Original Research Article   http://dx.doi.org/10.20546/ijcmas.2016.504.068

In silico Analysis of Physicochemical Properties of Hyperthermophillic and Thermophillic Nitrile Hydratases

Pawan Kumar Sharma[1]* and Tek Chand Bhalla[2]

[1]Department of Agricultural Biotechnology, CSKHPKV Palampur HP 176062, India
[2]Department of Biotechnology Himachla Pradesh University, Shimla HP171005, India
*Corresponding author

A B S T R A C T

Nitrile hydratases (NHases, EC 4.2.1.84) are non-heme iron-containing or non-corrinoid cobalt-containing enzymes that catalyze the hydration of nitriles to their corresponding amides. Present research a computational study with statistical correlation has been done using the sequences of thermophiles and hyper thermophiles. The number of amino acids were found in range of 249 to 288 for hyperthermophiles and 235 to 292 in thermophiles, molecular weight range was 26738.7 to 32032.4 and 26044.9 to 35529.7, theoretical pI estimated were 6.18 to 9.44 and 6.16 to 9.68, negatively charged residues were 28 to 41 in both cases, positively charged residues were 30 to 39 and 30 to 46, extinction coefficient estimated were 8940 to 37930 and 17880 to 40910, aliphatic index observed as 101.73 to 111.74 and 90.49 to 113.01, instability index values were 38.68 to 40.89 and 30.71 to 42.11, GRAVY values were positive and negative for hyperthermophiles and thermophiles, respectively. Statistical correlation was found among the cysteine as 5.8840 , Met (M) 6.9980 , Phe (F) 9.5972 , Thr (T) 5.3864 , Tyr (Y) 7.7134 . This shows sulphur containing, aromatic and bulky groups as significant amino acids.

Introduction

The past few decades have witnessed fast development of nitrile-amide converting enzymes, their reaction mechanisms and applications in manufacture of a series of pharmaceuticals and chemicals. Nowadays, with the discovery of numerous nitrile-amide converting microorganisms and their extremely fast development, these enzymes are becoming more and more appreciated by scientists and are showing competency with esterases and lipases (Chen *et al.,* 2009). Besides their synthetic value, these enzymes

play an important role in environmental protection, by removal of toxic nitrile compounds, which are detrimental to human beings, animals and plants. In order to exploit their full biotechnological potential, researchers are focusing on (1) Overcoming some disadvantages of the nitrile-converting biocatalysts, such as narrow substrate specificity, low thermostability and pH stability, low tolerance of substrate and product, undesired and unsatisfied enantioselectivity. (2) Screening and

discovery of new nitrile-amide converting enzymes with promising and attractive properties. As previously demonstrated, microorganisms producing nitrile amide converting enzymes turned out to be dominantly from prokaryotic ones and eukaryotic organisms constitute only a small part. The latter ones, however, were always neglected as an excellent source of nitrilase, nitrile hydratase and amidase.

Moreover, these organisms may provide some different properties like excellent thermostability, regio-, enantio- selectivity, and improved stability. (3) Employing genetic engineering to alter some undesired properties of wild type strain, hence it is possible to isolate novel enzymes with extremophilic characteristics. Beside the advantages of nitrile-amide converting enzymes, a few of them have achieved success in industrial application, with the commercial production of acrylamide and nicotinamide being the most successful. Among these enzymes Nitrile hydratases (NHases, EC 4.2.1.84) are non-heme iron-containing or non-corrinoid cobalt-containing enzymes that catalyze the hydration of nitriles to their corresponding amides. They have considerable practical importance as biocatalysts; a member of this class of enzymes is presently being used in a 30000 ton yr^{-1} acrylamide production process, and is used for environmental remediation i.e. effective in the removal of nitriles from waste streams (Pelt, 2010). Fascinatingly, one subset of the iron-containing NHases is activated by light in a reaction involving photodetachment of nitric oxide from an endogenous Fe–NO unit (Kubiak and Nowak, 2008). These enzymes comprise two subunits, α and β, not related in amino acid sequence, with a molecular mass around 23 and 24 kDa (Rzeznicka *et al.,* 2010). These iron-containing enzymes have one iron ion per α β unit, most

probably bound to α subunit (Chen *et al.,* 2009).

Nitrile hydratases are widespread among the prokaryotes, with the enzyme being found in genera in the Proteobacteria, Actinobacteria, Cyanobacteria and Firmicutes (Foerstner *et al.,* 2008). Here the nitrile-degrading biochemical pathway allows organisms to use nitriles as a nutritional source for obtaining nitrogen (in the form of ammonia) and carbon (in the form of organic acids). Prokaryotic nitrile hydratases and nitrile-related biochemistry have also been adapted for biotechnological roles. Uses include the bioremediation of toxic nitrile-containing compounds (such as herbicides or industrial effluents) from soil and aquatic environments (Banerjee *et al.,* 2002; Marron *et al.,* 2012), and the industrial synthesis of acrylamide, nicotinamide and 5-cyanovaleramide (Yamada *et al.,* 2001; Pelt, 2010). The economic importance of nitrile hydratases has led to increased research into their metabolic functions, mechanisms and phylogenetic distribution (Prasad & Bhalla, 2010). Keeping in view the above stated facts computational analysis of some thermostable and Co containing nitrile hydratses has been done.

Materials and Methods

The amino acid sequences of the nitrile hydratases (hyperthermophilic, thermophilic and cobalt containing) were downloaded from the National Center for Biotechnology Information (NCBI) server (Table 1). Physiochemical data were generated from Expert Protein Analysis System (EXPASy) i.e. the proteomic server of Swiss Institute of Bioinformatics (SIB). FASTA format of sequences were used for analysis.

Blastp (Protein BLAST) was performed to study the homology among the various

nitrile hydratases sequences and sequences each belonging to hyprthermophiles, thermophiles, α and β subunit of Co containing nitrile hydratase were selected (Table 1). Clustal W was used for multiple sequence alignment.

Various tools in the proteomic server (Clustal W, ProtParam, Protein calculator, Compute pI/Mw, ProtScale (Kyte and Dollitle, 1982**)** were applied to calculate/deduce different physiochemical properties of nitrile hydratase from the protein sequences.

The atomic composition of nitrile hydratase was derived using the ProtParam tool, available at ExPASy. The extinction coefficient of various nitrile hydratases were calculated using the following equation (Gill and von Hippel, 1989):

E(Prot) = Numb(Tyr)*Ext(Tyr) + Numb (Trp)*Ext(Trp) + Numb (Cystine)*Ext (Cystine)

The values of aliphatic index of various nitrile hydratase sequences were obtained using ProtParam (ExPASy) tool (Kyte and Doolittle, 1982). The instability index and grand average of hydropathicity (GRAVY) were estimated following the method of Guruprasad et al. (1990) and Kyte and Doolittle (1982), respectively.

Statistical Analysis

An analysis of variance (ANOVA) was conducted on various physiochemical parameter variables for each study with the statistical packages 'Asistat version-7.6 beta 2014'. F-tests were used to determine the statistical significance. When significant effects were detected, a Tukey test was applied for all pairwise comparisons of mean responses.

Results and Discussion

Present study focuses on comparison of amino acid sequences of hyperthermophilic and thermophillic microorganisms amino acid sequences as given in table 2. The no. of amino acids were found in range of 249 to 288 for hyperthermophiles and 235 to 292 in thermophiles, molecular weight range was 26738.7 to 32032.4 and 26044.9 to 35529.7, theoretical pI estimated were 6.18 to 9.44 and 6.16to 9.68, negatively charged residues were 28 to 41 in both cases, positively charged residues were 30 39 and 30 to 46, extinction coefficient estimated were 8940 to 37930 and 17880 to 40910, aliphatic index observed as 101.73 to 111.74 and 90.49 to 113.01, instability index values were 38.68 to 40.89 and 30.71 to 42.11, GRAVY values were both positive and negative for hyperthermophiles and thermophiles, respectively. None of the above estimated parameters found significant at probability values ($p < 0.01$ and $p < 0.05$).

Amino acid analysis shows that cysteine is present as 5.8840, Met (M) 6.9980, Phe (F) 9.5972, Thr (T) 5.3864, Tyr (Y) 7.7134. This shows that hyperthermophiles and thermophiles are composed of sulphur containing, aromatic and bulky groups which account for the thermostability of the enzyme. The results can be summarized and discussed in following manner:

Physico Chemical Parameters

No. of Amino Acids

Query was given by providing either the accession number of proteins or by providing amino acid sequence in FASTA format. Number of amino acids in a protein gives rough idea about the complication of the tertiary structure it achieves. It is well

known that more the amino acids number of protein, more beta turns and coiled coiled conformation it achieves during protein folding (Hirokawa *et al*. 1998). The number of amino acids was found in range of 249 to 288 for hyperthermophiles and 235 to 292 in thermophiles (Table 1). Similar results have been obtained by Rzeznicka *et al.,* 2010.

Molecular Weight

Present investigation shows molecular weight within a range of 26 KDa to 35 KDa. Determination of the molecular weight of protein gives a clear view about the method one should adopt for the purification of the protein. However, *in vitro* analysis of experimental weight of the α- and β-subunits have a molecular weight of around 26 and 28 kDa and the characteristic C**T**LCSC domain for Co-type nitrile hydratases was identified by a PCR based screening method using degenerate primers (Pelt, 2010; Rzeznicka *et al.,* 2010).

pI

Theoretical pI estimated were 6.18 to 9.44 and 6.16to 9.68. Table 1 show that the experimental NHase proteins have a pI value above 5. Which show proteins will be positively charged at pH 7 and negatively charged above pH 7. These predictions will be helpful in designing methodology for the purification of these proteins. The bacterial strains isolated by Pelt (2010) from moderately salt-tolerant alkalophiles with a pH range for growth from pH 7.0 to 8.5 up to 10.3 to 10.5. However, both their NHase and amidase activities had a near neutral pH optimum, indicating an intracellular localisation of these enzymes. So far, all known nitrile degrading microorganisms are neutrophilic, i.e., they grow optimally at neutral pH values (Sorokin and Kuenen, 2005).

Charged Residues

Negatively charged residues were found in a range of 28 to 41 in both cases, positively charged residues were ranged between 30 to 39 and 30 to 46, respectively. Determining the positively and negatively charged amino acids help in formulating the posttranslational modification.

Extinction Coefficient

Extinction coefficient values ranged between 8940 to 37930 and 17880 to 40910, it predicts the spectroscopic properties of the protein. Higher the extinction coefficient more is the absorbance. It is one of the parameter that characterize for the primary structure analysis of the protein. The extinction coefficient for a particular substance is a measure of how well it scatters and absorbs electromagnetic radiation. This will help in prediction of some spectroscopic properties of proteins. Use of Protparam tool provides accuracy for this parameter up to the level of $5e^{-15}$.

Aliphatic Index

Aliphatic index observed for hyperthermophillic and thermophillic nitrile hydratase sequences were in a range of 101.73 to 111.74 and 90.49 to 113.0, respectively. Lower aliphatic index determine the fact that particular protein would not be involved in electrophilic substitution reaction, which is characteristic of aromatic reactions. More is the aliphatic index more stable is the protein. Aliphatic indices of proteins are between 90–120 (Ikai, 1980; Sahay and Shakya, 2010). It suggests that these proteins undergo more of the nucleophilic substitution reaction than the electrophilic substitution reaction, as these proteins would have more of aliphatic characters than aromatic character. Most

results showed that the Fe-type NHases preferentially hydrate small aliphatic nitriles, whereas the Co-type NHases exhibit a relatively high affinity for aromatic nitriles (Miyanaga *et al.,* 2001).

Instability Index

Hyperthermophillic and thermophillic nitrile hydratase amino acid sequence values werein arrange of 38.68 to 40.89 and 30.71 to 42.11, respectively. Table 1 shows that all the proteins studied are fairly stable below 40, and after purification can be used for laboratory study. Instability index were calculated with respect of *E. coli* to test the stability of protein.

Gravy Index

GRAVY index of the protein predicts the hydrophobic character of the protein. If the value is below 1 it shows that the given protein is transmembrane protein with more than 4 helical turns. Most of the prokaryotic transmembrane proteins have GRAVY index with negative values as they possess more of beta strands. Low range of value indicates the possibility of better interaction with water (Sahay and Shakya, 2010).

Atomic Composition

The prediction of atomic composition of the protein is for using particular atom in the radioactive labeling to trace the physical mobility and metabolic route it follows in the cell. It helps in locating the region where the protein is present in the cell.

Statistical Significant Amino Acids and their Behavior

Thermophiles are under constant destabilizing effects of high temperatures. To counteract ill effects of temperature on proteins, thermophiles adopt strategies that

are characteristically different from mesophiles. Some of the strategies are replacement of Arg with Tyr, presence of sulfate ions. This is in accordance with Trivedi *et al.,* (2006), the type of metal ions as co-factors, thermamines, chaperones, substrate, flexibility and or rigidity of proteins (Vieille *et al.,* 2001), smaller size of protein and higher compactness ratio may also influence stability of proteins (Hickey and Singer, 2004). However, some reports suggest that compactness ratio of proteins in thermophiles and mesophiles is not significantly different (Kumar *et al.,* 2000). It appears that these strategies are important but not universally adopted by thermophilic prokaryotes.

Cys another significant amino acid observed in present study is modified to a cysteine sulfinic acid ~$Cys\text{-}SO_2H$ and a cysteine sulfenic acid ~Cys-SOH respectively (Woo and Rhee, 2010). Murakami *et al.,* 2000 suggested that Cys112 and Cys114 are spontaneously oxidized to $Cys\text{-}SO_2H$ and Cys-SOH, respectively, and $Cys112\text{-}SO_2H$ is responsible for the catalytic activity solely or in combination with aCys114-SOH.

The full reaction mechanism of nitrile hydratase has remained elusive, despite extensive theoretical and experimental studies. A novel reaction mechanism for nitrile hydratase based on large quantum-mechanical active site models, identify Cys-SO^- as the nucleophile, performing a direct nucleophilic attack on the metal-coordinated nitrile. This implies the formation of an intriguing cyclic intermediate, which subsequently cleaved through attack of the axial cysteine on the sulfenate, thereby forming a disulfide bond. In this mechanism, nitrile hydration occurs without directly involving a water molecule. Subsequent water-mediated disulfide cleavage regenerates the active site (Hopmann, 2014). Cysteine was considered to be a hydrophilic

amino acid based on the belief that the thiol group interacts well with water. However, the cysteine side chain participates in the hydrophobic bonding system of the micelle. The quantitative comparison of the values of the critical micelle concentrations leads to the conclusion that the CH_2SH group of amino acid has the same effect on the micelle stability as a methylene group of the hydrocarbon tail. Therefore, there is justification for the assumptions of the hydrophobic nature of the cysteine side chain, which enhances thermostability of the protein (Foerstner, 2008). The thiol side chain often participates in enzymatic reactions, serving as a nucleophile. The thiol is susceptible to oxidization to give the disulfide derivative cystine, which serves an important structural role in many proteins (Nagano *et al.,* 1999).

Cysteine's high sensitivity to oxidation at high temperature suggests that hyperthermophilic enzymes contain fewer cysteines than their mesophilic counterparts do. One would expect a high selection pressure against the presence of cysteines in proteins from aerobic hyperthermophiles (and the absence of such selection pressure in anaerobic hyperthermophiles). Cysteines that are present in proteins from aerobic hyperthermophiles are often involved in specific stabilizing interactions (e.g., disulfide bridges and metal liganding) and/or are inaccessible to the solvent (Kubiak and Nowak, 2008). Disulfide bridges are believed to stabilize proteins mostly through an entropic effect, by decreasing the entropy of the protein's unfolded state (Matsumura *et al.,* 1989). The entropic effect of the disulfide bridge increases in proportion to the logarithm of the number of residues separating the two cysteines bridged (Sahay and shakya, 2010).

Methionine is an α-amino acid, classified as nonpolar. Hence contributing towards the hydrophobicity and enhancing thermostability of protein.

The side chain is quite hydrophobic and methionine is usually found buried within proteins. Unlike cysteine, the sulfur of methionine is not highly nucleophilic, although it will react with some electrophilic centers.

Perutz and Raidt, 1975 suggested that ion pairs linking portions of the protein that are juxtaposed in the structure but nonadjacent in the sequence can significantly contribute to protein thermostability. Recent information accumulated on hyperthermophilic proteins strongly supports this hypothesis. Aromatic amino acids are relatively nonpolar**.** To different degrees, all aromatic amino acids absorb ultraviolet light. Tyrosine and tryptophan absorb more than do phenylalanine; tryptophan is responsible for most of the absorbance of ultraviolet light (ca. 280 nm) by proteins. Tyrosine is the only one of the aromatic amino acids with an ionizable side chain (Vieille *et al.,* 2001).

Threonine is a polar α-amino acid. Together with serine, threonine is one of two proteinogenic amino acids bearing an alcohol group (tyrosine is not an alcohol but a phenol, since its hydroxyl group is bonded directly to an aromatic ring. It is also one of two common amino acids that bear a chiral side chain, along with isoleucine (Lehninger *et al.,* 2000; Walsh *et al.,* 2005).

The similarities of the pre-edge and extended X-ray absorption fine structure spectrum suggest that the ligand environments of the metal ions in the cobalt- and iron-containing NHases are similar. The cobalt NHases have threonine in the -V-C-(T/S)-L-C-S-C- sequence as the active site (Payne et al. 1997), whereas the ferric NHases have serine (Banerjee *et al.,* 2002; Tucker, 2008).

Table.1 Hyperthermophillic and Thermophillic Microorganisms

S.No.	Microorganism	Hyperthermophillic / Thermophillic	Acession no.
1	*Pyrobaculum arseneticum*	Hyperthermophillic	NCBI Reference Sequence: YP_001153494.1
2	*Pyrobaculumcalidifontis* JCM 11548	Hyperthermophillic	GenBank: ABO09137.1
3	*Pyrolobus fumarii* 1A	Hyperthermophillic	NCBI Reference Sequence: YP_004781857.1
4	*Pyrolobus fumarii* DSM11204	Hyperthermophillic	GenBank: AEM39605.1
5	*Pyrolobus fumarii* 1A	Hyperthermophillic	GenBank: AEM38524.1
6	*Sulfolobus solfataricus* 98/2	Hyperthermophillic	GenBank: ACX91631.1
7	*Thermococcus gammatolerans* EJ3	Hyperthermophillic	NCBI Reference Sequence: YP_002959971.1
1	*Thermosipho africanus* TCF52B	Thermophillic	NCBI Reference Sequence: YP_002335118.1
2	*Thermosipho melanesiensis* BI429	Thermophillic	GenBank: ABR30911.1
3	*Thermus thermophilus* SG0.5JP17-16	Thermophillic	GenBank: AEG34295.1
4	*Thermococcus gammatolerans* EJ3	Thermophillic	NCBI Reference Sequence: YP_002959325.1
5	*Thermus thermophilus* SG0.5JP17-16	Thermophillic	GenBank: AEG33323.1
6	*Vulcanisaeta moutnovskia*	Thermophillic	NCBI Reference Sequence: YP_004244510.1
7	*Sulfolobus islandicus*	Thermophillic	GenBank: ACR42692.1

Table.2 *In Silico* Analysis of Physicochemical Properties of Hyperthermophillic and Thermophillic Nhase Sequences

Parameters		Micro-organisms(MO's)							Average	(p < 0.01)
		1	2	3	4	5	6	7		
No. of amino acid	H	260	249	263	288	263	288	288	276.50000	0.0902 ns
	T	318	264	292	263	235	263	289		
Mol. wt.	H	27861.1	26738.7	28967.5	32032.4	28967.5	32032.4	31722.9	30787.30000	0.4776 ns
	T	35529.7	30380.3	32278.0	29987.3	26044.9	28967.5	31834.0		
Theoretical PI	H	6.77	6.85	7.72	9.44	7.72	9.44	6.18	7.78500	0.0591 ns
	T	9.39	9.68	5.89	7.80	6.16	7.72	6.18		
Negative charge	H	30	28	29	32	29	32	41	34.50000	1.2033 ns
	T	35	28	38	37	32	29	41		
Positive charge	H	30	28	30	39	30	39	39	37.00000	0.9910 ns
	T	46	41	32	38	30	30	39		
Extinction coefficient	H	35410	33920	37930	11920	37930	11920	8940	24925.00000	0.0131 ns
	T	17880	20860	34380	21890	40910	37930	8940 24410		
Instability index	H	38.72	39.91	38.68	40.27	38.68	40.27	40.89	37.04500	1.1390 ns
	T	35.58	30.71	40.46	32.66	43.38	38.68	42.11		
Aliphatic index	H	107.31	101.73	102.93	111.70	102.93	111.70	111.74	101.75	3.1623 ns
	T	91.42	100.83	105.21	90.49	103.19	102.93	113.01		
Gravy	H	0.191	0.174	0.076	-0.065	0.076	-0.065	-0.111		
	T	-0.118	-0.071	0.037	-0.221	-0.193	0.076	-0.096		

Table 3 In silico analysis of amino acid contents of hyperthermophillic and thermophillic NHase sequences

Amino acid composition		1	2	3	4	5	6	7	Average	(p < 0.01)
				Micro organisms						
Ala (A)	H	13.5%	11.2%	8.0%	7.6%	8.0%	7.6%	5.6%	9.55000	0.0715 ns
	T	6.4%	6.4%	11.3%	9.9%	14.5%	4.6%	5.2%		
Arg (R)	H	9.6%	9.2%	9.9%	8.0%	9.9%	8.0%	3.5%	7.45000	0.1154 ns
	T	6.4%	8.0%	9.6%	11.4	10.2	5.5%	3.8%		
Asn (N)	H	1.9%	2.0%	2.7%	5.2%	2.7%	5.2%	3.8%	2.95000	0.0593 ns
	T	3.8%	4.2%	0.7%	4.9%	0.9%	3.8%	3.8%		
Asp (D)	H	3.5%	3.2%	3.8%	2.4%	3.8%	2.4%	4.5%	3.95000	3.8027 ns
	T	3.8%	3.4%	4.1%	4.2%	3.4%	5.5%	4.8%		
Cys (C)	H	1.2%	1.2%	1.5%	0.3%	1.5%	0.3%	0.3%	0.90000	5.8840 *
	T	0.4%	0.4%	0.3%	0.4%	0.4%	0.4%	0.3%		
Gln (Q)	H	0.8%	0.8%	1.1%	1.7%	1.1%	1.7%	2.1%	1.90000	3.7163 ns
	T	1.9%	1.9%	2.1%	0.8%	3.0%	1.7%	2.1%		
Glu (E)	H	8.1%	8.0%	7.2%	8.7%	7.2%	8.7%	9.7%	8.70000	1.3696 ns
	T	8.0%	7.2%	8.9%	9.9%	10.2%	8.4%	9.3%		
Gly (G)	H	11.5%	11.6%	8.7%	6.6%	8.7%	6.6%	7.6%	9.10000	0.2478 ns
	T	7.2%	7.2%	9.6%	7.2%	11.1%	8.0%	7.6%		
His (H)	H	0.8%	1.2%	1.9%	2.1%	1.9%	2.1%	1.4%	1.7500	0.2824 ns
	T	1.1%	1.5%	3.1%	0.8%	1.7%	0.4%	1.4%		
Ile (I)	H	3.8%	2.0%	3.0%	8.3%	3.0%	8.3%	12.5%	6.60000	0.3083 ns
	T	8.7%	8.3%	2.4%	6.1%	0.4%	10.9%	12.8%		
Leu (L)	H	11.9%	10.4%	11.4%	12.2%	11.4%	12.2%	11.1%	14.4000	0.2604 ns
	T	10.2%	11.4%	14.0%	8.4%	20.4%	10.9%	10.7%		
Lys (K)	H	1.9%	2.0%	1.5%	5.6%	1.5%	5.6%	10.1%	5.75000	1.1082 ns
	T	9.1%	7.6%	1.4%	3.0%	2.6%	8.0%	9.7%		
Met (M)	H	1.2%	0.8%	2.7%	2.8%	2.7%	2.8%	2.1%	1.60000	6.9980 *
	T	1.1%	1.1%	0.7%	1.9%	0.4%	1.3%	1.7%		
Phe (F)	H	2.3%	3.6%	2.3%	1.7%	2.3%	1.7%	2.4%	5.00000	9.5972 **
	T	8.3%	8.0%	4.1%	6.5%	3.0%	4.2%	2.4%		
Pro (P)	H	4.2%	4.8%	6.5%	5.2%	6.5%	5.2%	4.2%	4.75000	0.7477 ns
	T	3.0%	3.0%	5.5%	5.3%	6.0%	5.9%	4.2%		
Ser (S)	H	4.2%	6.0%	6.5%	5.6%	6.5%	5.6%	8.3%	5.00000	1.0768 ns
	T	7.6%	6.8%	4.1%	3.0%	1.7%	3.4%	8.3%		
Thr (T)	H	3.5%	2.4%	3.4%	4.9%	3.4%	4.9%	3.8%	3.30000	5.3864 *
	T	2.3%	2.3%	2.7%	3.4%	1.7%	2.1%	4.2%		
Trp (W)	H	1.5%	1.6%	1.9%	0.0%	1.9%	0.0%	0.0%	1.05000	0.6587 ns
	T	0.0%	0.0%	1.0%	0.4%	2.1%	0.8%	0.0%		
Tyr (Y)	H	3.5%	3.2%	2.7%	2.8%	2.7%	2.8%	2.1%	3.70000	7.7134 *
	T	4.2%	5.3%	4.1%	4.2%	3.8%	3.8%	2.1%		
Val (V)	H	11.2%	14.5%	13.3%	8.3%	13.3%	8.3%	4.9%	8.55000	4.1476 ns
	T	6.4%	6.1%	10.3%	8.4%	2.6%	10.5%	5.2%		

* **Significant at 5 % level of significance (p < 0.05); ** Significant at 1% level of significance (p <0.01), ns non-significant**

The difference in the metal co-factors may be ascribed to the different amino acid residues at this position. The crystal structure of Co (III)-containing NHase from *Pseudonocardia thermophila* JCM3095 revealed that both Co(III) and Fe(III) reside in similar environments. A tryptophan residue (Trp 72), which may be involved in substrate binding, in the Co (III)-containing enzyme replaces the tyrosine residue in the Fe (III)-containing enzyme. This is probably responsible for the preference of cobalt NHases for aromatic rather than aliphatic nitriles (Miyanaga et al. 2001; Kubiak and Nowak, 2008; Foerstner, 2008).

In all known NHases, each subunit has a highly homologous amino acid sequence CXYCSCX) that forms the metal-binding site. Cobalt-type NHases contain threonine and tyrosine in the –C (T/S) YCSC(Y/T) sequence of the active center, whereas iron-type NHases contain serine and threonine (Kovacs, 2004; Banerjee *et al.,* 2002).

Phenylalanine (Phe) is nonpolar because of the hydrophobic nature of the benzyl side chain. Due to its hydrophobicity, phenylalanine is nearly always found buried within a protein. The π electrons of the phenyl ring can stack with other aromatic systems and often do within folded proteins, adding to the stability of the structure. Mostly this is found in the side chain of the protein. Aromatic-aromatic interactions (aromatic pairs) are defined by a distance of less than 7.0 Å between the phenyl ring centroids (Burley and Petsko 1985; Kong and King, 2011).

Among the hyperthermophilic proteins whose structures have been solved, at least one might be stabilized by extra aromatic interactions (Dong *et al.,* 1997; Abou *et al.,* 2003; Eijsink *et al.,* 2004). The solvent-exposed aromatic pair, Tyr13-Tyr17, in *B. amyloliquefaciens* RNase was replaced with

Ala or Phe residues (single and double mutations). Both Tyr-Tyr and Phe-Phe pairs contributed approximately -1.3 kcal/mol toward thermodynamic stabilization (Hong *et al.,* 2009).

Another type of interaction involving aromatic residues exists in proteins. In cation-π interactions, positive charges (most often metal cations but possibly cationic side chains of Arg and Lys) typically interact with the center of the aromatic ring. The stabilization energy of the cation-π interaction does not decrease as a function of $1/r^3$ but, rather, exhibits a $1/r^n$ dependence with n < 2, which resembles more a Coulombic ($1/r$) than a hydrophobic interaction. The low dependence of the cation-π interaction on distance—and the fact that Phe, Tyr, and Trp do not have high desolvation energies and can easily be accommodated in hydrophobic environments—makes these interactions a potential stabilization mechanism (Dougherty, 1996; Scott, 2012).

In conclusion, the results obtained shows that thermophiles and hyper thermophiles contains two hydrophobic sulphur containing amino acids (Cys and Met), may be involved in disulphide bridge formation and may interact with some metal ions to enhance stability of the protein. The other statistically significant amino acids phenylalanine and tyrosine are aromatic and bulkier hydrophobic groups may be involved in the inner part of protein providing greater stability to protein. Threonine is the only polar α-amino acid found significant in this study which may be present in the active site of cobalt NHases.

References

Abou-Hachem, M., Olsson, F., Nordberg, K. E. 2003. Probing the stability of the modular family 10 xylanase from

Rhodothermus marinus. Extremophiles 7, 483-491.

Banerjee, A., Sharma, R., Banerjee, U.C. 2002. The nitrile-degrading enzymes: current status and future prospects. Applied Microbiology 60, 33–44.

Burley, S. K. and Petsko, G.A.1985. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. Science 229, 23-29.

Chen, J. ,Zheng, R.C., Zheng, Y.G., (2009) Microbial transformation of nitriles to high-value acids or amides. Advancesin Biochemical Engineering and Biotechnology 113, 33-77.

Dong, G., Vieille, C., Savchenko, A. and Zeikus, J.G. 1997. Cloning, sequencing, and expression of the gene encoding extracellular a-amylase from *Pyrococcus furiosus* and biochemical characterization of the recombinant enzyme. Applied Environmetal Microbiology 63,3569–3576.

Dougherty, D.A. (1996) Cation-p interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science 271,163–168.

Eijsink,V.G.H., Bjork, A., Gaseidnes, S., Sirevag, R., Synstad, B., van den Burg, B. and Vriend, G. 2004. Rational engineering of enzyme stability. Journal of Biotechnology 113,105–120.

Foerstner, K.U., Doerks, T., Muller, J., Raes, J., Bork, P. 2008. A nitrile hydratase in the eukaryote *Monosiga brevicollis*. PloS ONE 3, e3976. doi: 10.1371/journal.pone.0003976

Gill, S.C. and von Hippel, P.H. 1989. Calculation of protein extinction coefficients from amino acid sequence data. Analytical Biochemestry 182, 319-326.

Guruprasad, K., Reddy, B.V.B. and Pandit, M.W. 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Engineering 4, 155-161.

Heitmann, P. 1968. A Model for Sulfhydryl Groups in Proteins. Hydrophobic Interactions of the cysteine side chain in micelles. European Journal of Biochemistry 3 (3), 346–50.

Hickey, A. and Singer, G.A.C. .2004. Genomic and proteomic adaptations to growth at high temperature. Genome Biology 5, 117.

Hong, H., Joh, N.H., Bowie J.U., Tamm L.K. 2009. Methods for measuring the thermodynamic stability of membrane proteins. Methods in Enzymology 455,213-236.

Hopmann, K. H. (2014) Full Reaction Mechanism of Nitrile Hydratase: A Cyclic Intermediate and an Unexpected Disulfide Switch. Inorganic Chemistry 53(6), 2760–2762 DOI: 10.1021/ic500091k

Ikai, A. 1980. Thermostability and aliphatic index of globular proteins. Journal of Biochemistry 88, 1895-1898.

Kallwass, H. K., W. K. Surewicz, W. Parris, E. L. Macfarlane, M. A. Luyten, C. M. Kay, M. Gold, and J. B. Jones. 1992. Single amino acid substitutions can further increase the stability of a thermophilic L-lactate dehydrogenase. Protein Engineering 5, 769-774.

Kong, F. and King, J. 2011. Contributions of aromatic pairs to the folding and stability of long-lived human γD-crystallin. Protein Science 20(3), 513—528.

Kovacs, J.A. 2004. Structural and functional analysis of iron and non-corrinoid cobalt enzymes. Chemical Reviews 104, 825–848.

Kubiak, K. and Nowak, W. 2008. Molecular

dynamics simulations of the photoactive protein nitrile hydratase. Biophysical Journal 94, 3824–3838.

Kumar, S. 2000. Factors enhancing protein thermostability. Protein Engineering 13, 179-191.

Kyte, J and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology 157, 105-132.

Lehninger, A.L., Nelson, C.D.L. and Michael, M. 2000, Principles of Biochemistry (3rd ed.), New York: W. H. Freeman, ISBN 1-57259-153-6.

Marron, A.O. Akam, M. and Walker, G. 2012. Nitrile hydratase genes are present in multiple eukaryotic supergroups. PLoS ONE 7(4), e32867. DOI: 10.1371/journal.pone.0032867

Matsumura, M., Signor, G. and Matthews, B.W. 1989. Substantial increase of protein stability by multiple disulphide bonds. Nature 342, 291–293.

Miyanaga, A., Fushinobu, S., Ito, K., Wakagi, T. 2001. Crystal structure of cobalt-containing nitrile hydratase. Biochemistry Biophysics Research Communication 288, 1169–1174.

Murakami, T., Nojiri, M., Nakayama, H., Odaka, M., Yohda, M., Dohmae, N., Takio, K., Nagamune, T., Endo, I. 2000. Post-translational modification is essential for catalytic activity of nitrile hydratase. Protein Science 9, 1024–1030.

Nagano, N., Ota, M., Nishikawa, K. 1999. Strong hydrophobic nature of cysteine residues in proteins. FEBS Letters 458 (1), 69–71.

Nagamune, T., Kurata, H., Hirata, M., Honda, J., Koike, H., Ikeuchi, M., Inoue, Y., Hirata, A., Endo, I. 1990. Purification of inactivated photo responsive nitrile hydratase. Biochemistry Biophysics Research Communication 168,437–442.

Payne, M. 1997. A stereoselective cobalt containing nitrile hydratase. Biochemistry 36, 5447–5454.

Perutz, M. F. and Raidt, H. 1975. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. Nature 255, 256–259.

Prasad, S. and Bhalla, T.C. 2010. Nitrile hydratases (NHases): at the interface of academia and industry. Biotechnology Advances 28, 725–741.

Rzeznicka, K., Schatzle, S., Bottcher, D., Klein, J., Bornscheuer, U.T. 2010. Cloning and functional expression of a nitrile hydratase (NHase) from *Rhodococcus equi* TG328-2 in *Escherichia coli*, its purification and biochemical characterization. Applied microbiology and biotechnology 85 (5), 1417-1425.

Sahay, A. and Shakya, M. 2010. *In silico* analysis and homology modelling of antioxidant proteins of Spinach. Journal of Proteomics and Bioinformatics 3,148-154.

Scott, L.R. 2012. Aomatic interacions. Digital biology, pp. 195 -200. University of Chicago, pages 279.

Tucker, T.A. M. 2008. The effect of media composition on nitrile hydratase activity and stability, and on cell envelope components of *Rhodococcus* DAP 96253. Biology Dissertations Paper 56. Georgia State University.

Trivedi, S. 2006 Protein thermostability in Archaea and Eubacteria. Genetics and Molecular Research 5 (4), 816-827.

Vieille, C., Epting, K.L., Kelly, R.M., Zeikus, J.C. 2001. Bivalent cations and amino-acid composition contribute to the thermostability of *Bacillus licheniformis* xylose isomerase. European Journal of Biochemistry 268, 6291-6301.

Walsh, C.T. Garneau-Tsodikova. S., Gatto, G.J. 2005. Protein posttranslational modifications: The chemistry of proteome diversifications. Angewandte Chemie International Edition 44: 7342 – 7372.

Woo, H.A. and Rhee, S.G. 2010. Immunoblot detection of proteins that contain cysteine sulfinic or sulfonic acids with antibodies specific for the hyperoxidized cysteine-containing sequence. In Das, D. Ed, Methods in redox signaling. Mary Ann Liebert, Inc. pp. 19-23.

Yamada, H. et al. 2001 Hydratases involved in nitrile conversion: screening, characterization and application. Chemical Record. 1, 152–161.