Nitrilases are enzymes that hydrolyze the carbon-nitrogen triple bond (Pace and Brenner, 2001). This hydrolytic reaction generally yields ammonia and a carboxylic acid, depending upon the original substrate. Occasionally, nitrilases yield an amide product as well, in addition to, or instead of, ammonia and the acid product. Nitrilases tend to have preferences towards specific substrates or categories of substrates. Commonly found throughout nature, the nitrilases form one branch of the much larger nitrilase superfamily which have common features of an αββα fold and three catalytic residues: a cysteine, glutamate, and lysine that are readily identified by sequence alignment. The importance of the cysteine residue was first shown by Kobayashi et al. (1992).

The simplest of the nitrile substrates is cyanide, HCN, or frequently a CN salt with a different cation. Although a variety of microbial pathways for cyanide assimilation or remediation have been characterized, most require inputs of energy, co-factors or use of intact organisms. Nitrilases, however, do not require energy or co-factors to degrade their substrate and the purified enzymes have robust activity against their substrates. This makes the cyanide degrading nitrilases prime candidates for the large-scale bioremediation of cyanide waste, especially in waste-water streams but elsewhere as well (Banerjee et al., 2002; Gupta et al., 2010; Thuku et al., 2009).

The nitrilases that catalyze cyanide hydrolysis are of two types: cyanide hydratase or cyanide dihydratase. The first of these, cyanide hydratase (CHT), is found broadly throughout the fungal kingdom but best studied variants are from *Neurospora crassa*, *Gloeocercospora sorghi* and *Fusarium lateritium* (Basile et al., 2008; Fry and Munch 1975; Jandhyala et al., 2005). It converts cyanide into formamide as the product, HCN + H₂O → CH₃NO. In nature there are cellular amidases that normally convert the formamide into formate and ammonia, hence fungal cultures generate the carboxylic acid product, but purified nitrilase does not.

Analysis of the Genbank database indicates numerous variants that closely align with the archetypal *Gloeocercospora sorghi* sequence (M99044). The conservation at the protein level is generally greater than 80% identity across the entire protein length among the known or presumed cyanide hydratases (excluding the highly variant C-terminus), more than 60 such homologs are found in the database (BlastP against non-redundant protein database of December 2016). Homologs with somewhat lower sequence identity (between 70–80%) are also revealed by Blast searching although these have not been demonstrated to have activity on cyanide. Analyzing those matches with yet lower levels of identity reveal a variety of nitrilases with either known activity on other substrates or unknown activities. Not surprisingly the sequences are similarly conserved at the DNA level.

The second member of this family, cyanide dihydratase (CynD), is of bacterial origin. It is frequently referred to as a cyanidase and converts cyanide into ammonia and formate: HCN + 2H₂O → NH₃ + HCOOH (Jandhyala et al., 2003, 2005; Watanabe et al., 1998). Structural similarities between these two groups of cyanide-degrading nitrilases and other nitrilases are abundant. In addition to the αββα fold motif and the conserved catalytic residues, their higher-order structure is that of a left-handed spiral oligomer, generally with 10 or more subunits. Hence, the native proteins have molecular masses ranging from 300 kDa to over 1 mDa depending upon species and conditions. The sequence conservation is generally quite high with a notable exception at the protein C-terminus that is highly variable (Jandhyala et al., 2005). Even among proteins from the same organism, strain types with otherwise
Cyanide-degrading nitrilases in nature

Fig. 1. Phylogenetic tree of various nitrilases built using ClustalX (Larkin et al., 2007) and drawn using FigTree (http://tree.bio.ed.ac.uk). Sequences are, in order from top, benzonitrilase from *Aeribacillus pallidus* Dac521 ABH04285; *Bacillus* sp. Ox81 WP_041072590; aliphatic nitrilase from *Coomonas testosteroni* AAA82085; CynD from *B. pumilus* C1 AAN77004; CynD from *B. pumilus* 8A3 AAN77003; CynD from *P. stutzeri* BAA11653; CynD from *Alcaligenes* DN25 AOQ26352; aliphatic nitrilase J1 from *R. rhodochrous* Q03217; bromoxynil-specific nitrilase from *K. pneumoniae* ozaenae; and CHT from Aspergillus nidulans XP_682042; Gloeocercospora sorgii P32964; Leptosphaeria maculans Q9P8V3; Fusarium graminearum XP_011324397; Neurospora crassa XP_960160.

Fig. 2. Sequence alignment between *P. stutzeri* AK61 (PSt), *Alcaligenes* DN25 (Alk) and *B. pumilus* C1 (BPu) CynD proteins. Differing residues between the protein sequences are denoted by *.
nearly identical sequences can diverge across the terminal 30–50 residues. Surprisingly, the CynD protein is far more similar to other bacterial nitrilases with different substrates than it is to CHT proteins: CynD and CHT only have about 25–30% amino acid identity and are therefore quite dissimilar (Supplementary Fig. S1). In fact, CynD and CHT are more similar to other nitrilases than to each other. This can be clearly portrayed in a phylogenetic tree (Fig. 1). Although not closely related by sequence, the CynD and CHT enzymes have notable structural relatedness, as do many nitrilases. Aligning structural models of the B. pumilus CynD and the N. crassa CHT shows this relatedness (Supplementary Fig. S2).

Unlike the more commonly found CHT enzyme, CynD is found more rarely in nature and to date has only been characterized from three bacteria (protein sequences aligned in Fig. 2), B. pumilus (Jandhyala et al., 2003; Meyers et al., 1991, 1993), Pseudomonas stutzeri AK61 (Sewell et al., 2003; Watanabe et al., 1998) and Alcaligenes xylosoxidans (Banerjee et al., 2006; Ingversen et al., 1991). The limited distribution of the enzymes leads to a number of questions. Is there evidence for a single derivation of the gene followed by horizontal gene transfer? Or did the genes evolve independently, presumably from other more distant nitrilases.

Using either the B. pumilus C1 (AF492815), or the P. stutzeri AK61 (D82961), protein sequence to probe the current Genbank databases (December 2016), similar profiles were revealed. About 60 high identity protein sequences were found in the database. These are found primarily in the genomes of Gram-positive bacteria, and homologs in the Bacillus species were seen from Bacillus safensis and Bacillus thuringiensis, in addition to B. pumilus. Additional homologs were found in species of Clostridium, Brevibacillus, and Paenibacillus. These predominantly soil and rhizosphere residents are closely related. All 3 genomes published to date of B. pumilus, as well as a randomly selected B. pumilus type strain (8A3; AF492814) obtained from ATCC, all carry cynD. However, cynD is not found in many other Clostridium or Bacillus strains. This suggests that cynD is resident in some, but certainly not most, of these soil bacterial genomes.

In contrast, slightly more distantly related CynD proteins (Fig. 2) are found in the Gram-negative P. stutzeri, as well as a recently sequenced homolog from an Alcaligenes sp DH25 (KT965153) (whose relatedness to a patented strain Alcaligenes xylosoxidans subs denitrificans (Ingversen et al., 1991) can be inferred but not proven). Using the P. stutzeri protein sequence as a probe, there is a 99% identity with the Alcaligenes protein but only 82% with the next highest homolog, which is a Gram-positive Clostridium. This is out of more than 11,000 completed genomes plus hundreds of thousands of genes from sequences and incomplete genomes. Surprisingly, complete genomes of other P. stutzeri strains lack the cynD allele.

The orthologous P. stutzeri sequence, and the Alcaligenes sequence, are 99% identical at the DNA level (and at the protein level), so they are essentially identical genes. The rarity of cynD in Gram-negative bacteria, including its absence from most P. stutzeri and Alcaligenes genomes, suggests its presence is likely a result of specific environmental selection. In fact, P. stutzeri AK61 was isolated from cyanide contaminated wastewater from a metal plating plant. It is worth noting that neither Gram-negative strain carrying cynD has had its genome sequenced. Therefore, it is open to speculation whether the cynD allele is encoded within the genome or found on a plasmid. The latter would readily explain the near identity of the two genes found in P. stutzeri and Alcaligenes, but this point remains open.

The final question to address is the relatedness of the Gram-positive and Gram-negative cynD genes, typified by the P. stutzeri and B. pumilus alleles. A simple search using either a DNA sequence as a probe in Blast with default settings does not detect the other orthologue. Furthermore, the GC content of B. pumilus cynD is 39% (B. pumilus overall GC is 41%) and that of P. stutzeri cynD is 52% (P. stutzeri genomes are generally above 60%), supporting the conclusion that the presence of cynD in the Gram-negative bacteria is not the result of recent horizontal gene transfer, because the alleles differ significantly (Supplementary Fig. S3).

To distinguish between the possibilities of whether the Gram-negative cynD was the result of an older horizontal gene transfer, or the convergent evolution of an existing nitrilase found in each bacterium, towards a common substrate, we asked the simple question as to whether the P. stutzeri cynD gene is more similar to that from B. pumilus or to other nitrilases found in Gram-negative bacteria. The unambiguous finding is that the cynD alleles are more similar to each other than to other nitrilases, even those from the same phylum.

A brief summary of our conclusions would suggest the following points:
- CHT and CynD, despite using the same substrate, are notably dissimilar in sequence as one might expect from highly divergent nitrilases that yield different end products.
- CynD from Gram-positive bacteria is genome encoded with high amino acid identity in the more than 60 genomes where it has been found to date.
- CynD from Gram-negative bacteria is very rare and may well be plasmid encoded. The 2 existing gene sequences are identical at above 99%.
- The Gram-negative cynD alleles are relatively dissimilar to the Gram-positive cynD alleles, these differences can primarily be accounted for by differences in codon usage based on the differences in GC content of these organisms.
- Most striking is the relative absence of cynD from the thousands of Gram-negative genomes or environmental samples sequenced to date.

Acknowledgments

This work was supported by the Welch Foundation [A1310], the Texas Hazardous Waste Research Center [513TAM0032H], and the National Research Foundation of South Africa. All authors declare that they do not have conflicts of interest. This article does not contain any studies with human participants or animals performed by any of the authors.
Cyanide-degrading nitrilases in nature 93

Supplementary Materials

Supplementary figures are available in our J-STAGE site (http://www.jstage.jst.go.jp/browse/jgam).

References

Banerjee, A., Sharma, R., and Banerjee, U. C. (2002) The nitrile-degrading enzymes: current status and future prospects. Appl. Microbiol. Biotechnol., 60, 33–44.

Banerjee, A., Kaul, P., and Banerjee, U. C. (2006) Purification and characterization of an enantioselective arylacetonitrilase from Pseudomonas putida. Arch Microbiol., 184, 407–418.

Basile, L. J., Willson, R. C., Sewell, B. T., and Benedik, M. J. (2008) Genome mining of cyanide-degrading nitrilases from filamentous fungi. Appl. Microbiol. Biotechnol., 80, 427–435.

Fry, W. E. and Munch, D. C. (1975) Hydrogen cyanide detoxification by Gloeocercospora sorghi. Physiological Plant Pathology, 7, 23–33.

Gupta, N., Balomajumder, C., and Agarwal, V. K. (2010) Enzymatic mechanism and biochemistry for cyanide degradation: a review. J. Hazard Mater., 176, 1–13.

Ingversen, K., Højjer-Pedersen, B., and Godtfredsen, S. E. (1991) Novel cyanide-hydrolyzing enzyme from Alcaligenes xylosoxidans subsp. denitrificans. Appl. Environ. Microbiol., 57, 1783–1789.

Jandhyala, D., Berman, M., Meyers, P. R., Sewell, B. T., Willson, R. C. et al. (2003) CynD, the cyanide dihydratase from Bacillus pumilus: gene cloning and structural studies. Appl. Environ. Microbiol., 69, 4794–4805.

Jandhyala, D. M., Willson, R. C., Sewell, B. T., and Benedik, M. J. (2005) Comparison of cyanide-degrading nitrilases. Appl. Microbiol. Biotechnol., 68, 327–335.

Kobayashi, M., Komeda, H., Yanaka, N., Nagasawa, T., and Yamada, H. (1992) Nitrilase from Rhodococcus rhodochrous J1. Sequencing and overexpression of the gene and identification of an essential cysteine residue. J. Biol. Chem., 267, 20746–20751.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A. et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics, 23, 2947–2948.

Meyers, P. R., Gokool, P., Rawlings, D. E., and Woods, D. R. (1991) An efficient cyanide-degrading Bacillus pumilus strain. J. Gen. Microbiol., 137, 1397–1400.

Meyers, P. R., Rawlings, D. E., Woods, D. R., and Lindsey, G. G. (1993) Isolation and characterization of a cyanide dihydratase from Bacillus pumilus Ci. J. Bacteriol., 175, 6105–6112.

Pace, H. C. and Brenner, C. (2001) The nitrilase superfamily: classification, structure and function. Genome Biol., 2, RE-VIEWS0001.0001-0001.0009.

Sewell, B. T., Berman, M. N., Meyers, P. R., Jandhyala, D., and Benedik, M. J. (2003) The cyanide degrading nitrilase from Pseudomonas stutzeri AK61 is a two-fold symmetric, 14-subunit spiral. Structure, 11, 1413–1422.

Thuku, R. N., Brady, D., Benedik, M. J., and Sewell, B. T. (2009) Microbial nitrilases: versatile, spiral forming, industrial enzymes. J. Appl. Microbiol., 106, 703–727.

Watanabe, A., Yano, K., Ikebukuro, K., and Karube, I. (1998) Cyanide hydrolysis in a cyanide-degrading bacterium, Pseudomonas stutzeri AK61, by cyanidase. Microbiology, 144, 1677–1682.