International Journal of Current Advanced Research
ISSN: O: 2319-6475, ISSN: P: 2319 – 6505, Impact Factor: SJIF: 5.438
Available Online at www.journalijcar.org
Volume 6; Issue 2; February 2017; Page No. 2131-2138

Research Article

A COMPARATIVE ANALYSIS OF THERMOPHILIC AND PSYCHROPHILIC METALLOPEPTIDASES: INVOLVEMENT OF BIOINFORMATIC APPROACH

Anuprabha[1]., AbhigyanNath[2] and Radha Chaube[3]*

[1,2]Bioinformatic Division, MahilaMahavidyalya, Banaras Hindu University, Varanasi-221005, Uttar Pradesh, India
[3]Department of Zoology, Institute of Science, Banaras Hindu University, Varanasi-221005, Uttar Pradesh, India

A R T I C L E I N F O

Article History:

Received 9th November, 2016
Received in revised form 12thDecember, 2016
Accepted 5th January, 2017
Published online 28th February, 2017

Key words:

Thermophilic, Psychrophilic, Metallopeptidases, amino acid residues, analysis

A B S T R A C T

The thermolysin family of enzymes is classified as the M4 family of metallopeptidases. M4 family comprises numerous zinc-dependent metallopeptidases that hydrolyze peptide bonds. Zinc containing peptidases are widely distributed in nature and have important roles in many physiological processes. M4 family comprises numerous zinc-dependent metallopeptidases that hydrolyse peptide bonds. Metallopeptidases were studied for the reason of their great relevance to biology, medicine and biotechnology. In the present study, detailed comparative analyses of both thermophilic and psychrophilic metallopeptidases were performed. The comparative analysis of both thermophilic and psychrophilic metallopeptidases was performed by taking sequence parameters such as amino acid composition, amino acid property group composition, physico-chemical properties, secondary structure content. From comparison between the two groups, it was found that Gln, Asn, Ser, Thr, and His are significantly lower in thermophilic metallopeptidase. Positively charged residues (Lys, Arg and Glu) are more significant in thermophilic metallopeptidases than in psychrophilic metallopeptidases. By Scan Prosite tool it was found that VSAHEVSHGF and VIGHETHAV or VVGHETHV amino acid pattern were common in all psychrophilic and thermophilic metallopeptidases respectively. By studying the secondary structure, it was found that the content of helices is more in thermophilic metallopeptidase and strand is overrepresented in psychrophilic metalopeptidases. These findings provide the required knowledge that these factors are useful for differentiating the thermophilic and psychrophilic metallopeptidases and are helpful for comparative analysis of thermophilic and psychrophilic metallopeptidases.

INTRODUCTION

A protease also termed peptidase or proteinase, is any enzyme that conducts proteolysis, that is begins proteincatabolism by hydrolysis of the peptide bonds that link amino acids together in the polypeptide chain forming the protein. Proteases, also known as proteinases or proteolytic enzymes, are a large group of enzymes. Proteases belong to the class of enzymes known as hydrolases, which catalyzes the reaction of hydrolysis of various bonds with the participation of a water molecule. Proteases are currently classified into six broad groups: out of which, one is

Metallopeptidase

A Metallopeptidase is any protease enzyme whose catalytic mechanism involves a metal. Metallpeptidses are the most diverse of the four main types of protease, with more than 50 families identified to date. In these enzymes, a divalent cation, usually zinc, activates the water molecule. The metal ion is held in place by amino acid ligands, usually three in number. The known metal ligands are His, Glu, Asp or Lys and at least one other residue is required for catalysis, which may play an electrophilic role. Of the known around half contain an HEXXH motif, which has been shown in crystallographic studies to form part of the metal-binding site[2]. The HEXXH motif is relatively common, but can be more stringently defined for metalloproteases as 'abXHEbbHbc', where 'a' is most often valine or threonine and forms part of the S1' subsite in thermolysin and nephrilysin, 'b' is an uncharged residue, and 'c' a hydrophobic residue. Peptidases, their substrates and inhibitors are of great relevance to biology, medicine and biotechnology. The MEROPS database (http://merops.sanger.ac.uk) aims to fulfill the need for an integrated source ofinformation about these. The organizational principle of the database is a hierarchical classification in which homologous sets of the proteins of interest are grouped in families and the homologous families are grouped in clans.

The thermolysin family of enzymes is classified as the M4 family of metallopeptidases (http://merops.sanger.ac.uk/). The thermolysin family (M4) is a family of zinc metalloproteases in the MA(E) subclan of the MA clan (6) . The peptidases in this family are produced by bacteria from various habitats like cold and thermal adapted enzyme, cold adapted enzyme are produced by organism thriving in permanently cold

habitat,are characterized by high catalytic efficiencies at low temperature and low stability at high temperature however opposite was true for thermal adapted enzyme (4). Thermolysin from *Bacillus thermoproteolyticus* (sub-family M04.001), the M4 family prototype, was one of the first metallopeptidases to be sequenced and also one of the first metallopeptidases with known 3D structure.

A psychrophilic organism is grouped into sub family *Vibrio lysin* and Thermophilic organism are grouped into sub family *Thermolysin* in M4 Family of metallopeptidase of MEROPS database. M4 Family is also known as Thermolysin family.

In this study the 20 thermophile and psychrophile protein sequences having peptidase unit, was considered for the comparative analysis, the final dataset, listed in Table 1, contained 10 thermophile and 10 psychrophilemetallopeptidase protein, the structure in Table 1 are divided into thermolysin and vibriolysin sub-family according to M4 family within the MEROPS database. The parameters studied are phylogeny analysis, amino acid property group distribution, secondary structure distribution, composition of amino acid and its distribution in hydrophobic, charged residue. The Sequence Logo for their amino acids was also studied.

MATERIALS AND METHOD

There are various computational tools and databases

- MEROPS Database- Usedfor classification of metallopeptidases
- Uniprot database- Used for sequence of amino acid
- MEGA5 Software – Used for phylogenetic tree construction
- ExPASyprotparam tool - Used for analysing the amino acid composition
- Copid Tool- Usedfor analysing distribution of amino acid property group
- Prosite tool – For consensus pattern in Psychrophile & Thermophilic metallopeptidases.
- MLRC Tool – For Secondary structure distribution in Thermophile & Psychrophile
- MATLAB Software- For comparing group mean of two data.

RESULTS

Creation of Dataset

A dataset of 20 proteins of thermophilic and psychrophilic was considered for the comparative analysis of metallopeptidases. Among them, 10 are thermophiles and 10 are psychrophile that are constructed from Uniprot (http://www.uniprot) and Merops (http://www.sanger.ac.uk) Databases.

Thermophilic (E.C 3.4.24.27) and Psychrophilic (E.C 3.4.24.25) metallopeptidases belong to M4 (thermolysin) Family of peptidases. On the basis of M4-family, Peptidase units, Length of sequence and organism name thermophilic and psychrophilic metallopeptidases protein were classified as tabulated in (Table 1).

Phylogenetic tree relationship

The tree was constructed along Thermophilic & Psychrophilic metallopeptidases (in dataset) by neighbor-joining (NJ) method. By constructing phylogenetic tree of TE and PS metallopeptidases protein we found that both thermophile and psychrophile have same ancestor but they bifurcate into two branches, and the two branches also bifurcate into several branches.

The thermophile and psychrophile both have same ancestor, but they show different extremophile adapatation, the thermophile thrives at high temperature and psychrophile thrives at low temperature, for this we did amino acid composition and property analysis, secondary structure analysis.

Table 1 Table of metallopeptidase protein with sequence and its organism details

S.N0.	MEROPS Accession	UniprotAcc.No. /Prot_ID	Sources (Bacterial Species)	Peptidase Unit	Sub-family name
1.	MER03181	O32309	*Bacillus thuringiensis*	250-566	Thermolysin
2.	MER01030	P05806	*Bacillus cereus*	250-566	Thermolysin
3.	MER01354	Q48857	*Lactobacillus sp.*	250-566	Thermolysin
4.	MER01353	Q43880	*Alicyclobacillusacidocaldarius*	228-546	Thermolysin
5.	MER01927	Q59223	*Bacillus sp.*	228-546	Thermolysin
6.	MER01034a	P23384	*Bacillus caldolyticus*	226-544	Thermolysin
7.	MER01025	P06874	*Geobacillusstearothermophilus*	230-548	Thermolysin
8.	MER01026	P00800	*Bacillus thermoproteolyticus*	233-548	Thermolysin
9.	MER01028	P43263	*Brevibacillusbrevis*	224-527	Thermolysin
10.	MER01033	P29148	*Paenibacillus polymyxa*	P29148	Thermolysin
11.	MER19098	Q8RTZ9	*Pseudoalteromonas sp.*	210-502	Vibriolysin
12.	MER19099	Q8VLQ5	*Alteromonas sp.*	189-500	Vibriolysin
13.	MER12255	Q9KH34	*Antartic bacterium*	210-503	Vibriolysin
14.	MER01043	Q00971	*Vibrio proteolyticus*	202-499	Vibriolysin
15.	MER01044	P43147	*Vibrio angillarum*	205-502	Vibriolysin
16.	MER03353b	O06694	*Vibrio vulnificus*	202-499	Vibriolysin
17.	MER19097	Q8RR31	*Vibrio fluvialis*	206-503	Vibriolysin
18.	MER01041	P24153	*Vibrio cholerae*	202-499	Vibriolysin
19.	MER50231	-	*Shewanelladenitrificans*	181-492	Vibriolysin
20.	MER49928	-	*Shewanellaamazonensis*	153-482	Vibriolysin

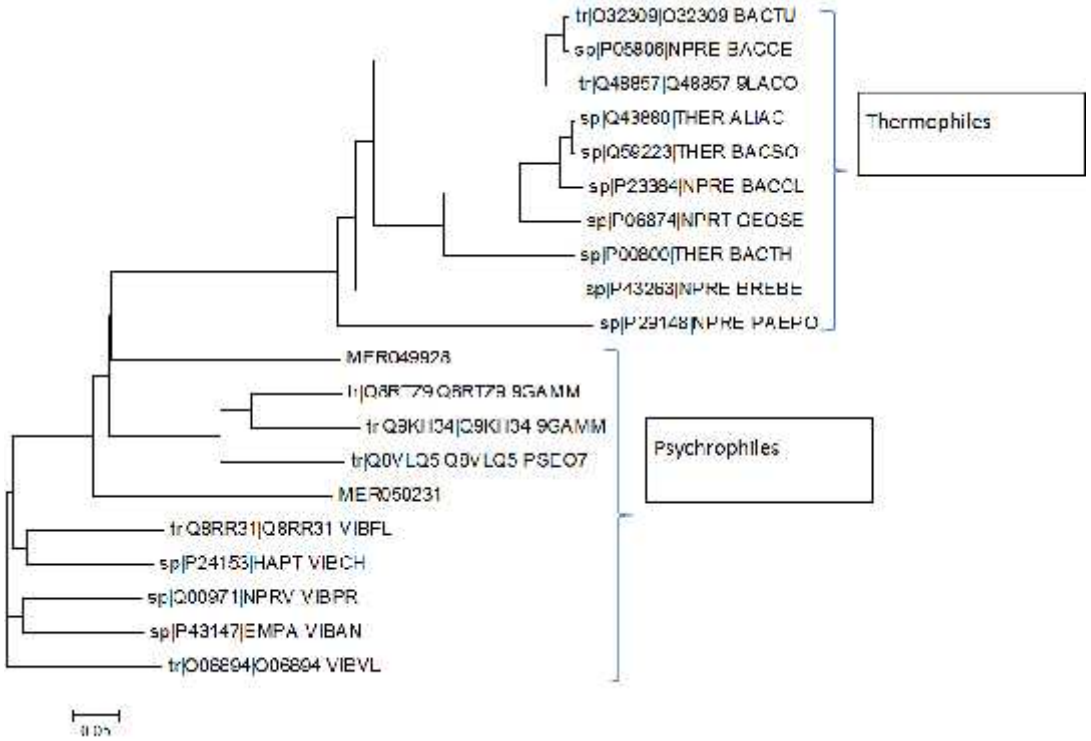


Table2 Table for amino acid distribution along thermophilic & psychrophilic metallopeptidase

Amino Acid(%)	Accession Number										
	O32309	**P05806**	**Q48857**	**Q43880**	**Q59223**	**P23384**	**P06874**	**P00800**	**P43263**	**P29148**	**Avg.**
Ala(A)	9.90%	10.10%	10.60%	9.30%	9.30%	9.70%	9.10%	8.40%	8.30%	9.80%	9.45%
Arg(R)	1.60%	1.60%	1.60%	4.60%	4.60%	4.60%	5.10%	2.20%	3.60%	2.10%	3.16%
Asn(N)	7.10%	6.70%	6.20%	5.10%	4.90%	5.30%	5.50%	6.60%	5.90%	8.50%	6.18%
Asp(D)	7.20%	7.20%	6.90%	6.20%	6.40%	6.10%	6.00%	6.80%	8.20%	5.70%	6.67%
Cys(C)	0.00%	0.00%	0.00%	0.20%	0.20%	0.60%	0.20%	0.00%	0.00%	1.10%	0.23%
Gln(Q)	3.50%	3.50%	3.20%	4.20%	4.20%	4.20%	4.40%	4.70%	3.80%	4.00%	3.97%
Glu(E)	4.10%	4.10%	4.60%	4.60%	4.60%	4.20%	4.20%	4.00%	5.10%	3.30%	4.28%
Gly(G)	9.50%	9.50%	9.40%	11.50%	11.40%	11.20%	12.20%	9.50%	9.10%	10.10%	10.34%
His(H)	1.60%	1.60%	1.60%	1.80%	1.80%	1.80%	2.20%	2.00%	3.40%	1.40%	1.92%
Ile(I)	3.70%	3.40%	2.80%	4.20%	4.20%	4.00%	4.60%	4.60%	4.00%	3.30%	3.88%
Leu(L)	5.80%	5.80%	6.40%	6.80%	6.80%	7.20%	6.90%	7.30%	6.30%	7.10%	6.64%
Lys(K)	8.10%	8.30%	8.70%	4.20%	4.20%	4.00%	3.30%	6.60%	5.90%	3.60%	5.69%
Met(M)	0.90%	0.90%	0.70%	2.20%	2.20%	2.40%	1.50%	1.10%	1.90%	2.20%	1.60%
Phe(F)	3.20%	3.20%	2.70%	3.10%	3.10%	3.10%	2.90%	4.20%	3.20%	3.80%	3.25%
Pro(P)	2.80%	2.80%	3.00%	3.70%	3.70%	3.30%	3.30%	3.10%	3.00%	3.00%	3.17%
Ser(S)	7.60%	7.60%	7.60%	6.20%	6.20%	6.60%	5.80%	6.80%	5.70%	10.10%	7.02%
Thr(T)	8.10%	8.30%	8.00%	5.90%	5.90%	5.70%	6.90%	7.30%	7.60%	8.60%	7.23%
Trp(W)	0.70%	0.70%	0.70%	1.50%	1.50%	1.50%	1.10%	0.90%	1.30%	1.40%	1.13%
Tyr(Y)	6.40%	6.40%	7.10%	7.10%	7.10%	6.80%	6.90%	6.60%	6.60%	4.90%	6.59%
Val(V)	8.10%	8.30%	8.50%	7.50%	7.70%	7.50%	7.80%	7.50%	7.00%	6.00%	7.59%
Ala(A)	9.80%	9.20%	8.40%	8.70%	8.20%	10.20%	9.00%	8.40%	11.10%	12.40%	9.45%
Arg(R)	2.10%	1.90%	1.50%	3.00%	2.30%	3.60%	2.60%	3.30%	2.70%	3.30%	2.63%
Asn(N)	8.50%	7.70%	8.00%	7.10%	6.50%	7.40%	5.70%	6.40%	6.80%	6.60%	7.07%
Asp(D)	5.70%	5.80%	5.80%	6.20%	5.40%	5.40%	6.10%	5.10%	6.70%	5.20%	5.74%
Cys(C)	1.10%	1.20%	1.10%	1.00%	1.00%	1.20%	1.00%	1.00%	1.00%	1.00%	1.06%
Gln(Q)	4.00%	4.70%	3.20%	3.80%	6.10%	3.80%	4.90%	5.40%	3.50%	2.90%	4.23%
Glu(E)	3.30%	3.70%	3.60%	4.30%	3.30%	3.50%	2.80%	3.60%	2.80%	4.30%	3.52%
Gly(G)	10.10%	9.10%	10.20%	9.00%	8.50%	8.90%	10.20%	10.50%	9.00%	11.50%	9.70%
His(H)	1.40%	1.80%	1.50%	2.00%	1.80%	2.30%	1.60%	1.80%	1.80%	3.10%	1.91%
Ile(I)	3.30%	3.30%	4.00%	2.60%	2.60%	2.50%	2.50%	3.00%	3.20%	2.10%	2.91%
Leu(L)	7.10%	6.90%	7.70%	5.40%	5.90%	6.10%	5.20%	5.60%	5.80%	5.60%	6.13%
Lys(K)	3.60%	5.00%	4.40%	4.40%	4.90%	4.30%	4.40%	3.10%	3.60%	3.90%	4.16%
Met(M)	2.20%	2.50%	2.10%	3.00%	3.10%	2.60%	2.50%	2.60%	2.30%	3.10%	2.60%
Phe(F)	3.80%	4.00%	3.40%	4.80%	4.70%	4.30%	4.60%	4.80%	4.90%	4.90%	4.42%
Pro(P)	3.00%	3.20%	2.90%	3.60%	3.40%	3.60%	3.00%	3.80%	3.60%	2.90%	3.30%
Ser(S)	10.10%	9.90%	11.80%	8.70%	9.50%	9.60%	10.50%	8.50%	9.80%	7.20%	9.56%
Thr(T)	8.60%	8.70%	7.40%	7.40%	6.90%	6.40%	7.40%	7.20%	7.30%	6.00%	7.33%
Trp(W)	1.40%	1.50%	1.50%	1.10%	1.30%	1.50%	1.10%	1.30%	1.50%	1.20%	1.34%
Tyr(Y)	4.90%	4.50%	5.10%	5.40%	5.20%	4.80%	5.70%	5.30%	4.60%	5.40%	5.09%
Val(V)	6.00%	5.50%	6.70%	8.50%	9.30%	7.90%	9.20%	9.40%	7.90%	7.40%	7.78%

Amino acid distribution in Thermophilic and Psychrophilic Metallopeptidases

Average residue distribution in thermophilic metallopeptidase and psychrophilic metallopeptidase is presented in table 2. Here it shows different values for the 20 amino acids. After studying theAmino acid composition of thermophilic and psychrophilic metallopeptidses, it was found that Gln (Q), Asn (N), Ser (S) and Cys (C) are significantly lower in thermophilic metallopeptidases and higher in psychrophiles, and Gly (G), Glu (E), Ile (I), Leu(L), Tyr(Y) higher in thermophiles.

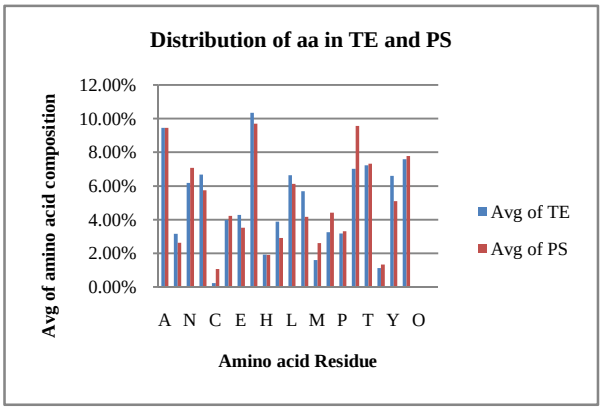


Fig. 1 Distribution of amino acids among TE and PS.

For sequences from thermophilic species the ratio of the amino acid I, V, Y, R, E, L were correlated to optimal growth temperature. The analysis indicated that M4 peptidases from cold-adapted species have fewer arginines and lysines, and also fewer tyrosines and more phenylalanines. However, the opposite was true for M4 peptidases from thermophilicspecies [3].

Physico-chemical property of TE and PS metallopeptidases

For thermophiic and psychrophilic metallopeptidases the physico-chemical property were computed, here it shows different values for TE & PS. After studying Physico-chemical properties such as molecularweight, atomic weight theoretical pI, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average hydropathy (GRAVY),

By physico-chemical parameter it was found that average molecular weight of thermophile is and 60371.18 Da psychrophile is 69921.97 Da. Isoeletric point (pI) is the pH at which surface of protein is covered with charge but net charge of protein is zero. At pI protein are stable and compact. The computed pI value of all psychrophile and thermophile show acidic character (pI<7). The computed isoelectric point (pI) will be useful for developing buffer system for purification by isoelectric point focusing method. On the basis of instability index Expasy'sProtparam classifies the all 20 thermophilic and psychrophilic metallopeptidases in dataset are stable (instability index>40).

Distribution of amino acid property group in Thermophilic and Psychrophilic metallopeptidases

Here, the calculation for all 20 proteins in thermolysin and vibriolysin are tabulated and computed Charged (DEKHR) Aliphatic (ILV) Aromatic (FHWY) Polar (DERKQN) Neutral (AGHPSTY) Hydrophobic (CFILMVW) + charged (KRH) - charged (DE) Tiny (ACDGST) Small (EHILKMNPQV) Large (FRWY) residue in TE and PS proteins.

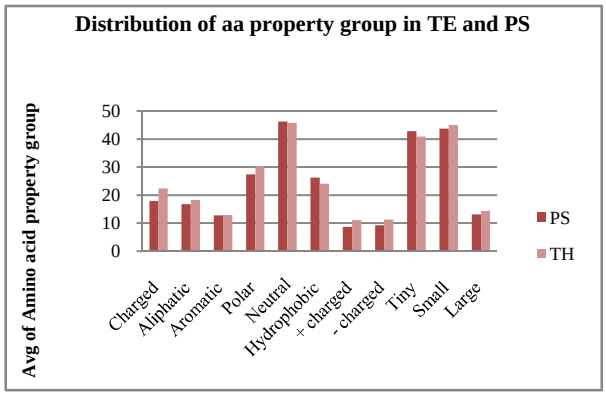


Fig.2 Distribution of amino acids among TE and PS.

By compairing amino acid in TE and PS metallopeptidases protein we found that charged Aliphatic, polar, small and large amino acid residue are high in comparision of Psychrophile counterpart. In order to comparative analysis of thermophilic and psychrophilic metallopeptidases.

Table3

Accession No.	Sequence Length	M. Wt.	pI	No. of Negative Residue	No.of Positive Residue	Extinction coefficient	Instability I ndex	Grand Average Hydropathy	Aliphatic Index
O32309	566	60974.3	5.60	64	55	75640	28.97	-0.483	70.67
P05806	566	60919.3	5.70	64	56	75640	29.23	-0.484	69.98
Q48857	566	60964.4	5.82	65	58	81600	30.35	-0.486	71.02
Q43880	546	59769.6	5.47	59	48	102110	29.95	-0.372	73.97
Q59223	546	59812.6	5.39	60	48	102110	29.08	-0.363	74.51
P23384	544	59413.3	5.64	56	47	99255	30.81	-0.317	75.33
P06874	548	59580.1	5.68	56	46	89260	29.10	-0.351	76.72
P00800	548	60103.8	5.53	59	48	81140	26.43	-0.392	76.35
P43263	527	58645.7	5.33	70	50	90650	29.44	-0.598	68.62
P29148	590	63528.7	4.88	74	50	86640	26.13	-0.430	75.10
Q8RTZ9	731	77820.5	4.73	66	41	109140	28.11	-0.308	67.85
Q8VLQ5	727	78398.5	5.15	69	50	110170	28.57	-0.405	64.87
Q9KH34	729	77681.7	4.79	68	43	116130	26.90	-0.249	73.33
Q00971	609	66362.5	5.07	64	45	88045	30.55	-0.359	68.84
P43147	611	66726.4	5.70	53	44	92055	30.83	-0.292	68.43
O06694	603	65613.8	6.17	54	48	93085	32.45	-0.342	66.67
Q8RR31	610	65549.5	5.42	54	43	91025	28.51	-0.277	65.69
P24153	609	65891.2	5.26	53	39	92055	25.18	-0.250	68.82
MER50231	777	83202.0	4.91	74	49	120140	28.38	-0.210	68.97
MER49928	485	51973.6	5.74	46	35	71990	24.58	-0.275	63.65

Table4-Table for distribution of amino acid property group along thermophilic and psychrophilic Metallopeptidases computed using Copid Tool.

Uniprot Accession No.	Amino Acid property group										
	Charged	Aliphatic	Aromatic	Polar	Neutral	Hydrophobic	+ charged	- charged	Tiny	Small	Large
O32309	22.66	17.70	11.86	31.68	45.84	22.48	11.33	11.33	42.30	45.84	11.86
P05806	22.79	17.49	11.84	31.45	46.29	22.26	11.48	11.31	42.76	45.41	11.84
Q48857	23.32	17.67	12.01	31.09	47.17	21.73	11.84	11.48	42.40	45.58	12.01
Q43880	21.43	18.50	13.55	28.94	45.61	25.46	10.62	10.81	39.38	44.32	16.30
Q59223	21.61	18.68	13.55	28.94	45.42	25.64	10.62	10.99	39.38	44.32	16.30
P23384	20.77	18.75	13.23	28.49	45.22	26.29	10.48	10.29	39.89	44.12	15.99
P06874	20.80	19.34	13.14	28.47	46.53	25.00	10.58	10.22	40.33	43.61	16.06
P00800	21.48	19.26	13.70	30.93	43.70	25.37	10.74	10.74	38.52	47.59	13.89
P43263	26.18	17.27	14.61	32.45	43.83	23.72	12.90	13.28	38.90	46.30	14.80
P29148	22.88	18.47	11.36	29.66	47.63	22.71	10.34	12.54	44.41	43.05	-
Q8RTZ9	16.01	16.42	11.49	27.09	48.02	24.90	6.98	9.03	45.55	42.27	12.17
Q8VLQ5	18.19	15.56	11.95	28.61	46.53	24.86	8.61	9.58	44.03	43.89	12.08
Q9KH34	16.74	18.38	11.52	26.34	47.19	26.47	7.41	9.33	44.58	43.90	11.52
Q00971	19.83	16.50	13.33	28.83	44.67	26.50	9.33	10.50	40.83	45.00	14.17
P43147	17.68	17.84	13.09	28.48	43.54	27.99	9.00	8.67	39.44	46.97	13.58
O06694	19.17	16.67	13.00	28.17	45.50	26.33	10.17	9.00	41.33	44.50	14.17
Q8RR31	17.50	16.67	13.17	26.67	47.33	26.00	8.67	8.83	44.00	42.00	14.00
P24153	16.85	17.59	13.52	27.78	44.45	27.78	7.96	8.89	39.45	45.93	-
MER50231	17.63	16.86	12.87	26.13	47.23	26.64	8.11	9.52	44.92	41.32	13.77
MER48811	19.63	15.08	14.46	26.24	48.35	25.41	10.12	9.50	43.39	41.74	-

Charged residues are polar and hydrophilic,they contribute to ion pair electrostatic interactions that are important binding force for maintaining conformational stability in surface of the proteins [12, 13, 14]. The more charged residues were found in thermophilic proteins than in psychrophilic proteins. Tiny and small amino acids are those with short side chains and are unable to participate in long range interactions among secondary structural elements and are usually confined to form local interactions. The tiny aa residue are more in psychrophile and small aa are more in thermophile.

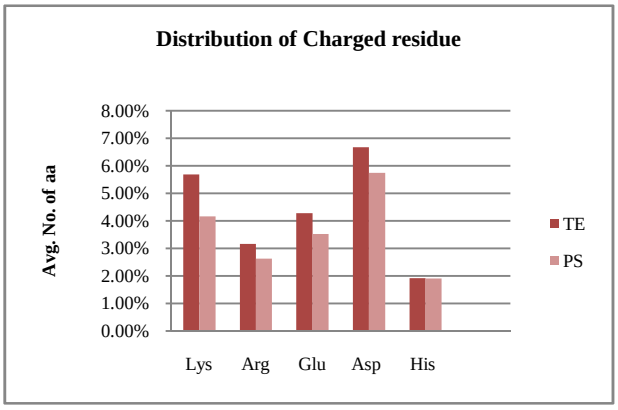
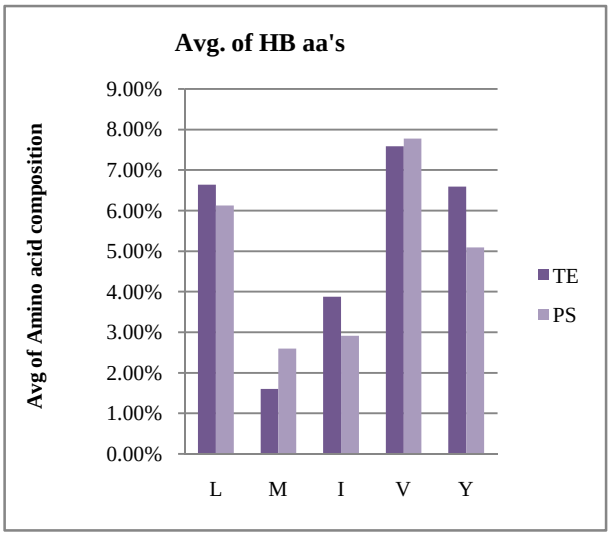


Fig. 2 Distribution of charged amino acids among TE and PS.

By compairing charged residues we found that (Lys, Arg, and Gln, Asp) are more significant in thermophiles than in psychrophiles metallopeptidases proteins. It can be concluded that thermophiles possesses more charged residue, such as Glu, and more hydrophobic residue, as compared to psychrophiles.

The hydrophobic residues were compared among two classes of protein called Thermophile and mesophile and was found that Lys, Tyr are more in thermophile and Met, Pro are more in psychrophile. Piechart of hydrophobic residues shows Lys, Met, Pro, Val, Tyr percentage and it resulted that Val has greater percentage while Met has significantly low frequency in dataset.



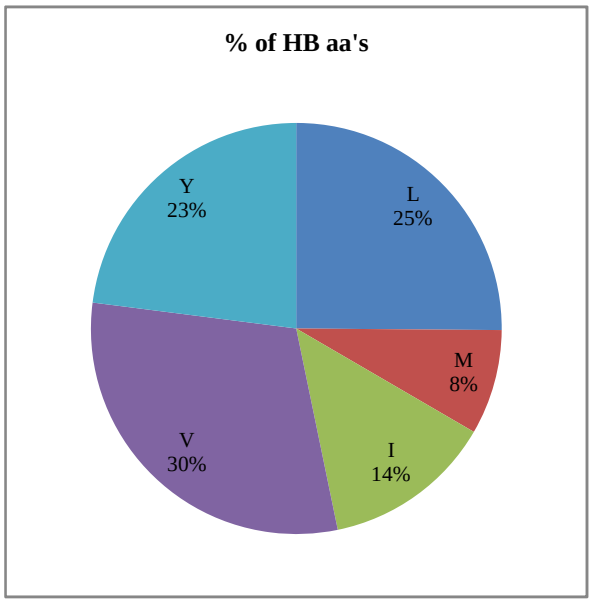
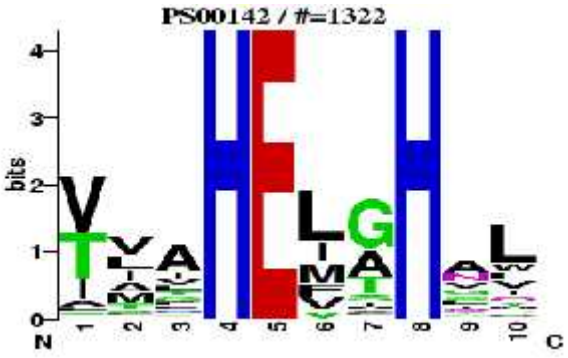


Fig. 3 Distribution of hydrophobic amino acids among TE and PS.

The substrate specificity is restricted to hydrophobic amino acids, which is used in functional and structural studies of proteins, among hydrophobic residue Ile, Leu, Val belong to aliphatic group of residues, which are significantly reduced to favor protein flexibility. It has been widely accepted that amino acids would contributeto the hydrophobic interaction for maintaining conformational stability and rigidity in core region of the proteins.

Sequence Logo-– For Zinc binding site in Thermophilic and Psychrophilic Metallopeptidases

By ScanProsite tool we also found that VSAHEVSHGF and VIGHETHV or VVGHETHVamino acid pattern were common in Psychrophiles and Thermophiles respectively.



Sequence logo is a graphical display of a multiple sequence alignment consisting of colour-coded stacks of letters representing amino acids at successive positions. The 'abXHEbbHbc' zinc binding motif are common in all metallopeptidases of M4 family, Hence by this logo we predict that Valine have maximum probability to come on first and second position, Alanine have maximum probability to come on fourth and fifth position and so on.

Secondary structure analysis

Here, the secondary structure calculation for all 20 proteins in thermolysin and vibriolysin are tabulated [Table 5,6] and computed Alpha helix, 3_{10} helix, Pi helix, Beta bridge, Extended strand, Beta turn, Bend region, Random coil were computed.

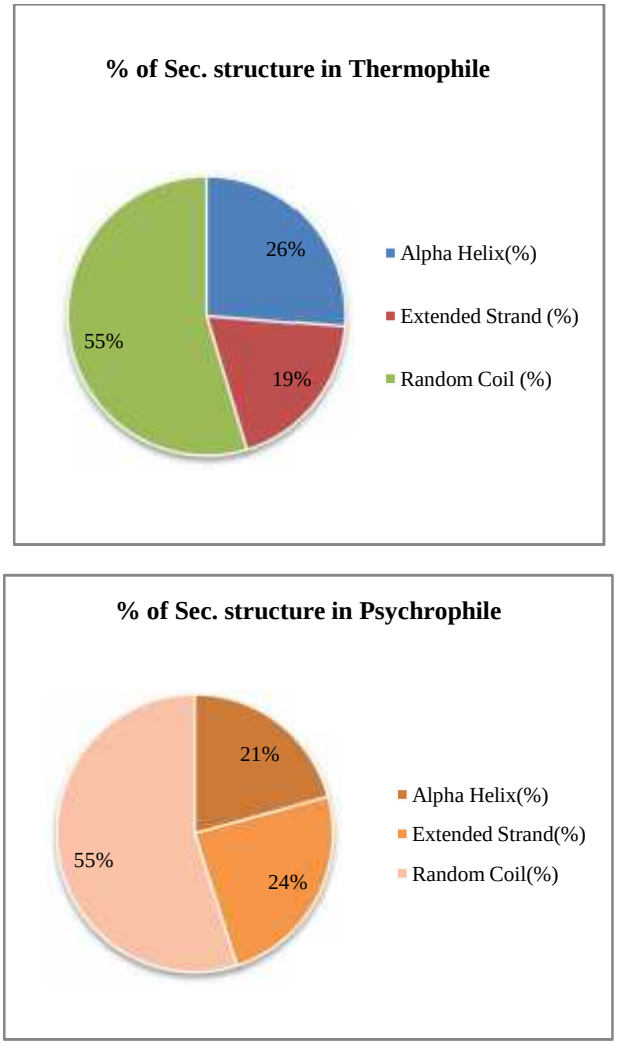


Fig. 4 Pi chart distribution of secondary structure

By compairing Secondary structure in TE and PS metallopeptidases protein we found that Alpha helix are more in thermophilic metallopeptidases than psychrophile counterpart, and psychrophilic metllopeptidases have more number of strand. There is no occurance of 3_{10}-Helix Pi-helix, Beta bridge and Bend region.

Table 5 Table for % of *Secondary structure along thermophilicand PsychrophilicMetallopeptidase computed using MLRC Tool.*

Accession No.	Alpha Helix(%)	310 Helix(%)	Pi Helix(%)	Beta Bridge(%)	Extended Strand (%)	Beta Turn (%)	Bend Region(%)	Random Coil (%)
O32309	21.06	0.00	0.00	0.00	19.47	0.00	0.00	59.47
P05806	22.08	0.00	0.00	0.00	20.32	0.00	0.00	57.60
Q48857	22.44	0.00	0.00	0.00	18.20	0.00	0.00	59.36
Q43880	27.84	0.00	0.00	0.00	19.41	0.00	0.00	52.75
Q59223	29.67	0.00	0.00	0.00	17.40	0.00	0.00	52.93
P23384	29.96	0.00	0.00	0.00	18.75	0.00	0.00	51.29
P06874	28.10	0.00	0.00	0.00	19.53	0.00	0.00	52.37
P00800	28.63	0.00	0.00	0.00	18.07	0.00	0.00	53.10
P43263	25.81	0.00	0.00	0.00	18.79	0.00	0.00	55.41
P29148	27.12	0.00	0.00	0.00	20.34	0.00	0.00	52.54
Q8RTZ9	19.29	0.00	0.00	0.00	24.62	0.00	0.00	56.09
Q8VLQ5	23.25	0.00	0.00	0.00	23.11	0.00	0.00	53.65
Q9KH34	16.74	0.00	0.00	0.00	26.47	0.00	0.00	56.79
Q00971	19.05	0.00	0.00	0.00	24.63	0.00	0.00	56.32
P43147	22.26	0.00	0.00	0.00	24.55	0.00	0.00	53.19
O06694	22.11	0.00	0.00	0.00	23.27	0.00	0.00	54.62
Q8RR31	18.03	0.00	0.00	0.00	25.74	0.00	0.00	56.23
P24153	20.83	0.00	0.00	0.00	25.33	0.00	0.00	53.83
MER50231	20.43	0.00	0.00	0.00	25.61	0.00	0.00	53.93
MER48811	25.67	0.00	0.00	0.00	19.47	0.00	0.00	54.77

t-test

t-test result for amino acid residue distribution –

Amino acid	H (hypothesis)	p-val	t-val
Ala(A)	0	0.8555	-0.1848
Arg(R)	0	0.3154	1.0327
Asn(N)	0	0.0565	-2.0384
Asp(D)	1	0.0041	3.2915
Cys(C)	1	1.2494e-006	-7.1143
Gln(Q	0	0.4718	-0.7350
Glu(E)	1	0.0032	3.4018
Gly(G)	0	0.1830	1.3848
His(H)	0	0.9668	0.0422
Ile(I)	1	0.0013	3.8194
Leu(L)	0	0.1200	1.6324
Lys(K)	1	0.0404	2.2085
Met(M	1	5.0828e-004	-4.2258
Phe(F)	1	3.7092e-005	-5.4283
Pro(P)	1	2.9707e-004	-4.4669
Ser(S)	1	2.9797e-004	-4.4669
Thr(T)	0	0.8198	-0.2312
Trp(W)	0	0.1054	-1.7052
Tyr(Y)	1	6.7323e-006	-6.2536
Val(V)	0	0.7035	-0.3867

To compare the means of two groups of data (TE and PS) t-test was done, and it was found that the Asp(D), Cys(C), Glu(D), Ile(I), Lys(K), Met(M), Phe(F), Ser(S), Tyr(Y) amino acids residue are significantly different among the TE & PS species, because the null hypothesis(h) is accepted for these amino acids.

t-test result for amino acid property group

Amino acid property group	H(hypothesis)	p-val	t-val
Charged	1	1.9394e-006	6.8846
Aliphatic	1	9.0279e-004	3.9675
Aromatic	0	0.9206	0.1011
Polar	1	1.3209e-004	4.8377
Neutral	0	0.4087	-0.8460
Hydrophobic	1	0.0023	-3.5490
+Charged	1	1.5828e-005	5.8350
-Charged	1	1.7304e-005	5.7920
Tiny	0	0.0598	-2.0091
Small	0	0.1061	1.7012
Large	0	0.1643	1.4622

By t-test computation we found that charged, Aliphatic, Polar, Hydrophobic, +charged, -Charged residue are significantly different in thermophilic and psychrophilic metallopeptidases.

t-test result for secondary structure

Secondary structure	H (hypothesis)	P-val	t-val
Helix	1	6.3158e-004	4.1280
Extended strand	1	5.9450e-007	-7.5108
Random coil	0	0.8080	-0.8080

By t-test computation we found that Helix and Extended strand are significantly different in thermophilic and psychrophilic metallopeptidases.

DISCUSSION

A method to discriminate thermophilic and psychrophilic metallopeptidases were studied and found that Arg(R), Ile(I), Glu(E), Gly (G), Leu(L), Lys(K), Tyr(Y), Asp(D) are more in thermophilic metallopeptidases and Asn(N), Cys(C), Met(M), Phe(F), Ser(S), Thr(T) more in psychrophilic metallopeptidases. The Arg(R) content is lower among peptidases from cold adapted species (Psychrophiles) and higher in thermal adapted species (Thermophile). Arg(R) has two terminal nitrogen atoms and may easily form salt bridges and H-bond that stabilize the 3D- Structure of thermophlilic metallopeptidases.The physico-chemical property of thermophilic and psychrophilic metallopeptidases are computed. The physico-chemical property like pI (isoelectric-point) value of all psychrophile and thermophile show acidic character (pI<7) and on the basis of instability index all the 20 thermophilic and psychrophilic proteins are stable (instability index>40). The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chains (A, V, I, AND L) is regarded as positive factor for increase of thermal stability of globular protein, the lower thermal stability is indicative of a more flexible structure. The aliphatic index of all 10 thermophiles are high in comparison of psychrophiles, hence the thermophiles show higher thermal stability, may be stable for a wide range of temperature. Grand Average Hydropathy (GRAVY) Index of thermophile and psychrophile are ranging from -0.598 to -0.210. The very low GRAVY Index of all sequences infers that metallopeptidases of TE and PS could result better interaction with water.

Peptidases from thermal adapted species also have more Tyr(Y) and fewer Phe(F) and Trp(W) than those from cold-adapted species, Tyr is a hydrophobic amino acid, but with potential of being both a hydrogen bond donor and acceptor in the 3D structure. Phe and Trp are large aromatic amino acids, but compared with Tyr, lack hydrogen bonding possibilities (lack an OH-group). An increased number of Tyr and decreased number of Phe and Trp may therefore be consistent with a more rigid 3D structure of the peptidases from thermal adapted species than of the M4 peptidases from other species.The peptidases from thermal adapted species also have fewer of the sulfur containing amino acids Met and Cys thanpsychrophiles. The metallopeptidases from thermal adapted species also have a higher content of the large hydrophobic amino acid Ile than other M4 peptidases.

The secondary structure predicted infers that TH have high Lys(L), Arg(R) and Glu(E) and mostly have alpha helices. A protein belonging to psychrophiles show there is a decrease in nonpolar amino acid group frequency in helices and a significant increase in loop regions, however opposite was true for thermophiles. This is in accordance with earlier findings that there are more nonpolar amino acids on the exposed surface area of the majority of psychrophilic proteinasas more loops are observed on surface regions. In psychrophiles, individual residue composition shows that there is significant preference for Ser(S) & Thr(T) content and significant avoidance of Glu(E), Leu(L) content and moderate preference for Gly(G) and avoidance for Phe(F) and Lys(K) content. All these residue preferences and avoidance directly show a strong correlation with respect to avoidance for helical content in psychrophiles as Ser(S), Pro(P) are helix breaker and Thr(T) is helix indifferent.

Likewise, the presence of Glu(E) tends to favor information of helical structure and Leu(L) to stabilize helical structure, that are highly avoided in psychrophiles and significantly more in thermophiles.

The t-test result of amino acid composition, amino acid property group, Secondary structure infers that, amino residue composition Ile(I), Glu(E), Cys(C), Met(M), Phe(F), Ser(S), Lys(K), Tyr(Y), Asp(D) are significantly different in

thermophilic and psychrophilic metallopeptidses, and amino acid property group Charged, Aliphatic, Polar, Hydrophobic, +Charged, -Charged aa residue significantly different in TE & PS Metallopeptidases, Secondary structure element Helix and Strand significantly different in Thermophilic & Psychrophilic Metallopeptidases.

CONCLUSION

We analyzed compositions of individual amino acid residues, amino acid groups and secondary structures. Significant differences in composition of amino acid residues were observed between the thermophilic and psychrophilic metallopeptidases as summarized below: (i) we observed an increase in frequency of individual amino acid like Lys, Arg, Glu, Asp which are significantly more in thermophiles; (ii) Hydrophobic amino acid Leu, Ile, Tyr are more in thermophiles and the Met, Val are more in psychrophiles (iv) Physico-chemical characterization studies gives idea about the properties such as pI, Extinction coefficient(EC), Aliphatic Index(AI), GRAVY and instability index that are essential and vital in providing data about the protein and their properties (v) The result from this analysis, especially significant t-values of individual amino acid for 20 protein can be used as a knowledge for differentiating psychrophile and thermophile on the basis of amino acids.

Amino acid groups were observed between thermophiles and psychrophiles and we concluded that charged, aliphatic, polar, tiny amino acids are more prevalent in thermophiles, while neutral and large amino acid are prevalent in psychrophiles. By secondary structure analysis we concluded that the helix percentages are more in thermophiles while the presence of strands is more in psychrophilic metallopeptidases. By ScanProsite tool was also found that VSAHEVSHGF and VIGHETHAV or VVGHETHV amino acid pattern were common in Psychrophilic and Thermophilic metallopeptidases respectively.

The information gained in the present study may be helpful for differentiating thermophilc and psychrophilic metallopeptidases, the identified parameter may also facilitate in understanding the extremophile adapatation of thermophilic and psychrophilic metallopeptidases. These finding would also help further efforts in selecting mutation for rational designs of proteins with enhanced thermophilic and psychrophilic metallopeptidase properties.

References

1. Neil D Rawlings, Alan J Barret, Alex Bateman (*2011)* Merops: the databases of proteolytic enzymes their substrate and inhibitors. *Nucleic Acid Research* 40: D343-D350.
2. F Xavier, Gomis Ruth, Tiago O Batelho, Wolfram Bode (2011) A Standard orientation for metallopeptidases. *BiochimicaetBiophysicaActa* 1824: 157–163.
3. Mahmud T Hassen Khan, IngebrigtSylte (2009) Determinant for Psychrophilic and Thermophilicfeautures of metallopeptidases of M4 Family. *In Silico Biology.*
4. Raghu P Rao, Boojala V Reddy (2009) Comparative proteome analysis of psychrophilic verses mesophilic bacterial species. *BMC Genomics* 10.
5. Bin-Bin Xie, FeiBian, Xiu-Lan, Chen Hai-Lun, He Jun Guo, Xiang Gao (2009) Cold Adaptation of Zinc Metalloproteases in the Thermolysin Family from Deep Sea and Arctic Sea Ice Bacteria Revealed by catalytic and Structural Properties and Molecular Dynamics. *Biological Chemistry* 284: 9257–9269.
6. Olayiwola A. Adekoya, IngebrigtSylte (2008) The Thermolysin Family (M4) of Enzymes: Therapeutic and biotechnological potential. *Chem Bio Biol Drug* 73:7-13
7. Manish K, T Varun, G Raghahava (2008) Copid: Composition based protein identification. In Silico Biology 8:11
8. Adekoya O. A, Helland R Willassen, I. Sylte (2006) Comparative sequence and structure analysis reveal features of cold adaptation of an enzyme in the thermolysin family. *Proteins* 62:435-449
9. Gasteiger E, Hoogland C, Duvaud S, Wilkins M.R, Appel R.D, Bairoch (2005) A Protein identification and Analysis Tools on the ExPASyServer:The proteomics protocols. *Handbook Humana Press* 571-67.
10. Donal A Hickey, Gregory A Singer (2004) Genomic and proteomic adaptation to growth at high temperature. *Genome Biology* 5.
11. Datta, P. P., Bhadra, R. K (2003) Cold shock response and major cold shock proteins of Vibrio cholerae. *Appl. Environ. Microbiol*. 69: 6361-6369.
12. Feller G, Gerday C (2003*)*.Psychrophilic enzyme:hot topics in cold adaptation*. Nat Rev.Microbiol* 3:200-208.
13. Turner A J (2003) Exploring the structure and function of zinc metallopeptidases. *BiochenSoc Trans* 3:723-7.
14. Szilagyi A, Zavodszky P (2000) Structural differences between mesophilic, moderatelythermophilic and extremely thermophilic protein subunits: results of a comprehensive. *Structure* 8(5): 493-504.
15. Gromiha M. Oobatake, M. Sarai (1999) Important amino acid properties for enhanced thermostability from mesophilic to thermophilicproteins. *Biophys Chem* 82(1):51-67.
16. Nigel M. Hopper (1994*)* Family of zinc metallopeptidases. *Biochem mol. & Biology* 354:1-6
17. Neil D Rawlings, Alan J Barret (1993) Evolutionary families of peptidases. *Biochem. J* 290: 205-218.
18. Naruya S, Mastoshi N (1987) The Neighbor-joining method: A new method for reconstructing Phylogenetic trees. *Mol. Biol*. 4:413-421.

Tools and Databases

1. PROTPARAM tool. (http://web.expasy.org/ protparam/).
2. Uniprot Protein database. (http://www.uniprot).
3. Prosite database. (http://prosite.expasy.org/).
4. MLRC tool (http://npsapbil.ibcp.fr/cgibin/ npsa_automat.pl?page=/NPSA/npsa_mlrc.html).
5. Copid Server (http://www.imtech.res.in/raghava/ copid/wholecomp.html).
6. Merops database (http://merops.sanger.ac.uk/).