WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 2, 2023 – 04:46 AM EDT

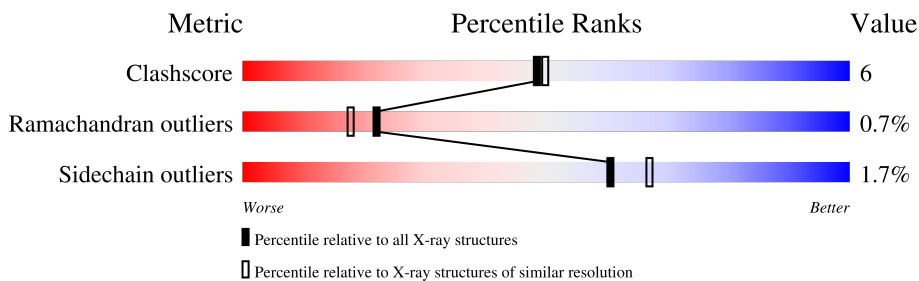|		
PDB ID	:	1BZX
Title	:	THE CRYSTAL STRUCTURE OF ANIONIC SALMON TRYPSIN IN COMPLEX WITH BOVINE PANCREATIC TRYPSIN INHIBITOR
Authors	:	Helland, R.; Leiros, I.; Berglund, G.I.; Willassen, N.P.; Smalas, A.O.
Deposited on	:	1998-11-05
Resolution	:	2.10 Å(reported)

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	Percentile Ranks	Value
Clashscore		6
Ramachandran outliers		0.7%
Sidechain outliers		1.7%

Worse   *Better*

▊ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (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 of 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%

Mol	Chain	Length	Quality of chain
1	E	222	84% 16%
2	I	58	81% 19%

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 ⓘ

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.

• Molecule 1: PROTEIN (TRYPSIN)

Chain E:
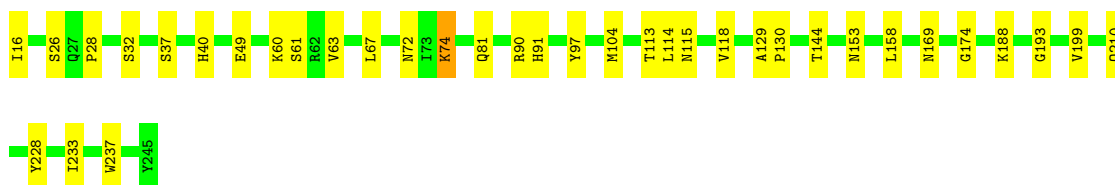
• Molecule 2: PROTEIN (BOVINE PANCREATIC TRYPSIN INHIBITOR)

Chain I:

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
$< I/\sigma(I) >$ [1]	1.59 (at 2.14Å)	Xtriage
Refinement program	X-PLOR 3.8	Depositor
R, R_{free}	0.206 , 0.238 0.462 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
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]	$< |L| > = 0.48, < L^2 > = 0.32$	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.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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	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 (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	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.

The worst 5 of 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

There are no symmetry-related clashes.

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	E	223/222 (100%)	215 (96%)	7 (3%)	1 (0%)	34	32
2	I	57/58 (98%)	56 (98%)	0	1 (2%)	8	4
All	All	280/280 (100%)	271 (97%)	7 (2%)	2 (1%)	22	18

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%)	53	59
2	I	47/46 (102%)	47 (100%)	0	100	100
All	All	237/233 (102%)	233 (98%)	4 (2%)	60	67

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	531	GLN
1	E	224	ASN
1	E	202	ASN
1	E	192	GLN
1	E	210	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 monosaccharides 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 ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	I	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	I	515:LYS	C	516:ALA	N	1.20

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.