



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2024 – 01:11 PM EDT

PDB ID : 2KAI  
Title : REFINED 2.5 ANGSTROMS X-RAY CRYSTAL STRUCTURE OF THE COMPLEX FORMED BY PORCINE KALLIKREIN A AND THE BOVINE PANCREATIC TRYPSIN INHIBITOR. CRYSTALLIZATION, PATTERSON SEARCH, STRUCTURE DETERMINATION, REFINEMENT, STRUCTURE AND COMPARISON WITH ITS COMPONENTS AND WITH THE BOVINE TRYPSIN-PANCREATIC TRYPSIN INHIBITOR COMPLEX  
Authors : Bode, W.; Chen, Z.  
Deposited on : 1984-05-21  
Resolution : 2.50 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (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) : 2.36.2

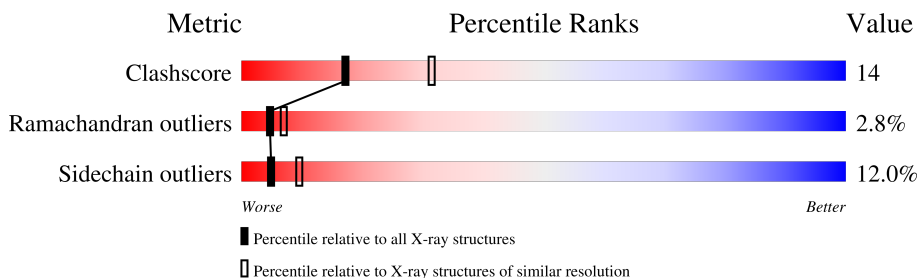
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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	80	59% 28% 11% .
2	B	152	60% 28% 9% .
3	I	58	57% 33% 7% . .

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2247 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
			Total	C	N	O	S			
1	A	80	645	414	113	115	3	60	0	0

- Molecule 2 is a protein called KALLIKREIN A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	152	1154	726	182	235	11	91	0	0

There are 4 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

- Molecule 3 is a protein called BOVINE PANCREATIC TRYPSIN INHIBITOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	I	57	438	275	80	76	7	41	0	1

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		
4	B	6	Total	O	0	0
			6	6		
4	I	3	Total	O	0	0
			3	3		

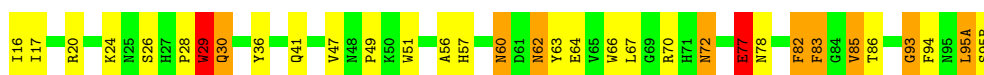
### 3 Residue-property plots [i](#)

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.

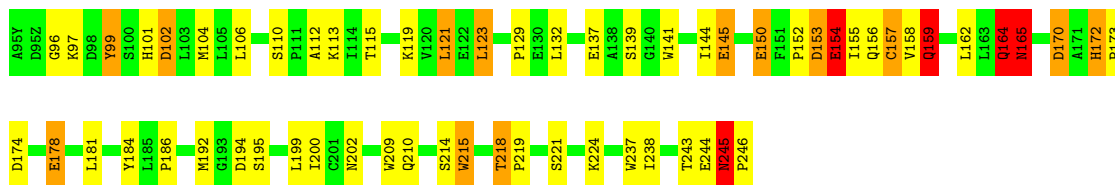
- Molecule 1: KALLIKREIN A

Chain A: 



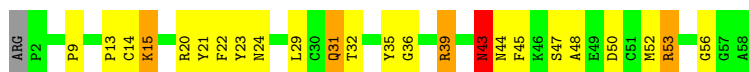
- Molecule 2: KALLIKREIN A

Chain B: 



- Molecule 3: BOVINE PANCREATIC TRYPSIN INHIBITOR

Chain I: 



## 4 Data and refinement statistics

Xtrriage (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, $\alpha$ , $\beta$ , $\gamma$	106.20Å 106.20Å 108.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.50)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	EREF	Depositor
R, $R_{free}$	0.224 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2247	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

## 5 Model quality i

### 5.1 Standard geometry i

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	1.23	4/668 (0.6%)	1.40	2/910 (0.2%)
2	B	1.13	5/1184 (0.4%)	1.37	2/1614 (0.1%)
3	I	1.07	1/449 (0.2%)	1.31	0/602
All	All	1.15	10/2301 (0.4%)	1.37	4/3126 (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	14
2	B	0	17
3	I	0	5
All	All	0	36

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	29	TRP	NE1-CE2	-8.37	1.26	1.37
1	A	66	TRP	NE1-CE2	-7.96	1.27	1.37
1	A	51	TRP	NE1-CE2	-7.81	1.27	1.37
2	B	209	TRP	NE1-CE2	-7.72	1.27	1.37
2	B	237	TRP	NE1-CE2	-7.70	1.27	1.37

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	178	GLU	CB-CA-C	-6.43	97.53	110.40
2	B	110	SER	N-CA-CB	-5.66	102.01	110.50
1	A	70	ARG	CD-NE-CZ	-5.40	116.03	123.60
1	A	77	GLU	OE1-CD-OE2	-5.33	116.91	123.30

There are no chirality outliers.

5 of 36 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	20	ARG	Mainchain
1	A	30	GLN	Sidechain
1	A	41	GLN	Sidechain,Mainchain
1	A	49	PRO	Mainchain
1	A	60	ASN	Sidechain

## 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	645	0	589	16	1
2	B	1154	0	1085	35	2
3	I	438	0	418	17	0
4	A	1	0	0	0	0
4	B	6	0	0	0	0
4	I	3	0	0	0	1
All	All	2247	0	2092	55	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 55 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:72:ASN:HD21	2:B:153:ASP:HB3	1.55	0.70
2:B:245:ASN:HB3	2:B:246:PRO:HA	1.73	0.70
3:I:50:ASP:HA	3:I:53:ARG:HB3	1.73	0.70
2:B:145:GLU:HG3	2:B:150:GLU:HB3	1.75	0.68
3:I:20:ARG:HE	3:I:44:ASN:HD21	1.44	0.64

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:170:ASP:OD2	4:I:108:HOH:O[4_554]	2.13	0.07
1:A:95(B):SER:OG	2:B:243:THR:CG2[7_555]	2.18	0.02

### 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	78/80 (98%)	69 (88%)	6 (8%)	3 (4%)	3	4
2	B	150/152 (99%)	135 (90%)	10 (7%)	5 (3%)	4	5
3	I	55/58 (95%)	51 (93%)	4 (7%)	0	100	100
All	All	283/290 (98%)	255 (90%)	20 (7%)	8 (3%)	5	7

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	96	GLY
2	B	172	HIS
1	A	95(A)	LEU
2	B	165	ASN
1	A	26	SER

#### 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%)	57 (84%)	11 (16%)	2	4

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	129/129 (100%)	115 (89%)	14 (11%)	6	12
3	I	45/46 (98%)	41 (91%)	4 (9%)	9	19
All	All	242/243 (100%)	213 (88%)	29 (12%)	5	9

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	123	LEU
3	I	43	ASN
2	B	164	GLN
2	B	218	THR
2	B	159	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	217	HIS
3	I	43	ASN
3	I	44	ASN
1	A	72	ASN
1	A	76	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 [i](#)

There are no ligands in this entry.

## 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.