



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 4, 2017 – 06:59 PM EDT

PDB ID : 2AOE  
Title : crystal structure analysis of HIV-1 protease mutant V82A with a substrate analog CA-P2  
Authors : Tie, Y.; Boross, P.I.; Wang, Y.F.; Gaddis, L.; Liu, F.; Chen, X.; Tozser, J.; Harrison, R.W.; Weber, I.T.  
Deposited on : unknown  
Resolution : 1.54 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](mailto: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)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030345

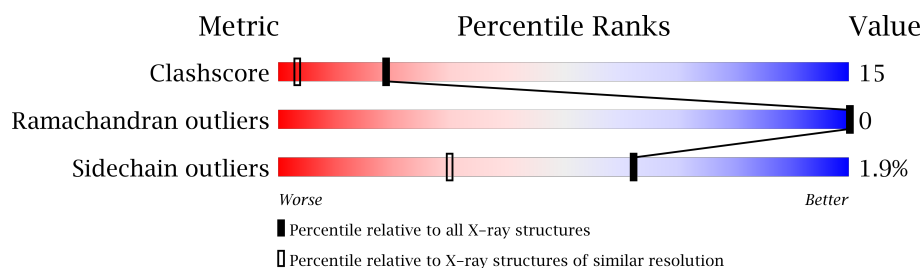
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.54 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))
Clashscore	112137	1845 (1.56-1.52)
Ramachandran outliers	110173	1810 (1.56-1.52)
Sidechain outliers	110143	1808 (1.56-1.52)

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  $\geq 3$ , 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	99	 90% 10%
1	B	99	 80% 18% .

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 1982 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 POL POLYPROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	99	Total	C	N	O	S	0	11	0
			803	522	139	139	3			
1	B	99	Total	C	N	O	S	0	15	0
			831	547	137	144	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	LYS	GLN	ENGINEERED	UNP P04587
A	33	ILE	LEU	ENGINEERED	UNP P04587
A	63	ILE	LEU	ENGINEERED	UNP P04587
A	67	ALA	CYS	ENGINEERED	UNP P04587
A	82	ALA	VAL	ENGINEERED	UNP P04587
A	95	ALA	CYS	ENGINEERED	UNP P04587
B	107	LYS	GLN	ENGINEERED	UNP P04587
B	133	ILE	LEU	ENGINEERED	UNP P04587
B	163	ILE	LEU	ENGINEERED	UNP P04587
B	167	ALA	CYS	ENGINEERED	UNP P04587
B	182	ALA	VAL	ENGINEERED	UNP P04587
B	195	ALA	CYS	ENGINEERED	UNP P04587

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Na	0	0
			1	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

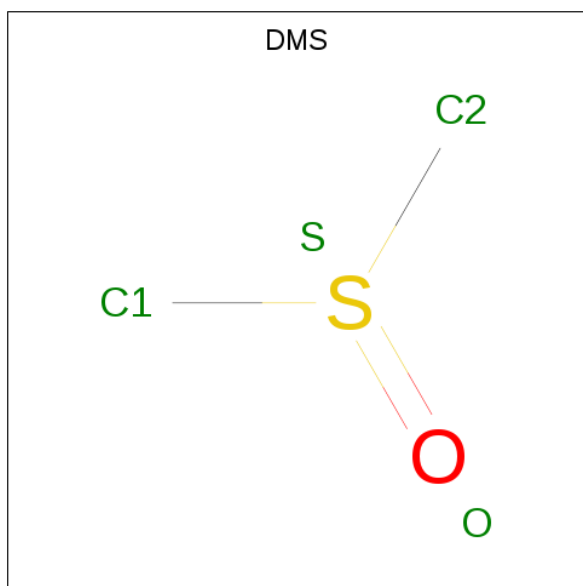
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Cl	0	0
			2	2		

*Continued on next page...*

*Continued from previous page...*

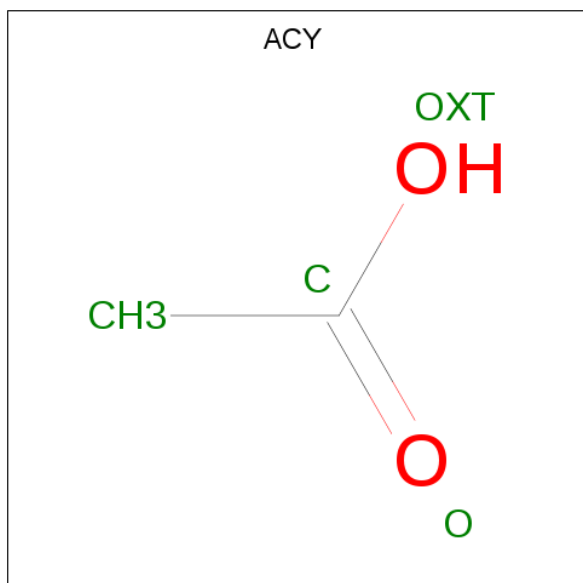
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Cl	0	0
			1	1		

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula:  $C_2H_6OS$ ).



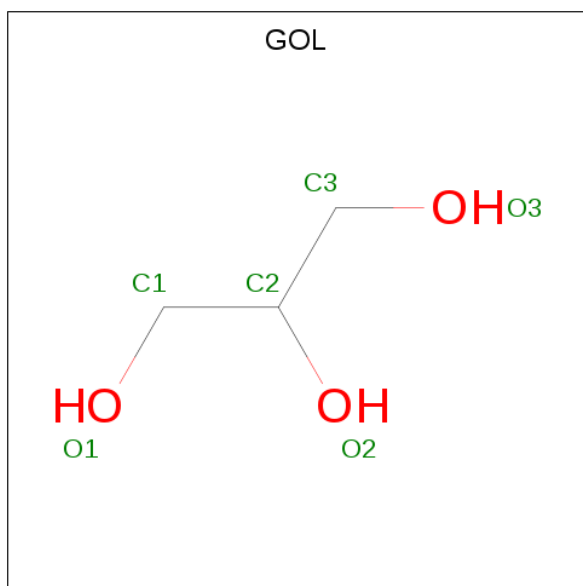
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 5 is ACETIC ACID (three-letter code: ACY) (formula:  $C_2H_4O_2$ ).



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

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		
6	A	1	Total	C	O	0	1
			12	6	6		
6	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is N-[(2R)-2-({N 5 -[amino(iminio)methyl]-L-ornithyl-L-valyl}amino)-4-methylpentyl]-L-phenylalanyl-L-alpha-glutamyl-L-alanyl-L-norleucinamide (three-letter code: 0Q4) (formula:  $C_{40}H_{70}N_{11}O_8$ ).



- Molecule 8 is water.

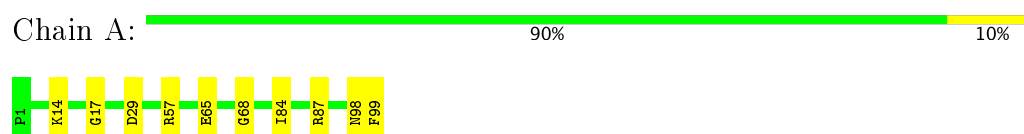
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	85	Total O 85 85	0	1
8	B	109	Total O 109 109	0	0

### 3 Residue-property plots [i](#)

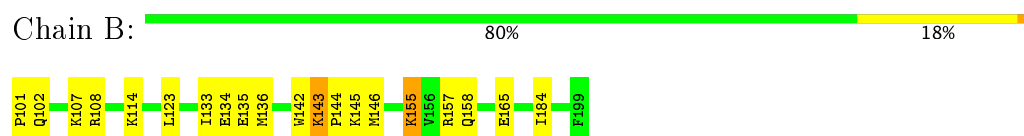
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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 ( $RSRZ > 2$ ). 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.

Note EDS was not executed.

- Molecule 1: POL POLYPROTEIN



- Molecule 1: POL POLYPROTEIN



## 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 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.89 Å 85.96 Å 46.19 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.54	Depositor
% Data completeness (in resolution range)	99.6 (10.00-1.54)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	SHELXL-97	Depositor
R, $R_{free}$	0.130 , 0.198	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1982	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP



## 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: GOL, CL, NA, DMS, 0Q4, ACY

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	0.51	0/872	1.09	2/1171 (0.2%)
1	B	0.55	0/920	1.14	1/1238 (0.1%)
All	All	0.53	0/1792	1.11	3/2409 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	108	ARG	NE-CZ-NH2	-9.40	115.60	120.30
1	A	57	ARG	CD-NE-CZ	7.90	134.66	123.60
1	A	87	ARG	NE-CZ-NH2	-6.00	117.30	120.30

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	803	0	871	6	0
1	B	831	0	908	23	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
3	B	2	0	0	0	0
4	A	4	0	6	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	4	0	3	0	0
6	A	18	0	24	1	0
6	B	6	0	8	0	0
7	B	118	0	137	30	0
8	A	85	0	0	0	0
8	B	109	0	0	23	0
All	All	1982	0	1957	56	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 15.

The worst 5 of 56 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)
7:B:301[A]:0Q4:HA6	8:B:1177:HOH:O	1.16	1.29
7:B:301[A]:0Q4:HD24	7:B:301[A]:0Q4:CD21	0.97	1.16
7:B:301[A]:0Q4:HD1	7:B:301[A]:0Q4:CD11	0.97	1.08
7:B:301[B]:0Q4:CD11	7:B:301[B]:0Q4:HD1	0.97	1.07
7:B:301[B]:0Q4:CD21	7:B:301[B]:0Q4:HD24	0.97	1.05

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	109/99 (110%)	108 (99%)	1 (1%)	0	100	100
1	B	115/99 (116%)	115 (100%)	0	0	100	100
All	All	224/198 (113%)	223 (100%)	1 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	92/80 (115%)	92 (100%)	0	100	100
1	B	98/80 (122%)	91 (93%)	7 (7%)	17	1
All	All	190/160 (119%)	183 (96%)	7 (4%)	62	9

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	143[B]	LYS
1	B	155[B]	LYS
1	B	143[C]	LYS
1	B	107[B]	LYS
1	B	155[A]	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2	GLN
1	B	102	GLN
1	B	198	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](#)

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

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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
4	DMS	A	601	-	3,3,3	0.57	0	3,3,3	0.27	0
5	ACY	A	602	-	1,3,3	4.83	1 (100%)	0,3,3	0.00	-
6	GOL	A	702	-	5,5,5	0.55	0	5,5,5	0.38	0
6	GOL	A	703[A]	-	5,5,5	0.58	0	5,5,5	0.62	0
6	GOL	A	703[B]	-	5,5,5	0.55	0	5,5,5	0.45	0
7	0Q4	B	301[A]	-	56,59,59	1.17	2 (3%)	66,77,77	1.34	7 (10%)
7	0Q4	B	301[B]	-	56,59,59	1.24	2 (3%)	66,77,77	1.37	6 (9%)
6	GOL	B	701	-	5,5,5	0.54	0	5,5,5	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
4	DMS	A	601	-	-	0/0/0/0	0/0/0/0
5	ACY	A	602	-	-	0/0/0/0	0/0/0/0
6	GOL	A	702	-	-	0/4/4/4	0/0/0/0
6	GOL	A	703[A]	-	-	0/4/4/4	0/0/0/0
6	GOL	A	703[B]	-	-	0/4/4/4	0/0/0/0
7	0Q4	B	301[A]	-	-	0/75/77/77	0/1/1/1
7	0Q4	B	301[B]	-	-	0/75/77/77	0/1/1/1
6	GOL	B	701	-	-	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)
7	B	301[B]	0Q4	C2-N3	-5.61	1.32	1.47
7	B	301[A]	0Q4	C2-N3	-4.39	1.35	1.47
5	A	602	ACY	CH3-C	4.83	1.55	1.48
7	B	301[A]	0Q4	C6-N7	6.83	1.46	1.32

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	301[B]	0Q4	C6-N7	6.89	1.46	1.32

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	301[A]	0Q4	NE-CZ-NH2	-2.50	116.08	120.58
7	B	301[B]	0Q4	CA6-C6-N7	-2.25	112.62	116.86
7	B	301[B]	0Q4	O2-C3-N4	-2.14	118.88	122.90
7	B	301[A]	0Q4	CA5-C5-N6	-2.10	111.92	116.86
7	B	301[B]	0Q4	C2-CA2-N2	2.17	114.31	109.29

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	703[B]	GOL	1	0
7	B	301[A]	0Q4	11	0
7	B	301[B]	0Q4	19	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 ⓘ

### 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.