



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 08:14 AM GMT

PDB ID : 3DW3  
Title : Proteinase K by Classical hanging drop method before high X Ray dose on ESRF ID 14-2 beamline  
Authors : Pechkova, E.; Tripathi, S.K.; Nicolini, C.  
Deposited on : 2008-07-21  
Resolution : 0.99 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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

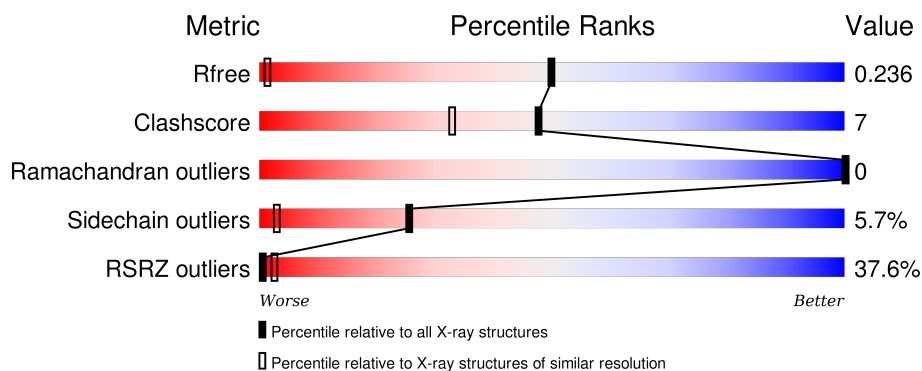
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 0.99 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))
$R_{free}$	91344	1235 (1.10-0.90)
Clashscore	102246	1333 (1.10-0.90)
Ramachandran outliers	100387	1247 (1.10-0.90)
Sidechain outliers	100360	1246 (1.10-0.90)
RSRZ outliers	91569	1239 (1.10-0.90)

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 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.

Mol	Chain	Length	Quality of chain
1	X	279	<div> <div>38%</div> <div>86%</div> <div>11%</div> <div>.</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2220 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 Proteinase K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	279	Total	C	N	O	S	0	1	0
			2021	1242	353	416	10			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	X	1	Total	Ca	0	0
			1	1		

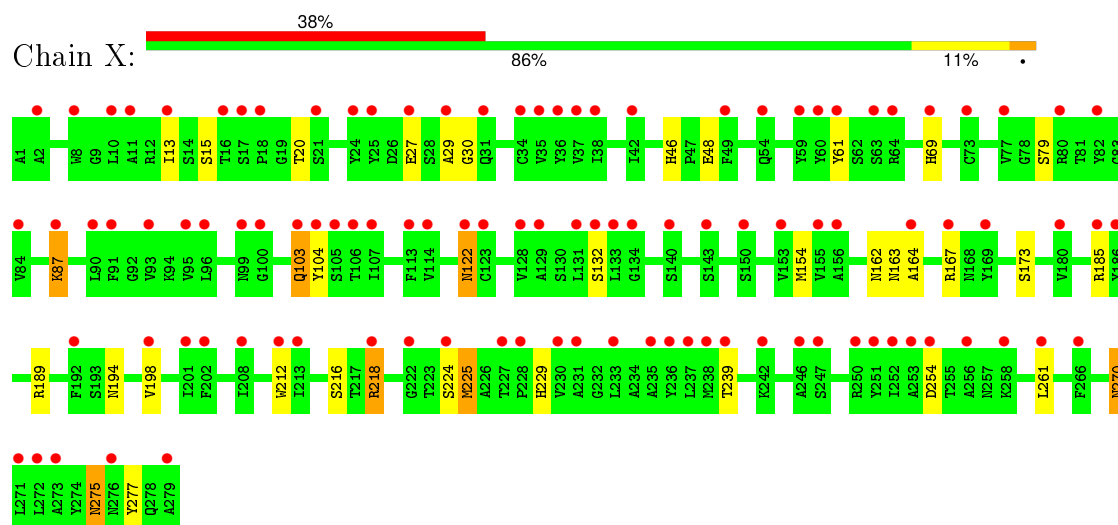
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	X	198	Total	O	0	0
			198	198		

### 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 errors 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.

#### • Molecule 1: Proteinase K



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.94Å 67.94Å 102.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	56.61 – 0.99 27.82 – 0.99	Depositor EDS
% Data completeness (in resolution range)	86.6 (56.61-0.99) 86.6 (27.82-0.99)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.00 (at 0.99Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.226 , 0.235 0.226 , 0.236	Depositor DCC
$R_{free}$ test set	5714 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	5.1	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 42.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 113835 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2220	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	6.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.56% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 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: CA

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	X	0.36	0/2065	0.61	2/2805 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	X	218	ARG	NE-CZ-NH2	-8.44	116.08	120.30
1	X	218	ARG	NE-CZ-NH1	6.92	123.76	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	X	2021	0	1920	29	2
2	X	1	0	0	0	0
3	X	198	0	0	3	1
All	All	2220	0	1920	29	2

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 7.

All (29) 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:X:29:ALA:HB3	1:X:87:LYS:HD3	1.49	0.94
1:X:20:THR:HG23	3:X:349:HOH:O	1.75	0.85
1:X:261:LEU:H	1:X:270:ASN:HD21	1.30	0.78
1:X:46:HIS:HD2	1:X:48:GLU:H	1.32	0.75
1:X:173:SER:HA	1:X:198:VAL:HG21	1.69	0.75
1:X:13:ILE:HD11	1:X:229:HIS:HB3	1.72	0.70
1:X:132:SER:HB3	1:X:224[A]:SER:OG	1.93	0.69
1:X:29:ALA:CB	1:X:87:LYS:HD3	2.24	0.65
1:X:30:GLY:C	1:X:239:THR:HG21	2.16	0.65
1:X:163:ASN:ND2	1:X:189:ARG:HH12	2.01	0.59
1:X:46:HIS:HE1	1:X:216:SER:O	1.90	0.54
1:X:103:GLN:NE2	1:X:103:GLN:H	2.06	0.54
1:X:167:ARG:HD3	3:X:381:HOH:O	2.08	0.53
1:X:13:ILE:CD1	1:X:229:HIS:HB3	2.37	0.53
1:X:162:ASN:HB2	1:X:194:ASN:HD21	1.75	0.51
1:X:225:MET:O	1:X:229:HIS:HD2	1.95	0.49
1:X:164:ALA:H	1:X:194:ASN:ND2	2.11	0.48
1:X:30:GLY:HA2	1:X:239:THR:HG21	1.96	0.48
1:X:275:ASN:ND2	1:X:277:TYR:H	2.12	0.48
1:X:30:GLY:CA	1:X:239:THR:HG21	2.44	0.47
1:X:163:ASN:HD22	1:X:189:ARG:HH12	1.63	0.44
1:X:164:ALA:H	1:X:194:ASN:HD22	1.66	0.44
1:X:46:HIS:CD2	1:X:48:GLU:H	2.21	0.43
1:X:103:GLN:CD	1:X:103:GLN:H	2.22	0.43
1:X:48:GLU:HB3	1:X:79:SER:HB2	2.01	0.41
1:X:212:TRP:HB2	1:X:218:ARG:HB2	2.03	0.41
1:X:270:ASN:C	1:X:270:ASN:HD22	2.23	0.41
1:X:162:ASN:H	1:X:194:ASN:ND2	2.19	0.41
1:X:218:ARG:HD3	3:X:316:HOH:O	2.21	0.40

All (2) 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:X:15:SER:OG	1:X:122:ASN:ND2[5_444]	1.93	0.27
1:X:254:ASP:OD2	3:X:407:HOH:O[3_454]	2.03	0.17

## 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	X	278/279 (100%)	271 (98%)	7 (2%)	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	X	212/213 (100%)	200 (94%)	12 (6%)	25	3

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	27	GLU
1	X	61	TYR
1	X	69	HIS
1	X	87	LYS
1	X	103	GLN
1	X	104	TYR
1	X	122	ASN
1	X	154	MET
1	X	185	ARG
1	X	225	MET
1	X	270	ASN
1	X	275	ASN



Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	46	HIS
1	X	54	GLN
1	X	99	ASN
1	X	162	ASN
1	X	163	ASN
1	X	194	ASN
1	X	229	HIS
1	X	257	ASN
1	X	270	ASN
1	X	275	ASN
1	X	276	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 [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	X	279/279 (100%)	2.00	105 (37%) 0 2	3, 5, 9, 13	0

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	61	TYR	15.7
1	X	279	ALA	6.8
1	X	69	HIS	6.5
1	X	54	GLN	6.2
1	X	80	ARG	5.7
1	X	266	PHE	5.7
1	X	103	GLN	5.5
1	X	276	ASN	5.1
1	X	13	ILE	5.1
1	X	104	TYR	4.9
1	X	99	ASN	4.8
1	X	239	THR	4.7
1	X	198	VAL	4.3
1	X	122	ASN	4.3
1	X	254	ASP	4.2
1	X	27	GLU	4.2
1	X	167	ARG	4.0
1	X	218	ARG	4.0
1	X	224[A]	SER	3.7
1	X	250	ARG	3.5
1	X	87	LYS	3.4
1	X	185	ARG	3.3
1	X	31	GLN	3.3
1	X	64	ARG	3.1
1	X	131	LEU	3.0
1	X	17	SER	3.0
1	X	247	SER	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	X	84	VAL	2.9
1	X	246	ALA	2.9
1	X	238	MET	2.9
1	X	150	SER	2.9
1	X	212	TRP	2.8
1	X	258	LYS	2.8
1	X	105	SER	2.8
1	X	272	LEU	2.7
1	X	230	VAL	2.7
1	X	134	GLY	2.7
1	X	140	SER	2.7
1	X	201	ILE	2.7
1	X	252	ILE	2.7
1	X	133	LEU	2.7
1	X	24	TYR	2.7
1	X	60	TYR	2.7
1	X	186	TYR	2.7
1	X	261	LEU	2.7
1	X	35	VAL	2.7
1	X	93	VAL	2.7
1	X	77	VAL	2.6
1	X	73	CYS	2.6
1	X	123	CYS	2.6
1	X	91	PHE	2.6
1	X	59	TYR	2.6
1	X	21	SER	2.6
1	X	155	VAL	2.5
1	X	180	VAL	2.5
1	X	227	THR	2.5
1	X	231	ALA	2.4
1	X	34	CYS	2.4
1	X	38	ILE	2.4
1	X	107	ILE	2.4
1	X	16	THR	2.4
1	X	90	LEU	2.4
1	X	37	VAL	2.4
1	X	156	ALA	2.4
1	X	8	TRP	2.4
1	X	233	LEU	2.4
1	X	143	SER	2.3
1	X	25	TYR	2.3
1	X	10	LEU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	X	271	LEU	2.3
1	X	42	ILE	2.3
1	X	29	ALA	2.3
1	X	132	SER	2.3
1	X	242	LYS	2.3
1	X	82	TYR	2.2
1	X	63	SER	2.2
1	X	213	ILE	2.2
1	X	251	TYR	2.2
1	X	237	LEU	2.2
1	X	129	ALA	2.2
1	X	36	TYR	2.2
1	X	228	PRO	2.2
1	X	95	VAL	2.1
1	X	114	VAL	2.1
1	X	236	TYR	2.1
1	X	113	PHE	2.1
1	X	208	ILE	2.1
1	X	128	VAL	2.1
1	X	106	THR	2.1
1	X	49	PHE	2.1
1	X	202	PHE	2.1
1	X	164	ALA	2.1
1	X	235	ALA	2.1
1	X	256	ALA	2.1
1	X	18	PRO	2.1
1	X	169	TYR	2.1
1	X	96	LEU	2.0
1	X	100	GLY	2.0
1	X	153	VAL	2.0
1	X	192	PHE	2.0
1	X	2	ALA	2.0
1	X	11	ALA	2.0
1	X	253	ALA	2.0
1	X	273	ALA	2.0
1	X	222	GLY	2.0

## 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 [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	X	280	1/1	0.23	1.46	-	111,111,111,111	0

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

There are no such residues in this entry.