



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 09:03 PM GMT

PDB ID : 3A2O  
Title : Crystal Structure of HIV-1 Protease Complexed with KNI-1689  
Authors : Adachi, M.; Tamada, T.; Hidaka, K.; Kimura, T.; Kiso, Y.; Kuroki, R.  
Deposited on : 2009-05-26  
Resolution : 0.88 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

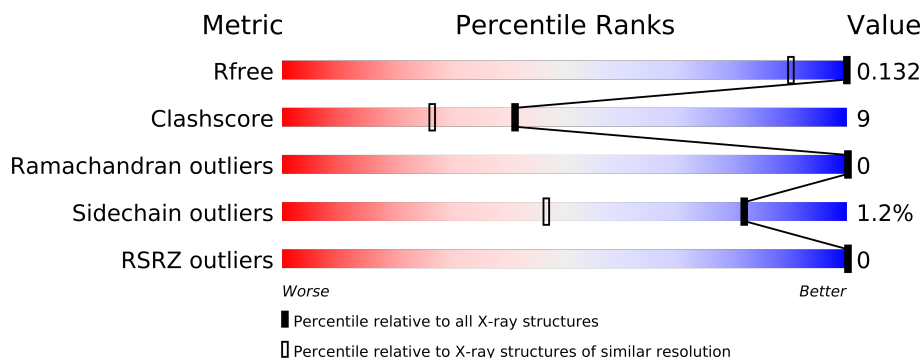
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 0.88 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))
$R_{free}$	66092	1083 (1.12-0.66)
Clashscore	79885	1125 (1.10-0.66)
Ramachandran outliers	78287	1050 (1.10-0.66)
Sidechain outliers	78261	1049 (1.10-0.66)
RSRZ outliers	66119	1083 (1.12-0.66)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	99	<div><div></div><div></div><div></div><div></div><div></div></div>
1	B	99	<div><div></div><div></div><div></div><div></div><div></div></div>

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	GOL	B	803[A]	-	X
2	GOL	B	803[B]	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4650 atoms, of which 2096 are hydrogens and 0 are deuterium.

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 PROTEASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	99	Total	C	H	N	O	S	0	32	0
			1922	593	1006	159	160	4			
1	B	99	Total	C	H	N	O	S	0	34	0
			1926	598	1012	151	162	3			

There are 10 discrepancies between the modelled and reference sequences:

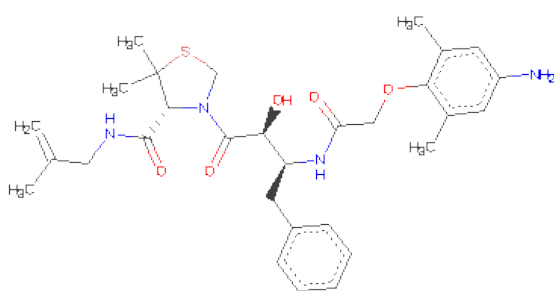
Chain	Residue	Modelled	Actual	Comment	Reference
A	7	LYS	GLN	ENGINEERED	UNP Q9EKL4
A	33	ILE	VAL	ENGINEERED	UNP Q9EKL4
A	63	ILE	LEU	ENGINEERED	UNP Q9EKL4
A	67	ALA	CYS	ENGINEERED	UNP Q9EKL4
A	95	ALA	CYS	ENGINEERED	UNP Q9EKL4
B	107	LYS	GLN	ENGINEERED	UNP Q9EKL4
B	133	ILE	VAL	ENGINEERED	UNP Q9EKL4
B	163	ILE	LEU	ENGINEERED	UNP Q9EKL4
B	167	ALA	CYS	ENGINEERED	UNP Q9EKL4
B	195	ALA	CYS	ENGINEERED	UNP Q9EKL4

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 3 is (4R)-3-[(2S,3S)-3-[[[(4-AMINO-2,6-DIMETHYLPHENOXY)ACETYL]AMINO}-2-HYDROXY-4-PHENYLBUTANOYL]-5,5-DIMETHYL-N-(2-METHYLPROP-2-EN-1-YL)-1,3-THIAZOLIDINE-4-CARBOXAMIDE (three-letter code: KNJ) (formula:  $C_{30}H_{40}N_4O_5S$ ).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	B	1	Total	C	H	N	O	S	0	1
			158	60	78	8	10	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	320	Total 320	O 320	0	104
4	B	312	Total 312	O 312	0	85



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.22Å 85.81Å 46.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 0.88 24.67 – 0.88	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-0.88) 92.3 (24.67-0.88)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.27 (at 0.88Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, $R_{free}$	0.101 , 0.118 0.117 , 0.132	Depositor DCC
$R_{free}$ test set	8498 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	7.9	Xtriage
Anisotropy	0.107	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 86.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 169747 reflections	Xtriage
$F_o, F_c$ correlation	0.99	EDS
Total number of atoms	4650	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.75% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, KNJ

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	0.76	0/982	1.27	14/1323 (1.1%)
1	B	0.80	1/1000 (0.1%)	1.33	14/1351 (1.0%)
All	All	0.78	1/1982 (0.1%)	1.30	28/2674 (1.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	141	ARG	CG-CD	-5.99	1.36	1.51

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	41	ARG	NE-CZ-NH1	10.50	125.55	120.30
1	B	102	GLN	CG-CD-NE2	10.05	140.83	116.70
1	A	46[A]	MET	CG-SD-CE	-9.90	84.35	100.20
1	A	46[B]	MET	CG-SD-CE	-9.90	84.35	100.20
1	A	46[C]	MET	CG-SD-CE	-9.90	84.35	100.20

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	916	1006	0	20	0
1	B	914	1012	0	15	0
2	B	12	0	0	0	0
3	B	80	78	0	4	0
4	A	320	0	0	19	3
4	B	312	0	0	10	3
All	All	2554	2096	0	37	3

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

The worst 5 of 37 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:70[B]:LYS:CE	4:A:1417:HOH:O	2.32	0.76
1:B:191:THR:HG22	4:B:1150:HOH:O	1.86	0.75
1:A:7[B]:LYS:HD3	4:A:1258:HOH:O	1.87	0.75
1:A:18:GLN:HB2	4:A:1372:HOH:O	1.87	0.74
1:B:182[A]:VAL:HG21	4:B:1456[A]:HOH:O	1.89	0.72

All (3) 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	Distance(Å)	Clash(Å)
4:A:1469[A]:HOH:O	4:B:1423:HOH:O[1_554]	0.51	1.69
4:A:1622[B]:HOH:O	4:B:1423:HOH:O[1_554]	2.07	0.13
4:A:1621[B]:HOH:O	4:B:1301:HOH:O[4_456]	2.17	0.03

## 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	125/99 (126%)	124 (99%)	1 (1%)	0	100	100
1	B	127/99 (128%)	127 (100%)	0	0	100	100
All	All	252/198 (127%)	251 (100%)	1 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	103/81 (127%)	103 (100%)	0	100	100
1	B	108/81 (133%)	103 (95%)	5 (5%)	37	5
All	All	211/162 (130%)	206 (98%)	5 (2%)	82	18

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	123[A]	LEU
1	B	123[B]	LEU
1	B	123[C]	LEU
1	B	146[A]	MET
1	B	146[B]	MET

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	A	58	GLN
1	B	102	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

4 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$
2	GOL	B	803[A]	-	5,5,5	1.85	1 (20%)	5,5,5	1.23	0
2	GOL	B	803[B]	-	5,5,5	1.13	1 (20%)	5,5,5	1.26	0
3	KNJ	B	999[A]	-	42,42,42	0.88	1 (2%)	60,60,60	1.07	3 (5%)
3	KNJ	B	999[B]	-	42,42,42	1.05	3 (7%)	60,60,60	1.81	7 (11%)

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
2	GOL	B	803[A]	-	-	0/4/4/4	0/0/0/0
2	GOL	B	803[B]	-	-	0/4/4/4	0/0/0/0
3	KNJ	B	999[A]	-	-	0/34/50/50	0/3/3/3
3	KNJ	B	999[B]	-	-	0/34/50/50	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	803[A]	GOL	O2-C2	-3.67	1.32	1.43
3	B	999[B]	KNJ	C27-C28	3.35	1.53	1.50
3	B	999[B]	KNJ	C22-S2	-3.09	1.81	1.84
3	B	999[A]	KNJ	C22-S2	-2.68	1.81	1.84
2	B	803[B]	GOL	O3-C3	-2.35	1.32	1.42

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	999[B]	KNJ	C30-C28-C29	7.02	139.33	121.75
3	B	999[B]	KNJ	C27-N5-C26	5.77	134.06	121.29
3	B	999[B]	KNJ	C30-C28-C27	-5.74	102.55	115.25
3	B	999[A]	KNJ	C4-C3-C2	3.62	123.29	121.34
3	B	999[A]	KNJ	C4-C5-C7	3.37	123.16	121.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	99/99 (100%)	-0.70	0 100 100	4, 9, 18, 49	0
1	B	99/99 (100%)	-0.75	0 100 100	4, 9, 23, 30	0
All	All	198/198 (100%)	-0.72	0 100 100	4, 9, 21, 49	0

There are no RSRZ outliers to report.

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	GOL	B	803[B]	6/6	0.08	2.99	5,11,13,14	6
2	GOL	B	803[A]	6/6	0.08	2.85	7,9,12,20	6
3	KNJ	B	999[B]	40/40	0.06	0.27	4,6,14,28	79
3	KNJ	B	999[A]	40/40	0.06	0.25	4,6,14,17	79

## 6.5 Other polymers ⓘ

There are no such residues in this entry.