	0
1	T	197	0	0	1	1
All	All	3952	0	0	25	1

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 6.

All (25) 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:C:156:GLY:CA	1:I:156:GLY:CA	2.58	0.81
1:C:184:ALA:CA	1:R:181:VAL:CA	2.58	0.80
1:G:147:PRO:CA	1:G:148:THR:CA	2.89	0.50
1:L:121:ASN:CA	1:L:122:PRO:CA	2.90	0.50
1:Q:121:ASN:CA	1:Q:122:PRO:CA	2.90	0.49
1:B:121:ASN:CA	1:B:122:PRO:CA	2.90	0.49
1:G:121:ASN:CA	1:G:122:PRO:CA	2.90	0.49
1:E:121:ASN:CA	1:E:122:PRO:CA	2.91	0.48
1:T:121:ASN:CA	1:T:122:PRO:CA	2.91	0.48
1:O:121:ASN:CA	1:O:122:PRO:CA	2.91	0.48
1:J:121:ASN:CA	1:J:122:PRO:CA	2.91	0.48
1:L:156:GLY:CA	1:R:156:GLY:CA	2.92	0.48
1:R:121:ASN:CA	1:R:122:PRO:CA	2.92	0.48
1:C:121:ASN:CA	1:C:122:PRO:CA	2.92	0.48
1:H:121:ASN:CA	1:H:122:PRO:CA	2.92	0.47
1:L:147:PRO:CA	1:L:148:THR:CA	2.92	0.47
1:I:121:ASN:CA	1:I:122:PRO:CA	2.93	0.47
1:A:121:ASN:CA	1:A:122:PRO:CA	2.93	0.47
1:P:121:ASN:CA	1:P:122:PRO:CA	2.93	0.47
1:D:121:ASN:CA	1:D:122:PRO:CA	2.93	0.47
1:M:121:ASN:CA	1:M:122:PRO:CA	2.92	0.47
1:S:121:ASN:CA	1:S:122:PRO:CA	2.93	0.46
1:K:121:ASN:CA	1:K:122:PRO:CA	2.93	0.46
1:F:121:ASN:CA	1:F:122:PRO:CA	2.93	0.46
1:N:121:ASN:CA	1:N:122:PRO:CA	2.93	0.46

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	Interatomic distance (Å)	Clash overlap (Å)
1:E:184:ALA:CA	1:T:184:ALA:CA[1_556]	2.06	0.14

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.