Review

The nitrilase superfamily: classification, structure and function

Helen C Pace and Charles Brenner

Address: Structural Biology and Bioinformatics Program, Kimmel Cancer Center, 233 S Tenth Street, Philadelphia, PA 19107, USA.

Correspondence: Charles Brenner. E-mail: brenner@dada.jci.tju.edu

Published: 15 January 2001

Genome ***Biology*** 2001, **2(1)**:reviews0001.1–0001.9

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2001/2/1/reviews/0001

Abstract

The nitrilase superfamily consists of thiol enzymes involved in natural product biosynthesis and post-translational modification in plants, animals, fungi and certain prokaryotes. On the basis of sequence similarity and the presence of additional domains, the superfamily can be classified into 13 branches, nine of which have known or deduced specificity for specific nitrile- or amide-hydrolysis or amide-condensation reactions. Genetic and biochemical analysis of the family members and their associated domains assists in predicting the localization, specificity and cell biology of hundreds of uncharacterized protein sequences.

Plants, animals and fungi perform a wide variety of nonpeptide carbon-nitrogen hydrolysis reactions using members of the nitrilase superfamily of enzymes. These nitrilase [1,2] and amidase [3,4] reactions, which produce auxin, biotin, β-alanine and other natural products, and which result in deamination of protein and amino acid substrates, all involve attack of a cyano or carbonyl carbon by a conserved cysteine [5,6]. Many bacteria and archaea, particularly those with an ecological relationship to plants and animals, encode members of the nitrilase superfamily and utilize the enzymes for chemically similar nitrile or amide hydrolysis reactions or for condensation of acyl chains to polypeptide amino termini.

On the basis of global and structure-based sequence analysis, members of the nitrilase superfamily can now be classified into 13 branches and the substrate specificity of members of nine branches can be anticipated. Despite historical classification of all of these sequences as nitrilase-related, only one branch is known to have nitrilase activity, whereas eight branches have apparent amidase or amide-condensation activities. Members of seven branches of the nitrilase superfamily have participated in domain fusion events that alter the localization of the nitrilase-related domain, link ammonia production to ammonia consumption, or potentially link proteins involved in cellular signaling. For example, fusion of domains we expect to have glutamine amidohydrolase (GAT) activity to some bacterial and all eukaryotic nicotinamide adenine dinucleotide (NAD) synthetases can account for the previously unsolved problem that only some NAD synthetases use glutamine as a source of ammonia [7-9]. Remarkably, these fusions contain the fourth apparent GAT domain involved in coupled amide transfer reactions as they are unrelated to other GAT-domain-containing families: the amino-terminal nucleophile (Ntn) hydrolases and triad amidotransferases [10], and the amidase signature family [11]. Crystal structures of two nitrilase superfamily members - worm NitFhit [12] and a bacterial *N*-carbamyl-D-amino acid amidohydrolase [13] - reveal that nitrilase-related proteins are multimeric α-β-β-α sandwich proteins that have a conserved Glu-Lys-Cys catalytic triad responsible for covalent catalysis. Mutating catalytic triad residues may allow substrates to be trapped and identified for the branches that remain to be characterized biochemically.

Evolution and classification

Members of the nitrilase superfamily appear to be found in all plants, animals and fungi, and many of these organisms have multiple nitrilase-related proteins from more than one branch of the superfamily. Nitrilase-related sequences are

also found in phylogenetically isolated prokaryotes that appear to have an ecological relationship to plants and animals. The nitrilase superfamily therefore probably emerged prior to the separation of plants, animals and fungi, radiated into families, and then spread laterally to bacteria and archaea. Some branches of the nitrilase superfamily are found only in prokaryotes; members of these branches may constitute rational antibiotic targets.

Automated sequence searching easily identifies predicted polypeptides as members of the nitrilase superfamily, but many database annotations have been applied haphazardly. Because members of the nitrilase superfamily are reported to be nitrilases, aliphatic amidases, β-ureidopropionases, β-alanine synthases, *N*-carbamyl-D-amino acid amidohydrolases and so on, these designations appear in the sequence-definition lines of multiple databases, often irrespective of the activity of the most closely related characterized enzyme.

The reactions performed by nitrilases, amidases, carbamylases and N-acyl transferases within the nitrilase superfamily are shown schematically in Figure 1. It should be noted that the nitrilase branch of the nitrilase superfamily may be the only branch that contains members that perform nitrile hydrolysis (from a nitrile to the corresponding acid plus ammonia); at least eight branches appear to be either amidases of various specificities or enzymes that condense acyl chains to amino groups. Nitrile hydratases, metal-containing enzymes that convert a nitrile to the corresponding amide [14], are not members of the nitrilase superfamily. Additionally, despite the fact that most branches of the nitrilase superfamily are actually amidases, there are many amidases including Ntn and triad hydrolases [10], amidase signature enzymes [15] and thiol proteases [16] that are unrelated to the nitrilase superfamily. Because of the historical observation that aliphatic amidases are related to nitrilases [4,6], we retain 'nitrilase' as the superfamily designation and as a branch designation, and embrace several families of homologous Glu-Lys-Cys amidases as branches of the nitrilase superfamily.

We performed a large number of BLASTp (version 2.1.2) [17] and manual searches to identify prototypical members of branches of the nitrilase superfamily and we currently classify the superfamily as having 13 branches**,** shown in Table 1. For the data uniquely classifying nitrilase sequences into 13 branches, see the Additional data file available with the online version of this article. Examination of the E-values of sequences aligned with a prototype guided the classification of each of the 176 identified sequences as a member of only one branch. Within most branches, there is a relatively sharp cutoff in E-values such that sequences with E-values greater than 1 x 10-25 can be identified as belonging to another branch. In the 13th branch, definition of a prototype - a sequence to which all branch members can be easily compared - was less straightforward as the sequences are relatively diverse. With more data, it would not be surprising to find further ways to divide and to classify members of the nitrilase superfamily.

Most members of each branch can be assigned to the branch not only by virtue of an E-value cutoff, but also by virtue of signature sequences surrounding active-site residues, providing further confidence in the classification scheme. Essentially all members of the nitrilase superfamily have a conserved, apparent catalytic triad of glutamate, lysine and cysteine (only three apparently truncated sequences lack the glutamate). The motif that most highly correlates with E-value cutoffs consists of the two residues carboxy-terminal to the cysteine nucleophile. For example, members of the nitrilase branch of the nitrilase superfamily have a Cys-Trp-Glu motif at the active-site cysteine, whereas β-ureidoproprionases have a Cys-Tyr-Gly motif. Consensus sequences for the glutamate-, lysine- and cysteine-surrounding residues of each branch of the nitrilase superfamily are shown in Figure 2.

Domain fusions in the nitrilase superfamily

In seven branches of the nitrilase superfamily, a nitrilase-related domain is fused to at least one additional conserved domain (Figure 3). In three branches, the domain fusion appears to be constitutive; that is, all members of that branch (defined by BLAST E-value and signature sequences within the nitrilase-related domains) contain the additional domain. In four branches, the additional domain(s) are not found in every member. Some of the domain-fusion events can be considered 'Rosetta Stone' fusions, in that separate polypeptides appear to be fused to coordinate biochemical reactions or cellular functions [18,19]. Other domain-fusion events appear more likely to affect cellular localization. The significance of domain fusions in branches 7 and 8, the prokaryotic and eukaryotic NAD synthetases, is discussed below.

Two independently derived families of GAT domains have been found in a variety of two-domain polypeptides that couple ammonia hydrolysis from glutamine to ammonia consumption at a second active site [10]. Glutamine phosphoribosylpyrophosphate (PRPP) amidotransferase is prototypical of enzymes that utilize a GAT domain related to the Ntn hydrolases [20], whereas GMP synthetase is prototypical of enzymes that use a triad amidotransferase domain to perform the GAT function [21]. The second active site of GMP synthetase contains a nucleotide-binding domain similar to that of ammonia-dependent NAD synthetase [22]. It has been known for more than 30 years that *Escherichia coli* NAD synthetase differs from eukaryotic NAD synthetases in that it cannot use glutamine as an ammonia source [7]. Although yeast [8] and some prokaryotic NAD synthetases [23] are glutamine dependent, they do not contain an Ntn or triad GAT domain. The glutamine-dependent NAD synthetase from *Mycobacterium*

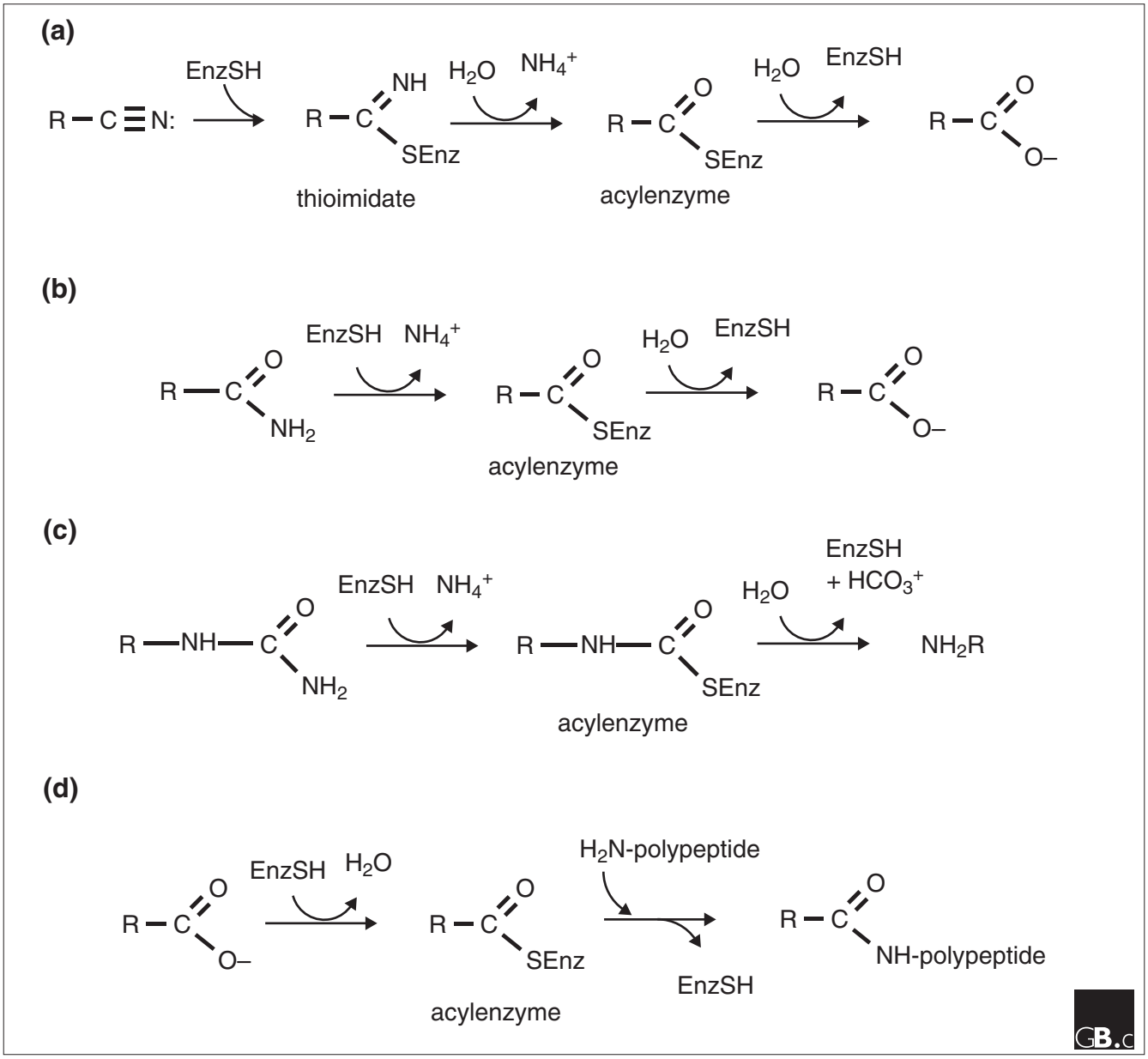

Figure 1
Four types of reaction carried out by nitrilase superfamily members. **(a)** The nitrilase reaction is performed by branch 1 enzymes. In plants, the substrate is indole-3-acetonitrile and the product is indole-3-acetic acid. **(b)** The amidase reaction is the most frequently observed activity in the superfamily. Branch 2-4 enzymes are amidases and nitrilase-related domains of branch 7 and 8 enzymes are proposed to be amidases specific for glutamine. **(c)** The carbamylase reaction is a special case of the amidase reaction, carried out by branch 5 and 6 enzymes. **(d)** Branch 9 N-acyltransferases perform the amidase reaction in reverse, transferring a fatty acid from phospholipid (not shown) to a polypeptide amino terminus. The polypeptide acceptor usually contains an amino-terminal diacylglyceride-modified cysteine (not shown). All nitrilase-related reactions are thought to proceed through acylenzyme intermediates.

tuberculosis, however, contains an amino-terminal domain [23] not present in the *E. coli* enzyme [24]. After the discovery that the multiprotein *Bacillus* glutamyl-tRNAGln amidotransferase contains yet a third type of GAT activity [11] related to the amidase signature family [15], it was hypothesized that the amino terminus of the prokaryotic glutamine-dependent NAD synthetase is related to the amidase signature family [23].

In contrast, we find that the amino terminus of prokaryotic glutamine-dependent NAD synthetase and the amino-terminal domains of all eukaryotic NAD synthetases are branches

Table 1

Summary of the enzyme activities of the nitrilase superfamily

Nitrilase branch	Nitrilase	Amidase		Reverse amidase	Carbamylase	Protein substrate
	R-C≡N	R−C(=O)−NH_2	R−C(=O)−NHR'		R−NH−C(=O)−NH_2	
1 - Nitrilase	Yes					
2 - Aliphatic amidase		Yes				
3 - Amino-terminal amidase		Yes				Yes
4 - Biotinidase		Yes	Yes			Sometimes
5 - β-Ureidopropionase		Yes			Yes	
6 - Carbamylase		Yes			Yes	
7 - Prokaryote NAD synthetase		Predicted				
8 - Eukaryote NAD synthetase		Predicted				
9 - Apolipoprotein N-acyltransferase			Yes	Yes		Yes
10 - Nit and Nitfhit						
11 - NB11						
12 - NB12						Predicted
13 - Nonfused outliers						

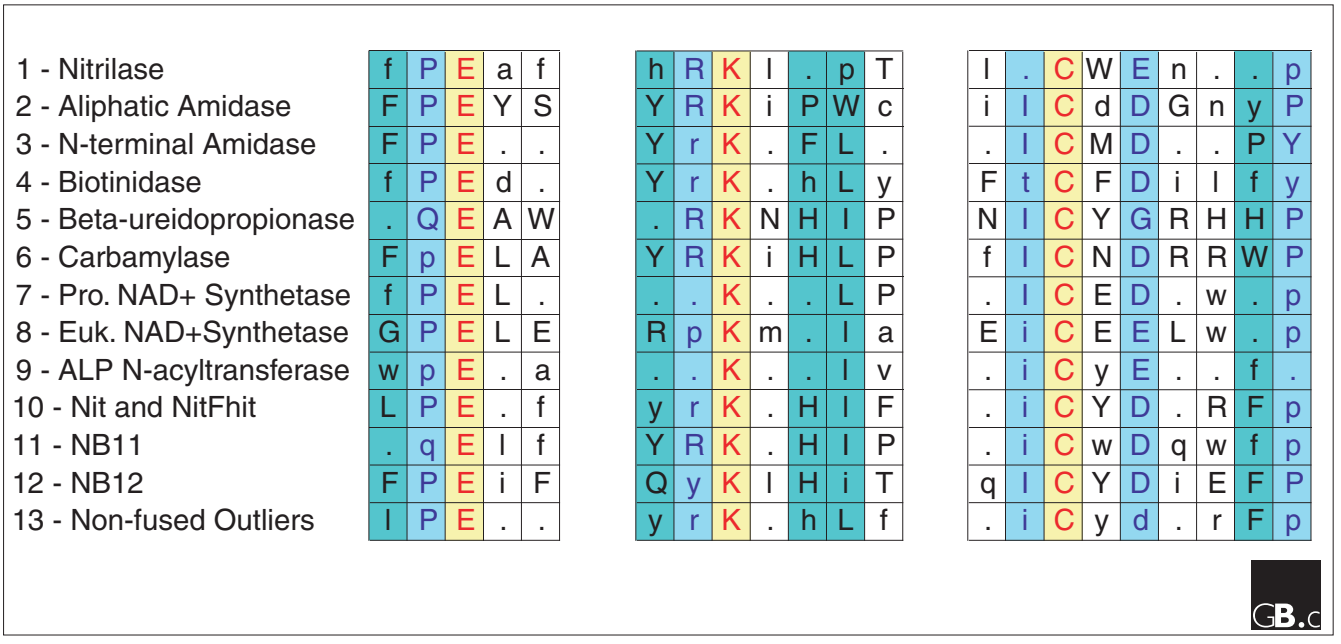

Figure 2
The nitrilase superfamily catalytic triad motifs. Consensus sequences flanking the invariant catalytic triad residues, glutamate, lysine and cysteine, were obtained by doing multiple sequence alignments within each branch [54]. Red letters on a yellow background indicate the same residue is conserved in all branches. Dark blue letters on light blue background indicate the residue is conserved in nine or more branches. Green background shows positions in which the conserved amino acid is found in six to eight of the branches. Upper case letters indicate 90% or greater consensus levels within a branch, whereas lower case are 50% or greater. Residue numbers are shown for the prototypical members of branches 1 to 12 and for the first listed member of branch 13.

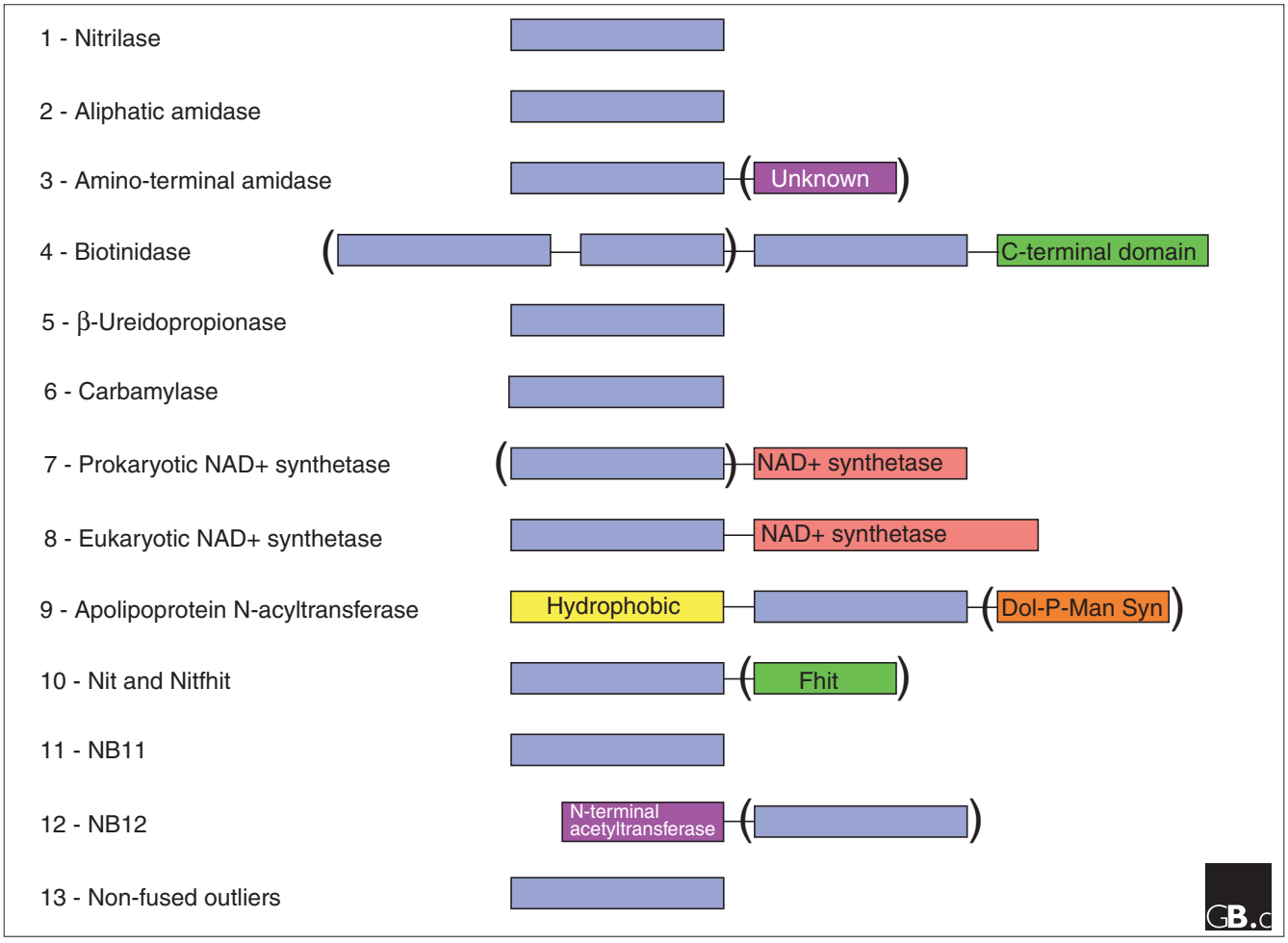

Figure 3
Domain structures for 13 branches of the nitrilase superfamily. Additional domains are found in members of seven branches. Parentheses denote domains found in only some members of the branch. In branch 4, vanins and biotinidases have carboxy-terminal domains unique to these two sub-branches and one vanin has additional full and partial nitrilase-related domains. The NAD synthetase domains of eukaryotes are always fused with a nitrilase-related domain. In contrast, only some prokaryotic NAD synthetases are fusion proteins with a nitrilase-related domain. This led to the prediction that branch 7 and 8 nitrilase domains are glutamine amidotransferases for the associated NAD synthetases (see text for details). Apolipoprotein N-acyltransferases (branch 9) always have a hydrophobic amino-terminal domain and one member is fused to an apparent dolichol phosphate mannose synthetase, which underscores the proposed function of branch 9 enzymes in post-translational modification. Nit proteins, branch 10, are found as fused Rosetta Stone proteins with Fhit in invertebrates and are coordinately expressed with separate Fhit proteins in mammals. Branch 12 enzymes are predicted to have protein substrates as they are fused to a homolog of an amino-terminal acetyltransferase.

7 and 8 of the nitrilase superfamily, respectively. We deduce that branch 7 and 8 nitrilase-related domains have substrate specificity as glutamine amidases, and that branch 7 and 8 enzymes utilize these novel GAT domains to confer glutamine dependence to the associated carboxy-terminal NAD synthetase domains. We therefore expect to find that the presence of branch 7 nitrilase-related domains will correlate with the ability of purified prokaryotic NAD synthetases to use glutamine, and we expect that the glutamine dependence of prokaryotic and eukaryotic glutamine-dependent NAD synthetases will depend on nitrilase-homologous active-site residues. If this is confirmed, branch 7 and 8 nitrilase domains will constitute the fourth independent type of GAT domain to participate in coupled amino-transfer reactions.

Enzymology

Nonenzymatic hydrolysis of a nitrile of the form R-C≡N would produce the corresponding acid amide, R-C=O(NH2), with one water addition and the corresponding acid, R-CO2-, with the second water addition. Nitrilases are interesting, however, in that the substrates are nitriles but the reaction does not involve release of, or reaction with, a substantial amount of the corresponding amide [1,25]. Nitrilases

produce the acid without the production or release of an acid amide by virtue of covalent, thiol-mediated catalysis [5,25]. As illustrated in Figure 1, the enzyme attacks a nitrile substrate covalently, producing ammonia with the first water addition, and producing acid and a regenerated enzyme with the second water addition. The geometric constraints of this reaction suggest that nitrilase facilitates interaction with a linear (approximately 180°) substrate, planar (approximately 120°) thioimidate and acylenzyme intermediates, and tetrahedral (approximately 109.5°) water-bonded intermediates. In contrast, serine and thiol proteases and amidases are confined to interacting with planar substrates and tetrahedral intermediates. We speculate that most nitrilases bind strongly to a bulky substrate R group in a conformation that places the 2 carbon closer to 120° than to 180° from the cyano nitrogen. Fitting a distorted substrate nitrile would push the substrate toward thioimidation and would reduce the geometric sweeps required of enzyme complexes. In support of this view, most nitrilases prefer bulky substrates to nonsubstituted acetonitrile [1,25-28]. Cyanide hydratase, a member of the nitrilase branch, may be the exception that proves the rule: the R-group free substrate does not stay bound to produce acid but rather is decomposed to formamide after one water addition [29,30].

As we have discussed, most branches of the nitrilase superfamily do not contain nitrilases but rather amide-hydrolyzing or amide-condensing enzymes. Although activation of water to attack planar intermediates is expected to be shared by all enzymatically active members of the superfamily, the biochemical basis for nitrile versus amide attack within the nitrilase superfamily is not yet understood. Biotinidases, branch 4 of the nitrilase superfamily, are amidases specific for hydrolysis of biotinamides such as biocytin to biotin plus lysine [31]. For this branch, leaving group specificity allows biotinylated peptides, biocytin, simple biotinamide and biotin esters to be substrates [31]. As alcohols are better leaving groups than amines, it would not be surprising if other members of the nitrilase superfamily have a biological function as esterases. Although no member of the nitrilase superfamily has been reported to have protease activity, members of branches 3 and 4 act on sidechains of polypeptides and members of branch 9 perform a condensation to polypeptide amino termini. Because branch 12 enzymes are fused to a probable amino-terminal acetyltransferase, they may have protein substrates as well. Protease activities may remain to be discovered in the superfamily. The enzyme activities of the nitrilase superfamily are summarized in Table 1.

Structural features

Crystal structures of an *N*-carbamyl-D-amino acid amidohydrolase from *Agrobacterium* [13] (a carbamylase; branch 6) and the *Caenorhabditis elegans* NitFhit Rosetta Stone protein [12] (branch 10) have been determined. The nitrilase-homologous domain of NitFhit and the carbamylase have similar three-dimensional structures, conserved chemical features, and were independently interpreted as utilizing the conserved glutamate residue as a general base for the cysteine nucleophile [12,13]. The Nit domain of NitFhit and the carbamylase can be described as α-β-β-α sandwich proteins, both of which assemble as tetramers. Nit and the carbamylase are unrelated to other enzymes with known structures such as Ntn and triad hydrolases [10], and thiol proteases [16]. Figure 4 shows the geometry of the Nit active site, highlighting residues that are absolutely conserved in the superfamily (Glu54, Lys127 and Cys169) and residues at positions that are highly conserved (Tyr125, His129, Tyr170, Asp171, Arg173 and Phe174), as aligned in Figure 2.

Branches of the nitrilase superfamily

Branch 1: nitrilase

Members of the nitrilase branch (EC 3.5.5.1) are found in plants, animals (*C. elegans*), fungi (*Saccharomyces cerevisiae*'s frequently inactivated *NIT1* gene), and many types of bacteria. The best evidence that nitrilase functions *in vivo* to convert indoleacetonitrile to the plant growth factor indole-3-acetic acid (auxin) comes from *Arabidopsis,* in which it was shown that recessive mutations in a nitrilase gene resulted in reduced sensitivity to the auxin-like effects of indoleacetonitrile and that overexpression of a nitrilase caused increased sensitivity to indoleacetonitrile [32]. Bacterial nitrilases are often exploited for biochemical syntheses and for environmental remediation [33]. It is not clear whether bacterial nitrilases primarily function in ecological relationships with plants or whether they benefit isolated microbes.

Branch 2: aliphatic amidase

Aliphatic amidases (EC 3.5.1.4) [3,4] comprise a small branch of nearly identical proteins found in *Pseudomonas, Bacillus, Brevibacteria* and *Helicobacteria*. They hydrolyze substrates such as the carboxamide sidechains of glutamine

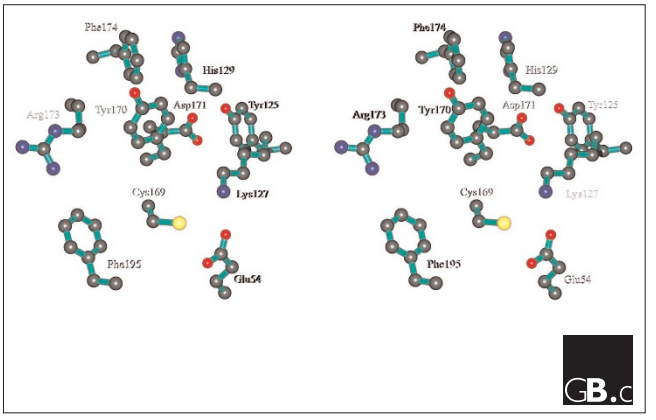

Figure 4
Nitrilase-related active site of *C. elegans* NitFhit. Stereoview of sidechains of invariant and highly conserved residues from the crystal structure of NitFhit [12].

and asparagine utilizing the conserved cysteine within the nitrilase superfamily.

Branch 3: amino-terminal amidase

The N-end rule is a means by which the rates of ubiquitin-dependent protein degradation is regulated. The *S. cerevisiae* Nta1 protein deaminates amino-terminal asparagine and glutamine residues to aspartate and glutamate, which lead to more rapid rates of protein turnover [34]. Nta1 has fungal homologs but mammalian amino-terminal amidases appear to be unrelated.

Branch 4: biotinidase

Biotinidases (EC 3.5.1.12) utilize specific amidase/esterase activity to release biotin from biotinamide, biotin-lysine and biotin-peptide conjugates and biotin methylester [35]. Biotinidase deficiency can result in an inability to recycle biotin that is manifested in neurological and cutaneous abnormalities in humans [36]. Biotinidases are secreted into serum and have a unique, conserved carboxy-terminal domain. Vanins [37] and GPI-80 [38] are members of the biotinidase branch that contain a similar carboxy-terminal domain containing, in addition, a GPI anchor and are involved in T-cell thymic homing and neutrophil adherence and migration. One member of this branch contains repeated nitrilase-related domains. Recently, porcine panthetheinase (EC 3.5.1.-), an amidase that converts pantetheine to panthothenate plus cysteamine in the dissimilative pathway of CoA, was sequenced and found to be nearly identical to vanins [39]. Although the biologically important substrate of vanins remains unproven, sequence and enzymatic similarity with biotinidases suggest that an amine molecule at least the size of an amino acid (that is, bigger than ammonia) may be the leaving group. Branch 4 enzymes are the only amidases in the nitrilase superfamily known to prefer secondary amine substrates of the form R-C=O(NHR') as opposed to simple acid amides. An extensive archive of vanins, including 118 expressed sequence tag (EST) sequences is available [40,41].

Branch 5: β-ureidopropionase

The β-ureidopropionases (EC 3.5.1.6) are enzymes involved in the catabolism of pyrimidine bases and the production of β-alanine [42]. Substrates of this enzyme are of the carbamylase type (see Figure 1c) and the amine product is usually a non-standard amino acid.

Branch 6: carbamylase

A variety of bacteria express hydrolases specific for the decarbamylation of D-amino acids. These enzymes have been exploited in the production of semisynthetic β-lactam antibiotics [43] and are now represented by the structure of the *Agrobacterium* enzyme [13].

Branches 7 and 8: glutamine-dependent NAD synthetase

As discussed earlier, the presence of a nitrilase-related domain appears to correlate with the ability of bacterial NAD synthetase (EC 6.3.5.1) to utilize glutamine as an ammonia source. Eukaryotic NAD synthetases always contain this novel, putative GAT domain and exhibit glutamine dependence. Substrate specificity of nitrilase-related proteins as glutamine amidases is not surprising given the specificity of the branch 2 and 3 enzymes. It remains to be seen how glutamine-dependent NAD synthetase may channel ammonia from the nitrilase-related active site to the NAD active site.

Branch 9: apolipoprotein N-acyltransferase

The modification and processing of Braun's lipoprotein, a major component of the outer membrane of *E. coli*, has been studied for decades [44]. Defects in this post-translational modification pathway are associated with copper sensitivity [45]. The apolipoprotein becomes proteolized, exposing an amino-terminal cysteine. After the cysteine is modified by diacylglycerol, branch 9 enzymes condense a fatty acid to the amino terminus of the modified cysteine residue.

Branch 10: Nit

Nit was originally identified as an approximately 300 amino acid amino-terminal extension on fly and worm homologs [46] of the human [47] and murine [48] Fhit tumor suppressor protein. Nit homologs are found in organisms with Fhit homologs [12] and, in the mouse, *Nit1* and *Fhit* mRNA levels are highly correlated in seven of eight tissues examined [46]. Satisfaction of these criteria suggested that NitFhit is a Rosetta Stone protein, whose fusion might decode a previously unsuspected interaction between the proteins [18,19]. As Fhit is part of a cell-death pathway that is not clearly connected to known apoptotic players [49,50], identification of Nit as a Fhit-interacting protein was welcomed. The Fhit active site of NitFhit has been characterized and the structure of worm NitFhit has been elucidated [12], but the Nit substrate, cell biology and relationship to tumor suppression are not known.

The most striking feature of the Nit-Fhit interaction apparent from the crystal structure of the worm protein is that the complex assembles with a central Nit tetramer binding two Fhit dimers [12]. The carboxy-terminal β strands of Nit-conserved polypeptide sequences exit the compact Nit tetramer domain and physically interact with Fhit dimers. Fhit dimers are bound in a way that allows them to expose diadenosine polyphosphate-binding sites opposite from the Nit interaction surface [12]. Futhermore, the nucleotide kinetics of NitFhit active sites [12] were extremely similar to those of human Fhit dimers in the absence of Nit [51].

Concord between the phylogenetic profiles [52] of Fhit and Nit breaks down slightly with the discovery of Nit-related sequences in a small number of prokaryotes that have no Fhit homolog (see the Additional data file). The idea that nitrilase-related proteins spread from animals and plants to prokaryotes is, however, supported by the animal-associated ecology of these microbes.

Branches 11-13

Branches 11 and 12 contain distinct similarity groups with no characterized member. Branch 12 may contain Rosetta Stone [18,19] proteins in that a distinctive nitrilase-related domain is found fused to an amino-terminal domain of approximately 210 amino acids. The branch-12-associated domain is related to the RimI [53] superfamily of amino-terminal acetyltransferases, suggesting that branch 12 enzymes are involved in post-translational modifications. Branch 13 contains uncharacterized, nonfused nitrilase-related proteins that are difficult to place in a distinct similarity group.

Conclusions

On the basis of newly obtained structures of nitrilase-related proteins and the available literature, we have provided a classification of all available nitrilase-related sequences. Every activity appears to work through a thiol acylenzyme intermediate and depend on a novel Glu-Lys-Cys catalytic triad. No activity forms or hydrolyzes a peptide bond, yet several affect post-translational modifications of lysine or carboxyamide sidechains or polypeptide amino termini. Other activities are involved in natural product biosynthesis and other metabolic pathways. Activities on amide substrates are found in at least eight branches of the superfamily. Activity on nitrile substrates has only been found in one branch. Membership in branches, based on BLAST E-value and structure-based signature sequence analysis, appears to correlate well with distinct substrate specificity and biological activities in all branches for which experimental data are available. Fusions between nitrilase-related domains and other conserved sequences are extremely common in the nitrilase superfamily. Fusions with NAD synthetase domains are here interpreted as solving a 30 year old problem: two branches of the nitrilase superfamily are posited to be novel GAT domains that account for the glutamine dependence of some bacterial and all eukaryotic NAD synthetases.

Additional data

The following additional data file is available (in HTML format): links for the 176 sequences in the 13-branch classification system of the nitrilase superfamily. The additional data file can be accessed from: http://genomebiology.com/2001/2/1/reviews/0001/gb-2001-2-1-reviews0001-S1.asp

Acknowledgements

We thank Janet L. Smith for helpful comments. Work was supported by National Cancer Institute grant P01CA77738.

References

1. Harper DB: **Characterization of a nitrilase from Nocardia sp. (Rhodochrous group) N.C.I.B. 11215, using p-hydroxybenzonitrile as sole carbon source.** *Intl J Biochem* 1985, **17:**677-683.
2. Harper DB: **Microbial metabolism of aromatic nitriles.** *Biochem J* 1977, **165:**309-319.
3. Ambler RP, Auffret AD, Clarke PH: **The amino acid sequence of the aliphatic amidase from *Pseudomonas aeruginosa*.** *FEBS Lett* 1987, **215:**285-290.
4. Novo C, Tata R, Clemente A, Brown PR: ***Pseudomonas aeruginosa* aliphatic amidase is related to the nitrilase/cyanide hydratase enzyme family and Cys166 is predicted to be the active site nucleophile of the catalytic mechanism.** *FEBS Lett* 1995, **367:**275-279.
5. Stevenson DE, Feng R, Storer AC: **Detection of covalent enzyme-substrate complexes of nitrilase by ion-spray mass spectroscopy.** *FEBS Lett* 1990, **277:**112-114.
6. Bork P, Koonin EV: **A new family of carbon-nitrogen hydrolases.** *Protein Sci* 1994, **3:**1344-1346.
7. Spencer RL, Preiss J: **Biosynthesis of diphosphopyridine nucleotide. The purification and the properties of diphospyridine nucleotide synthetase from *Escherichia coli* b.** *J Biol Chem* 1967, **242:**385-392.
8. Yu CK, Dietrich LS: **Purification and properties of yeast nicotinamide adenine dinucleotide synthetase.** *J Biol Chem* 1972, **247:**4794-4802.
9. Zalkin H: **NAD synthetase.** *Methods Enzymol* 1985, **113:**297-302.
10. Zalkin H, Smith JL: **Enzymes utilizing glutamine as an amide donor.** *Advances in Enzymology and Related Areas of Molecular Biology* 1998, **72:**87-144.
11. Curnow AW, Kw H, Yuan R, Si K, Martins O, Winkler W, Henkin TM, Soll D: **Glu-tRNAGln amidotransferase: a novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation.** *Proc Natl Acad Sci USA* 1997, **94:**11819-11826.
12. Pace HC, Hodawadekar SC, Draganescu A, Huang J, Bieganowski P, Pekarsky Y, Croce CM, Brenner C: **Crystal structure of the worm NitFhit Rosetta Stone protein reveals a Nit tetramer binding two Fhit dimers.** *Curr Biol* 2000, **10:**907-917.
13. Nakai T, Hasegawa T, Yamashita E, Yamamoto M, Kumasaka T, Ueki T, Nanba H, Ikenaka Y, Takahashi S, Sato M, *et al.*: **Crystal structure of *N*-carbamyl-D-amino acid amidohydrolase with a novel catalytic framework common to amidohydrolases.** *Structure* 2000, **8:**729-737.
14. Huang W, Jia J, Cummings J, Nelson M, Schneider G, Lindqvist Y: **Crystal structure of nitrile hydratase reveals a novel iron centre in a novel fold.** *Structure* 1997, **5:**691-699.
15. Patricelli MP, Cravatt BF: **Clarifying the catalytic roles of conserved residues in the amidase signature family.** *J Biol Chem* 2000, **275:**19177-19184.
16. Rawlings ND, Barrett AJ: **MEROPS: the peptidase database.** *Nucleic Acids Res* 2000, **28:**323-325.
17. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25:**3389-3402.
18. Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D: **Detecting protein function and protein-protein interactions from genome sequences.** *Science* 1999, **285:**751-753.
19. Marcotte E, Pellegrini M, Thompson M, Yeates T, Eisenberg D: **A combined algorithm for genome-wide prediction of protein function.** *Nature* 1999, **402:**83-86.
20. Smith JL, Zaluzec EJ, Wery JP, Niu L, Switzer RL, Zalkin H, Satow Y: **Structure of the allosteric regulatory enzyme of purine biosynthesis.** *Science* 1994, **264:**1427-1433.
21. Tesmer JJ, Klem TJ, Deras ML, Davisson VJ, Smith JL: **The crystal structure of GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two enzyme families.** *Nat Struct Biol* 1996, **3:**74-86.
22. Rizzi M, Nessi C, Mattevi A, Coda A, Bolognesi M, Galizzi A: **Crystal structure of NH3-dependent NAD+ synthetase from *Bacillus subtilis*.** *EMBO J* 1996, **15:**5125-5134.
23. Cantoni R, Branzoni M, Labo M, Rizzi M, Riccardi G: **The MTCY428.08 gene of *Mycobacterium tuberculosis* codes for NAD+ synthetase.** *J Bacteriol* 1998, **180:**3218-3221.
24. Willison JC, Tissot G: **The *Escherichia coli* efg gene and the Rhodobacter capsulatus adgA gene code for NH3-dependent NAD synthetase.** *J Bacteriol* 1994, **176:**3400-3402.
25. Stevenson DE, Feng R, Dumas F, Groleau D, Mihoc A, Storer AC: **Mechanistic and structural studies on Rhodococcus ATCC 39484 nitrilase.** *Biotech Appl Biochem* 1992, **15:**283-302.

26. Stalker DM, Malyj LD, McBride KE: **Purification and properties of a nitrilase specific for the herbicide bromoxynil and corresponding nucleotide sequence analysis of the bxn gene.** *J Biol Chem* 1988, **263:**6310-6314.
27. Kobayashi M, Nagasawa T, Yamada H: **Nitrilase of *Rhodococcus rhodochrous* J1, Purification and characterization.** *Eur J Biochem* 1989, **182:**349-356.
28. Schmidt RC, Muller A, Hain R, Bartling D, Weiler EW: **Transgenic tobacco plants expressing the Arabidopsis thaliana nitrilase II enzyme.** *Plant J* 1996, **9:**683-691.
29. Wang P, VanEtten HD: **Cloning and properties of a cyanide hydratase gene from the phytopathogenic fungus *Gloeocercospora sorghi*.** *Biochem Biophys Res Commun* 1992, **187:**1048-1054.
30. Cluness MJ, Turner PD, Clements E, Brown DT, O'Reilly C: **Purification and properties of cyanide hydratase from *Fusarium lateritium* and analysis of the corresponding chy1 gene.** *J Gen Microbiol* 1993; **139**:1807-1815.
31. Hymes J, Wolf B: **Biotinidase and its roles in biotin metabolism.** *Clinica Chimica Acta* 1996, **255:**1-11.
32. Normanly J, Grisafi P, Fink GR, Bartel B: **Arabidopsis mutants resistant to the auxin effects of indole-3-acetonitrile are defective in the nitrilase encoded by the NIT1 gene.** *Plant Cell* 1997, **9:**1781-1790.
33. Cowan D, Cramp R, Pereira R, Graham D, Almatawah Q: **Biochemistry and biotechnology of mesophilic and thermophilic nitrile metabolizing enzymes.** *Extremophiles* 1998, **2:**207-216.
34. Baker RT, Varshavsky A: **Yeast N-terminal amidase. A new enzyme and component of the N-end rule pathway.** *J Biol Chem* 1995, **270:**12065-12074.
35. Cole H, Reynolds TR, Lockyer JM, Buck GA, Denson T, Spence JE, Hymes, Wolf B: **Human serum biotinase: cDNA cloning, sequence, and characterization.** *J Biol Chem* 1994, **269:**6566-6570.
36. Pomponio RJ, Hymes J, Reynolds TR, Meyers GA, Fleischhauer K, Buck GA, Wolf B: **Mutations in the human biotinidase gene that cause profound biotinidase deficiency in symptomatic children: molecular, biochemical, and clinical analysis.** *Pediatric Research* 1997, **42:**840-848.
37. Aurrand-Lions M, Galland F, Bazin H, Zakharyev VM, Imhof BA, Naquet P: **Vanin-1, a novel GPI-linked perivascular molecule involved in thymus homing.** *Immunity* 1996, **5:**391-405.
38. Suzuki K, Watanabe T, Sakurai S, Ohtake K, Kinoshita T, Araki A, Fujita T, Takei H, Takeda Y, Sato Y, Yamashita T, Araki Y, Sendo F: **A novel glycosylphosphatidyl inositol-anchored protein on human leukocytes: a possible role for regulation of neutrophil adherence and migration.** *J Immunol* 1999, **162:**4277-4284.
39. Maras B, Barra D, Dupre S, Pitari G: **Is pantetheinase the actual identity of mouse and human vanin-1 proteins?** *FEBS Lett* 1999, **461:**149-152.
40. Granjeaud S, Naquet P, Galland F: **An ESTs description of the new Vanin gene family conserved from fly to human.** *Immunogenetics* 1999, **49:**964-972.
41. **Vanin project** [http://tagc.univ-mrs.fr/pub/vanin]
42. Kvalnes-Krick KL, Traut TW: **Cloning, sequencing, and expression of a cDNA encoding beta-alanine synthase from rat liver.** *J Biol Chem* 1993, **268:**5686-5693.
43. Louwrier A, Knowles CJ: **The aim of industrial enzymic amoxycillin production: characterization of a novel carbamoylase enzyme in the form of a crude, cell-free extract.** *Biotechnol Appl Biochem* 1997, **25:**143-149.
44. Tokunaga M, Tokunaga H, Wu HC: **Post-translational modification and processing of *Eschericia coli* prolipoprotein *in vitro*.** *Proc Natl Acad Sci USA* 1982, **79:**2255-2259.
45. Rogers SD, Bhave MR, Mercer JFB, Camakaris J, Lee BTO: **Cloning and characterization of *cutE*, a gene involved in copper transport in *Escherichia coli*.** *J Bacteriol* 1991, **173:**6742-6748.
46. Pekarsky Y, Campiglio M, Siprashvili Z, Druck T, Sedkov Y, Tillib S, Draganescu A, Wermuth P, Rothman JH, Huebner K, Buchberg AM, Mazo A, Brenner C, Croce CM: **Nitrilase and Fhit homologs are encoded as fusion proteins in *Drosophila melanogaster* and *Caenorhabditis elegans*.** *Proc Natl Acad Sci USA* 1998, **95:**8744-8749.
47. Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, *et al.*: **The *FHIT* gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers.** *Cell* 1996, **84:**587-597.
48. Fong LYY, Fidanza V, Zanesi N, Lock LF, Siracusa LD, Mancini R, Siprashvili Z, Ottey M, Martin SE, Dolsky R, Druck T, McCue PA, Croce CM, Huebner K: **Muir-Torre-like syndrome in *FHIT* deficient mice.** *Proc Natl Acad Sci USA* 2000, **97:**4742-4747.
49. Sard L, Accornero P, Tornielli S, Delia D, Bunone G, Campiglio M, Colombo MP, Gramegna M, Croce CM, Pierotti MA, *et al.*: **The tumor-suppressor gene *FHIT* is involved in the regulation of apoptosis and in cell cycle control.** *Proc Natl Acad Sci USA* 1999, **96:**8489-8492.
50. Ji L, Fang B, Yeh N, Fong K, Minna JD, Roth JA: **Induction of apoptosis and inhibition of tumorigenicity and tumor growth by adenovirus vector-mediated fragile histidine triad (*FHIT*) gene overexpression.** *Cancer Res* 1999, **59:**3333-3339.
51. Draganescu A, Hodawadekar SC, Gee KR, Brenner C: **Fhit-nucleotide specificity probed with novel fluorescent and fluorigenic substrates.** *J Biol Chem* 2000, **275:**4555-4560.
52. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO: **Assigning protein functions by comparative genome analysis: protein phylogenetic profiles.** *Proc Natl Acad Sci USA* 1999, **96:**4285-4288.
53. Yoshikawa A, Isono S, Sheback A, Isono K: **Cloning and nucleotide sequencing of the genes rimI and rimJ which encode enzymes acetylating ribosomal proteins S18 and S5 of Esherichia coli K12.** *Mol Gen Genet* 1987, **209:**481-488.
54. Corpet F: **Multiple sequence alignment with hierarchical clustering.** *Nucleic Acids Res* 1988, **16:**10881-10890.