



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 17, 2021 – 01:34 AM EDT

PDB ID : 1TSQ  
Title : CRYSTAL STRUCTURE OF AP2V SUBSTRATE VARIANT OF NC-P1  
DECAMER PEPTIDE IN COMPLEX WITH V82A/D25N HIV-1 PRO-  
TEASE MUTANT  
Authors : Prabu-Jeyabalan, M.; Nalivaika, E.A.; King, N.M.; Schiffer, C.A.  
Deposited on : 2004-06-21  
Resolution : 2.00 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](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

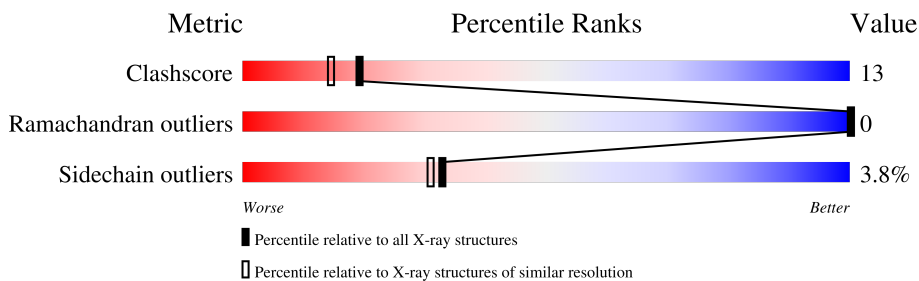
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	99	 79% 19%
1	B	99	 83% 17%
2	P	10	 60% 40%

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
3	ACT	B	505	-	-	X	-
3	ACT	B	508	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1722 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
			Total	C	N	O	S			
1	A	99	744	478	129	133	4	0	0	0
1	B	99	736	472	127	133	4	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	LYS	GLN	engineered mutation	UNP P03369
A	25	ASN	ASP	engineered mutation	UNP P03369
A	82	ALA	VAL	engineered mutation	UNP P03369
B	7	LYS	GLN	engineered mutation	UNP P03369
B	25	ASN	ASP	engineered mutation	UNP P03369
B	82	ALA	VAL	engineered mutation	UNP P03369

- Molecule 2 is a protein called AP2V NC-P1 SUBSTRATE PEPTIDE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	P	10	64	41	11	12	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	4	VAL	ALA	see remark 999	UNP P04591
P	11	ASN	TRP	engineered mutation	UNP P04591

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	68	Total O 68 68	0	0
4	B	65	Total O 65 65	0	0

*Continued on next page...*

*Continued from previous page...*


<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	P	5	Total	O	0	0
			5	5		

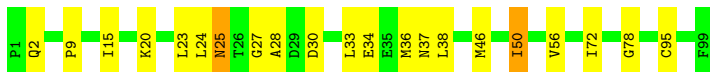
### 3 Residue-property plots [i](#)

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.


Note EDS was not executed.

- Molecule 1: Pol polyprotein

Chain A:  79% 19%



- Molecule 1: Pol polyprotein

Chain B:  83% 17%



- Molecule 2: AP2V NC-P1 SUBSTRATE PEPTIDE

Chain P:  60% 40%



## 4 Data and refinement statistics

Xtrriage (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, $\alpha$ , $\beta$ , $\gamma$	50.77Å 57.26Å 60.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.99 – 2.00	Depositor
% Data completeness (in resolution range)	99.6 (37.99-2.00)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5, CNS 1.0	Depositor
R, $R_{free}$	0.195 , 0.231	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	1722	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	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: ACT

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	0.30	0/758	0.62	0/1030
1	B	0.32	0/750	0.63	0/1022
2	P	0.45	0/64	0.62	0/85
All	All	0.31	0/1572	0.62	0/2137

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	744	0	771	23	0
1	B	736	0	749	26	0
2	P	64	0	47	9	0
3	A	16	0	12	0	0
3	B	24	0	18	6	0
4	A	68	0	0	2	0
4	B	65	0	0	2	0
4	P	5	0	0	0	0
All	All	1722	0	1597	41	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 13.

All (41) 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:50:ILE:HG23	1:B:49:GLY:HA3	1.47	0.96
1:B:47:ILE:HD12	2:P:7:LEU:HD11	1.57	0.85
1:A:25:ASN:ND2	1:A:27:GLY:H	1.88	0.72
1:A:50:ILE:HD11	1:B:54:ILE:HG12	1.74	0.69
1:A:25:ASN:HD22	1:A:27:GLY:H	1.44	0.65
1:B:47:ILE:HG22	2:P:10:ILE:HA	1.83	0.61
1:A:50:ILE:CG2	1:B:49:GLY:HA3	2.29	0.56
1:A:50:ILE:HD11	1:B:47:ILE:HG13	1.87	0.55
1:A:25:ASN:HD22	1:A:25:ASN:C	2.13	0.52
1:A:95:CYS:HA	1:B:99:PHE:HB3	1.93	0.51
1:B:14:ARG:NH2	3:B:508:ACT:OXT	2.43	0.51
1:A:15:ILE:HD13	1:A:33:LEU:CD1	2.41	0.50
1:B:31:THR:HG23	1:B:84:ILE:O	2.12	0.50
1:B:14:ARG:HG2	1:B:19:LEU:CD2	2.42	0.49
1:A:30:ASP:HB3	4:A:517:HOH:O	2.12	0.47
1:B:30:ASP:HB3	4:B:540:HOH:O	2.14	0.47
1:B:50:ILE:HD12	2:P:4:VAL:HB	1.95	0.47
1:A:9:PRO:HD2	1:A:23:LEU:HD11	1.96	0.47
1:B:28:ALA:HA	2:P:7:LEU:HB3	1.95	0.47
1:A:56:VAL:HG12	1:A:78:GLY:HA3	1.96	0.46
1:A:20:LYS:HE2	1:A:20:LYS:HB3	1.81	0.46
1:B:14:ARG:HB2	1:B:65:GLU:HB3	1.98	0.46
1:B:69:HIS:HE1	3:B:502:ACT:OXT	1.99	0.46
1:B:57:ARG:HE	3:B:505:ACT:CH3	2.29	0.45
1:A:46:MET:HE1	4:A:511:HOH:O	2.16	0.45
1:A:50:ILE:CD1	1:B:47:ILE:HG13	2.48	0.44
1:B:14:ARG:HG2	1:B:19:LEU:HD23	2.00	0.44
1:A:28:ALA:HB1	2:P:4:VAL:HG13	1.99	0.44
1:A:50:ILE:HG23	1:B:49:GLY:CA	2.34	0.44
1:A:36:MET:HG3	1:A:38:LEU:HG	1.98	0.43
1:B:46:MET:O	2:P:11:ASN:N	2.48	0.43
1:A:50:ILE:CD1	1:B:54:ILE:HG12	2.46	0.43
1:A:2:GLN:HE21	1:A:2:GLN:HB2	1.66	0.42
1:A:72:ILE:O	1:A:72:ILE:HG23	2.21	0.41
1:B:14:ARG:HH22	3:B:508:ACT:C	2.32	0.41
1:B:45:LYS:CB	4:B:569:HOH:O	2.68	0.41
1:A:28:ALA:CB	2:P:4:VAL:HG13	2.51	0.41
1:B:57:ARG:HE	3:B:505:ACT:H3	1.85	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:28:ALA:HA	2:P:4:VAL:HA	2.02	0.40
1:B:14:ARG:NH2	3:B:508:ACT:H1	2.36	0.40
1:B:28:ALA:CB	2:P:7:LEU:HB3	2.51	0.40

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	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
1	B	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
2	P	8/10 (80%)	7 (88%)	1 (12%)	0	100	100
All	All	202/208 (97%)	199 (98%)	3 (2%)	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	79/82 (96%)	74 (94%)	5 (6%)	18	13
1	B	77/82 (94%)	76 (99%)	1 (1%)	69	74
2	P	4/9 (44%)	4 (100%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	160/173 (92%)	154 (96%)	6 (4%)	33 31

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	LEU
1	A	25	ASN
1	A	34	GLU
1	A	37	ASN
1	A	50	ILE
1	B	88	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2	GLN
1	A	18	GLN
1	A	25	ASN
1	A	37	ASN
1	B	18	GLN
1	B	37	ASN
1	B	61	GLN
1	B	69	HIS
1	B	88	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 monosaccharides in this entry.

## 5.6 Ligand geometry

10 ligands 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	ACT	B	508	-	1,3,3	2.77	1 (100%)	0,3,3	-	-
3	ACT	B	509	-	1,3,3	2.79	1 (100%)	0,3,3	-	-
3	ACT	A	503	-	1,3,3	3.17	1 (100%)	0,3,3	-	-
3	ACT	A	506	-	1,3,3	3.23	1 (100%)	0,3,3	-	-
3	ACT	B	510	-	1,3,3	2.89	1 (100%)	0,3,3	-	-
3	ACT	B	501	-	1,3,3	2.65	1 (100%)	0,3,3	-	-
3	ACT	B	505	-	1,3,3	2.70	1 (100%)	0,3,3	-	-
3	ACT	A	507	-	1,3,3	2.78	1 (100%)	0,3,3	-	-
3	ACT	A	504	-	1,3,3	2.99	1 (100%)	0,3,3	-	-
3	ACT	B	502	-	1,3,3	3.33	1 (100%)	0,3,3	-	-

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	502	ACT	CH3-C	3.33	1.53	1.48
3	A	506	ACT	CH3-C	3.23	1.52	1.48
3	A	503	ACT	CH3-C	3.17	1.52	1.48
3	A	504	ACT	CH3-C	2.99	1.52	1.48
3	B	510	ACT	CH3-C	2.89	1.52	1.48
3	B	509	ACT	CH3-C	2.79	1.52	1.48
3	A	507	ACT	CH3-C	2.78	1.52	1.48
3	B	508	ACT	CH3-C	2.77	1.52	1.48
3	B	505	ACT	CH3-C	2.70	1.52	1.48
3	B	501	ACT	CH3-C	2.65	1.52	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	508	ACT	3	0
3	B	505	ACT	2	0
3	B	502	ACT	1	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](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.