Biotechnol. Bioprocess Eng. **2002**, 7: 155-162

Characterization and Evaluation of a Distinct Fusion Ability in the Functionally Related Cyclic Amidohydrolase Family Enzymes

Geun-Joong Kim[1], Dong-Eun Lee[2], and Hak-Sung Kim[2]*

[1] Department of Molecular Science and Technology, Ajou University, Suwon 442-749, Korea
[2] Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

Abstract The cyclic amidohydrolase family enzymes, which include allantoinase, dihydroorotase, dihydropyrimidinase and (phenyl)hydantoinase, are metal-dependent hydrolases and play a crucial role in the metabolism of purine and pyrimidine *in vivo*. Each enzyme has been independently characterized, and thus well documented, but studies on the higher structural traits shared by members of this enzyme family are rare due to the lack of comparative study. Here, we report upon the expression in *E. coli* cells of maltose-binding protein (MBP)- and glutathione S-transferase (GST)-fused cyclic amidohydrolase family enzymes, facilitating also for both simple purification and high-level expression. Interestingly, the native quaternary structure of each enzyme was maintained even when fused with MBP and GST. We also found that in fusion proteins the favorable biochemical properties of family enzymes such as, their optimal pHs, specific activities and kinetic properties were conserved compared to the native enzymes. In addition, MBP-fused enzymes showed remarkable folding ability *in-vitro*. Our findings, therefore, suggest that a previously unrecognized trait of this family, namely the ability to functional fusion with some other protein but yet to retain innate properties, is conserved. We described here the structural and evolutionary implications of the properties in this family enzyme.

Keywords: fusion, cyclic amidohydrolase, hydantoinase, dihydroorotase, allantoinase

INTRODUCTION

The prediction of the structural, functional, and evolutionary features of unidentified proteins generally depends on how their sequences are related to known protein. The reasonable classification of a protein family based on the functional and structural traits is expected to stimulate progress in terms of allowing the prediction of the structures and biochemical functions of unknown proteins [1,2]. Current protein families are vast and provide a huge amount of information on the structure-function relationships of newly isolated and unidentified proteins [2,3]. It is believed that the structure of proteins from a common origin can be conserved through the evolutionary process, despite a high sequence divergence [4]. Therefore, proteins belonging to a structural family with functional relatedness are presumed to share a structural scaffold, permitting different substrate specificities or stabilities within family proteins even given subtle changes in their common scaffolds [5,6]. As a consequence, the unification of protein families deduced from higher structural resemblances and relevant functions, especially as recently explored by the structural remote homology approach, leads to the broadening of the scope of a protein family beyond its primary protein structure [7]. The uncovering these kinds of traits is valuable and useful for the elucidation of protein architecture, and may implicate an undisclosed structural feature in a higher structural view, such as tolerable structure in the case of protein fusion.

Cyclic amide bonds are ubiquitously found in a variety of biomolecules in living organisms [8]. Evidence suggesting that enzymes acting on cyclic amide bonds are structurally and functionally related at the molecular level, primarily in terms of participating in nucleotide metabolism, continued to accumulate [7,9,10]. Direct experiments on dihydroorotase and hydantoinase led to the elucidation of the nature of the metal binding sites and the signature sequence of these enzymes [11,12]. With this signature sequence and rigidly conserved regions, a recent study on related family enzymes, including dihydroorotase, allantoinase, dihydropyrimidinase and hydantoinase, proposed a cyclic amidohydrolase (CA) family [11]. This family was further extended by a remote homology search based on the substructure of urease [7]. These findings support the hypothesis that in various organisms, conserved signature sequences and residues might be useful probes for identifying related enzymes. It has also been suggested that the higher architecture of the CA enzyme family is shared as a common scaffold. We noticed that dihydroorotase, as a pivotal enzyme for this family, constituted the CAD multienzyme complex [13] and thus possesses

* **Corresponding author**
Tel: +82-42-869-2616 Fax: +82-42-869-2610
e-mail: hskim@mail.kaist.ac.kr

a structure tolerable to the fusion of a protein. Hence, it seems to be a structural trait of the CA family, if this folding property has been conserved throughout the evolutionary process.

A recent study showing the distinct fusion ability of hydantoinase also provided a clue that related enzymes may be able to functionally tolerate fusion to other proteins [14]. More recently, we identified two CA genes from *E. coli* encoding allantoinase and a novel phenylhydantoinase that were found to be functional even in the fusion state [15]. Furthermore, human adenine deaminase [16], a novel member of the CA family, was reported to functionally separable from an MBP fusion protein, implying that the functional fusion can be extended to other enzymes in the CA family.

In this paper, to explore further the structural and functional trait shared by enzymes of the CA family, we first constructed the fusion proteins by simple end-to-end fusion of the whole open reading frames. An investigation of the resulting fusion proteins in terms of their structural and biochemical characteristics and terms of their *in vitro*-refolding led us to believe that high fusion ability might be shared as a structural trait in the CA enzyme family. Relevant studies of the functional fusion led to the interesting implication that the functional expressions of the fusion proteins with correct folding could broaden the potential use of CA family members.

MATERIALS AND METHODS

Strains and Media

Derivatives of *E. coli* K-12 were used as sources of the CA genes, allantoinase [15,17], dihydroorotase [18] and phenylhydantoinase [15], and *Bacillus thermocatenulatus* GH2 was used for hydantoinase [19]. *E. coli* strain JM109 was used as a cloning host of the MBP- and GST-fusion proteins. The dihydroorotase-deficient *E. coli* strain X7014 (CGSC5358) was obtained from the Yale *E. coli* Genetic Stock Center. *E. coli* were transformed by electroporation, and the resulting transformants were grown at 30°C or 37°C in Luria-Bertani (LB) medium. Ampicillin (50 μg/mL) and/or chloramphenicol (25 μg/mL) were supplemented when needed.

Construction and Expression of MBP- and GST-CA Fusion Proteins

Chromosomal DNA was isolated from *E. coli,* or *B. thermocatenulatus* GH2, using a genomic DNA purification kit (Promega), and the genes encoding the corresponding family enzymes were cloned using standard recombinant DNA techniques. Three sets of primers were designed to span the genes encoding allantoinase (1362 bp), dihydroorotase (1047 bp) and phenylhydantoinase (1398 bp) from *E. coli*, and a set of primers were design for hydantoinase (1419 bp) from *B. thermocatenulatus* GH2 (Table 1). A restriction site for *Bam*HI was inserted into the C-terminal primer of *E. coli* allantoinase. In other cases, each restriction site for *Eco*RI and *Pst*I was introduced into the N- and C-terminal primers, respectively. The amplified DNA fragments encoding the family enzymes were cloned into the *Eco*RI/*Pst*I (or *Eco*RI/*Bam*HI) site of pMAL-c2X.

Table 1. Primers used in this study

Primer name	Sequence[a]
ALLN	5′-AGGAATTCGTTATGTCTTTTGATTTAATCATT-3′
ALLC	5′-GGGGATCCTTACTGCTGATGTTTAAGGATAA-3′
DHON	5′-ATGAATTCACTGCACCATCCCAGGTATTAAAGAT-3′
DHOC	5′-ATCTGCAGTTATTGTTTAACGGACCAGCGTACC-3′
HYDN	5′-TAGAATTCATGACAAAATTGATAAAAAATG-3′
HYDC	5′-TACTGCAGTTAGGACATTTTCACCACATCT-3′
PHN	5′-GGAGAATTCTTGGAGTTTGCTATGCGCGTA-3′
PHC	5′-TGGCTGCAGTTAGAGCACGGGAGGGACAAA-3′

[a] Underlines indicate the restriction sites introduced in the synthetic oligonucleotides.

The pGEX-5X-1 vector (Amersham Pharmacia Biotech) was used to express the family enzymes as fusion proteins with glutathione S-transferase (GST) according to the manufacturer's instructions. The family genes were amplified by PCR from the chromosomal DNA, as described in the MBP fusion system, and then blunted by using the Klenow Fragment. The resulting DNA fragments were digested with *Eco*RI, and then cloned into between the *Eco*RI and blunt-ended site of pGEX-5X-1. All constructs were confirmed by DNA sequencing.

E. coli JM109 cells were transformed by electroporation in 10% glycerol with the plasmids harboring MBP- or GST fusion enzymes, and maintained in LB medium containing ampicillin (50 μg/mL) at 30°C. Expressions of each fusion protein were induced by adding 0.5 mM of isopropyl-β-D-thiogalactopyranoside (IPTG) at 30°C when the optical density at 600 nm reached between 0.45 and 0.6.

Purification of Fusion- and Wild-type Proteins

After cultivating induced cells (300 mL) for 3 hr the cells were harvested by centrifugation at 6,000 × g for 10 min. The resulting pellets were resuspended in a total volume of 10 mL of Tris-HCl buffer (20 mM, pH 8.0) containing 0.5 mM of dithiotreitol (DTT), 1 mM of phenylmethylsulfonyl fluoride (PMSF) and 0.1% of Triton X-100. The suspended cells were sonicated, and the supernatant obtained by centrifuging at 18,000 × g for 30 min was analyzed directly for enzyme activity and protein expression by SDS-polyacrylamide gel electrophoresis (PAGE). For biochemical characterizations, the protein was further purified to apparent homogeneity using standard chromatographic procedures. MBP- and GST-fused CA proteins in the soluble fraction were absorbed onto their appropriate affinity resins, amylose- and glutathione Sepharose 4B, respectively, and then

washed three or four times with a buffer containing 20 mM of Tris (pH 8.0) and 200 mM of NaCl. After complete washing, proteins were eluted with glutathione (5 mM) or maltose (10 mM) for GST- or MBP fusion proteins, respectively. The affinity column purified fusion proteins were further concentrated using a Centriprep 10 (Amicon). The fused domain of the CA enzyme was separated from the MBP- or GST domain by treatment with factor X_a for 20 h at 8-10°C. Then the cleaved enzyme was isolated from its fusion partners, MBP or GST, by reapplying it to the affinity resin. Pooled enzyme solutions were concentrated by dialysis against 20 mM Tris-HCl (pH 8.0) buffer containing 5% glycerol, and then stored at –20°C.

CA enzymes were also purified in their native states as previously reported, with slight modification [11,15, 20]. All purification steps were conducted at room temperature. Briefly, the supernatant (10 mL) was loaded onto a column of Resource Q (5 mL) equilibrated with 20 mM of Tris-HCl (pH 8.0) buffer containing 5% glycerol, on a fast protein liquid chromatography (FPLC) system (Amersham Pharmacia Biotech.). The column was washed with 10 volumes of the same buffer and eluted with a linear gradient of 0–0.5 M NaCl. The active fractions were pooled and concentrated using a Centriprep 10 (Amicon). The concentrated protein solution was loaded onto a Superose-12 gel filtration column equilibrated with 20 mM Tris-HCl (pH 8.0) buffer containing 150 mM NaCl. The eluted enzyme was concentrated by dialysis against 20 mM Tris-HCl (pH 8.0) buffer containing 5% glycerol and stored at –20°C.

Gel Filtration Chromatography

The oligomeric structures of the enzymes were determined on a FPLC system fitted with a gel filtration column (Superose-12 HR10/30). The flow rate of the mobile phase, containing 20 mM Tris-HCl and 150 mM NaCl, was 0.3 mL/min. The column was calibrated using native protein markers (Amersham Pharmacia Biotech.). A molecular mass standard curve was established by plotting the elution profile of the protein markers versus the known molecular masses on semi-log paper. Aliquots of each eluted fraction were analyzed by SDS/PAGE and assayed for each enzyme activity.

Enzyme Activity Assay

The activities of the CA enzymes were determined at 40°C for 30 min with shaking. The enzyme reaction was carried out with the purified enzyme (10 μg) in 1 mL of reaction mixture containing 100 mM Tris-HCl (pH 8.0), 0.5 mM DTT, and an appropriate concentration of cyclic ureide (allantoin, 10 mM; dihydroorotate, 1 mM; hydroxyphenyl- or phenylhydantoin, 15 mM). Decreases in the concentrations of the cyclic ureide used as a substrate and increases in the concentrations of the corresponding N-carbamyl product were analyzed by high performance liquid chromatography (HPLC) and thin layer chromatography (TLC) [21,22]. The amount of product formed was also determined by HPLC or by using the color reagent *p*-dimethylaminobenzaldehyde. All assays were carried out in duplicate. One unit of enzyme activity was defined as the amount of enzyme required to hydrolyze 1 μ-mole of cyclic ureide per min under the specified conditions. Kinetic parameters (K_m and V_{max}) were determined according to the reported methods [15,21].

RESULTS

Deduced Fusion Capability of the Family Enzymes

We recently reported that a microbial hydantoinase, as a typical member of the CA enzyme family, forms a distinct homotetramer irrespective of the presence of a larger polypeptide fused at its N-terminus, also displaying a bifunctional activity [14]. In line with these fining, we analyzed the primary and secondary structures of the CA family enzymes at the molecular level. We found that, although the apparent homology of the enzymes was less than 20%, the structural similarities of the higher structures strongly suggested a close evolutionary relationship within the CA enzyme family. We also found that 7 regions in the primary and secondary structures were rigidly conserved in identical positions over the entire sequences [11]. The striking structural identities of the regions shared by the family members suggested that the conserved regions might play an important role in the structure or function of the CA family. Consistent with this view, more evidence was obtained by scanning the conserved regions on the protein database (SwissProt). Query sequences were designed from regions showing a high level of alignment, *i.e*., the percentage of fixed residues in the sequence. Gaps were avoided to reduce fortuitous matching results. This analysis, along with multiple sequence alignment [23], identified the seven regions in the majority of the known CA enzymes. Thus, it is reasonable to suggest that the existence of such similar amino acid arrangements through a significant length of the protein backbone is not the result of coincidence, but rather that it provides a evidence of a structurally important region in the enzyme family [24]. Therefore, functional fusion ability, as shown by a typical family member hydantoinase [14], may be also founded in other CA family members, expecting that their structural and biochemical properties were also retained in a fusion state.

Retained Nascent Properties of CA Family Members for MBP Fusion

To determine the fusion ability of other member of this family, we constructed fusion proteins of four typical members of this family, namely, a hydantoinase (HYD) from *B. thermocatenulatus* GH2, and three enzymes, allantoinase (ALN), phenylhydantoinase (PH) and dihydroorotase (DHO) from *E. coli*, at their N-termini with maltose binding protein (MBP). We used

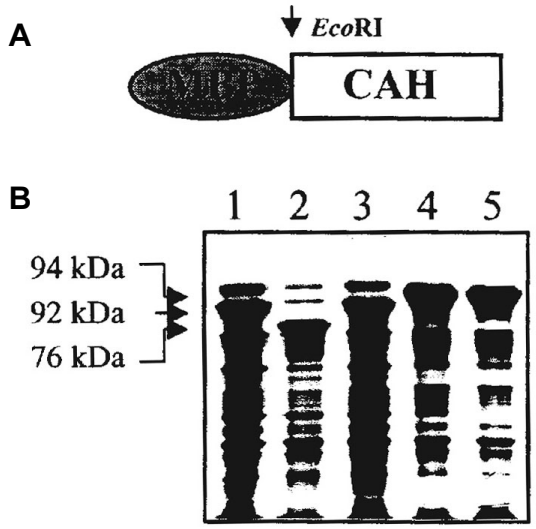

Fig. 1. Construction and expression of the MBP-fused cyclic amidohydrolase enzymes. (A) Organization of the cyclic amidohydrolase enzyme family fused with maltose-binding protein at their N-terminus by simple end-to-end fusion. Arrows indicate the *Eco*RI-cleavage site. (B) SDS/PAGE analyses of crude extracts of *E. coli* cells expressing the MBP-fused proteins. Lane 1, MBP-ALN (92 kDa); lane 2, MBP-DHO (76 kDa); lane 3, MBP-PH (93 kDa); lane 4, MBP-HYD (94 kDa); and lane 5, a control fusion protein reported previously. Crude extracts were analyzed after induction with 0.5 mM IPTG at 30°C for 2 h.

simple end-to-end fusion without a peptide linker to evaluate their direct fusion abilities, and expressed the fusion proteins in a general host, *E. coli* JM109, at 30°C. When the *E. coli* cells harboring fusion proteins were plated on the selective plate containing each typical substrate [24], we detected enzyme activities comparable to those of wild-type enzymes, whereas growth inhibition was negligible in each case. Similar results were also observed at 37°C. By SDS/PAGE analysis, the four MBP-fused proteins were mainly founded in the soluble fraction (>80%) with the expected molecular masses, and these accounted for about 20-30% of the total cell protein (Fig. 1). The fragments derived from proteolysis were negligible.

To compare the structural and biochemical properties with those of the wild-type enzymes, we further purified the fusion proteins using amylose resin to apparent homogeneity. Wild-type enzymes were also purified, as has been reported previously [11,15,20]. About 2–3 mg of the fusion proteins were recovered from a 300 mL culture of the induced cells and showed a reasonable migration rate calculated by deduced amino acid residues at the molecular level. As a preliminary experiment, we tried to cleave the fusion proteins with factor Xa and found that the enzymes were readily separated into their respective domains, which indicated that both fusion partners fold independently and are accessible to site-specific protease factor X_a. In general, misfolded or aggregated fusion protein is resistant to protease attack.

To analyze whether kinetic properties differed in the fused state, we compared MBP-fused enzymes with their nascent wild-type counterparts. The activities of fused proteins toward their individual substrates was determined and found to be similar to those of the native enzymes (Table 2), except for MBP-DHO. The specific activity of MBP-DHO toward dihydroorotate was about 50% lower than that of the native enzyme, whereas two other enzymes (MBP-HYD and MBP-PH) had activities that were very similar to those of the wild-type enzymes, and MBP-ALN maintained its full activity. Km values were also obtained to determine the effect of MBP fusion on substrate affinity. Most of the fusion proteins were not substantially different in terms of their affinity values versus the wild-type enzymes (Table 2). In the case of MBP-ALN, about two folds increase in the K_m value was observed, but the overall catalytic efficiency of the fusion enzyme was similar to that of the wild-type enzyme. Interestingly, MBP-PH showed a marginal increase in its affinity for the typical substrate phenylhydantoin. We also determined the optimum pH for the hydrolysis of the corresponding substrate. pH dependency was analyzed in 0.1mM boric acid-NaOH, 0.1 mM Tris-HCl and 0.1 mM sodium phosphate buffer at pHs ranging from 8.5-10.5, 7.5-8.5 and 5.5-7.5, respectively. The fusion proteins showed optimum pHs at about 7.5-8.5, as was the case for the corresponding wild-type enzymes [15].

Table 2. Comparison of the biochemical properties between the fusion and free enzymes

	ALN		DHO		HYD		PH	
	Fusion	Native	Fusion	Native	Fusion	Native	Fusion	Native
Specific activity[a] (U/mg protein)	21.0	19.9	16.6	34.9	34.7	49.2	2.83	4.64
K_m[b] (mM)	8.9	4.2	0.083	0.053	27.2	25.6	7.8	11.3
Oligomer[c]	Tet	Tet	Di	Di	Tet	Tet	Tet	Tet

[a] Enzyme activity was determined using the purified enzyme (10 μg) in the presence of the corresponding substrate (allantoin, 10 mM; dihydroorotate, 1 mM; hydroxyphenylhydantoin and phenylhydantoin, 15 mM) in a 1 mL of 100 mM Tris-HCl (pH 8.0) buffer. The reaction products were analyzed by either HPLC or using a color reagent described in Materials and Methods.

[b] The Km values were determined in a 1 mL of 100 mM Tris-HCl (pH 7. 8) buffer, using the appropriate substrate concentrations, ranging from 0.5-15 mM for allantoinase, 0.01-0.5 mM for dihydroorotase, and 0.5-15 mM for phenylhydantoinase and hydantoinase.

[c] Tet and Di denote tetramer and dimmer.

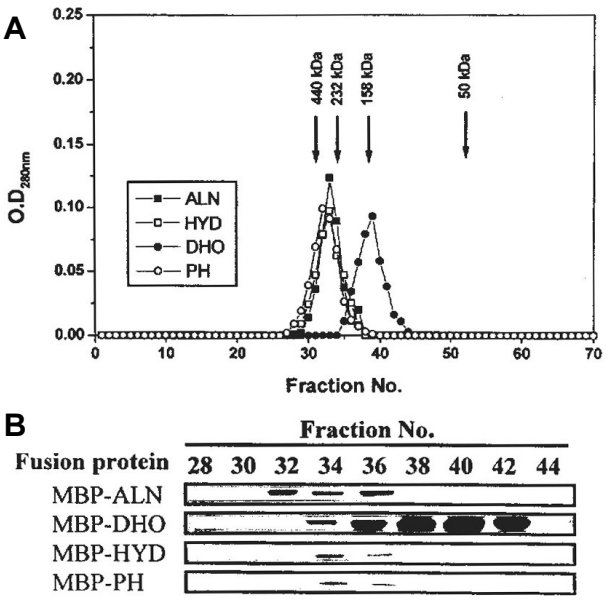

Fig. 2. Oligomeric structures of MBP-fused family enzymes. Purified fusion proteins (50-125 μg) were analyzed on a Superose-12 gel filtration column. (A) Elution profiles of the MBP-fused proteins. The molecular mass of each fusion protein was estimated from the elution profile of the standard protein markers: blue dextran, 2000 kDa; ferritin, 440 kDa; catalase, 232 kDa; aldolase, 158 kDa; Fab fragment, 50 kDa; and MBP, 42 kDa. All experiments were repeated three times at different protein concentrations. The shift in elution time was negligible (< 0.2 min). (B) Fractions showing the corresponding enzyme activity were collected after gel filtration chromatography and analyzed on 9% SDS-PAGE.

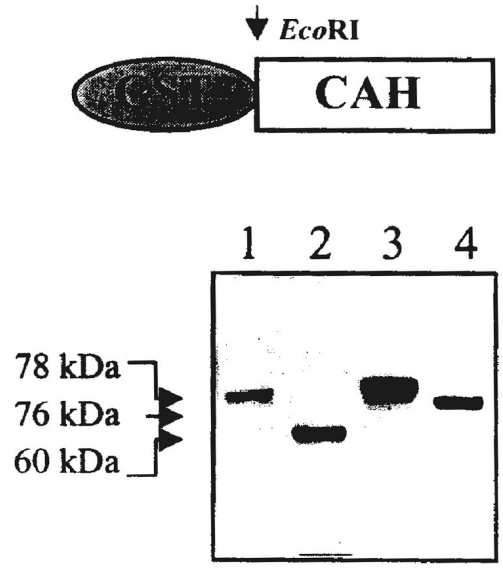

Fig. 3. Construction and expression of the GST-fused cyclic amidohydrolase family enzymes. (A) Organization of the GST-fused enzymes at their N-terminus by simple end-to-end fusion. (B) SDS/PAGE analysis of crude extracts of *E. coli* cells expressing the fusion proteins. Lane 1, GST-ALN (76 kDa); lane 2, GST-DHO (60 kDa); lane 3, GST-HYD (78 kDa); and lane 4, GST-PH (77 kDa). Crude extracts were analyzed on 9% SDS/PAGE after 2 h induction with 0.5 mM of IPTG at 30°C.

Gel filtration column chromatography was used to examine the quaternary structures of the MBP-fused proteins. From the gel filtration column chromatography elution profiles, the four fused enzymes were found to maintain the innate quaternary structures of the corresponding wild-type ones (Table 2), as was also observed for a typical case of CA enzyme family in our previous work [14]. Three fusion proteins, MBP-ALN, MBP-HYD and MBP-PH, coeluted with an apparent molecular mass between 340- and 380 kDa, whereas MBP-DHO was detected in fractions corresponding to molecular masses between 140-160 kDa (Fig. 2A). SDS-PAGE analysis of the eluted fractions also revealed the corresponding fusion proteins (Fig. 2B). The respective subunit masses were calculated to be around at 92-94 kDa for three of the fusion enzymes (MBP-ALN, MBP-HYD and MBP-PH), whereas that of MBP-DHO was estimated to be about 76 kDa. Thus, the quaternary structures of these enzymes were predicted to be homotetrameric for MBP-ALN, MBP-HYD and MBP-PH, and homodimeric for MBP-DHO. Cross-linking experiments with SMCC (sulfosuccinimidyl 4-[*N*-maleimidomethyl]-cyclohexane-1-carboxylate) and ultracentrifugation also confirmed that the quaternary structures of the enzymes were conserved even when fused with a larger protein at their N-terminus.

The nascent quaternary structures were retained in the fusion protein, and thus improper bands resulting from high molecular weight aggregates or other oligomeric structures were negligible. However, MBP-HYD had a high tendency to self-aggregate at the moderate concentration of 1.2 mg/mL. Other fusion proteins showed a negligible tendency to aggregate even at protein concentrations up to 2 mg/mL. Generally, aggregation is the main cause of the decreased refolding yield and loss of activity. However, for MBP-HYD, this was not found to be the case, because although MBP-HYD exhibited a high aggregation tendency, its activity was unaffected when it was incubated at a low protein concentration (< 0.05 mg/mL) at 50°C for 1 h. This suggests that the fusion enzyme readily refolds into an active form under reaction conditions (pH 8.0, 50 mM Tris and 1 mM DTT).

Favorable Properties are Also Conserved in GST-fusion Proteins

In the general experiments, MBP alone did not oligomerize even at the high concentrations employed in this study, emphasizing that the oligomerization of the fusion proteins are due to interactions between the enzyme subunits. This result suggests the possibility that

the enzymes are self-oligomerizable even when different fusion partners are fused to the *N*-terminus of the gene encoding the enzyme subunit. To test this possibility of self-oligomerization with different fusion partners *in-vivo*, we constructed another fusion enzyme by employing glutathione S-transferase (GST) as a fusion partner, using the same procedure as used for the MBP fusion system (Fig. 3A). Four GST fusion proteins were also well-expressed and accounted for 15-20% of the total cell protein (data not shown), which corresponded to about 40-60% of the MBP fusion proteins. For further analysis, the fusion proteins were purified by affinity column chromatography. Approximately 1-2 mg of purified proteins was obtained from a 300 mL of cell culture, which was almost identical to that obtained from the MBP fusion proteins. The resulting proteins appeared homogeneous when analyzed by SDS-PAGE (Fig. 3B), and had migration rates that corresponded to their calculated molecular masses. In addition, the factor Xa cleavage pattern and elution profile on gel-filtration column chromatography also revealed that the GST-fusion proteins are fully functional and retain their nascent oligomeric structures (data not shown), although their aggregation tendency was slightly higher than MBP-fusion proteins [25]. In a set of observations *in vivo*, when the expression of MBP- and GST-fusion proteins were induced with 1 mM IPTG at 37°C, the GST-fusion proteins showed a reduced tendency to oligomerize in their native structural forms in the soluble fraction, whereas the oligomerization of the MBP-fusion proteins was more distinct. Therefore, affinity-precipitated or purified protein solutions of MBP-fused proteins revealed a high portion of the correct oligomer than those of GST-fusion proteins.

DISCUSSION

Here we demonstrate that the four typical enzymes from a CA family of enzymes possess a distinct ability to form fusion proteins. When members of the enzyme family were fused by simple end-to-end fusion with large proteins, such as MBP or GST, they retained their nascent properties in the fused state. It is generally known that MBP and GST, although MBP is far more effective than GST as a fusion partner [25], are permissible to structural changes within fusion proteins, and thus, well address the correct folding of the fused domain. However, as was found for the typical family hydantoinase [26], it is plausible that the ability to fuse shown by an enzyme of the CA family may originate from their intrinsic folding ability and the stability of their oligomerization. Although this characteristic might be closely related to the structure and/or function of the enzyme family, little is known of the oligomerization. In this regard, a previous report provided a clue that the oligomerization of the hydantoinase enzyme family is readily mediated in a metal dependent manner, and is the most important for enzyme activity and stability [26].

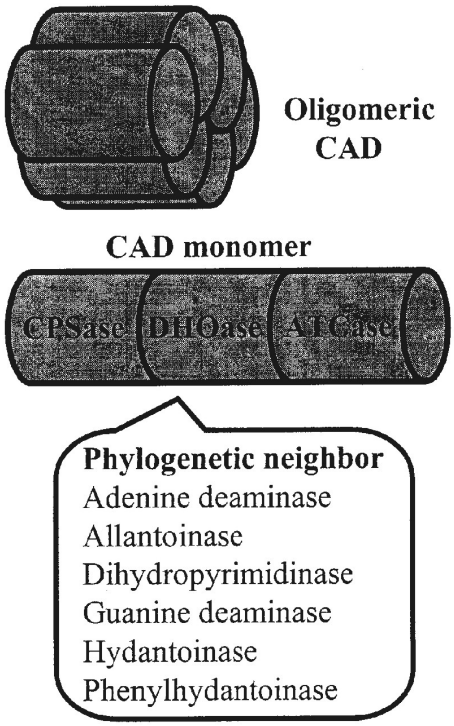

Fig. 4. Evolutionary possible neighbor in the cyclic amidohydrolase enzyme family. Each member of the cyclic amidohydrolase family family is composed of similar primary and secondary structural organizations over the entire open reading frame, and is involved in the metabolism of pyrimidines and purines. All of the cyclic amidohydrolase enzymes share various biochemical and structural properties, with respect to their optimum pHs, substrate structure, metal dependency, oligomerization, metal-binding site and even in terms of the size of the open reading frame. Dihydroorotase is a possible pivotal enzyme, and is found in eucaryotic multifunctional CAD, which an oligomeric protein complex composed of carbamoyl-phosphate synthetase, dihydroorotase, and aspartate transcarbamylase.

To ensure that the CA enzymes have the ability to fuse, studies on the functionalities of fusion constructs with other fusion partners will be required. In this respect, we tested other fusion proteins using green fluorescent protein (GFP) as a fusion partner, and employing the same procedures as described for MBP and GST fusion systems, without a linker peptide, to ascertain the intrinsic ability of the enzymes to fuse with a whole protein [14]. Both the GFP and the CA enzyme family were observed to function distinctly in the fusion proteins. In addition, the four cyclic amidohydrolases were also found to be active when fused with GFP, although this resulted in a lower structural stability than MBP or GST fusion. These results strongly suggest the possibility that the CA enzyme family is useful for generating bifunctional fusion proteins. This was supported by a recent result that a novel bifunctional enzyme could be constructed by end-to-end fu-

sion of D-hydantoinase and *N*-carbamylase [14].

With complete genome sequences available from various sources, it is very interesting that the putative cyclic amidohydrolases are ubiquitously distributed in eukaryotic and prokaryotic cells, but the functional roles of most of these enzymes remain unknown [15,27]. The elucidation of the exact role of the enzyme, and the identification of functional defects in engineered enzymes would unearth some crucial information regarding the structural and functional relationships of the CA enzyme family. Therefore, the finding that the CA enzyme family tolerate partner proteins fused, and thus can be functionally expressed with a fusion partner, could be applied to the isolation of the family enzymes by simple affinity chromatographic procedures. In addition, fusion with other polypeptides or domains to probe or capture interactive molecules would provide a powerful new route to analytical tool, because the self-oligomerization ability of the enzymes is distinct and the direct detection of these enzymes is simple and accurate as has been reported previously [24].

The evolutionary relatedness of the CA family, as defined by experimental and computational methods may be depicted as shown in Fig. 4, based on their structural and functional properties as well as their actions in the context of metabolic pathways [7,10,11]. The CA enzyme family members possess similar primary and secondary structures, and in addition, most of them are involved in the metabolism of pyrimidines and purines. Recently, having being credited with predicted three-dimensional fold and active site architecture, the CA family was extended to include dihydroorotases, allantoinases, hydantoinases, AMP-, adenine and cytosine deaminases, and imidazolonepropionase [7]. The family members share a common motif and structural region, which provide clues for the identification of unidentified proteins [15]. From current knowledge, a structural family with high functional similarity is proposed to have a shared enzymatic trait, and thus provide evidence of an uncovered structural or functional trait. These uncovered traits remains to be discovered in the related cyclic amidohydrolases. Naturally occurring multifunctional enzymes may have evolved from separated enzymes by gene fusion. In contrast, it is also feasible in nature that the functional separation and thus evolution of a respective domain from a multifunctional protein complex in nature [28,29]. In this context, we noticed that dihydroorotase, the origin of which cab be traced back to the early states of evoluton, is found both in the multi-enzyme complex CAD (carbamoyl phosphate synthetase, aspartate transcarbamoylase and dihydroorotase) [13] and as a separate domain [9,18,30,31].

Dihydroorotase, which catalyzes the third step in pyrimidine biosynthesis, is presumed to be a possible ancestor of the cyclic amidohydrolase enzyme family, because this enzyme is common to all living organisms. Furthermore, dihydroorotase is considered to be more important as a synthetic enzyme for nucleotides than other cyclic amidohydrolase family members participating in the catabolic pathway [7]. A more stringent and crucial role for dihydroorotase in nucleotide synthesis strongly supports that this enzyme might have been present at an early stage of evolution [32]. This hypothesis is reinforced by the fact that dihydroorotase is present as a part of a fused tri-functional enzyme (CAD) in eukaryotes. Interestingly, the dihydroorotase domain can be functionally separated from this multienzyme complex [33,34], and retain its activity in a dimeric state as in the separately evolved dihydroorotase [35,36]. This strongly suggests that an independent folding unit of the dihydroorotase domain is well conserved in the multi-enzyme complex, and hence separable. Related substructures and similar arrangements are also found for other members of the cyclic amidohydrolase family [7,11,24]. These observations are consistent with the hypothesis that other family members also may possess a high ability to fuse and form bifunctional fusion protein. Moreover, this ability is strongly supported by the results of the present study and those reported in a previous paper [14]. Although the specific configurations required to facilitate fusion have yet to be verified experimentally, this study that found the cyclic amidohydrolase enzyme family share a trait should extend the scope and use of member and related enzymes.

REFERENCES

[1] Holm, L. and C. Sander (1995) DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem. Sci.* 20: 345-347.

[2] Todd, A. E., C. A. Orengo, and J. M. Thornton (1999) Evolution of protein function, from a structural perspective. *Curr. Opin. Chem. Biol.* 3: 548-556.

[3] Zarembinski, T. I., L. W. Hung, H. J. Mueller-Dieckmann, K. K. Kim, H. Yokota, R. Kim, and S. H. Kim (1998) Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics. *Proc. Natl. Acad. Sci. USA* 95: 15189-15193.

[4] Chothia, C. (1992) Proteins: One thousand families for the molecular biologist. *Nature* 357: 543-544.

[5] Pujadas, G., F. M. Ramirez, R. Valero, and J. Palau (1996) Evolution of beta-amylase: patterns of variation and conservation in subfamily sequences in relation to parsimony mechanisms. *Proteins* 25: 456-472.

[6] Jez, J. M., M. J. Bennett, B. P. Schlegel, M. Lewis, and T. M. Penning (1997) Comparative anatomy of the aldo-keto reductase superfamily. *Biochem. J.* 326: 625-636.

[7] Holm, L. and C. Sander (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. *Proteins* 28: 72-82.

[8] Vogels, G. D. and C. van der Drift (1970) Degradation of purines and pyrimidines by microorganisms. *Bacteriol. Rev.* 40: 403-468.

[9] Aleksenko, A., W. Liu, Z. Gojkovic, J. Nielsen, and J.

Piskur (1999) Structural and transcriptional analysis of the pyrABCN, *pyr*D and *pyr*F genes in *Aspergillus nidulans* and the evolutionary origin of fungal dihydroorotases. *Mol. Microbiol*. 33: 599-611.

[10] May, O., A. Habenicht, R. Mattes, C. Syldatk, and M. Siemann (1998) Molecular evolution of hydantoinases. *Biol. Chem*. 379: 743-747.

[11] Kim, G. J. and H. S. Kim (1998) Identification of the structural similarity in the functionally related amidohydrolases acting on the cyclic amide ring. *Biochem. J.* 330: 295-302.

[12] Zimmermann, B. H., N. M. Kemling, and D. R. Evans (1995) Function of conserved histidine residues in mam-malian dihydroorotase. *Biochemistry* 34: 7038-7046.

[13] Simmer, J. P., R. E. Kelly, A. G. Jr. Rinker, B. H. Zimmermann, J. L. Scully, H. Kim, and D. R. Evans (1990) Mammalian dihydroorotase: nucleotide sequence, peptide sequences, and evolution of the dihydroorotase domain of the multifunctional protein CAD. *Proc. Natl. Acad. Sci. USA* 87: 174-178.

[14] Kim, G. J., D. E. Lee, and H. S. Kim (2000) Construction and evaluation of a novel bifunctional *N*-carbamylase-D-hydantoinase fusion enzyme. *Appl. Environ. Microbiol*. 66: 2133-2138.

[15] Kim, G. J., D. E. Lee, and H. S. Kim (2000) Functional expression and characterization of the two cyclic amidohydrolase enzymes, allantoinase and a novel phenylhydantoinase, from *E. coli*. *J. Bacteriol*. 182: 7021-7028.

[16] Yuan, G., J. C. Bin, D. J. McKay, and F. F. Snyder (1999) Cloning and characterization of human guanine deaminase. Purification and partial amino acid sequence of the mouse protein. *J. Biol. Chem.* 274: 8175-8180.

[17] Cusa, E., N. Obradors, L. Baldoma, J. Badia, and J. Aguilar (1999) Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in *Escherichia coli*. *J. Bacteriol*. 181: 7479-7484.

[18] Wilson, H. R., P. T. Chan, and C. L. Turnbough Jr (1987) Nucleotide sequence and expression of the *pyr*C gene of *Escherichia coli* K-12. *J. Bacteriol*. 169: 3051-3058.

[19] Kim, G. J. and H. S. Kim (1998) C-Terminal regions of D-hydantoinases are nonessential for catalysis, but affect the oligomeric structure. *Biochem. Biophys. Res. Commun*. 243: 96-100.

[20] Brown, D. C. and K. D. Collins (1991) Dihydroorotase from *Escherichia coli.* Substitution of Co(II) for the active site Zn(II). *J. Biol. Chem.* 266: 1597-1604.

[21] Soong, C. L., J. Ogawa, M. Honda, and S. Shimizu (1999) Cyclic-imide-hydrolyzing activity of D-hydantoinase from *Blastobacter sp.* strain A17p-4. *Appl. Environ. Microbiol*. 65: 1459-1462.

[22] Watabe, K., T. Ishikawa, Y. Mukohara, and H. Nakamura (1992) Cloning and sequencing of the genes involved in the conversion of 5-substituted hydantoins to the corresponding L-amino acids from the native plasmid of *Pseudomonas sp*. strain NS671. *J. Bacteriol*. 174: 962-969.

[23] Thompson, J. D., D. G. Higgins, and T. J. Gibson (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice**.** *Nucleic Acids Res*. 22: 4673-4680.

[24] Kim, G. J., J. H. Park, D. C. Lee, H. S. Ro, and H. S. Kim (1997) Primary structure, sequence analysis, and expression of the thermostable D-hydantoinase from *Bacillus stearothermophilus* SD1. *Mol. Gen. Genet*. 255: 152-156.

[25] Kapust, R. B. and D. S. Waugh (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci*. 8: 1668-1674.

[26] Mukohara, Y., T. Ishikawa, K. Watabe, and H. Nakamura (1994) A thermostable hydantoinase of *Bacillus stearothermophilus* NS1122A: Cloning, sequencing, and high expression of the enzyme gene, and some properties of the expressed enzyme**.** *Biosci. Biotech. Biochem.* 58: 1621-1626.

[27] Hamajima, N., K. Matsuda, S. Sakata, N. Tamaki, M. Sasaki, and M. Nonaka (1996) A novel gene family defined by human dihydropyrimidinase and three related proteins with differential tissue distribution. *Gene* 180: 157-163.

[28] Maley, J. A. and J. N. Davidson (1988) The aspartate transcarbamylase domain of a mammalian multifunctional protein expressed as an independent enzyme in *Escherichia coli*. *Mol. Gen. Genet*. 213: 278-284.

[29] Murley, L. L. and R. E. MacKenzie (1995) The two monofunctional domains of octameric formiminotransferase-cyclodeaminase exist as dimmers. *Biochemistry* 34: 10358-10364.

[30] Banks, G. R. and S. Y. Taylor (1988) Cloning of the PYR3 gene of *Ustilago maydis* and its use in DNA transformation. *Mol. Cell. Biol*. 8: 5417-5424.

[31] Ogawa, J. and S. Shimizu (1995)). Purification and characterization of dihydroorotase from *Pseudomonas putida. Arch. Microbiol*. 164: 353-357.

[32] Davidson, J. N., K. C. Chen, R. S. Jamison, L. A. Musmanno, and C. B. Kern (1993) The evolutionary history of the first three enzymes in pyrimidine biosynthesis. *BioEssays* 15: 157-164.

[33] Grayson, D. R., L. Lee, and D. R. Evans (1985) Immunochemical analysis of the domain structure of CAD, the multifunctional protein that initiates pyrimidine biosynthesis in mammalian cells. *J. Biol. Chem*. 260: 15840-15849.

[34] Williams, N. K., R. J. Simpson, R. L. Moritz, Y. Peide, L. Crofts, E. Minasian, S. J. Leach, R. G. Wake, and R. I. Christopherson (1990) Location of the dihydroorotase domain within trifunctional hamster dihydroorotate synthetase. *Gene* 94: 283-288.

[35] Kelly, R. E., M. I. Mally, and D. R. Evans (1986) The dihydroorotase domain of the multifunctional protein CAD. Subunit structure, zinc content, and kinetics. *J. Biol. Chem*. 261: 6073-6083.

[36] Williams, N. K., Y. Peide, K. K. Seymour, G. B. Ralston, and R. I. Christopherson (1993) Expression of catalytically active hamster dihydroorotase domain in *Escherichia coli*: Purification and characterization. *Protein Eng.* 6: 333-340.

[Received February 7, 2002; accepted May 28, 2002]