Full wwPDB X-ray Structure Validation Report (i)

May 13, 2020 – 05:05 am BST

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 Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *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 (i) symbol.

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

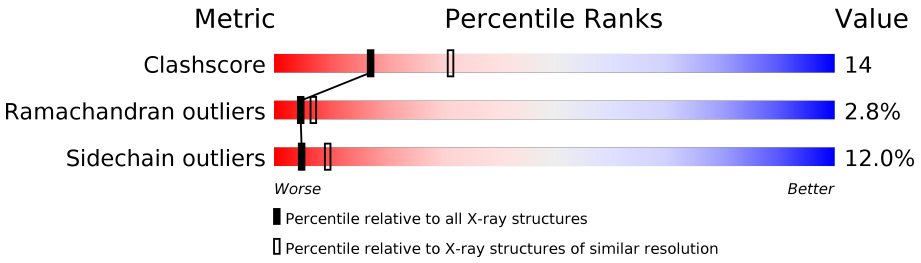
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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 on the lower bar indicate the fraction of residues that contain outliers for >=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 <=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

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
1	A	80	Total	C	N	O	S	60	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	91	0	0
			1154	726	182	235	11			

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
3	I	57	Total	C	N	O	S	41	0	1
			438	275	80	76	7			

- 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		

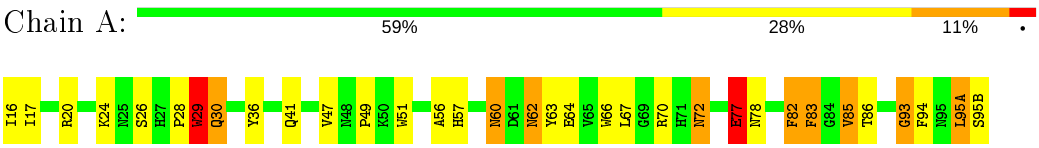
3 Residue-property plots (i)

These plots are drawn for all protein, RNA 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. 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.

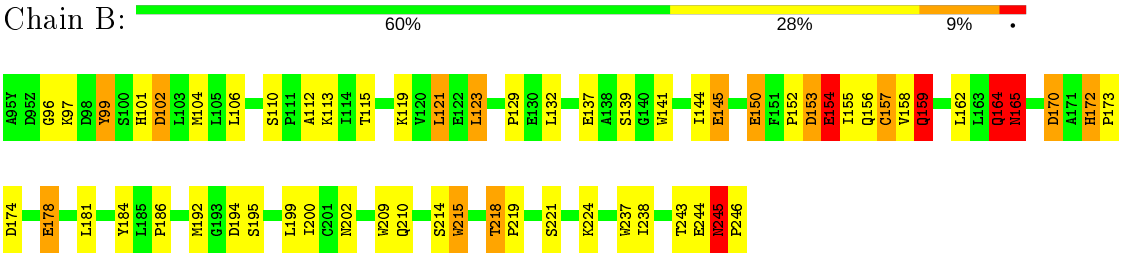
- Molecule 1: KALLIKREIN A

Chain A:



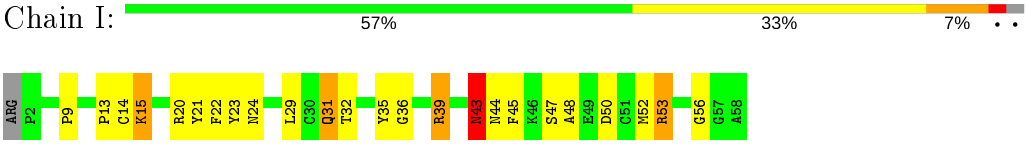
- Molecule 2: KALLIKREIN A

Chain B:



- Molecule 3: BOVINE PANCREATIC TRYPSIN INHIBITOR

Chain I:



4 Data and refinement statistics (i)

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, α, β, γ	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.	Xtriage
Total number of atoms	2247	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

All (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
2	B	141	TRP	NE1-CE2	-7.60	1.27	1.37
2	B	215	TRP	NE1-CE2	-7.45	1.27	1.37
1	A	36	TYR	CZ-OH	6.31	1.48	1.37
2	B	99	TYR	CZ-OH	5.71	1.47	1.37
3	I	35	TYR	CZ-OH	5.13	1.46	1.37

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
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.

All (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,Mainchain,Peptide
1	A	62	ASN	Sidechain
1	A	72	ASN	Sidechain
1	A	77	GLU	Sidechain
1	A	78	ASN	Sidechain
1	A	83	PHE	Mainchain
1	A	93	GLY	Mainchain
2	B	102	ASP	Sidechain
2	B	145	GLU	Sidechain
2	B	150	GLU	Sidechain
2	B	153	ASP	Sidechain
2	B	154	GLU	Sidechain
2	B	156	GLN	Sidechain
2	B	157	CYS	Mainchain
2	B	159	GLN	Sidechain
2	B	164	GLN	Sidechain
2	B	165	ASN	Sidechain
2	B	173	PRO	Mainchain
2	B	178	GLU	Sidechain
2	B	181	LEU	Mainchain
2	B	194	ASP	Sidechain
2	B	210	GLN	Sidechain
2	B	244	GLU	Sidechain
2	B	245	ASN	Sidechain
3	I	15	LYS	Mainchain
3	I	24	ASN	Sidechain
3	I	31	GLN	Sidechain
3	I	43	ASN	Sidechain
3	I	53	ARG	Mainchain

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.

All (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
2:B:145:GLU:CG	2:B:150:GLU:HB3	2.29	0.63
2:B:154:GLU:H	2:B:154:GLU:CD	2.01	0.62
2:B:221:SER:HB2	2:B:224:LYS:HG3	1.82	0.61
2:B:99:TYR:HE2	3:I:39:ARG:HG3	1.68	0.58
2:B:121:LEU:HD11	2:B:200:ILE:HD11	1.87	0.57
1:A:57:HIS:HD1	2:B:102:ASP:CG	2.09	0.57
1:A:83:PHE:CZ	2:B:112:ALA:HA	2.39	0.57
1:A:57:HIS:HE1	2:B:214:SER:HA	1.69	0.55
1:A:67:LEU:HD22	1:A:83:PHE:CE1	2.43	0.52
2:B:113:LYS:O	2:B:115:THR:HG23	2.10	0.52
3:I:21:TYR:CE2	3:I:48:ALA:HB2	2.45	0.52
3:I:23:TYR:H	3:I:43:ASN:ND2	2.08	0.52
3:I:20:ARG:O	3:I:32:THR:HA	2.09	0.52
2:B:144:ILE:HG23	2:B:152:PRO:HG3	1.91	0.51
3:I:13:PRO:HD2	3:I:39:ARG:HB2	1.93	0.51
3:I:9:PRO:HG3	3:I:22:PHE:CZ	2.46	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:21:TYR:HB2	3:I:45:PHE:HB2	1.92	0.50
3:I:22:PHE:CE2	3:I:31:GLN:HB2	2.46	0.50
2:B:139:SER:HB3	2:B:157:CYS:SG	2.52	0.50
2:B:184:TYR:CZ	2:B:186:PRO:HG2	2.47	0.49
3:I:21:TYR:HA	3:I:31:GLN:O	2.13	0.49
3:I:52:MET:O	3:I:56:GLY:HA3	2.13	0.49
2:B:145:GLU:HG3	2:B:150:GLU:CB	2.41	0.48
2:B:129:PRO:HB2	2:B:162:LEU:CD2	2.43	0.48
2:B:99:TYR:HB2	2:B:215:TRP:CD1	2.49	0.48
2:B:218:THR:HA	2:B:219:PRO:C	2.34	0.48
2:B:195:SER:OG	3:I:15:LYS:C	2.53	0.47
1:A:30:GLN:HG2	2:B:155:ILE:CD1	2.46	0.46
1:A:64:GLU:HB3	1:A:82:PHE:HE2	1.82	0.45
2:B:137:GLU:HB2	2:B:159:GLN:OE1	2.16	0.45
3:I:14:CYS:HB2	3:I:36:GLY:O	2.17	0.45
2:B:123:LEU:HD21	2:B:238:ILE:HG21	1.98	0.44
1:A:16:ILE:CG2	2:B:158:VAL:HB	2.47	0.44
2:B:132:LEU:HD13	2:B:164:GLN:HB3	1.99	0.44
1:A:56:ALA:HA	2:B:104:MET:HB2	2.00	0.43
1:A:28:PRO:HB2	2:B:119:LYS:H	1.84	0.43
1:A:67:LEU:HD22	1:A:83:PHE:HE1	1.84	0.43
1:A:63:TYR:HB2	1:A:85:VAL:HG13	2.01	0.43
1:A:29:TRP:CD1	2:B:121:LEU:HG	2.54	0.42
3:I:43:ASN:ND2	3:I:45:PHE:HE1	2.18	0.42
2:B:164:GLN:HB2	2:B:164:GLN:HE21	1.56	0.42
1:A:93:GLY:HA3	2:B:101:HIS:CD2	2.55	0.42
1:A:63:TYR:HB2	1:A:85:VAL:CG1	2.50	0.41
2:B:245:ASN:HB3	2:B:246:PRO:CA	2.48	0.41
1:A:28:PRO:HB2	2:B:119:LYS:HB2	2.01	0.41
3:I:43:ASN:HD22	3:I:43:ASN:HA	1.49	0.41
2:B:123:LEU:HD12	2:B:123:LEU:HA	1.91	0.41
3:I:9:PRO:HG3	3:I:22:PHE:CE1	2.56	0.41
2:B:184:TYR:CE2	2:B:186:PRO:HG2	2.56	0.40
2:B:145:GLU:HG2	2:B:150:GLU:HB3	2.02	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)
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

All (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
1	A	94	PHE
2	B	170	ASP
2	B	245	ASN

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
2	B	129/129 (100%)	115 (89%)	14 (11%)	6 12
3	I	45/46 (98%)	41 (91%)	4 (9%)	9 19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	242/243 (100%)	213 (88%)	29 (12%)	5 9

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	ILE
1	A	24	LYS
1	A	29	TRP
1	A	47	VAL
1	A	60	ASN
1	A	62	ASN
1	A	77	GLU
1	A	82	PHE
1	A	85	VAL
1	A	86	THR
1	A	95(A)	LEU
2	B	97	LYS
2	B	106	LEU
2	B	121	LEU
2	B	123	LEU
2	B	154	GLU
2	B	159	GLN
2	B	164	GLN
2	B	165	ASN
2	B	172	HIS
2	B	174	ASP
2	B	192	MET
2	B	199	LEU
2	B	202	ASN
2	B	218	THR
3	I	29	LEU
3	I	39	ARG
3	I	43	ASN
3	I	47	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	GLN
1	A	48	ASN
1	A	62	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	72	ASN
1	A	76	ASN
2	B	164	GLN
2	B	165	ASN
2	B	217	HIS
3	I	43	ASN
3	I	44	ASN

5.3.3 RNA

There are no RNA molecules 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

There are no ligands in this entry.

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.