wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 07:41 PM GMT

PDB ID	:	1GT9
Title	:	HIGH RESOLUTION CRYSTAL STRUCTURE OF A THERMOSTABLE SERINE-CARBOXYL TYPE PROTEINASE, KUMAMOLISIN (KSCP)
Authors	:	Comellas-Bigler, M.; Fuentes-Prior, P.; Maskos, K.; Huber, R.; Oyama, H.; Uchida, K.; Dunn, B.M.; Oda, K.; Bode, W.
Deposited on	:	2002-01-14
Resolution	:	1.38 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
A user guide is available at
http://wwpdb.org/validation/2016/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:

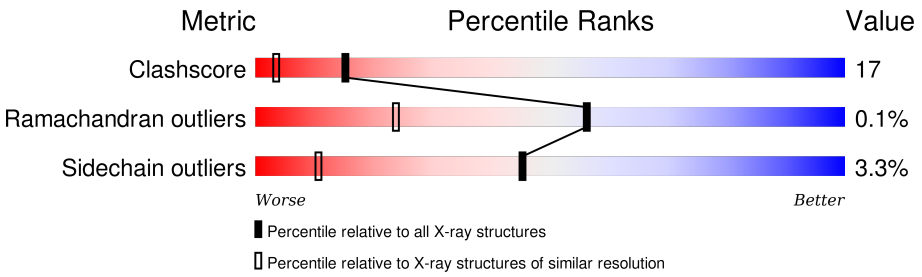
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 1.38 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	102246	2042 (1.40-1.36)
Ramachandran outliers	100387	1993 (1.40-1.36)
Sidechain outliers	100360	1992 (1.40-1.36)

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 <=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	1	357	78% 20% •
1	2	357	82% 16% ••

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5928 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.

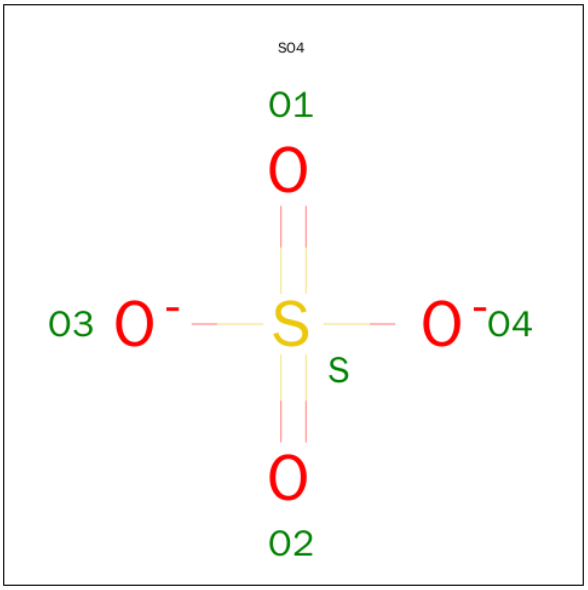
- Molecule 1 is a protein called KUMAMOLYSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	1	357	2663	1674	453	531	5	20	14	0
1	2	357	2687	1689	455	538	5	25	17	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	2	1	1	1	0	0
2	1	1	1	1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
3	1	1	5	4	1	0	0
			Total	O	S		
3	2	1	5	4	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	1	285	285	285	0	0
			Total	O		
4	2	281	281	281	0	0

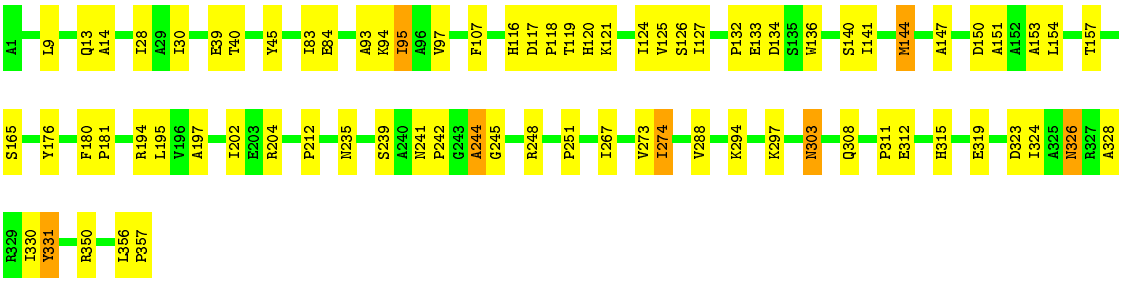
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

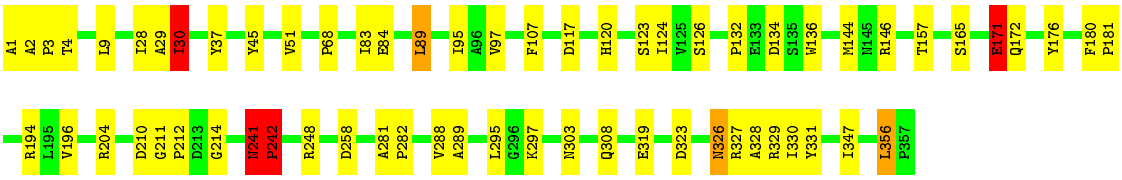
- Molecule 1: KUMAMOLYSIN

Chain 1:



- Molecule 1: KUMAMOLYSIN

Chain 2:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	54.86Å 78.25Å 73.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.52 – 1.38	Depositor
% Data completeness (in resolution range)	95.4 (19.52-1.38)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.194 , 0.208	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5928	wwPDB-VP
Average B, all atoms (Å^2)	25.0	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: CA, SO4

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	1	0.89	10/2729 (0.4%)	0.95	12/3747 (0.3%)
1	2	1.58	18/2754 (0.7%)	2.66	19/3784 (0.5%)
All	All	1.29	28/5483 (0.5%)	2.00	31/7531 (0.4%)

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	1	0	4
1	2	1	4
All	All	1	8

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	2	171[A]	GLU	CD-OE1	36.98	1.66	1.25
1	2	171[B]	GLU	CD-OE1	36.98	1.66	1.25
1	2	242[A]	PRO	N-CD	22.17	1.78	1.47
1	2	242[B]	PRO	N-CD	22.17	1.78	1.47
1	2	242[A]	PRO	CA-CB	18.34	1.90	1.53

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	2	146[A]	ARG	NH1-CZ-NH2	-104.76	4.17	119.40
1	2	146[B]	ARG	NH1-CZ-NH2	-104.76	4.17	119.40
1	2	242[A]	PRO	N-CA-CB	-18.05	81.64	103.30
1	2	242[B]	PRO	N-CA-CB	-18.05	81.64	103.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	2	242[A]	PRO	CA-N-CD	-16.25	88.76	111.50

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	2	242[B]	PRO	CA

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1	165[B]	SER	Mainchain
1	1	331	TYR	Mainchain
1	1	95[A]	ILE	Mainchain
1	1	95[B]	ILE	Mainchain
1	2	30[A]	ILE	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	1	2663	0	2559	90	83
1	2	2687	0	2581	92	54
2	1	1	0	0	0	0
2	2	1	0	0	0	0
3	1	5	0	0	0	0
3	2	5	0	0	0	0
4	1	285	0	0	32	4
4	2	281	0	0	14	7
All	All	5928	0	5140	182	87

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 17.

The worst 5 of 182 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)
1:2:241[A]:ASN:CA	1:2:241[A]:ASN:CB	1.75	1.64
1:1:95[A]:ILE:CG1	1:1:95[A]:ILE:CD1	1.76	1.59
1:2:242[A]:PRO:N	1:2:242[A]:PRO:CA	1.68	1.56
1:2:242[B]:PRO:CD	1:2:242[B]:PRO:N	1.75	1.49
1:2:144[B]:MET:SD	1:2:144[B]:MET:CE	2.03	1.47

The worst 5 of 87 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)
1:1:39:GLU:CA	1:1:357:PRO:C[2_556]	0.43	1.77
1:1:153:ALA:O	1:2:329:ARG:NH2[1_556]	0.46	1.74
1:1:154:LEU:N	1:2:329:ARG:NH1[1_556]	0.48	1.72
1:1:150:ASP:OD1	1:2:327:ARG:CG[1_556]	0.51	1.69
1:1:40:THR:N	1:1:357:PRO:CB[2_556]	0.53	1.67

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	1	369/357 (103%)	364 (99%)	5 (1%)	0	100	100
1	2	371/357 (104%)	362 (98%)	7 (2%)	2 (0%)	34	10
All	All	740/714 (104%)	726 (98%)	12 (2%)	2 (0%)	56	19

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	2	242[A]	PRO
1	2	242[B]	PRO

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	1	271/260 (104%)	263 (97%)	8 (3%)	48 12
1	2	274/260 (105%)	264 (96%)	10 (4%)	42 9
All	All	545/520 (105%)	527 (97%)	18 (3%)	45 11

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	2	89	LEU
1	2	107	PHE
1	2	248	ARG
1	1	308	GLN
1	1	326	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	2	24	GLN
1	2	67	GLN
1	2	303	ASN
1	2	13	GLN
1	2	315	HIS

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)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

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	SO4	1	1359	-	4,4,4	0.51	0	6,6,6	0.18	0
3	SO4	2	1359	-	4,4,4	0.42	0	6,6,6	0.21	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	SO4	1	1359	-	-	0/0/0/0	0/0/0/0
3	SO4	2	1359	-	-	0/0/0/0	0/0/0/0

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 (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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.