



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 12, 2017 – 10:50 pm GMT

PDB ID : 1BZX
Title : THE CRYSTAL STRUCTURE OF ANIONIC SALMON TRYPSIN IN COM-
PLEX WITH BOVINE PANCREATIC TRYPSIN INHIBITOR
Authors : Helland, R.; Leiros, I.; Berglund, G.I.; Willassen, N.P.; Smalas, A.O.
Deposited on : 1998-11-05
Resolution : 2.10 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

<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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

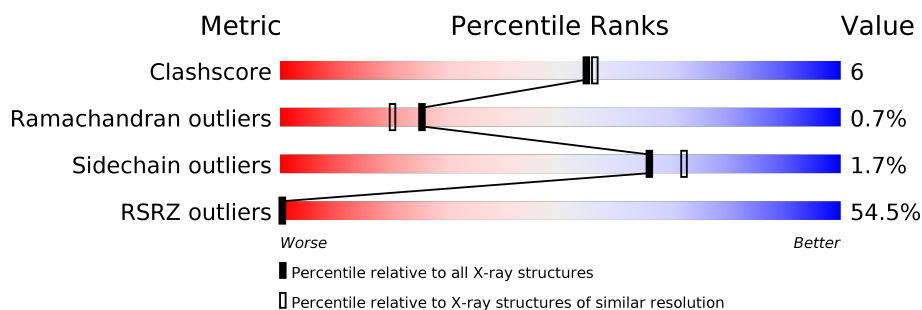
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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. 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	E	222	<div> <div>55%</div> <div>84%</div> <div>16%</div> </div>
2	I	58	<div> <div>47%</div> <div>81%</div> <div>19%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2219 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 PROTEIN (TRYPSIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	222	Total	C	N	O	S	58	3	0
			1675	1046	280	330	19			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	24	PRO	ALA	CONFLICT	UNP P35031
E	28	PRO	THR	CONFLICT	UNP P35031

- Molecule 2 is a protein called PROTEIN (BOVINE PANCREATIC TRYPSIN INHIBITOR).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	58	Total	C	N	O	S	19	1	0
			458	287	84	79	8			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	1	Total	Ca	0	0
			1	1		

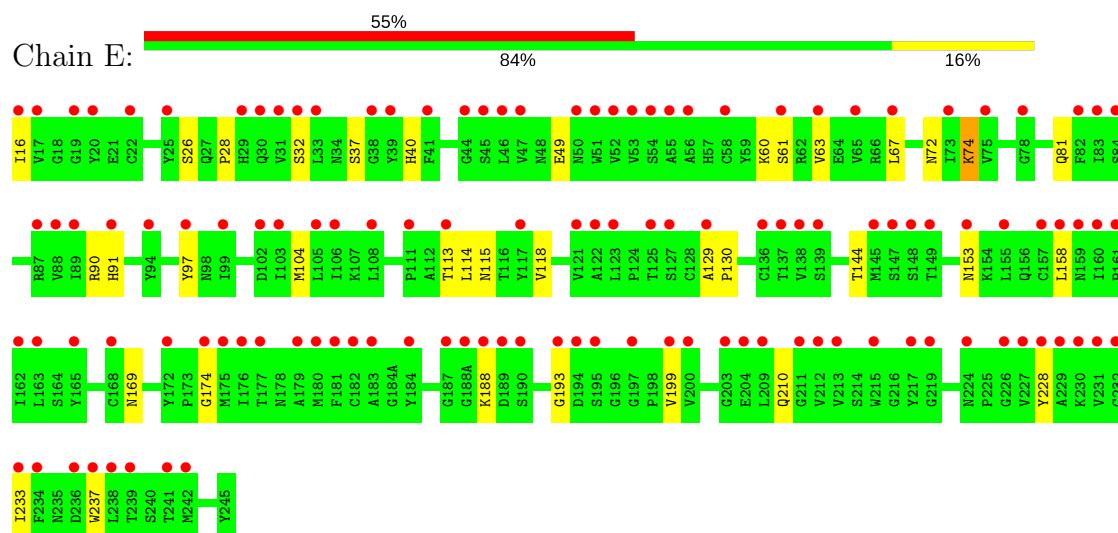
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	E	66	Total	O	0	0
			66	66		
4	I	19	Total	O	0	0
			19	19		

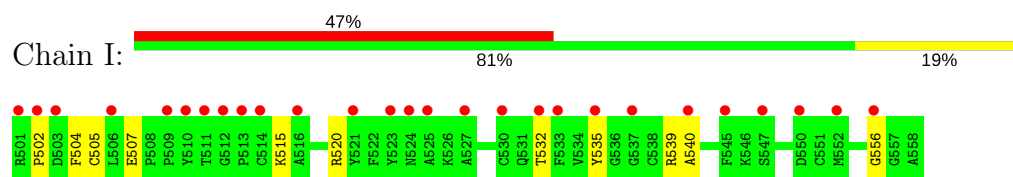
3 Residue-property plots [i](#)

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 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: PROTEIN (TRYPSIN)



• Molecule 2: PROTEIN (BOVINE PANCREATIC TRYPSIN INHIBITOR)



4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	84.12Å 84.12Å 222.15Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	8.00 – 2.10 27.33 – 2.13	Depositor EDS
% Data completeness (in resolution range)	97.3 (8.00-2.10) 95.4 (27.33-2.13)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.59 (at 2.14Å)	Xtriage
Refinement program	X-PLOR 3.8	Depositor
R, R_{free}	0.206 , 0.238 0.440 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	38.3	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 94.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.73	EDS
Total number of atoms	2219	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	E	0.49	0/1728	0.70	0/2351
2	I	0.59	1/473 (0.2%)	0.71	0/632
All	All	0.52	1/2201 (0.0%)	0.71	0/2983

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	515	LYS	C-N	-6.07	1.20	1.34

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	E	1675	0	1582	20	0
2	I	458	0	439	5	0
3	E	1	0	0	0	0
4	E	66	0	0	1	0
4	I	19	0	0	0	0
All	All	2219	0	2021	24	0

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 6.

All (24) 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:E:74:LYS:HE3	1:E:74:LYS:HA	1.69	0.74
2:I:504:PHE:O	2:I:507:GLU:HG2	2.00	0.62
2:I:502:PRO:HG2	2:I:505:CYS:SG	2.40	0.60
1:E:169:ASN:HD21	1:E:174:GLY:H	1.53	0.55
2:I:535:TYR:CZ	2:I:540:ALA:HB2	2.42	0.55
1:E:114:LEU:HA	1:E:118:VAL:O	2.11	0.49
1:E:97:TYR:O	2:I:539:ARG:NH2	2.46	0.49
1:E:90:ARG:HA	1:E:104:MET:HB2	1.95	0.49
1:E:210:GLN:NE2	4:E:781:HOH:O	2.46	0.49
1:E:26:SER:C	1:E:28:PRO:HD3	2.32	0.48
1:E:32:SER:OG	1:E:40:HIS:HD2	1.97	0.47
1:E:113:THR:O	1:E:115:ASN:ND2	2.48	0.47
1:E:60:LYS:HB2	1:E:63[B]:VAL:CG1	2.47	0.45
1:E:129:ALA:HA	1:E:130:PRO:HD3	1.67	0.45
1:E:74:LYS:HE3	1:E:74:LYS:CA	2.43	0.44
1:E:67:LEU:HD12	1:E:81:GLN:O	2.18	0.44
1:E:91:HIS:HA	1:E:237:TRP:CZ2	2.53	0.43
2:I:520:ARG:O	2:I:532:THR:HA	2.18	0.43
1:E:158:LEU:HD11	1:E:188:LYS:HB3	2.01	0.43
1:E:16:ILE:O	1:E:144:THR:HA	2.19	0.42
1:E:72:ASN:HA	1:E:153:ASN:O	2.19	0.42
1:E:40:HIS:HE1	1:E:193:GLY:O	2.02	0.42
1:E:199:VAL:HG21	1:E:228:TYR:CD2	2.54	0.42
1:E:169:ASN:ND2	1:E:174:GLY:H	2.16	0.40

There are no symmetry-related clashes.

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	E	223/222 (100%)	215 (96%)	7 (3%)	1 (0%)	38	35
2	I	57/58 (98%)	56 (98%)	0	1 (2%)	10	4
All	All	280/280 (100%)	271 (97%)	7 (2%)	2 (1%)	25	20

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	49	GLU
2	I	556	GLY

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	E	190/187 (102%)	186 (98%)	4 (2%)	59	64
2	I	47/46 (102%)	47 (100%)	0	100	100
All	All	237/233 (102%)	233 (98%)	4 (2%)	66	72

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	37	SER
1	E	61	SER
1	E	74	LYS
1	E	233	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	30	GLN
1	E	40	HIS
1	E	169	ASN
1	E	192	GLN
1	E	202	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	210	GLN
1	E	224	ASN
2	I	531	GLN

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 ⓘ

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, 95th 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 ²)	Q<0.9
1	E	219/222 (98%)	2.34	123 (56%) 0 0	18, 35, 57, 75	8 (3%)
2	I	56/58 (96%)	2.15	27 (48%) 0 0	17, 32, 57, 67	3 (5%)
All	All	275/280 (98%)	2.30	150 (54%) 0 0	17, 34, 57, 75	11 (4%)

All (150) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	22	CYS	6.4
1	E	242	MET	5.8
1	E	172	TYR	5.8
1	E	229	ALA	5.1
1	E	203	GLY	5.1
2	I	525	ALA	5.1
1	E	117	TYR	5.0
1	E	165	TYR	5.0
1	E	53	VAL	4.8
2	I	530	CYS	4.6
2	I	540	ALA	4.6
2	I	514	CYS	4.5
2	I	552[A]	MET	4.4
1	E	153	ASN	4.4
1	E	193	GLY	4.4
1	E	228	TYR	4.4
1	E	89	ILE	4.3
1	E	55	ALA	4.3
1	E	197	GLY	4.3
1	E	181	PHE	4.3
2	I	523	TYR	4.3
1	E	61	SER	4.2
1	E	215	TRP	4.1
1	E	175[A]	MET	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	16	ILE	4.0
1	E	162	ILE	4.0
1	E	239	THR	4.0
1	E	121	VAL	4.0
1	E	195	SER	4.0
1	E	148	SER	4.0
1	E	187	GLY	3.9
1	E	51	TRP	3.9
1	E	237	TRP	3.9
1	E	234	PHE	3.9
1	E	94	TYR	3.9
1	E	217	TYR	3.9
1	E	67	LEU	3.8
1	E	84	SER	3.8
1	E	213	VAL	3.8
1	E	136	CYS	3.7
1	E	212	VAL	3.7
1	E	125	THR	3.7
1	E	183	ALA	3.7
1	E	219	GLY	3.7
1	E	209	LEU	3.7
2	I	509	PRO	3.7
1	E	97	TYR	3.6
2	I	506	LEU	3.6
2	I	502	PRO	3.6
1	E	163	LEU	3.6
1	E	155	LEU	3.5
1	E	113	THR	3.5
1	E	47	VAL	3.4
2	I	527	ALA	3.4
1	E	20	TYR	3.4
1	E	190	SER	3.4
1	E	139	SER	3.4
1	E	63[A]	VAL	3.3
1	E	231	VAL	3.3
1	E	179	ALA	3.3
1	E	108	LEU	3.3
1	E	184	TYR	3.3
1	E	39	TYR	3.2
2	I	545	PHE	3.2
1	E	147	SER	3.1
1	E	204	GLU	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	33	LEU	3.1
1	E	182	CYS	3.1
1	E	226	GLY	3.1
1	E	25	TYR	3.1
1	E	188	LYS	3.1
2	I	535	TYR	3.1
1	E	19	GLY	3.1
2	I	550	ASP	3.1
1	E	29	HIS	3.0
1	E	194	ASP	3.0
1	E	211	GLY	3.0
2	I	524	ASN	3.0
1	E	158	LEU	3.0
1	E	41	PHE	3.0
2	I	511	THR	3.0
1	E	188(A)	GLY	3.0
2	I	537	GLY	3.0
1	E	111	PRO	2.9
1	E	137	THR	2.9
1	E	138	VAL	2.9
1	E	52	VAL	2.9
1	E	103	ILE	2.9
1	E	78	GLY	2.9
1	E	199	VAL	2.8
2	I	501	ARG	2.8
1	E	105	LEU	2.8
1	E	233	ILE	2.8
1	E	232	CYS	2.8
2	I	556	GLY	2.8
2	I	516	ALA	2.7
1	E	91	HIS	2.7
2	I	510	TYR	2.7
1	E	56	ALA	2.7
1	E	160	ILE	2.7
1	E	38	GLY	2.7
1	E	241	THR	2.7
1	E	180	MET	2.6
1	E	32	SER	2.6
1	E	123	LEU	2.5
1	E	46	LEU	2.5
1	E	145	MET	2.5
1	E	174	GLY	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	161	PRO	2.4
2	I	503	ASP	2.4
1	E	87	ARG	2.4
1	E	176	ILE	2.4
1	E	17	VAL	2.4
1	E	99	ILE	2.4
1	E	149	THR	2.4
2	I	533	PHE	2.4
1	E	200	VAL	2.3
1	E	54	SER	2.3
1	E	50	ASN	2.3
1	E	31	VAL	2.3
1	E	58	CYS	2.3
1	E	168	CYS	2.3
1	E	127	SER	2.3
2	I	512	GLY	2.3
1	E	122	ALA	2.2
1	E	227	VAL	2.2
1	E	65	VAL	2.2
1	E	75	VAL	2.2
1	E	82	PHE	2.2
2	I	547	SER	2.2
1	E	44	GLY	2.2
1	E	83	ILE	2.2
1	E	189	ASP	2.2
2	I	532	THR	2.2
1	E	30	GLN	2.1
1	E	73	ILE	2.1
1	E	102	ASP	2.1
1	E	224	ASN	2.1
1	E	45	SER	2.1
1	E	129	ALA	2.1
1	E	230	LYS	2.1
1	E	157	CYS	2.1
1	E	238	LEU	2.1
1	E	88	VAL	2.1
1	E	177	THR	2.1
2	I	521	TYR	2.1
1	E	106	ILE	2.1
2	I	513	PRO	2.1
1	E	236	ASP	2.0
1	E	159	ASN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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, 95th 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(\AA^2)	Q<0.9
3	CA	E	700	1/1	0.47	0.14	-3.72	52,52,52,52	0

6.5 Other polymers [i](#)

There are no such residues in this entry.