



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 08:45 am GMT

PDB ID : 2PKA  
Title : REFINED 2 ANGSTROMS X-RAY CRYSTAL STRUCTURE OF PORCINE PANCREATIC KALLIKREIN A, A SPECIFIC TRYPSIN-LIKE SERINE PROTEINASE. CRYSTALLIZATION, STRUCTURE DETERMINATION, CRYSTALLOGRAPHIC REFINEMENT, STRUCTURE AND ITS COMPARISON WITH BOVINE TRYPSIN  
Authors : Bode, W.; Chen, Z.  
Deposited on : 1984-05-21  
Resolution : 2.05 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) : recalc28949

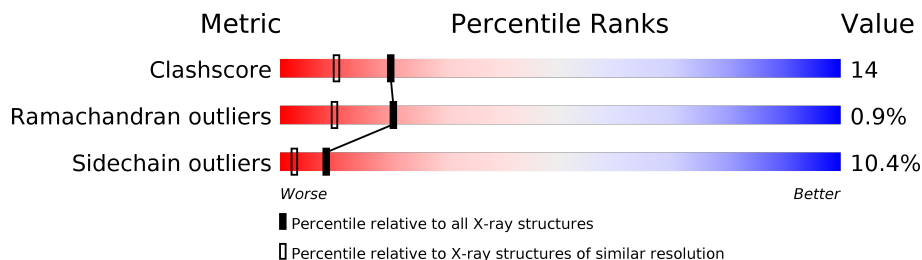
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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	1394 (2.04-2.04)
Ramachandran outliers	110173	1383 (2.04-2.04)
Sidechain outliers	110143	1383 (2.04-2.04)

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	80	
1	X	80	
2	B	152	
2	Y	152	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3701 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 KALLIKREIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	80	Total	C	N	O	S	0	0	0
			645	414	113	115	3			
1	X	80	Total	C	N	O	S	0	0	0
			645	414	113	115	3			

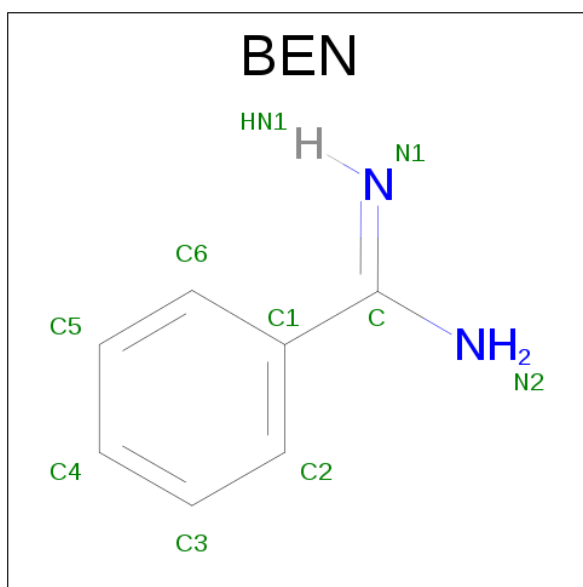
- Molecule 2 is a protein called KALLIKREIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	152	Total	C	N	O	S	0	0	0
			1154	726	182	235	11			
2	Y	152	Total	C	N	O	S	0	0	0
			1154	726	182	235	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	148	ASP	-	INSERTION	UNP P00752
B	170	ASP	-	INSERTION	UNP P00752
B	174	ASP	-	INSERTION	UNP P00752
B	239	ASP	ASN	CONFLICT	UNP P00752
Y	148	ASP	-	INSERTION	UNP P00752
Y	170	ASP	-	INSERTION	UNP P00752
Y	174	ASP	-	INSERTION	UNP P00752
Y	239	ASP	ASN	CONFLICT	UNP P00752

- Molecule 3 is BENZAMIDINE (three-letter code: BEN) (formula: C<sub>7</sub>H<sub>8</sub>N<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	N	0	0
			9	7	2		
3	Y	1	Total	C	N	0	0
			9	7	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	16	Total	O	0	0
			16	16		
4	B	34	Total	O	0	0
			34	34		
4	X	7	Total	O	0	0
			7	7		
4	Y	28	Total	O	0	0
			28	28		

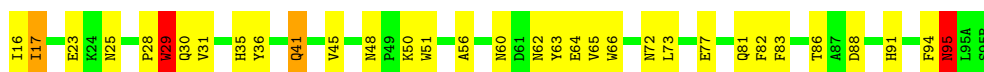
### 3 Residue-property plots

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.

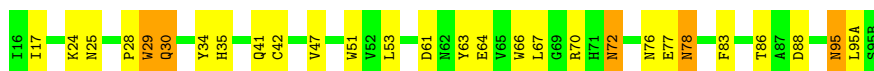
#### • Molecule 1: KALLIKREIN A

Chain A: 



#### • Molecule 1: KALLIKREIN A

Chain X: 



#### • Molecule 2: KALLIKREIN A

Chain B: 



#### • Molecule 2: KALLIKREIN A

Chain Y: 



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.20Å 90.20Å 159.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.05	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.05)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	EREF	Depositor
R, $R_{free}$	0.220 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3701	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	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: BEN

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	1.11	3/668 (0.4%)	1.33	3/910 (0.3%)
1	X	1.12	3/668 (0.4%)	1.28	1/910 (0.1%)
2	B	1.02	4/1184 (0.3%)	1.26	1/1614 (0.1%)
2	Y	1.02	4/1184 (0.3%)	1.25	2/1614 (0.1%)
All	All	1.05	14/3704 (0.4%)	1.27	7/5048 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	8
1	X	0	9
2	B	0	10
2	Y	0	10
All	All	0	37

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	215	TRP	NE1-CE2	-8.36	1.26	1.37
2	B	141	TRP	NE1-CE2	-8.18	1.26	1.37
2	Y	209	TRP	NE1-CE2	-7.93	1.27	1.37
1	A	29	TRP	NE1-CE2	-7.73	1.27	1.37
2	Y	237	TRP	NE1-CE2	-7.70	1.27	1.37

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	95	ASN	O-C-N	8.88	136.91	122.70
1	A	95	ASN	CA-C-N	-7.61	100.46	117.20
1	X	70	ARG	NE-CZ-NH2	-6.53	117.03	120.30
2	Y	150	GLU	OE1-CD-OE2	-5.90	116.22	123.30
2	B	181	LEU	CB-CA-C	5.62	120.89	110.20

There are no chirality outliers.

5 of 37 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	25	ASN	Mainchain
1	A	30	GLN	Sidechain
1	A	41	GLN	Sidechain
1	A	63	TYR	Mainchain
1	A	77	GLU	Sidechain

## 5.2 Too-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 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	645	0	589	23	0
1	X	645	0	589	9	0
2	B	1154	0	1085	36	2
2	Y	1154	0	1085	51	2
3	B	9	0	7	0	0
3	Y	9	0	7	0	0
4	A	16	0	0	1	0
4	B	34	0	0	0	0
4	X	7	0	0	0	0
4	Y	28	0	0	2	0
All	All	3701	0	3362	100	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 14.

The worst 5 of 100 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)
2:Y:95(Z):ASP:HB3	2:Y:98:ASP:CB	1.23	1.60
2:Y:95(Z):ASP:CB	2:Y:98:ASP:HB2	1.42	1.49
2:Y:95(Z):ASP:CB	2:Y:98:ASP:CB	2.03	1.32
2:Y:95(Z):ASP:CG	2:Y:175:LYS:HD3	1.50	1.30
2:Y:95(Z):ASP:OD2	2:Y:175:LYS:HD3	1.29	1.30

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)
2:B:113:LYS:NZ	2:Y:148(A):ASP:OD1[7_566]	1.06	1.14
2:B:113:LYS:CE	2:Y:148(A):ASP:OD1[7_566]	1.97	0.23

## 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	78/80 (98%)	72 (92%)	6 (8%)	0	100	100
1	X	78/80 (98%)	72 (92%)	5 (6%)	1 (1%)	14	5
2	B	150/152 (99%)	140 (93%)	9 (6%)	1 (1%)	25	14
2	Y	150/152 (99%)	140 (93%)	8 (5%)	2 (1%)	14	5
All	All	456/464 (98%)	424 (93%)	28 (6%)	4 (1%)	20	10

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	95	ASN
2	B	96	GLY
2	Y	165	ASN
2	Y	96	GLY

### 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	68/68 (100%)	63 (93%)	5 (7%)	16	8
1	X	68/68 (100%)	58 (85%)	10 (15%)	3	1
2	B	129/129 (100%)	120 (93%)	9 (7%)	18	9
2	Y	129/129 (100%)	112 (87%)	17 (13%)	5	1
All	All	394/394 (100%)	353 (90%)	41 (10%)	8	2

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	72	ASN
1	X	95	ASN
2	Y	212	ILE
1	X	76	ASN
1	X	77	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	165	ASN
2	B	210	GLN
1	X	78	ASN
2	B	127	GLN
2	B	164	GLN

### 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](#)

2 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	BEN	B	1	-	9,9,9	2.19	1 (11%)	9,11,11	1.09	1 (11%)
3	BEN	Y	1	-	9,9,9	1.86	1 (11%)	9,11,11	0.55	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
3	BEN	B	1	-	-	0/4/4/4	0/1/1/1
3	BEN	Y	1	-	-	0/4/4/4	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1	BEN	C1-C	-6.32	1.36	1.47
3	Y	1	BEN	C1-C	-5.29	1.38	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1	BEN	C1-C-N2	2.36	121.75	118.12

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

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.