CHARACTERIZATION OF BETAINE:HOMOCYSTEINE METHYLTRANSFERASE,

CYSTATHIONINE β-SYNTHASE AND CYSTATHIONINE γ-LYASE: PROTEINS

CONTROLLING HOMOCYSTEINE METABOLISM

by

NANDITA BOSE

(Under the direction of Dr. Cory Momany)

ABSTRACT

The objective of this dissertation was to produce three of the human homocysteine

metabolizing enzymes, cystathionine β-synthase (CBS), cystathionine γ-lyase (CGL) and

betaine:homocysteine methyltransferase (BHMT) for use in structural studies. A

modified pET28b vector was designed to produce N-terminal poly-histidine tagged

proteins with a simple construction scheme having broad applicability by the use of rare

Sap I cloning sites. With CBS, the main focus was to address the aggregation problems of

the full-length protein. After various futile attempts to solubilize CBS, denaturation and

renaturation of the protein did yield soluble protein, but the protein was inactive. Finally,

soluble, full-length and active CBS was obtained by the manipulation of factors such as

pH, medium of growth and addition of detergent. Chromatography and SDS-PAGE

analysis revealed the presence of an inseparable, proteolytically cleaved form of CBS

along with the full-length protein, confirming the protein's strong tendency towards

aggregation. The insolubility problems of CGL were tackled by the usage of the designed expression vector and modification of the induction-time. Preliminary X-ray studies of BHMT crystals revealed that BHMT is a homotetramer, which was confirmed by gel-filtration analysis. An earlier study of the activating effect of SAMe on rat BHMT was challenged when it was found that SAMe had no effect on the recombinant BHMT's ability to methylate homocysteine, nor did the enzyme appear to bind SAMe when examined by microcalorimetry. Crystallization trials of both CGL and BHMT were marred by oxidation-sensitivity and reproducibility problems.

INDEX WORDS: Homocysteine, Cystathionine, Betaine, Synthase, Lyase,
            Methyltransferase, Crystallization

CHARACTERIZATION OF BETAINE:HOMOCYSTEINE METHYLTRANSFERASE,

CYSTATHIONINE β-SYNTHASE AND CYSTATHIONINE γ-LYASE: PROTEINS

CONTROLLING HOMOCYSTEINE METABOLISM

by

NANDITA BOSE

B.Pharm., Bombay College of Pharmacy, India, 1997

A Dissertation submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2002

CHARACTERIZATION OF BETAINE:HOMOCYSTEINE METHYLTRANSFERASE,

CYSTATHIONINE β-SYNTHASE AND CYSTATHIONINE γ-LYASE: PROTEINS

CONTROLLING HOMOCYSTEINE METABOLISM

by

NANDITA BOSE

Approved:

Major Professor: Dr Cory Momany

Committee: Dr. Will Taylor
                   Dr. Phillip Greenspan
                   Dr. James Hargrove
                   Dr. Diane Hartle

Electronic Version Approved:

Gordhan L. Patel
Dean of the Graduate School
The University of Georgia
May 2002

DEDICATION

To my husband, Jay Parakkuth

With admiration and love

would like to thank my husband, Jay for loving me and patiently waiting for me to finish my studies.

Thank you all for being there......

TABLE OF CONTENTS

INTRODUCTION AND LITERATURE REVIEW

Homocysteine Metabolism and Cardiovascular Disease, Alzheimer's Disease,

AIDS...What next? Cardiovascular disease causes 44% of all deaths in the United States. Alzheimer's dementia affects 4 million Americans and will increase sharply as the population ages. Human Immunodeficiency virus (HIV) has infected about 800 to 900 thousand people in the Unites States, as per statistics from the Centers for Disease Control and Prevention (1). Various studies carried out to study the biochemical basis of these diseases have at one point or the other identified defects in a common metabolic pathway – the homocysteine metabolism pathway.

Donald Jacobsen, director of the Laboratory for Homocysteine research at the Cleveland Clinic said in an article in *The Scientist*, "In the future, a homocysteine test will be as common as a cholesterol test". The American Medical Association's (AMA) 17th Annual Science Reporters Conference reported that people with Alzheimer's disease have moderately elevated blood levels of homocysteine. Research carried out in the area of AIDS has shown that patients infected with HIV exhibit low plasma cysteine and glutathione levels, which again are the products of the transsulfuration pathway of homocysteine metabolism. Thus, homocysteine and or the products of its metabolism are no longer the hidden causes or effects of various diseases.

Homocysteine Metabolism. Homocysteine is a nonessential sulfur amino acid produced from methionine, an essential amino acid derived from dietary protein. Homocysteine is formed through a series of steps (Figure 1), the first being the conversion of methionine to S-

1

adenosylmethionine (SAMe). S-adenosylmethionine is then demethylated to S-adenosylhomocysteine, which is then quickly hydrolyzed to homocysteine and adenosine. Intracellular homocysteine is either salvaged to methionine through remethylation or is converted to cysteine via the transsulfuration pathway. The remethylation pathway is catalyzed by two enzymes: methionine synthase (MS), which requires 5-methyltetrahydrofolate as a methyl donor, and methylcobalamin as cofactor; betaine:homocysteine methyltransferase (BHMT), which requires betaine as a methyldonor (2, 3). The first step in the transsulfuration pathway is catalyzed by the enzyme cystathionine β-synthase (CBS) (4), which forms cystathionine. Cystathionine is then degraded to alpha-ketobutyrate and cysteine through the action of cystathionine-gamma-lyase (CGL) (5). Both these reactions require vitamin B_6 as a cofactor. Cysteine then leads to glutathionine biosynthesis. The transsulfuration reaction provides a direct link between homocysteine and glutathione, the major redox buffer in mammalian cells. Homocysteine undergoes metabolic reactions depending on the status of the organism. Intracellular homocysteine is probably exported into the extracellular medium when the rate of homocysteine formation exceeds the metabolic capacity. This may occur either due to increased homocysteine formation or some aberration in homocysteine metabolism. Such an imbalance between production and metabolism of homocysteine can increase the amount of homocysteine in extracellular media like plasma or urine. Normal homocysteine concentration ranges between 5 to 10 μmol/L when measured in the fasting state (6). In plasma, homocysteine is extensively bound to various plasma proteins by disulfide and peptide bonds (7). Almost no unbound homocysteine is found circulating in plasma. Currently, homocysteine concentrations are measured as total plasma concentrations.

Protein

H₄folate

CH_2-H_4folate

CH_2-H_4folate
Reductase

CH_3-H_4folate

Methionine Adenosyl
Transferase

Methionine

AdoMet

Methionine
Synthase
(B₁₂ cofactor)

Betaine
Homocysteine
Methyltransferase

Methyl Transferases

Homocysteine

AdoHcy

Adenosyl
Homocysteinase

Serine

Cystathionine-β-synthase
(PLP-cofactor)

Cystathionine

α-ketobutyrate

Cystathionine-γ-lyase
(PLP-cofactor)

Cysteine

ATP +
Glutamate

γ-Glutamyl Cysteine
Synthetase

γ-Glu-Cys

ATP +
Glycine

Glutathione Synthetase

Glutathione (GSH)

Glutathione Peroxidase

Glutathione Reductase
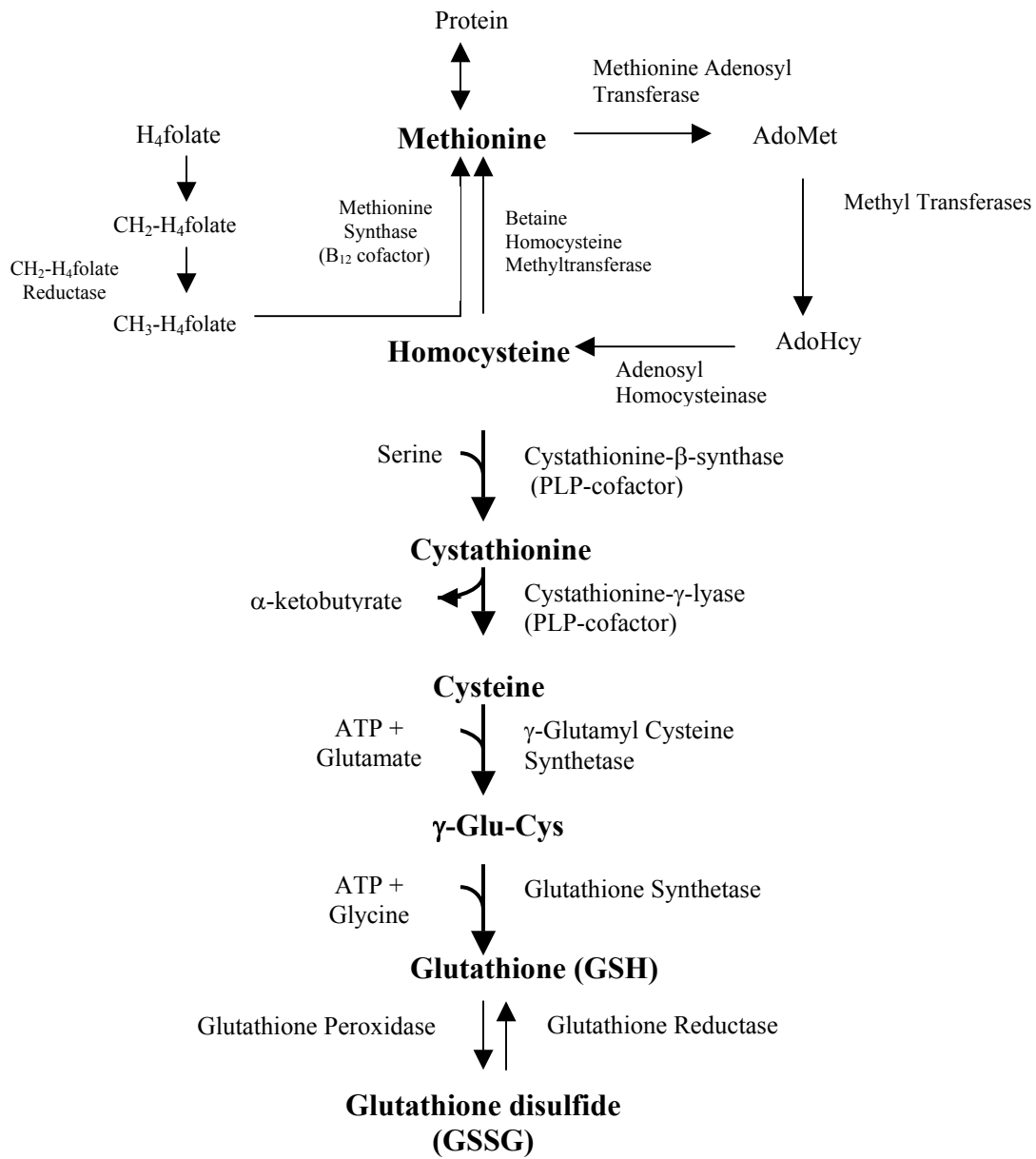
**Glutathione disulfide
(GSSG)**

Figure1: Intracellular Metabolism of homocysteine

Factors influencing homocysteine metabolism. Nutritional deficiency of vitamins is the primary cause of hyperhomocysteinemia. The vitamins that act as cofactors for the enzymes of homocysteine metabolism are folate, vitamin B_6 and vitamin B_{12} (8, 9). The most common deficiency *i.e.*, the folate deficiency, has the most profound influence on homocysteine concentration. Low folate concentrations have been associated with increased plasma homocysteine concentration (10). Studies with vitamin B_6 and B_{12} have shown non-linear correlation between homocysteine concentrations and these vitamin concentrations (11, 12). Administration of either of these vitamins alone is ineffective for treatment of the disease, but when co-administered with folic acid, homocysteine concentrations return to normal.

Homozygous hyperhomocysteinemia caused by deficiency of cystathionine β-synthase results in homocysteine concentrations being as high as 200 μmol/L (13). Heterozygous hyperhomocysteinemia is caused by deficiencies of the enzymes CBS, methyltransferase, methylenetetrahydrofolate reductase, or the substrates/cofactors folate, cobalamin, betaine or choline (14, 15, 16). Defects in the remethylation pathway, like impaired $N^{5,10}$-methylenetetra-hydrofolate reductase (MTHFR) or methionine synthase, can lead to compromised homocysteine metabolism and hence hyperhomocysteinemia (17, 17, 19, 20).

Demographic characteristics like age, gender and race can also affect homocysteine metabolism. Homocysteine concentrations increase with age in both genders. Men compared to age-matched women generally have hyperhomocysteinemia. Whites compared to African-Americans are prone to develop hyperhomocysteinemia (21, 22, 23).

Certain drugs can also cause elevated homocysteine concentration. Anticonvulsants like phenytoin and lipid lowering drugs like cholestyramine interfere in folate metabolism.

Methotrexate and nitrous oxide hinder remethylation pathway enzymes. Theophylline may hamper the transsulfuration pathway by hindering vitamin B_6 synthesis (6, 11, 21, 22, 24, 12).

Various disease states can be associated with hyperhomocysteinemia, the most common being chronic renal failure (13, 25, 26). Elevated homocysteine concentrations have also been found in diseases like psoriasis, pernicious anemia, hypothyroidism and various cancers (13, 22, 25, 26, 28).

Thus there is an increasing interest in monitoring plasma homocysteine because the level increases during several diseases or deficiency states. Homocysteine is not only linked to cardiovascular disease but it could also serve as a biomarker for various other broad ranges of diseases like diabetes, Alzheimer's disease, AIDS etc.

HOMOCYSTEINE METABOLISM AND HIV INFECTION

A disturbed sulfhydryl homeostasis in HIV-positive individuals was initially reported by Eck et al (29). The sulfhydryl pool, mainly comprising of glutathione and cysteine, are the by-products formed during the metabolic pathway of homocysteine breakdown by transsulfuration (Fig1). Glutathione exists almost exclusively intracellularly, where reductase enzymes keep it in a reduced form. Cysteine is found mainly in the plasma and exists in reduced, oxidized and protein-bound forms. Homocysteine, which is present mainly in the oxidized and protein-bound form and a small fraction in the reduced form, is also considered to play an important role in the maintenance of extracellular redox thiol status (4). A slight tilt of the delicately balanced redox-thiol status in favor of the pro-oxidant state can lead to the catastrophic progression of HIV infection into full blown acquired immunodeficiency syndrome (AIDS).

Epidemiology. Decreased concentrations of acid-soluble thiols (glutathione and cysteine) in plasma and peripheral blood mononuclear cell (PBMC) lysates was found in a small group of HIV-infected individuals (29, 30). This was followed by a confirmation of low plasma glutathione (GSH) levels in lung epithelial fluid (31). The Herzenberg group, by measuring GSH levels in PBMC subsets in 134 seropositive subjects in different stages of disease, showed that CD4 and CD8 T cells with high glutathione levels were selectively lost during HIV infection (32). In another study of HIV-infected persons, the mean plasma cysteine level was found to be 5 μmol/L, the normal concentration range being 10-20 μmol/L (33). Later in 1996, it was found that HIV-infected individuals showed high plasma levels of the reduced form of homocysteine. A study carried out in Netherlands in 1998, on 35 HIV-infected patients and an equal number of healthy controls, proposed that glutathione redox balance in CD4+ lymphocytes could be important in the pathogenesis of HIV infection and have implications for therapy (34).

Pathophysiology of HIV infection. The beginning of Human Immunodeficiency Virus (HIV) infection is the introduction of retroviral genome into T lymphocytes and mononuclear phagocytes. After an initial viremic stage, free virus is substantially reduced and the infection becomes latent. At this stage, most of the virions are integrated into the DNA of T cells (35). This early infection is characterized by a low number of infected cells and a low level of viral expression (36). The progress from this asymptomatic stage to active, late-stage acquired immunodeficiency syndrome (AIDS) is characterized by higher levels of viremia and p24 antigenemia, activation of HIV expression in the infected cells, an increased number of infected cells, a marked immune dysfunction (35, 36) and massive loss of skeletal muscle mass (cachexia) (37, 38).

The evolution of HIV infection and the progression of immunosupression, which is associated with an increased activation of latent virus, are influenced by various factors. These factors can induce HIV expression by acting at the transcriptional and posttranscriptional level. All the factors have a common *modus operandi, i.e*., to create oxidative stress, which leads to the activation of the nuclear transcription factor, (NF-κB), which then further induces HIV expression with the disease progressing into AIDS.

Role of NF-κB in HIV life cycle. The target genes under the regulation of NF-κB include a variety of cellular and viral genes (39, 40). These include genes encoding cytokines like TNF-α (tumor necrosis factor), GM-CSF (granulocyte macrophage-colony stimulating factor), interleukin-1, interleukin-6, interleukin-8, the interleukin-2 receptor alpha chain, cell adhesion molecules such as ICAM-1 and E-selectin, inducible nitric oxide synthase and viruses such as human immunodeficiency virus and cytomegalovirus (41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51).

NF-κB consists of two subunit molecules, p65 and p50, and usually exists as a molecular complex with an inhibitory molecule, IκB, in the cytosol (52, 53, 54, 55). Activators of NF-κB induce the dissociation of IκB from the NF-κB-IκB-complex, and positively charged nuclear location sequences (NLS) in the two subunits are unmasked (56). This activation cascade includes at least two independent pathways: kinase pathway and redox-signaling pathway. The kinase pathways include NF-κB kinase, a serine kinase (57) and IκB kinase, (58, 59, 60, 61, 62, 63, 64, 65, 66) known to phosphorylate these two molecules and activate the primary step of NF-κB dissociation. After dissociation from IκB, NF-κB undergoes a redox regulation by cellular reducing catalyst, thioredoxin (TRX). NF-κB is unable to bind to the κB DNA sequence of the

target genes until it is reduced by TRX (67, 68, 69, 70, 71, 72, 73). This activated NF-κB then binds to the target DNA element within the promoter region of HIV long terminal repeat region (LTR) (74) and then initiates HIV gene expression. This is followed by the production of virus-encoded trans-activator Tat, which triggers explosive viral replication (75, 76, 77, 78). But what are the factors that trigger these kinase and redox pathways? Involvement of reactive oxygen intermediates (ROI) was suggested to act in the upstream of the NF-κB activation pathway. This clue was given by the various studies of inhibition of NF-κB signaling by antioxidants such as N-acetylcysteine or α-lipoic acid (79, 80, 81, 82, 83, 84, 85).

HIV infection and oxidative stress. The generation of free radicals in cases of excessive or uncontrolled oxidation disturbs the pro-oxidant/anti-oxidant balance in favor of the pro-oxidant and this leads to a condition known as "oxidative stress" (86). Free radical damage plays an important role in the dysfunction of the immune system because the activation and proliferation of T lymphocytes requires a delicately orchestrated action of oxidizing and reducing substances (86, 87). One source of ROIs is provided by activated polymorphonuclear neutrophils and monocytes at the very early stages of HIV infections (88). These reactive oxygen species are released for killing bacteria and preventing opportunistic infections; but, when not properly compensated by antioxidant molecules, oxidative stress builds up. HIV also modulates superoxide anion production by human monocytes (89). Tat, a viral protein, is also specifically responsible for an endogenous cellular increase of ROI (90). ROI, which includes hydrogen peroxide and other oxygen radicals, act as messengers mediating directly or indirectly the activation of NF-κB, which then further induces the expression and replication of HIV in a human T-cell line. There are many other ways in which NF-κB could be activated but all these

different pathways of signal transduction involving NF-κB apparently merge on a ROI dependent step (79).

Some of the factors that may not be the direct cause but may be aggravators of oxidative stress and thereby NF-κB activators are:

a) *T cell mitogens*. Phorbol 12-myristate 13-acetate (PMA), a T cell mitogen, activates NF-κB via protein kinase C (PKC) activation (59). Various other studies however suggested that PKC does not directly activate NF-κB. PKC could activate NF-κB by directly enhancing the activity of an NADPH oxidase-like enzyme. These enzymes generate superoxide anions and hydrogen peroxide from oxygen and NADPH. Within seconds after PMA stimulation, PKC was found to activate the plasma membrane bound NADPH oxidase, causing the oxidative burst reaction (91). Treatment of Jurkat cells with PMA rapidly decreased intracellular thiol levels, supporting the idea that there is depletion of GSH by oxidants (92). Thus these results, strengthened by the observation of inhibitory effect of antioxidant N-acetylcysteine (NAC) on the activation by PMA of NF-κB (80), strongly suggested that the effect of phorbol ester treatment relies on the production of ROI.

b) *Cytokines*. The ability of HIV to replicate in CD4+ T lymphocytes and mononuclear phagocytes (MP) is strongly influenced by immunoregulatory cytokines (93). Some of the cytokines whose genes are switched on by activated NF-κB, such as TNF-α (tumor necrosis factor), IL-2 (interleukin), IL-6, IL-8 and GM-CSF (granulocyte macrophage-colony stimulating factor), are themselves activators of NF-κB. TNF-α has been shown to trigger virus expression both in T-cells and MP whereas other cytokines like colony stimulating factors, IL-1, IL-3 and IL-6 have demonstrated upregulatory effects on HIV replication in MP (94). It is important to

note the fact that increased TNF-α production had been reported in infected individuals both in plasma (95, 96) and /or freshly isolated peripheral blood mononuclear cells, circulating monocytes, tissue macrophages and B lymphocytes (94). The results showing that TNF-α produces superoxide anions and hydrogen peroxide in granulocytes, fibroblasts and other cell types (97, 98, 99) and that the antioxidant NAC inhibited the activation of NF-κ B by TNF-α, confirmed the hypothesis that ROI was involved in the cytokine-mediated stimulation of NF-κB. Increased levels of another cytokine, IL-6 was present in either the plasma or cerebrospinal fluid of infected individuals (100, 101). *In vitro* HIV infection of MP also led to the secretion of this cytokine (102). All these observations suggested that IL-6 alone or combination of IL-6 and GM-CSF induced HIV expression by acting both at the posttranscriptional level as well as at the transcriptional level in synergy with TNF-α. This confirmed that IL-6 is an important factor in the regulation of virus expression. Interleukin-1, a pleiotropic cytokine, was also shown to stimulate virus replication in primary monocyte-derived macrophage (MDM) cells (103). T-cell activation and proliferation induced by mitogen and IL-2, triggered the completion of reverse transcription followed by integration of the proviral DNA and active replication of HIV (104). Colony stimulating factors and IL-3 have also been shown to upregulate HIV replication and expression in cells of the MP lineage (105, 106, 107).

   c) Homocysteine and Oxidative stress. Redox dysregulation, resulting in oxidative stress was found to play an important pathophysiological role in patients with HIV. Various investigators have monitored different parameters that cause a redox imbalance and thereby oxidative stress. Reduced, oxidized and protein-bound forms of homocysteine, cysteine and cysteinylglycine in plasma interact via redox and disulfide exchange reactions, and these

aminothiol species comprise a dynamic system referred to as redox thiol status. This status can be influenced by an elevation of plasma homocysteine level (108). Protein bound homocysteine increases the protein binding saturation point and then there is a concurrent displacement of protein bound cysteine along with substantial increase in the free oxidized and reduced homocysteine. This resulting increase in the ratio of reduced:total homocysteine causes a similar change in this ratio for other aminothiols. These dynamics were observed during HIV infection that was further elucidated by Muller *et al.* (109). When the total, reduced, and protein bound forms of the thiols homocysteine, cysteine, cysteinylglycine and methionine in plasma from 21 HIV-infected patients and 15 healthy control subjects were quantified, it was found that the HIV-infected patients had significantly higher concentrations of reduced homocysteine in plasma compared with control subjects. Compared with control subjects, the HIV-infected patients had lower concentrations of methionine in their plasma. High homocysteine levels and low methionine concentrations could be associated with the below-normal serum cobalamin concentrations found in HIV patients (110, 111).

An elevated concentration of reduced homocysteine is relevant as a causative factor of oxidative stress in the immunopathogenesis of HIV infection because it has shown to be a contributive factor in the production of ROI. The sulfhydryl group of homocysteine is capable of acting catalytically with cupric or ferric ions to generate hydrogen peroxide and various homocysteinyl radicals (112, 113, 114). Homocysteine coupled with an increasing concentration of copper led to hydrogen peroxide production in a dose-dependent manner. Ceruloplasmin was also demonstrated to catalyze homocysteine auto-oxidation since the majority of plasma copper exists in the form of ceruloplasmin. This result could be related to the longitudinal study, which

showed that patients progressing to AIDS have significantly higher serum copper concentrations compared with the non-progressors (115). Moreover in the presence of metal ions, the hydrogen peroxide released by homocysteine can react to form the highly reactive hydroxyl radical (112) that is responsible for the NF-κB activation. Copper can also induce an increased release of homocysteine in the extracellular medium (116). Another important result that contributed to homocysteine's role in oxidative stress is that homocysteine inhibits the antioxidant enzyme, glutathione peroxidase *in vitro* and leads to a dramatic reduction in steady state mRNA levels for the intracellular isoform (117). This effect was unique to homocysteine and could explain the surprising result that cysteine is non-toxic, despite being present in plasma at concentration at least three-to-fourfold greater than that of homocysteine and equally capable of generating free radicals during auto-oxidation (118). Due to the pro-oxidant properties of homocysteine, it was speculated that elevated circulating concentrations of reduced homocysteine in HIV-infected patients could be one of the several factors contributing to an enhanced production of ROI, which in turn could lead to NF-κB activation and further stimulation of HIV replication. This speculation was corroborated when it was demonstrated that homocysteine indeed increased NF-κB-DNA binding activity in a concentration dependent manner by causing a reduction in the expression of I-κB (119). In an attempt to study the effects of glutathione precursors on HIV replication, it was found that homocysteine, even though a natural cysteine precursor, stimulated PMA-mediated HIV-LTR transactivation in Jurkat T-cells and GM-CSF mediated HIV expression (223).

Homocysteine is also capable of indirectly activating NF-κB by causing the upregulation of IL-6. Incubation with homocysteine but not methionine nor cysteine resulted in a dose

dependent increase in IL-6 accumulation (121). This specific action of homocysteine could contribute to the oxidative stress since IL-6 is capable of NF-κB stimulation as described earlier. Taking into consideration these harmful effects of homocysteine, it is not surprising that there are concerns over the speculation that elevated levels of reduced homocysteine might contribute to the pathogenesis of ischaemic heart disease in HIV patients (122).

Role of Cysteine and Glutathione in HIV infection. Various reasons have been suggested for the abnormally low levels of plasma cysteine and glutathione levels in HIV-1 seropositive individuals. One proposal is that the HIV viral infection, through the production and release of TAT (a protein encoded by HIV for efficient viral replication), creates additional oxidative stress, at least in part by decreasing γ-glutamylcysteine synthetase mRNA and protein content (123) and thus hindering the synthesis of glutathione. This deficiency of intracellular glutathione could relate directly to reduced cysteine levels in plasma for two reasons. First, a decrease in hepatic glutathione biosynthesis has been shown to reduce the export of thiols to the peripheral circulation (124). Second, glutathione breakdown is essential for the formation of cysteinylglycine, which is further cleaved into cysteine and glycine (122). Another explanation that could contribute to the cause of low plasma cysteine levels in HIV patients is the result of a very recent study that revealed that patients with AIDS had a deficiency of hepatic cystathionine γ-lyase, the enzyme that converts cystathionine into cysteine (125).

Reduced glutathione (GSH), the main intracellular antioxidant, is an important immunomodulator that is required for T-cell activation (126). GSH, which constitutes greater than 90% of the cellular non-protein thiols provides cells with their reducing equivalents and serves as the major cellular antioxidant (127). Intracellular glutathione levels play an important

role in modulating signal transduction and the cytokine-stimulated expression of genes controlled by the HIV LTR. Reduced glutathione reduces the reactive oxidative intermediates generated during oxidative stress and in the process is oxidized. The oxidative stress created by TNF-α and PMA in certain kinds of cells (128, 129, 130) causes a rapid decrease in the GSH levels, which indirectly causes a significant augmentation of stimulation by TNF-α since a low glutathione level is necessary to permit signal transduction by cytokines (131). This result was particularly relevant in terms of maintaining latency in HIV-infected individuals, since GSH levels in these individuals are consistently below normal (30, 31, 132, 133). Glutathione depletion in HIV infection is also caused partly by decreased availability of cysteine that is rate limiting in glutathione synthesis (30, 134). HIV infection could lead to liver problems, and this liver dysfunction could contribute to a systemic glutathione deficiency (135). Another hypothesis derived from the fact that HIV-infected patients have glutathione deficiency is that AIDS might be the consequence of lower glutathione disulfide (GSSG) (136). But this speculation was challenged by the results demonstrating increased levels of oxidized glutathione in CD+ lymphocytes associated with HIV infection (137). A very important result, recently obtained by Mosharov *et al.*, demonstrated that homocysteine flux through the transsulfuration pathway, which eventually leads to the generation of cysteine and glutathione, increased approximately two fold under oxidative conditions, thus indicating that any defective enzyme in this pathway could create or aggravate the havoc created by the oxidative stress during HIV infection (138).

Increasing evidence suggests that abnormal cysteine and glutathione metabolism are important in the development of catabolic conditions and related immunological dysfunctions. Elevated plasma glutamate levels in HIV infected persons (29, 139) also play a decisive role in

the lowering of cysteine and glutathione concentration (29). Glutamate levels are important in this context because elevated extracellular concentrations of glutamate were found to decrease intracellular cysteine and glutathione levels in macrophages and peripheral blood mononuclear cells (29) by inhibiting the membrane transport activity for cystine (136). The indirect activation of tumor necrosis factor-α (TNF-α) caused by HIV-induced cysteine deficiency is also commonly believed to be largely responsible for cachexia (33).

Therapeutic Intervention with N-acetylcysteine. N-acetylcysteine (NAC), a pro-drug that replenishes cysteine, is the most studied drug after reverse transcriptase inhibitors for HIV infection. Reverse transcriptase (RT) inhibitors are therapeutically relevant only up to the initial stage of integration of viral DNA into the human genome *i.e.*, they are ineffective at inhibiting viral production once latency has been established. But N-acetylcysteine, the thiol-enhancing drug, has a very broad therapeutic range, from the asymptomatic stage up to the symptomatic stage of HIV infection. NAC can inhibit the stimulation of viral production during latency (32) as well as during the active HIV replication cycle.

The primary function of NAC is as an antioxidant. NAC is a direct scavenger of ROI (140) like, hydrogen peroxide, hydroxyl radical and hypochlorous acid (141). Thus NAC, by reducing these free radical species can inhibit their activation of NF-κB. NAC may also exert its antioxidant effect indirectly by facilitating glutathione biosynthesis and supplying glutathione for GSH-peroxidase-catalyzed reactions (142). An 8-week double blind study suggested that long-term high-dose NAC treatment by virtue of its glutathione replenishment is safe and is associated with longer survival in HIV-infected individuals (143).

N-acetylcysteine not only inhibits the direct activation of NF-κB by free radicals but also

blocks cytokine stimulated activation of NF-κB and HIV replication in an acutely infected T-cell line and peripheral blood mononuclear cells (80). Restoration of GSH levels in persons with latent HIV infection by treatment with NAC could help inhibit signal transduction from inflammatory cytokines, hence, prevent cytokine-stimulation of HIV replication in latent-infected cells (132).

N-acetylcysteine also acts at a posttranscriptional stage of HIV replication by inhibiting HIV replication in chronically infected monocyte and T-cell lines, which are used as models for latent infection in AIDS (133). This inhibition of the production of virions from chronically infected cell lines was due to the inhibition of synthesis of RNA as directed by the HIV LTR. Those results suggest that NAC could be used as an adjunct in the therapy of HIV positive, but asymptomatic patients. NAC may slow or stall the change from latency to the later stages of AIDS. This observation was supported by a study demonstrating the down-regulation of HIV expression in a chronically infected cell line (U1) by using GSH and NAC as therapeutic agents and relating them to the levels of RT, cell-associated viral proteins and mRNA, which are the markers of progressive HIV expression (144).

Altogether the clinical data on use of NAC as a cysteine and glutathione replenisher confirm the massive loss of sulfur in HIV infected patients. The immunoreconstituting effect of cysteine supplementation adds to the importance of transsulfuration pathway, which converts homocysteine to cysteine and glutathione. Homocysteine is the source from which approximately half of the intracellular glutathione pool in human liver cells is derived via the transsulfuration pathway (138). The redox sensitivity of the transsulfuration pathway can be justified as an autocorrective response that leads to an increased level of glutathione synthesis in cells

challenged by oxidative stress. Aberrations in this pathway could compromise the redox buffering capacity of cells, which may in turn be related to the pathophysiology of not only AIDS, but also different homocysteine-related diseases like atherosclerosis and Alzheimer's disease.

HYPERHOMOCYSTEINEMIA AND ALZHEIMER'S DISEASE

The medical establishment became aware of a possible danger related to homocysteine when various epidemiological studies revealed the association between hyperhomocysteinemia and Alzheimer's disease (AD). Hyperhomocysteinemia was already known to be an independent risk factor for vascular disease and AD was considered to be independent of vascular involvement. But the deluge of data accumulated through various clinical studies established a vascular component in the etiology of AD.

Dementia, defined as the progressive decline in mental function, memory and acquired intellectual skills (145, 146) is classified into two categories: vascular dementia and Alzheimer's dementia. Vascular dementia is mainly caused by multifocal stroke due to cerebral arteriosclerosis. The evidence that could logically link the vascular dementia to that of the Alzheimer's is that individuals commonly had both AD and vascular dementia (147) and that the prevalence of AD was shown to be increased with the degree of atherosclerosis. (148). Another important observation was the presence of apolipoprotein E allele E4 (apo E4) as a common risk factor for both AD and vascular disease (148, 149). The most convincing evidence was the finding in the Nun Study that there was a higher prevalence of cognitive impairment and dementia in individuals with both AD and brain infarcts as compared to those with only AD

(150). This evidence when compounded with the epidemiological data supported the hypothesis that hyperhomocysteinemia, recognized as a risk factor in cardiovascular, peripheral vascular and cerebrovascular disease (151), also plays an important role in the pathogenesis of AD.

Epidemiology. According to neuropsychologists, mild cognitive impairment, where the neurons in the brain start degenerating, occurs before the disturbances reach the level of dementia, in which the neurons can get completely destroyed and the brain's reserve capacities get exhausted. Chiu *et al.* have coined the term 'dysmentia' for a stage of cognitive impairment that is too severe to be considered part of normal aging, but not severe enough to fulfill the criteria for dementia, which is the extreme case of cognitive impairment (152). The risk factors known for the Alzheimer's type of dementia may have been operating over a long time before the clinical onset of the Alzheimer syndrome and therefore it is useful to investigate the preclinical markers like nutritional status, amino acid concentrations, genetic vulnerability, and psychological markers for detection of cognitive impairment (153). Goodwin *et al.* were among the first researchers to find an association between nutritional status and cognitive functioning in healthy elderly population (154). Their evaluation of the relationship between the vitamin status of B_{12} and folic acid and cognitive impairment in 260 men and women older than 60 years, showed that subjects with low levels of vitamin B_{12} and folic acid scored worse on different memory-related tests. It was also found that significant improvement occurred in neuropsychiatric functions among cobalamin-deficient patients after vitamin B_{12} supplementation (155). A study carried out in 1992, reported cognitive recovery after B_{12} supplementation in patients with cobalamin deficiency states of short duration (156). Another important study demonstrated that the elderly depressed patients, who had lower cognitive

screening test scores had significantly higher homocysteine concentrations than did either younger depressed patients or elderly depressed patients with normal cognitive screening test scores (157). Individual vitamins and homocysteine levels are important determinants of patterns of cognitive impairment (158). This suggestion was based on the results obtained from an investigation of the relation between plasma concentrations of folate, vitamin B_{12}, vitamin B_6 and homocysteine with the scores on battery of cognitive tests in 70 men, aged 54-81 years, participating in the Normative Aging Study. It was found that lower folate and vitamin B_{12} concentrations associated with poorer spatial copying skills. Also plasma homocysteine proved to be a strong predictor of spatial copying performance than either vitamin B_{12} or folate (158). Cognitive-function tests when performed (159) in 336 patients diagnosed with early AD, senile dementia, vascular dementia, minor cognitive impairment (dysmentia), led to the finding that homocysteine can be declared as an early marker of cognitive impairment in the elderly (159). A test of possible relation between serum total homocysteine concentrations and performance on two delayed-recall tests administered to subjects aged greater than 60 years showed that hyperhomocysteinemia is related to poor recall (160).

As the evidence for hyperhomocysteinemia and its ramifications on cognitive function in the elderly were emerging, various studies were carried out to quantify the homocysteine level and associated vitamin status in Alzheimer patients. The initial result of low levels of vitamin B_{12} observed in the cerebrospinal fluid of AD patients, obtained by Ikeda *et al.* (160) was followed by a study, which examined the relationship between folate, vitamin B_{12} and severity of cognitive impairment in patients with AD (162). This study also suggested a negative relationship between B_{12} levels and severity of cognitive impairment. A study published in Journal of Gerontology

and Biological Sciences by Joosten *et al.* (163) showed that homocysteine levels were significantly elevated in Alzheimer's disease patients compared to controls. It also found that folic acid and vitamin B_{12} deficiencies were present in both the Alzheimer's disease patients and the age-matched case controls. Later, in 1998, another group was able to confirm that patients diagnosed with senile dementia of the Alzheimer's type had significantly elevated levels of homocysteine compared to age-matched controls (164). The most convincing study was performed by Carke *et al.* wherein an inverse relation between serum folate and vitamin B_{12} levels and the risk of AD was established. The strength of this study was that the AD diagnosis for a large subset of the patients was confirmed upon postmortem analysis, which dispelled any notion that the differences between AD patients and controls in the homocysteine level and vitamin status could be the result of misdiagnosis (166). In the same year, David Snowden, known for his study of the brains of old-aged nuns, found that the sisters with high folate levels showed little evidence of Alzheimer's-type damage in their brain after death (150).

This wealth of epidemiological data slowly formed a definitive link between vascular disease and AD. However, there is a need to unravel which and how vascular risk factors contribute to dementia of AD (165). Even though hyperhomocysteinemia might seem to increase the risk for AD's dementia, it is impossible to distinguish, at least based on the epidemiological evidence, what underlies the associations. There could be various possibilities, vascular disease and AD may have some common etiologies, or vascular disease is indeed involved in the etiology of AD (167). Also, according to some authors, these biochemical changes involved in vascular disease could be a consequence, rather than a cause, of Alzheimer's disease, and that further work is required to distinguish between these two interpretations (168).

Hyperhomocysteinemia as a Consequence. The reason why homocysteine levels are so high in people suffering from dementia of the Alzheimer's type is not fully understood. Taking together the information obtained from clinical studies indicating inadequate B vitamin status in AD, a logical conclusion that can be derived is that hyperhomocysteinemia is a result of these inadequacies, the B vitamin intake and status being a major determinant of serum homocysteine (169). Most of the studies showed a deficiency of folic acid and vitamin B_{12} in AD patients, and because tetrahydrofolate and cyanocobalamin (vitamin B_{12}) are required by enzymes that metabolize homocysteine, there is a ready biochemical explanation for this epidemiological association. The circulating homocysteine concentrations reflect the vitamin B status and deficiency, which is also known to be a significant risk factor for cognitive impairment and dementia (160). Some scientists speculate that severe aberrations in the methylation cycle might be involved in the disease process and that elevated homocysteine could be a sign of the breakdown of the methylation cycle. The methylation cycle involves the conversion of homocysteine to methionine with the help of two enzymes, methionine synthase, which requires N^5-methyltetrahydrofolate as substrate and vitamin B_{12} as cofactor and betaine:homocysteine methyltransferase, which requires betaine as substrate. Any defect in this pathway could lead to the accumulation of homocysteine and S-adenosylhomocysteine and depletion of methionine. Methionine is believed to be the most crucial part of the pathway for the health of brain tissue because a considerable proportion of methionine is converted by ATP:1-methionine S-adenosyltransferase (EC 2.5.1.6, MAT) to form S-adenosylmethionine (SAMe), which serves primarily as a methyl donor in a variety of reactions. Since methylation is required for the maintenance of the myelin sheath and the repair of DNA in the brain, a hypothesis was proposed

that loss of neurocognitive function in the elderly is due in part to impaired methylation reactions in the brain tissue (170). Impaired methylation reactions could involve aberrations at several steps of the methylation cycle. First, inadequate methionine formation due to vitamin B_{12} and folate deficiencies found in AD patients; second, low SAMe levels in AD patients (171, 172) that could be attributed to reduced methionine availability or decreased methionine adenosyltransferase activity as found in the erythrocytes of patients with dementia disorders (173). A study carried out by Bottigileri *et al.* in 1994 confirmed the role of methylation in cognitive impairment by demonstrating that SAMe improved cognition in some Alzheimer's patients (174). Also upon transfer of its methyl group, SAMe is converted to S-adenosylhomocysteine (SAH), followed by hydrolysis to homocysteine and adenosine, the hydrolysis being a reversible reaction that favors SAH synthesis (169). Therefore in a state of folate or vitamin B_{12} deficiency, inability to methylate homocysteine leads to SAH accumulation and SAH being a potent inhibitor of various SAMe-dependent methylations, aggravates the dangers of under-methylation reactions in the brain (169).

Another reason for hyperhomocysteinemia was an abnormal amino acid metabolism in the early stages of AD. Significantly reduced levels of amino acids including methionine in the AD patients were found as compared to the control subjects (175). This also strengthened the hypothesis that the methylation deficiency is due to limited availability of methionine in AD patients.

Thus, while some scientists speculated that homocysteine represented an important measurable biomarker of vitamin deficiencies, methylation deficit and abnormal amino acid metabolism, a lot of research was carried out to determine whether homocysteine itself

contributes to Alzheimer's disease.

Hyperhomocysteinemia as a Cause. Oxidative stress has been implicated in the pathogenesis of AD and other neurodegenerative disorders. A number of factors contribute to making the neurons in the brain particularly vulnerable to free radical induced damage, the important factors being their low glutathione content, their membranes containing a high proportion of polyunsaturated fatty acids and the brain metabolism requiring substantial quantities of oxygen (176, 177). Evidence for free radical involvement include: substantial increase in free radical production in homogenates of AD frontal cortex as a result of incubation with ferrous sulphate (178); lesions in the brain that are typically associated with free radical attacks such as DNA damage (179), protein oxidation (180), lipid peroxidation (181, 182); presence of metals having catalytic activity to produce free radicals (183, 184, 185) and production of free radicals mediated by aggregated beta amyloid in the presence of free radicals (186).

Homocysteine and copper can aggravate AD in two ways, first, by production of reactive oxygen species and second, by reducing Cu (II) to Cu (I) and potentiating copper neurotoxicity. Homocysteines's ability to cause endothelial cell injury due to copper catalyzed hydrogen peroxide generation from homocysteine was already known (187). Another free radical produced during auto-oxidation of homocysteine is superoxide anion (112). Zhang *et al*. elucidated the exact mechanism by which homocysteine can exert profound cerebrovascular effects mediated by copper catalyzed production of superoxide anion (188). Homocysteine selectively impaired cerebrovascular responses mediated by nitrous oxide (NO) via superoxide-dependent mechanisms. Homocysteine in the presence of copper (Cu II) reduces the resting cerebral blood

flow, which strengthens the hypothesis that homocysteine impairs cerebrovascular regulation by decreasing the ability of the brain circulation to compensate for the reduction in blood flow that occurs during ischaemia. It was later proposed that homocysteine could compromise cerebrovascular regulation by interfering with the action of endothelial vasodilators such as NO. This possibility was confirmed by showing that superoxide anion, the homocysteine auto-oxidation product scavenges NO via rapid formation of peroxynitirite and thereby reducing the amount of NO available for vasodilation.

Homocysteine-copper toxicity could also depend on the ability of homocysteine to reduce Cu (II). This was based on the observation of inhibition of toxicity with the Cu (I)-specific chelator, bathocuproine disulphonate (188). Copper itself has been shown to be neurotoxic because of its ability to generate free radicals via a Fenton type of reaction (190) and to cause peroxidation of mitochondrial membranes (191). Also the aggregated beta amyloid protein, which is present in senile plaques of AD patients, can associate with Cu (I) and lead to hydrogen peroxide formation (192, 193). Thus, homocysteine by readily reducing Cu (II) to Cu (I), potentiates copper neurotoxicity. This action is unique to homocysteine because it is more effective in reducing Cu (II) as compared to other thiols like methionine and cysteine (188).

Evidence for altered mitochondrial functions has been reported in AD patients. According to a post-mortem study, cytochrome-*c* oxidase activity was approximately 25-30% lower than normal in the cerebral cortex and in the platelets of AD patients (194). A decline in cytochrome-c oxidase activity is associated with a decreased expression of messenger RNA molecules (195). Homocysteine has been shown to play a role in this altered mitochondrial gene expression, function and structure. Homocysteine in the presence of copper, known to produce

hydrogen peroxide, significantly decreased mitochondrial RNA levels, caused gross

morphological changes in mitochondrial ultrastructure, inhibited both cell growth and

mitochondrial respiration rates (196). AD has also been associated with a low level of oxidative

phosphorylation, this being expressed by energy deficits, *i.e.,* low ATP levels (197, 177).

Depletion of cellular ATP could result from the conversion of excess intracellular homocysteine

to homocysteine thiolactone, an ATP-dependent reaction catalyzed by methionyl-tRNA synthase

(198). Thus homocysteine, by causing mitochondrial anomalies, potentiates the production of

free radicals and a reduction in energy resources.

Homocysteine was reported to have neurotoxic properties due to its ability to act as an

excitotoxin. Human cell lines, derived from separate discrete areas of the brain, when exposed to

homocysteine and its derivative homocysteic acid and then assayed for lactate dehydrogenase

activity as a measure of cell death, showed that neurotoxic responses are cell-type specific for

homocysteine and its metabolites (199). Homocysteine exhibited its neurotoxicity through

overstimulation of the N-methyl-D-aspartate receptor (NMDA). Homocysteine acts as partial

antagonist at the glycine-binding site and as an agonist at the glutamate-binding site of the

NMDA receptor. Under pathological conditions such as stroke or trauma, homocysteine

outweighs its neurotoxic action (agonist) against its neuroprotective (antagonist) activity and

causes neurodegeneration (200). Homocysteine is also involved in apoptosis, which is commonly

seen in Alzheimer's disease. Increasing results have shown DNA damage in Alzheimer's disease

(201), which could trigger a cell death cascade involving activation of poly-ADP-ribose

polymerase (PARP) (202) and induction of the tumor suppressor protein p53 (203).

Homocysteine, which is rapidly taken up by the neurons via a synaptosomal plasma membrane

transporter (204), induce DNA strand breaks by disturbing the DNA methylation cycle. This DNA damage is followed by PARP activation, caspase activation, p53 activation, mitochondrial membrane potential decline, and nuclear disintegration. PARP inhibition was shown to prevent these apoptotic steps, suggesting a major role for PARP activation in homocysteine-induced neuronal apoptosis and increased neuronal vulnerability to excitotoxicity (205).

In summary, the published literature provides compelling evidence that elevated homocysteine has a definitive link to Alzheimer's disease. Some scientists speculate that hyperhomocysteinemia itself may not cause Alzheimer's disease and is just a consequence of the pathological conditions of the disease. But considering the pro-oxidant and neurotoxic nature of homocysteine, it can be definitely considered as a dangerous causative factor of the disease. Either way, elevated homocysteine appears to represent an important measurable biomarker of the methylation deficits, oxidative stress and the progression of dementia of the Alzheimer's disease and therefore it presents itself as an important therapeutic target to escape the worst ravages of this heartbreaking illness.

HOMOCYSTEINE METABOLISM AND CARDIOVASCULAR DISEASE

Atherosclerosis is marked by the initiating event of impairment of endothelium, thickening of coronary arterial wall by the deposition of plaques, and reduction of blood flow due to narrowing of its lumen. Several risk factors are associated with atherosclerosis, each factor having its own ability to impair the endothelium (206)

High homocysteine levels have been shown to be one of the significant risk factors in atherosclerosis. The first evidence of linkage between plasma homocysteine and atherosclerotic

vascular disease was provided by McCully when the postmortem analysis of his patients with elevated plasma homocystiene concentration showed severe atherosclerosis (28). Recent studies have reinforced this tie between high homocysteine concentrations and increased risk of heart disease. In one study, among men and women younger than age 60, there was a 2.2 times higher risk of coronary and other vascular diseases when plasma homocysteine levels were in the top fifth of the normal range (207). Further, a Norwegian study found that death after four to five years for patients with coronary heart disease was proportional to plasma total homocysteine levels (20).

Although the exact relationship between atherogenecity and hyperhomocysteinemia is not established, various probable molecular and cellular mechanisms for homocysteine-induced endothelial injury have been proposed.

Homocysteine and free radical-mediated cellular damage. One of the stress-inducing effects of homocysteine has been attributed to its readily oxidizable thiol group resulting in reactive oxygen species. Homocysteine gets rapidly auto-oxidized in the plasma forming homocystine, mixed disulfides and homocysteine thiolactone (208). These produce superoxide, hydrogen peroxide and hydroxyl radicals (209). An *in vitro* relationship between endothelial cell injury and the concentrations of homocysteine was established by measuring cell detachment from confluent monolayers (210). The toxicity of homocysteine on endothelial cells was inhibited by catalase and not by superoxide dismutase, thereby indicating that hydrogen peroxide must be playing an important role. This was then confirmed when it was found that in cultured human and bovine endothelial cells, homocysteine and copper could lyse the cells in a dose-dependent manner, an effect that was completely inhibited by catalase. Ceruloplasmin was also

found to be equivalent to copper in catalyzing the oxidation of homocysteine and generation of hydrogen peroxide. (211). The mechanism of homocysteine auto-oxidation is explained by the following equations: (212, 213)

$$HCYS + M^{n+} \longrightarrow M^{(n-1)+} + HCYS^{\cdot}$$

$$M^{(n-1)+} + O_2 \longrightarrow M^{n+} + O_2^{\cdot-}$$

$$HCYS^{\cdot} + HCYS \longrightarrow HCY\text{-}SS\text{-}HCYS^{\cdot-}$$

$$HCYS\text{-}SS\text{-}HCYS^{\cdot-} + O_2 \longrightarrow HCY\text{-}SS\text{-}HCYS + O_2^{\cdot-}$$

$$O_2^{\cdot-} + O_2^{\cdot-} + 2H^+ \longrightarrow O_2 + H_2O_2$$
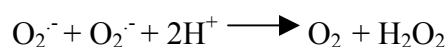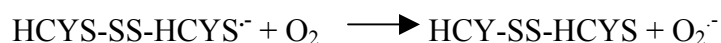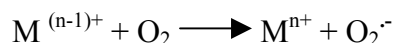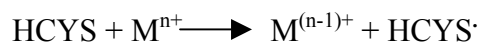
(HCYS – homocysteine; M – transition metal ion)

A different idea was put forth by a study that examined the effect of cysteine on auto-oxidation of homocysteine. It was found that homocysteine auto-oxidizes at a slower rate than cysteine but a very small concentration of cysteine/cystine is capable of accelerating homocysteine oxidation. It was therefore proposed that homocysteine only plays a role of reducing cystine to cysteine and it is the cysteine auto-oxidation that causes oxidative stress (214). However various studies have challenged this idea by demonstrating selective toxicity of homocysteine.

Homcysteine not only creates oxidative stress, but also aggravates the stress by attenuating the production of stress response proteins like heat shock proteins (mainly HSPs) and antioxidant enzyme like glutathione peroxidase (GPx). Under normal conditions, endothelial cells increase their expression and synthesis of heat shock/stress proteins in response to oxidative stress (215). The cells also increase the activity of certain antioxidant enzyme to limit oxidative damage, GPx being one of them (216). In addition to generating hydrogen peroxide,

homocysteine also compromises the cell's ability to detoxify the same by specifically impairing the intracellular isoform of GPx at the transcriptional level (217). A cDNA microarray study was performed to analyze the changes in gene expression in cultured human umbilical-vein endothelial cells (HUVEC) exposed to homocysteine. As per the microarray results, homocysteine blocked the hydrogen-induced expression of HSP70. Homocysteine elevated mRNA levels of GRP78 (glucose-regulated protein) and GADD153, which are stress-response genes, induced by agents or conditions that alter the function of endoplasmic reticulum (218). These effects of homocysteine on GPx and stress-response genes were very specific for homocysteine because other thiol containing compounds like cysteine failed to elicit such a response.

The reactive oxygen species produced during homocysteine auto-oxidation causes lipid peroxidation at the level of endothelial cell membrane (219) and also within lipoprotein particles (220). Homocysteine and iron play an important role in the oxidative modification of low density lipoprotein (LDL) (221) which is the first step in the macrophage-derived foam cell formation. GPx reduces the lipid peroxides to their corresponding alcohols as a defense mechanism but homocysteine interferes with GPx too (222). Homocysteine-mediated endothelial cell injury is due to the exposure of the underlying matrix and smooth-muscle cells to hydrogen peroxide, which thereby promotes the activation and adhesion of platelets and leukocytes (223).

Homocysteine and maintenance of vascular tone. Under normal physiological conditions, nitric oxide (NO) is produced by the endothelial cells to regulate the vascular tone. Homocysteine reacts with NO and endothelial-derived relaxing factor (EDRF) to form S-nitroso-homocysteine, which also shares the vasodilatory and antiplatelet effects of NO (223, 225). But

under the conditions of endothelial injury, the balance between the production of NO and formation of nitroso compound is destroyed. Lesser NO is available for production of S-nitroso-homocysteine, which leads to unopposed homocysteine-induced oxidative stress to endothelial cells. Homocysteine can also alter the bioavailability of NO by the reaction of superoxide and NO to form peroxy nitrite and this could explain the attenuation of endothelium-dependent vasodilator responses by homocysteine exposure (226, 227). GPx, by scavenging the reactive oxygen species can prevent the inactivation of NO, (228, 229), but homocysteine aborts this preventive action by interfering in the transcription of this enzyme. Homocysteine also reduces the redox pool of vascular cells by oxidizing glutathione (GSH) and disrupting the redox ratio (GSH/GSSG). Thus a co-substrate deficiency is developed for GPx, further hampering endothelial oxidative defense mechanisms (230).

Homocysteine and thrombosis. Homocystine, the oxidized form of homocysteine activates Hageman factor (factor XII), which is an important clotting factor for intrinsic clotting process (231). An increase in the procoagulant factor V (factor Leiden) was observed in homocysteine exposed vascular endothelial cells. This increase was attributed to the induction of an activator of factor V (232). This activator of factor V was found to be protein C. The maintenance of normal hemostasis is performed by anticoagulant proteins like thrombomodulin, protein C and protein S. Their main function is to inactivate factors Va and VIIIa (233). This pathway gets activated by the binding of thrombin to thrombomodulin, which in turn activates protein C. With protein S as an activator, the protein C inactivates factors Va and VIIIa. These factors then inhibit the prothrombinase activity. Rodgers *et al.* investigated the action of homocysteine on protein C and found that 90% inhibition of protein C activation occurred with

30

7.5 to 10 mmol/L homocysteine after 6 to 9 hours of incubation of arterial and venous

endothelial cells. Homocysteine did not induce an inhibitor to protein C but it acted as a

competitive inhibitor to thrombin (234). Homocysteine irreversibly impaired thrombomodulin

surface expression when tested in human umbilical vein endothelial cells and CV-1 (18A) cells

that express recombinant human thrombomodulin (235). Homocysteine was also found to induce

tissue factor (TF) procoagulant activity. TF, an integral membrane protein, plays an important

role in the initiation of extrinsic pathway by interacting with factor VIIa. Homocysteine, by

activating this factor at the transcriptional level, sets off the blood coagulation cascade (236,

237). Another mechanism by which homocysteine plays a role in thrombosis is by suppressing

the expression of heparan sulfate proteoglycans (HSPGs). HSPGs are important determinants of

anticoagulant properties of the blood vessel. They bind to antithrombin III and potentiate its

action as an inhibitor of the reactions of the coagulation cascade. Nishinaga *et al*. discovered that

there was a substantial reduction in the antithrombin III binding capacity of heparan sulfate

isolated from homocysteine exposed endothelial cells. This effect of homocysteine was inhibited

by catalase suggesting that hydrogen peroxide releasing action of homocysteine might be a

probable cause (238). Thus homocysteine can contribute the thrombotic tendency in patients

with atherosclerosis by various mechanisms.

      Homocysteine and activation of NF-κB.   NF-κB is a nuclear transcription factor and is

responsible for the stimulation of various genes pertinent to the pathophysiology of the vessel

wall. These genes include leukocyte adhesion moleucules (VCAM-1, ICAM-1 and E-selectin),

chemokines, cytokines and chemoattractant proteins. The promoters of leukocyte adhesion

molecules contain the consensus NF-κB sites and are induced by the action of inflammatory

cytokines on endothelial and smooth muscle cells (239, 240). Another important feature of

formation of atheromatous plaques is the recruitment of monocytes into the injured arterial wall.

This migration is stimulated by monocyte chemoattractant protein-1 (MCP-1). Homocysteine is

also capable of stimulating MCP-1 expression in human aorta vascular smooth-muscle cells.

This stimulation was associated to activation of protein kinase C (PKC) which, in turn inhibits

the expression of IκBα, the regulatory protein of NF-κB (241). This suggested that

homocysteine induces the expression of the adhesion molecules and the chemokines via the

activation of NF-κB. Homocysteine has these inflammatory effects on smooth muscle cells too.

It induces the expression of cyclin D1 via activation of NF-κB and this leads to a marked

increase in vascular smooth muscle proliferation. Homocysteine stimulates the NF-κB mediated

expression of inducible NO synthase in aortic smooth muscle and hence increases the production

of nitric oxide followed by oxidative stress (242).

Homocysteine in relation to cholesterol metabolism. Hyperlipidemia (total cholesterol,

low density lipoprotein (LDL) cholesterol, and the triacylglyceride (TG) levels) has been one of

the most important risk factors for atherosclerosis (243, 244). Absorption of dietary cholesterol

and synthesis of cholesterol in the liver are two sources of cholesterol in the human body.

Cholesterol levels in the human body are regulated by various ways. One important strategy of

regulation of cholesterol is by regulation of the levels and activity of the enzyme HMG-CoA

reductase (hydroxymethyl glutaryl -coenzyme A), which converts HMG- CoA to mevalonate,

the rate-limiting step in the biosynthesis of cholesterol (245). Very low density lipoproteins

(VLDL), the transporter of triglyceride and cholesterol from liver, contain newly synthesized

apolipoprotein B-100 (apoB-100), apo E, cholesterol esters, triglycerides and phospholipids. The

synthesis of VLDLs is regulated by apoB-100 and hence any inhibition in the degradation of apoB-100 would increase VLDL secretion, ultimately yielding LDL. LDLs, which carry cholesterol, are removed from the blood by endocytosis. Thus any mutations in the LDL receptor or apo-B gene would hinder the breakdown of LDL, leading to high levels of cholesterol. High density lipoproteins (HDL) also scavenge cholesterol and are endocytosed by specific receptors (246, 247). High cholesterol can cause atherosclerosis by, i) impairment of endothelium-dependent relaxation factor, ii) modification of the expression of nitric-oxide synthase activity (248), and iii) reduction in the velocity of blood flow by increasing the viscosity, which thereby would decrease the release of EDRF by the endothelial cells (249). Thus, any factor that would increase the cholesterol level in the plasma would definitely be considered as a risk factor for atherosclerosis.

Homocysteine has been demonstrated to influence cholesterol metabolism. In hyperhomocysteinemic patients, elevated plasma homocysteine levels were correlated with the increased levels of cholesterol and triglycerides. Supplementation of nutrients in the diet of patients with acute myocardial infarction produced a significant reduction in the plasma levels of homocysteine, cholesterol, triglycerides and apolipoprotein B-100 (250). The other evidence that linked homocysteine to occlusive arterial disease is that the administration of homocysteine to rats was accompanied with higher levels of plasma triglycerides (251). A very interesting study (252) directly linked homocysteine to cholesterol synthesis. Incubation of human hepatoma cells (HepG2) with 4 mM homocysteine enhanced the production of cholesterol, secretion of apolipoprotein B-100 and the activity of HMG-CoA reductase. Homocysteine not only abolished the feed-back inhibition of HMG-CoA reductase by LDL but also enhanced the enzyme activity.

The elevated enzyme activity is not caused by the direct interaction of homocysteine with the enzyme, but by a cascade mechanism through upregulation of the enzyme at the levels of transcription and translation. The fact that cholesterol stimulates apoB secretion (253) and elevation of apoB-100 is found in plasma of patients with both hypercholesterolemia and hyperhomocysteinemia led to a plausible relationship between hyperhomocysteinemia and atherosclerosis. Homocysteine thus has the ability to produce an increased amount of cholesterol, which in turn promotes the secretion of apoB, both of which are important for the development of atherosclerosis (247)

Even though homocysteine is an independent risk factor for atherosclerosis, it is intertwined with other risk factors like age, gender, smoking, blood pressure and lack of exercise (10, 20, 254, 255, 6 and 256). High homocysteine levels have also been associated with atherosclerosis seen in diabetes patients (257, 258). Higher levels of lipid peroxides and attenuated activities of antioxidant enzymes like GPx and superoxide dismutase have been reported in diabetes patients (259). This effect could be attributed to the auto-oxidation property of homocysteine. Hyperhomocysteinemia has also been found in end-stage renal disease (260). Despite the plethora of evidence provided by the epidemiological data and experimental results, there is still a fair amount of doubt that modest elevation of plasma homocysteine is benign and is a consequence rather than a cause of various diseases (261).

Research rationale: Having presented an overview of the deleterious roles of homocysteine in the pathology of a broad range of diseases (Figure 2 and Figure 3), the necessity of characterizing all the enzymatic pathways that regulate homocysteine metabolism should be apparent. The objective of the research project covered in the dissertation is to biochemically

characterize three human proteins in the homocysteine catabolic pathway, 1) cystathionine β-synthase (CBS) – the first enzyme involved in the transsulfuration pathway, 2) cystathionine γ-lyase (CGL) – the second enzyme involved in the trannssulfuration pathway and 3) betaine:homocysteine methyltransferase (BHMT) – the enzyme which has an equal share in the remethylation pathway of homocysteine along with methionine synthase. The only omission is methionine synthase. *The hypothesis of this research project is that the atomic structures of the enzymes associated with homocysteine metabolism can provide an important basis for understanding the pathology of homocysteine, can be used in the future to predict the effects of mutations on clinical outcomes, and finally, can provide a framework for therapeutic intervention.*

BETAINE: HOMOCYSTEINE METHYLTRANSFERASE

Why study BHMT? *In vitro* studies have indicated that homocysteine remethylation is shared equally between BHMT and methionine synthase (271) and large oral doses of betaine have been used therapeutically to reduce homocysteine levels (272). BHMT becomes a pharmacologically interesting target in the case of patients suffering from deficiencies of folate, vitamin B_6 or any enzymatic deficiencies in the tetrahydrofolate pathway. BHMT shares limited homology to the N-terminal region of *E.coli* methionine synthase (273), which is the only domain in methionine synthase that has not been structurally studied. BHMT, by virtue of its homology to the homocysteine and zinc-binding domain, may prove useful as a model.

BHMT - Specific aim 1: Design a simple expression vector for BHMT, using which, the protein can be expressed with few superfluous amino acids attached and, can be purified rapidly

(Chapter I). Rationale: Commercial vectors tend to introduce extraneous protein in addition to simple purification tags at the ends of the target protein. This could alter the native conformation of the target protein and thus hinder the likelihood of crystallization. Therefore the earlier methods of expressing BHMT as a fusion protein with beta-galactosidase was not considered appropriate for crystallization studies. The "seamless strategy" was therefore used to make an expression construct. This construct simplified cloning for not only BHMT, but it also proved to be a construct with broad applicability. The use of *Sap* I restriction enzyme, ensures that most genes can be inserted into this vector because of the rare occurrence of the 7 base restriction recognition sequence.

BHMT - Specific aim 2: Perform crystallization trials on BHMT (Chapter II & III). Rationale: Structural information obtained by X-ray crystallographic studies could be helpful to identify the critical residues involved in its mechanism and predict mutations in BHMT that could make individuals pre-disposed to heart disease. A structural study could also serve as a guide in the discovery of better methyl donors since they could be more potent than betaine in lowering plasma homocysteine. Activation of BHMT alone may be adequate to significantly reduce homocysteine levels since this will increase SAMe levels, and thus, in turn, activate CBS.

BHMT – Specific aim 3: Clarify the ambiguity about the oligomerization property of BHMT (Chapter I). Rationale: Human BHMT isolated from liver was reported to be a hexamer in solution (273). The preliminary X-ray studies on BHMT crystals showed that the recombinant human BHMT is a tetramer. Therefore, in order to confirm the X-ray results, it was necessary to perform studies like gel-filtration chromatography on the recombinant enzyme.

BHMT – Specific aim 4: Study the effect of SAMe on human recombinant BHMT

(Chapter I). Rationale: It seemed illogical that BHMT should have a SAMe binding site based on its sequence similarity arguments with the N-terminal of methionine synthase. The down-regulation action of BHMT by SAMe as shown in the rat liver enzyme is in contrast to its action on CBS (274). One consequence of the possible opposing actions of SAMe on CBS and BHMT is that any therapeutic approach that utilizes a SAMe-like molecule will in turn activate CBS while inhibiting BHMT. Because of the involvement of the remethylating enzymes in various metabolic pathways, it seemed important not to inhibit BHMT. Therefore it was necessary to understand the regulatory action of SAMe on the human recombinant BHMT. This aspect of BHMT was studied by SAMe radiochemical inhibition assay and qualitative isothermal titration calorimetry.

CYSTATHIONINE γ-LYASE

Why study CGL? CGL is the second enzyme involved in the transsulfuration of homocysteine to cysteine. In humans, the enzyme is associated with a broad spectrum of diseases like cystathionuria, cystinosis and cancer. Since the availability of cysteine is the rate-limiting step in the formation of glutathione, cystathionine γ-lyase plays an important role in the pathological and toxicological conditions associated with low glutathione levels. Abnormal CGL levels have been related to low glutathione levels found in patients with AIDS, alcoholic liver injury and atherosclerosis (125, 268, 269). Even though CGL is associated with various diseases, there is a lack of biochemical and structural understanding of the enzyme. The only structural data that is available is that of the homologous bacterial enzymes, cystathionine γ-synthase (CGS) and cystathionine β-lyase (CBL). CGS and CBL have been studied as antibacterial targets

and it has been extremely challenging to design antibacterials that can distinguish between the active sites of CBL, CGS and CGL. Thus a structural study of human CGL would greatly help in the identification of the subtle structural differences between these enzymes and development of more potent and specific inhibitors of CGS and CBL.

CGL - Specific aim 1: To develop an efficient expression system for human cystathionine γ-lyase (Chapter III). Rationale: Earlier study on recombinant human CGL had demonstrated that about 90% of the heterologous gene product was insoluble, and renaturation experiments from purified inclusion bodies met with limited success (270). Therefore the aim was to design a simple expression construct that could give an adequate amount of soluble and active CGL and also make purification a simple one-step process.

CGL – Specific aim 2: To perform various crystallization trials of CGL in order to find a suitable condition that would give large, diffraction-quality and reproducible crystals (Chapter III). Rationale: Obtaining diffraction-quality and reproducible crystals of a protein is the rate-limiting step in the structrual study of that protein. Therefore, screening for various crystallization conditions and then narrowing down to one optimum condition can really make the structrual study of proteins, a very rapid process.

CYSTATHIONINE β-SYNTHASE

Why study CBS? Computer modeling studies of the enzyme fluxes involved in homocysteine metabolism have indicated that the transsulfuration pathway is vital for clearing homocysteine (262). The methylation pathways involving methionine synthase and BHMT serve to salvage homocysteine and simply cycle it rather than remove the homocysteine from the

system. CBS, which catalyzes the first irreversible step in homocysteine transsulfuration, plays a very important role in preventing the recycling of homocysteine. Also, abnormal CBS expression levels have been associated with homocystinuria, premature peripheral and cerebral occlusive arterial disease and Down's syndrome (263, 264). Previous biochemical work performed on CBS has shown that the C-terminal of CBS is invloved in oligomerization and S-adenosylmethionine (SAMe) activation (265). An interesting study that gave a new dimension to the search of activators of CBS was that SAMe is not absolutely essential for activation of CBS, but any compound that can displace the autoinhibitory C-terminal domainof CBS from the active site can lead to the activation of the enzyme (266). Thus a structural study of CBS can resolve the complicated regulatory mechanism of CBS. It can also help to identify lead compounds through virtual drug screening and structure-based drug design and provide some alternatives to the expensive SAMe.

CBS - Specific aim: To obtain human, full-length, soluble and active CBS with no fusion protein attached (Chapter IV). Rationale: Although the human CBS has been previously cloned and expressed, the protein was either linked to β-galactosidase (265) and cleaved with protease or expressed with the C-terminus truncated (267). The kinetic and structural studies have been performed on the truncated CBS, but it is really doubtful whether the evaluated kinetic parameters and structural conformations can be extrapolated to the full-length enzyme. Also, expression of CBS along with the fusion protein can confer undesired conformational flexibility to the protein. Therefore the focus was to obtain soluble and homogeneous protein by addressing the aggregation problems of the full-length CBS that is a major hindrance to the crystallization and X-ray crystallographic studies of this enzyme.

Auto-oxidation in
the presence of
copper of ferric
ions

TNF
IL-2, IL-3, IL-1, IL-8
H_2O_2

Vitamin deficiency or
defective metabolic
pathways

Inhibits glutathione
peroxidase

NF-kB-
IkB

RN
A

RT

Provirus (ds DNA)

p50   p65

IkB

+

Inhibits
IkappaB
expresssion

High Homocysteine levels
in HIV-infection

+

Reactive
Oxygen
Intermediates

DNA
Integration

Stimulates PMA
mediated HIV-
LTR
transactivation

Activated
NF-kB

Binds
DNA

Latent
in host
genome

HIV
transcription

RNA

-

NAC

Stimulates GM-
CSF mediated HIV
expression

+

HIV
translation

HIV
protein

-

GSH

Upregulates IL-6
levels

+

Assembly

-

HIV replication

Figure 2: Summary of various factors contributing in the regulation of HIV infection
TNF – tumor necrosis factor; IL- interleukin; RT – reverse transcriptase; NAC – N-acetylcysteine; GSH -
glutathione

HOMOCYSTEINE

activate fV

activate fVII

activate TF

inhibits protein C

inhibit TM

inhibit HSPG

activate HMG-CoA

Increased amounts of cholesterol, triglycerides and apo-B 100

Auto-oxidation

Inhibition of stress response proteins like HSP70

Inhibition of antioxidant enzyme, GPx

Reactive oxgen species

PROTHROMBOTIC ACTION

OXIDATIVE STRESS

Impairs vasodilation by reducing the bioavailability of NO by forming peroxy nitrates

Activation of NF-κB

Activation of VCAM-1, ICAM-1, E-selectin, and MCP-1

Activation of NOS and cyclin D1 in vascular smooth cells

Lipid peroxidation

oxidation of LDL

macrophage derived foam cell

endothelial injury

ATHEROSCLEROSIS

Figure 3: Postulated mechanisms of homocysteine mediated atherosclerosis

fV – factor V; fXII – factor XII, PC – protein C, TF-tissue factor, TM – thrombomodulin, HSPG – heparan sulfate proteoglycan; HMG-CoA – hydroxy – methyl glutaryl coenzyme A; HSP70 – heat shock protein 70; GPx – glutathione peroxidase; NO – nitric oxide; NOS – nitric oxide synthase; MCP-1 – monocyte chemoattractant protein, LDL – low density lipoprotein

REFERENCES

1. Centers for Disease Control and Prevention (CDC). 2001. HIV and United States. MMWR 50:430-434.

2. Finkelstein JD. 1990. Methionine metabolism in mammals. *J Nutr Biochem* 1:228-37.

3. Storch KJ, Wagne DA, Burke JF, *et al*. 1990. [1-13C] methyl-2H3]Methionine kinetics in humans: methionine conservation and cystine sparing. *Am. J. Physiol*. 258:E790-98.

4. Ueland PM, Refsum H, Stabler SP, Malinow MR, Andersson A, Allen RH.1993. Total homocysteine in plasma or serum: methods and clinical applications. *Clin Chem* 39:1764-79.

5. Finkelstein JD, Martin JJ, Harris BJ. 1998. Methionine metabolism in mammals: the methionine-sparing effect of cystine. *J. Biol. Chem* 263:11750-54.

6. Malinow MR, Boston AG, Krauss RM.1999. Homocysteine, diet and cardiovascular diseases. A healthcare statement for healthcare professionals from the nutrition committee. American Heart Association. *Circulation* 99:178-82.
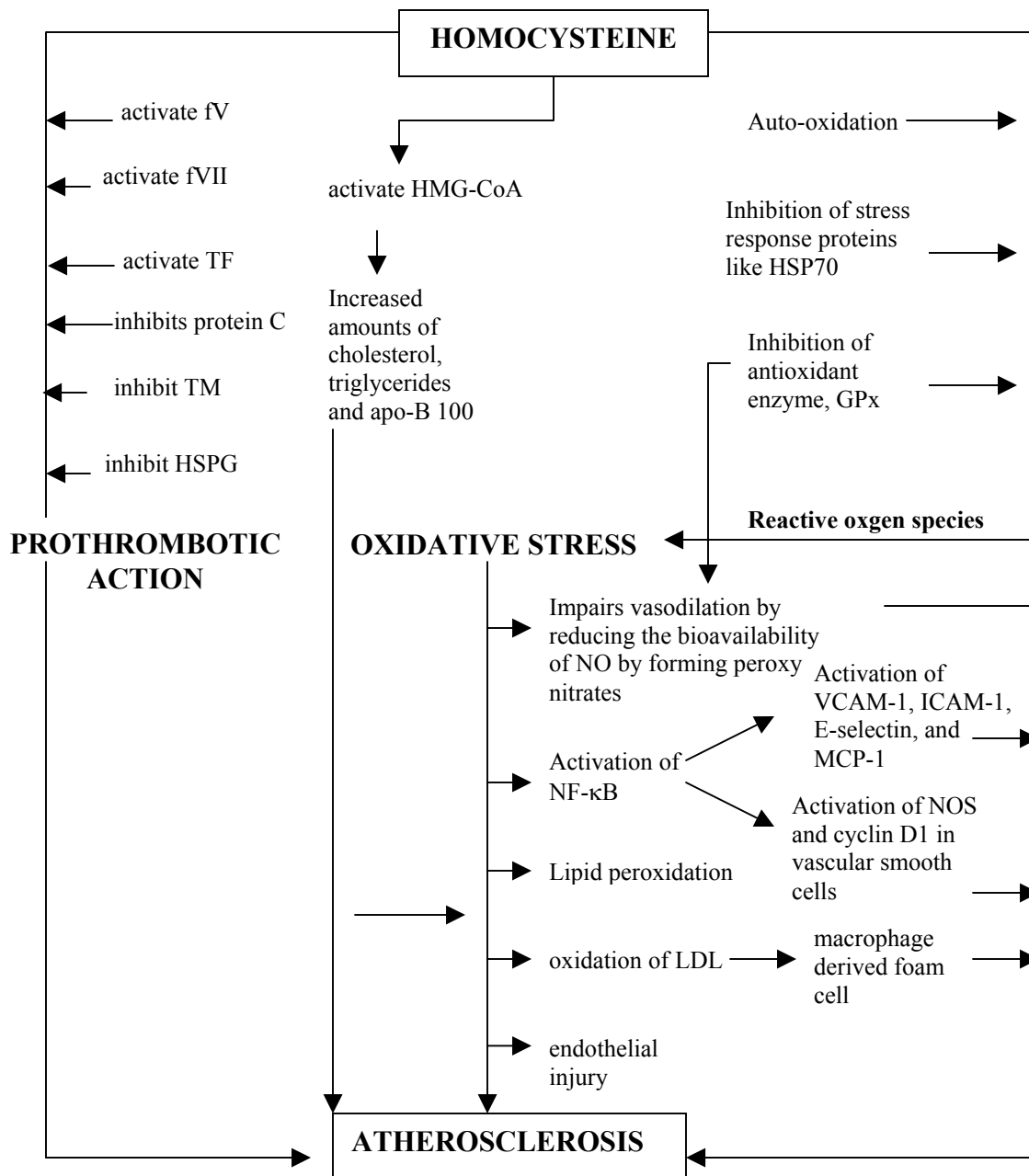
7. Olszewski AJ, McCully KS.1993. Homocysteine metabolism and the oxidative modification of proteins and lipids. *Free Rad Biol Med* 14:683-93.

8. Ubbink JB.1994. Vitamin nutrition status and homocysteine: an atherogenic risk factor. *Nutr Rev* 52:383-93.

9. Guttormsen A, Nesthus I, Uleland P, Nygard O, Schneede J, Vollset SE, *et al*.1996. Determinants and vitamin responsiveness of intermediate hyperhomocysteinemia (>40 mmol/L). *J Clin Invest* 98:2174-83.

10. Boushey CJ, Beresford SAA, Omenn GS, Motulsky AG.1995. A quantitative assessment of plasma homocysteine as a risk factor for vascular disease: probable benefits of increasing folic acid intakes. *JAMA* 274:1049-57.

11. Dalery K, Cacan S, Selhub J, Davignon J, Latour Y, Genest J.1995. Homocysteine and coronary artery disease in French Canadian subjects: relation with vitamins B6, B12, pyridoxal phosphate and folate. *Am J Cardiol* 75:1107-11.

12. Robinson K, Mayer EL, Miller DP, Green R, Van-Lente F, Gupta A, *et al*. 1995. Hyperhomocysteinemia and low pyridoxal phosphate. *Circulation* 92:2825-30.

13. Chauveau P, Chadefaux B, Coude M, Aupetit J, Hannedouche T, Kamoun P *et al*. 1993. Hyperhomocysteinemia: a risk factor for atherosclerosis in chronic uremic patients. *Kid Int* 43:72-7

14. Ubbink JB. 1995. Homocysteine – an atherogenic and a thrombogenic factor? *Nutr Rev* 53: 323-5.

15. Aronson DC, Onkenhout W, Raben AMTJ, Oudenhoven LFIJ, Brommer EJP, vanBrockel JH. 1994. Impaired homocysteine metabolism: a risk factor in young adults with atherosclerotic arterial occlusive disease of the leg. *Br J Surg* 81: 1114-8

16. Temple ME, Luzier AB, Kazierad DJ. 2000. Homocysteine as a risk factor for atherosclerosis. *Ann Pharmacother* 34:57-65.

17. Nordstrom M, Kellstrom T. 1992. Age dependency of cystathionine β-synthase activity in human fibroblasts in homocysteinemia and atherosclerotic vascular disease. *Atherosclerosis* 94:213-21.

18. Franken DG, Boers GHJ, Bloom HJ, Cruysberg JRM, Trijbels FJM, Hamel BCJ. 1996. Prevalence of familial mild hyperhomocysteinemia. *Atherosclerosis* 125:71-80.

19. Schwartz SM, Siscovick DS, Malinow R, Rosendaal FR, Beverly K, Hess DL, *et al*. 1997. Myocardial infarction in young women in relation to plasma total homocysteine, folate and a common variant in the methylenetetrahydrofolate reductase gene. *Circulation* 96:412-7.

20. Nygard O, Nordrehaug JE, Refsun H, Ueland PM, Farstad M, Vollset SE. 1997. Plasma homocysteine levels and mortality in patients with coronary artery disease. *N Engl J Med* 337:230-236.

21. Welch GN, Upchurch GR, Loscalzo J. 1997. Homocysteine, oxidative stress, and vascular disease. *Hosp Pract* 70:87-92.

22. Mayer EL, Jacobsen DW, Robinson K.1996. Homocysteine and coronary atherosclerosis. *J Am Coll Cardiol* 27:517-27.

23. Van Der Berg M, Boers GHJ, Franken DG, Blom HJ, Van Kamp GJ, Jakobs C *et al*. 1995. Hyperhomocysteinemia and endothelial dysfunction in young patients with peripheral arterial occlusive disease. *Eur J Clin Invest* 25:176-81.

24. D'Angelo A, Selhub J. 1997. Homocysteine and thrombotic disease. *Blood* 90:1-11.

25. Wilcken DE, Gupta VJ. 1979. Sulfur containing amino acids in chronic renal failure with particular reference to homocysteine and cysteine-homocysteine mixed disulfide. *Eur J Clin Invest* 9:301-7.

26. Ballal RS, Jacobsen DW, Robinson K. 1997. Homocysteine: update on a new risk factor. *Cleve Clin J Med* 64:543-9.

27. Robillon J, Canivet B, Candito M, Sadoul J, Jullien D, Morand P *et al*. 1994. Type I diabetes mellitus and homocysteine. *Diabetes Metab* 20:494-6.

28. McCully KS. 1969. Vascular pathology of homocysteinemia: implications for the development of arteriosclerosis. *Am J Pathol* 56:111-128.

29. Eck, HP, Gmunder H, Hartmann M, Petzoldt D, Daniel V, and Droge W. 1989. Low concentrations of acid soluble thiol (cysteine) in the blood plasma of HIV-1 infected patients. *Biol.Chem* 370:101-108.

30. Eck HP, and Droge W. 1989. Influence of the extracellular glutamate concentration on the intracellular cysteine concentration in macrophages and on the capacity to release cysteine. *Biol.Chem* 370:109-113.

31. Buhl R, Holroyd KJ, Mastrangeli A, Cantin AM, Jaffe HA, Wells FB, Saltini C, and Crystal RG. 1979. Systemic glutathione deficiency in asymptomatic HIV-seropositive individuals. *Lancet* 2:1294

32. Staal FJ, Ela SW, Roederer M, Anderson MT, Herzenberg LA, Herzenberg LA. 1992. Glutathione deficiency and human immunodeficiency virus infection. *Lancet* 339:909-12.

33. Droge W. 1993. Cysteine and Glutathione Deficiency in AIDS patients: A rationale for the treatment with N-Acetyl-Cysteine. *Pharmacology* 46:61-65.

34. van der Ven AJ, Blom HJ, Peters W, Jacobs LE, Verver TJ, Koopmans PP, Demacker P, van der Meer JW. 1998. Glutathione homeostasis is disturbed in CD4-positive lymphocytes of HIV-seropositive individuals. *Eur J Clin Invest* 28:187-93.

35. Baltimore D & Feinberg MB.1989. HIV revealed: toward a natural history of the infection. *N Engl J Med* 321:673-1675.

36. Schnittman SM, Psallidopoulos MC, Lane HC, Thompson L, Baseler M, Massari F, Fox CH, Salzman NP & Fauci AS 1990. *Science* 245:305-308.

37. Kotler DP, Wang J, and Pierson R. 1985. Body composition in patients with the acquired immunodeficiency syndrome. *Am J Clin Nutr* 42:1255-1265.

38. Grunfeld C. 1991. Mechanisms of wasting in infection and cancer. An approach to cachexia in AIDS. In "*Gastrointestinal and Nutritional Manifestations of AIDS*" (D.P. Kotler, ed.), pp. 207-229. Raven Press, New York.

39. Baeuerle PA, Henkel T. 1994. Function and activation of NF-kappa B in the immune system. *Ann Rev Immunol* 12:141-179.

40. Thanos D, Maniatis T. 1995. NF-kappa B: a lesson in family values. *Cell* 80:529-532.

41. Schindler U, Baichwal VR. 1994. Three NF-kappa B binding sites in the human E-selectin gene required for maxmal tumor necrosis factor alpha-induced expression. *Mol Cell Biol* 14:5820-5831.

42. Okamoto T, Matsuyama T, Mori S, Hamamoto Y, Kobayashi N, Yamamoto N, Josephs SF, Wong-Staal F, Shimotohno K. 1989. Augmentation of human immunodeficiency virus type 1 gene expression by tumor necrosis factor alpha. *AIDS Res Hum Retrovir* 5:131-138.

43. Maekawa T, Itoh F, Okamoto T, Kurimoto M, Imamoto F, Shii S. 1989. Identification and purification of the enhancer-binding factor of human immunodeficiency virus-1. Multiple preteins and binding to other enhancers. *J Biol Chem* 264:2826-2831.

44. Stade BG, Messer G, Riethmuller G., Johnson JP. 1990. Structural characteristics of the 5' region of the human ICAM-1 gene. *Immunobiol* 182:79-87.

45. Mukaida N, Mahe Y, Matsushima K. 1990. Cooperative interaction of nuclear factor-kappa B- and cis-regulatory enhancer binding protein-like factor binding elements in activating the interleukin-8 gene by pro-inflammatory cytokines. *J Biol Chem* 265:21128-21133

46. Wakasugi N, Tagaya Y, Wakasugi H, Mitsui A, Maeda M, Yodoi J, Tursz T. 1990. Adult T-cell leukemia-derived factor/thioredoxin, produced by both human T-lymphotropic virus type 1- and Epstein-Barr virus-transformed lymphocytes, acts as an autocrine growth factor and synergizes with interleukin 1 and interleukin 2. *Proc Natl Acad Sci USA* 87:8282-8286.

47. Roebuck KA, Rahman A, Lakshminarayanan V, Janakidevi K, Malik AB. 1995. H202 and tumor necrosis factor-alpha activate intercellular adhesion molecule 1 (ICAM-1) gene transcription through distinct cis-regulatory elements within the ICAM-1 promoter. *J Biol Chem* 270:18966-18974.

48. Donnelly RP, Crofford LJ, Freeman SL, Buras J, Remmers E, Wilder RL, Fenton MJ. 1993. Tissue-specific regulation of IL-6 production by IL-4. Differential effects of IL-4 on nuclear factor-kappa B activity in monocytes and fibroblasts. *J Immunol* 151:5603-5612.

49. Schreck R, Baeuerle PA. 1990. NF-kappa B as inducible transcriptional activator of the granulocyte-macrophage colony-stimulating factor gene. *Mol Cell Biol* 10:12811286.

50. Staynov DZ, Cousins DJ, Lee TH. 1995. A regulatory element in the promoter of the human granulocyte-macrophage colony-stimulating factor gene that has related sequences in other T-cell-expressed cytokine genes. *Proc Natl Acad Sci USA* 92: 3606-3610.

51. Nabel G, Baltimore D. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* 326:711-713.

52. Sen R, Baltimore D.1986.Inducibility of kappa immunoglobulin enhancer-binding protein NF-kappa B by a posttranslational mechanism. *Cell* 46:705-716.

53. Baeuerle PA, Baltimore D. 1988A. Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-kappa B transcription factor. *Cell* 53: 211-217

54. Baeuerle PA, Baltimore D. 1988B. I-kappa B: a specific inhibitor of the NF-kappa B transcription factor. *Science* 242:540-546.

55. Ghosh S, Gifford AM, Riviere L R, Tempst P, Nolan G P, Baltimore D. 1990. Cloning of the p50 DNA binding subunit of NF-kappa B: homology to Rel and dorsal. *Cell* 62:1019-1029.

56. Winyard PG, Blake DR. 1997. Antioxidants, redox-regulated transcription factors, and inflammation. *Adv Pharmacol* 38:403-21

57. Hayashi T, Sekine T, Okamoto T. 1993A. Identification of a new serine kinase that activates NF-kappa B by direct phosphorylation. *J Biol Chem* 826:26790-26795.

58. Ghosh S, Baltimore D. 1990. Activation in vitro of NF-kappa B by phosphorylation of its inhibitor I-kappa B. *Nature* 344: 678-682.

59. Shirakawa F, Mizel SB. 1989. In vitro activation and nuclear translocation of NF-kappa B catalyzed by cyclic AMP-dependent protein kinase and protein kinase C. *Mol Cell Biol* 9:2424-2430.

60. Meichle A, Schutze S, Hensel G, Brunsing D, Kronke M. 1990. Protein kinase C-independent activation of nuclear factor kB by tumor necrosis factor. *J Biol Chem* 265:8339-8343.

61. Feuillard J, Gouy H, Bismuth G, Lee LM, Debre P, Korner M. 1991. Nf-kappa B activation by tnmor necrosis factor alpha in the Jurkat T cell line is independent of protein kinase A, protein kinase C, and Ca (2+)-regulated kinase. *Cytokine* 3: 257-265.

62. Ostrowski J, Sims JE, Sibley CH, Valentine MA, Dower SK, Meier KE, Bomsztyk K. 1991. A serine/threonine kinase activity is clsely associated with a 65-kDa phosphoprotein specifically recognized by the kappa B enhancer element. *J Biol Chem* 266:12722-12733.

63. Schutze S, Potthoff K, Machleidt T, Berkovic D, Wiegmann K, Kronke M. 1992. TNF activates NF-kappa B by phosphatidylcholine-specific phospholipase C-induced "acidic" sphingomyelin breakdown. *Cell* 71:765-776.

64. Brown K, Gerstberger S, Carlson L, Franzoso G, Siebenlist U. *1995*. Control of I-kappa B-alpha proteolysis by site-specific, signal-induced phosphorylation. Science 267:1485-1488.

65. Cao Z, Henzel WJ, Gao X. 1996. IRAK: a kinase associated with the interleukin-1 receptor. *Science* 271:1128-1131.

66. Chen Z J, Parent L, Maniatis T. 1996. Site-specific phosphorylation of IkBa by a novel ubiquitination-dependent protein kinase activity. *Cell* 84:853-862.

67. Okamoto T, Ogiwara H, Hayashi T, Mitsui A, Kawabe T, Yodoi J. 1992. Human thioredoxin/adult T cell leukemia-derived factor activates the enhancer binding protein of human immunodeficiency virus type 1 bt thiol redox control mechanism. *Int Immunol* 4:811-819.

68. Hayashi T, Ueno Y, Okamoto T. 1993B. Oxidoreductive regulation of nuclear factor kappa B. Involvement of a cellular reducing catalyst thioredoxin. *J Biol Chem* 268:11380-11388.

69. Tagaya Y, Maeda Y, Mitsui A, Kondo N, Matsui H, Hamuro J, Brown J, Arai K I, Yokota T, Wakasugi H, Yodoi J. 1989. ATL-derived factor (ADF), an IL-2 receptor/Tac inducer homologous to thioredoxin; possible involvement of dithiol-reduction in the IL-2 receptor induction. *EMBO J* 8:57-764.

70. Schreck R, Rieber P, Baeuerle PA. 1991. Reactive oxygen intermediates as apparently widely used messengers in the activation of the NF-kappa B transcription factor and HIV-1. *EMBO J* 10:2247-2258.

71. Molitor JA, Ballard DW, Greene WC. 1991. Kappa-B-specific DNA binding proteins are differentially inhibited by enhancer mutations and biological oxidation. *New Biol* 3:987-996.

72. Toledano MB, Leonard WJ. 1991. Modulation of transcription factor NF-kappa B binding activity by oxidation-reduction in vitro. *Proc Natl Acad Sci USA* 88:4328-4332.

73. Matthews JR, Wakasugi N, Virelizier JL, Yodoi J, Hay RT. 1992. Thioredoxin regulates the DNA binding activity of NF-kappa B by reduction of a disulfide bond involving cysteine 62. *Nucleic Acids Res* 20:3821-3830.

74. Okamoto T, Benter T, Josephs SF, Sadaie MR, Wong-Staal F. 1990. Transcriptional activation from the long-terminal repeat of human immunodeficiency virus in vitro *Virology* 177:606-614.

75. Arya SK, Guo C, Josephs SF, Wong-Staal F. 1985. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). *Science* 229:69-73.

76. Sodroski J, Patarca R, Rosen C. 1985. Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. *Science* 229:74-77.

77. Okamoto T, Wong-Staal F. 1986. Demonstration of virus-specific transcriptional activator(s) in cells infected with HTLV-III by an in vitro cell-free system. *Cell* 147: 9-35.

78. Peterlin BM, Luciw PA, Barr PJ, Walker MD. 1986. Elevated levels of mRNA can account for the trans-activation of human immunodeficiency virus. *Proc Natl Acad Sci USA* 83:9734-9738.

79. Schreck R, Rieber P, Baeuerle PA. 1991. Reactive oxygen intermediates as apparently widely used messengers in the activation of the NF-kappa B transcription factor and HIV-1. *EMBO J* 10:2247-2258.

80. Roederer M, Staal FJT, Raju PA, Ela SW, Herzenberg LA, Herzenberg LA. 1990. Cytokine-stimulated human immunodeficiency virus replication is inhibited by N-acetyl-L-cysteine. *Proc Natl Acad Sci USA* 87:4884-4888.

81. Suzuki YJ, Aggarwal BB, Packer L. 1992. Alpha-lipoic acid is a potent inhibitor if NF-kappa B activation in human T cells. *Biochem Biophys Res Commun* 189:1709-1715.

82. Meyer M, Schreck R, Baeuerle PA. 1993. H_2O_2 and antioxidants have opposite effects on activation of NF-kappa B and AP-1 in intact cells: AP-1 as secondary antioxidant - responsive factor. *EMBO J* 12:2005-2015.

83. Biswas DK, Dezube BJ, Ahlers CM, Pardee AB. 1993. Pentoxifylline inhibits HIV-1 LTR-driven gene expression by blocking NF-kappa B action. *J AIDS* 6:778-786.

84. Suzuki YJ, Packer L. 1994. Signal transduction for nuclear factor-kappa B activation. Proposed location of antioxidant-inhibitable step. *J Immunol* 153:5008-5015.

85. Packer L, Witt EH, Tritschler HJ. 1995. Alpha-lipoic acid as a biological antioxidant. *Free Rad Biol Med* 19:227-250.

86. van der Ven AJ, Boers GH. 1997. Oxidative stress in immunodeficiency. *Eur J Clin Invest* 9:731-2.

87. Droge W, Eck HP, Gmunder H, Mihm S. 1991. Modulation of lymphocyte functions and immune responses by cysteine and cysteine derivatives. *Am J Med* 91:140S-144S

88. Elbim C, Pillet S, Prevost MH, Preira A, Girard PM, Rogine N, Hakim J, Israel N, Gougerot-Pocidalo MA. 2001. The role of phagocytes in HIV-related oxidative stress. *J Clin Virol* 20:99-109.

89. Smith PD, Gartner S, Popovic M, Lamerson CL, Wahl SW. 1988. Human immunodeficiency virus (HIV-1) modulates superoxide anion production by human leukocytes. IV International Conference on AIDS, Stockholm (abstract 2079).

90. Israel N, Gougerot-Pocidalo MA. 1997. Oxidative stress in human immunodeficiency virus infection. *Cell Mol Life Sci* 53:864-70.

91. Gennaro R, Florio C, and Romeo D. 1986. *Biochem Biophys Res Commun* 134: 305-312.

92. Staal FJ, Roederer M, Herzenberg LA, Herzenberg LA. 1990. Intracellular thiols regulate activation of nuclear factor kappa B and transcription of human immunodeficiency virus. *Proc Natl Acad Sci U S A* 87:9943-7

93. Poli G, Fauci AS. 1992. The effect of cytokines and pharmacologic agents on chronic HIV infection. *AIDS Res Hum Retroviruses* 1992 8:191-7.

94. Poli G, and Fauci AS. Cytokines and immunoregulatory factors in the pathogenesis of HIV infection: TNF effects. In: Workshop on Immunodeficiency in HIV Infections. G. Janossy, B.Autran, F. Miedema (eds.). Karger, Basel, CH, in press.

95. Reddy MM, Sorrell SJ, Lange M, Grieco MH. 1988. Tumor necrosis factor and HIV P24 antigen levels in serum of HIV-infected populations. *J Acquir Immune Defic Syndr* 1:436-40.

96. Mintz M, Rapaport R, Oleske JM, Connor EM, Koenigsberger MR, Denny T, Epstein LG. 1989.Elevated serum levels of tumor necrosis factor are associated with progressive encephalopathy in children with acquired immunodeficiency syndrome. *Am J Dis Child* 143:771-4.

97. Rosen GM, Freeman BA. 1984. Detection of superoxide generated by endothelial cells. *Proc Natl Acad Sci U S A* 81:7269-73.

98. Matsubara T, Ziff M. 1986. Increased superoxide anion release from human endothelial cells in response to cytokines. *J Immunol* 137:3295-8

99. Klebanoff SJ, Vadas MA, Harlan JM, Sparks LH, Gamble JR, Agosti JM, Waltersdorph AM. 1986. Stimulation of neutrophils by tumor necrosis factor. *J Immunol* 136:4220-5.

100. Breen EC, Rezai AR, Nakajima K, Beall GN, Mitsuyasu RT, Hirano T, Kishimoto T, Martinez-Maza O. 1990. Infection with HIV is associated with elevated IL-6 levels and production. *J Immunol* 144:480-4.

101. Gallo P, Frei K, Rordorf C, Lazdins J, Tavolato B, Fontana A. 1989. Human immunodeficiency virus type 1 (HIV-1) infection of the central nervous system: an evaluation of cytokines in cerebrospinal fluid. *J Neuroimmunol* 23:109-16.

102. Nakajima K, Martinez-Maza O, Hirano T, Breen EC, Nishanian PG, Salazar-Gonzalez JF, Fahey JL, Kishimoto T. 1989. Induction of IL-6 (B cell stimulatory factor-2/IFN-beta 2) production by HIV. *J Immunol* 142:531-6.

103. Schuitemaker H, Kootstra NA, Groenink M, Meyaard L, de Goede REY, Husiman H, Miedema F, and Tersmett M. 1991. Viral and cellular requirements for replication of HIV in primary monocytes. VII International Conference on AIDS, Florence 16-21, vol.1 p.100 (abstr).

104. Zack JA, Arrigo SJ, Weitsman SR, Go AS, Haislip A, Chen IS. 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell* 61:213-22.

105. Koyanagi Y, O'Brien WA, Zhao JQ, Golde DW, Gasson JC, Chen IS. 1988. Cytokines alter production of HIV-1 from primary mononuclear phagocytes. *Science* 241:1673-5.

106. Folks TM, Justement J, Kinter A, Dinarello CA, Fauci AS. 1987. Cytokine-induced expression of HIV-1 in a chronically infected promonocyte cell line. *Science* 38:800-2.

107. Gendelman HE, Orenstein JM, Martin MA, Ferrua C, Mitra R, Phipps T, Wahl LA, Lane HC, Fauci AS, Burke DS, *et al*. 1988. Efficient isolation and propagation of human immunodeficiency virus on recombinant colony-stimulating factor 1-treated monocytes. *J Exp Med* 167:1428-41.

108. Ueland PM, Mansoor MA, Guttormsen AB, Muller F, Aukrust P, Refsum H, Svardal AM. 1996. Reduced, oxidized and protein-bound forms of homocysteine and other aminothiols in plasma comprise the redox thiol status--a possible element of the extracellular antioxidant defense system. *J Nutr* 126:1281S-4S.

109. Muller F, Svardal AM, Aukrust P, Berge RK, Ueland PM, Froland SS. 1996. Elevated plasma concentration of reduced homocysteine in patients with human immunodeficiency virus infection. *Am J Clin Nutr* 63:242-8.

110. Remacha AF, Cadafalch J. 1999. Cobalamin deficiency in patients infected with the human immunodeficiency virus. *Semin Hematol* 36:75-87.

111. Paltiel O, Falutz J, Veilleux M, Rosenblatt DS, Gordon K. 1995. Clinical correlates of subnormal vitamin B12 levels in patients infected with the human immunodeficiency virus. *Am J Hematol* 4:318-22.

112. Munday R. 1989. Toxicity of thiols and disulphides: involvement of free-radical species. *Free Radic Biol Med* 7:659-73.

113. Wall RT, Harlan JM, Harker LA, Striker GE. 1980. Homocysteine-induced endothelial cell injury in vitro: a model for the study of vascular injury. *Thromb Res* 18:113-21.

114. Starkebaum G, Harlan JM. 1986. Endothelial cell injury due to copper-catalyzed hydrogen peroxide generation from homocysteine. *J Clin Invest* 77:1370-6.

115. Graham NM, Sorensen D, Odaka N, Brookmeyer R, Chan D, Willett WC, Morris JS, Saah AJ. 1991. Relationship of serum copper and zinc levels to HIV-1 seropositivity and progression to AIDS. *J Acquir Immune Defic Syndr* 4:976-80.

116. Hultberg B, Andersson A, Isaksson A. 1997. The effects of homocysteine and copper ions on the concentration and redox status of thiols in cell line cultures. *Clin Chim Acta* 262:39-51.

117. Upchurch GR Jr, Welch GN, Freedman, J.E., Loscalzo J. 1996. Homocysteine attenuates endothelial glutathione peroxidase and thereby potentiates peroxide-mediated cell injury. *Circulation* 92:I-228.

118. Loscalzo J. 1996. The oxidant stress of hyperhomocyst(e)inemia. *J Clin Invest* 98:5-7.

119. Wang G, Siow YL, OK. 2000. Homocysteine stimulates nuclear factor kappa B activity and monocyte chemoattractant protein-1 expression in vascular smooth-muscle cells: a possible role for protein kinase C. *Biochem J* 3:817-26.

120. Simon G, Moog C, Obert G. 1994. Effects of glutathione precursors on human immunodeficiency virus replication. *Chem Biol Interact* 217-24.

121. van Aken BE, Jansen J, van Deventer SJ, Reitsma PH. 2000. Elevated levels of homocysteine increase IL-6 production in monocytic Mono Mac 6 cells. *Blood Coagul Fibrinolysis* 11:159-64

122. Naisbitt DJ, Vilar FJ, Stalford AC, Wilkins EG, Pirmohamed M, Park BK. 2000. Plasma cysteine deficiency and decreased reduction of nitrososulfamethoxazole with HIV infection. *AIDS Res Hum Retroviruses* 16:1929-38.

123. Choi J, Lui RM, Kundu RK, Sangiorgi F, Wu W, Maxson R, and Forman HJ. 1999. Molecular mechanism of decreased glutathione content in human immunodeficiency virus type-1 tat-transgenic mice. *J Biol Chem* 275:3693-3698.

124. Helbling B, von Overbeck J, Lauterburg BH. 1996. Decreased release of glutathione into the systemic circulation of patients with HIV infection. *Eur J Clin Invest* 26:38-44.

125. Martin JA, Juan S, *et.al*. 2001. Hepatic gamma-cystathionase deficiency in patients with AIDS. JAMA, Research Letters 285:1444-5.

126. Droge W, Pottmeyer-Gerber C, Schmidt H, Nick S. Glutathione augments the activation of cytotoxic T lymphocytes in vivo. *Immunobiology* 1986 Aug;172(1-2):151-6.

127. Meister A, Anderson ME. Glutathione. *Annu Rev Biochem* 1983;52:711-60.

128. Kebanoff, SJ, Vadas MA, Harlan JM, Sparks LH, Gamble JR, Agosti JM and Waltersdorph AM. 1986. Stimulation of neutrophils by tumor necrosis factor. *J Immunol* 136:4220.

129. Yoshie O, Majima T, and Saito H. 1989. Membrane oxidative metabolism of human eosinophilic cell line EoL-1 in response to phorbol diester and formyl peptide: synergistic augmentation by interferon-gamma and tumor necrosis factor. *J leukocyte Biol* 45:10.

130. De Chatelet LR, Shirley PS, and Johnson RB. 1976. Effect of phorbol myristate acetate on the oxidative metabolism of human polymorphonuclear leukocytes. *Blood* 47:545.

131. Staal FJT, Roederer M, and Herzenberg LA. 1990. Intracellular thiols regulate activation of nuclear factor κB and transcription of human immunodeficiency virus. *Proc Natl Acad Sci* 87:9943.

132. Roederer M, Staal FJT, Osada H, Herzenberg LA and Herzenberg LA. 1991. CD4 and CD8 T cells with high intracellular glutathione levels are selectively lost as the HIV infection progresses. *Int Immunol* 3:933-7.

133. Roederer M, Staal FJT, Raju PA, Herzenberg LA and Herzenberg LA. 1991. N-acetylcysteine inhibits latent HIV expression in chronically infected cells. *AIDS Research and Human Retroviruses* 7:563-567

134. Jahoor F, Jackson A, Gazzard B, Philips G, Sharpstone D, Frazer ME, Heird W. 1999. Erythrocyte glutathione deficiency in symptom-free HIV infection is associated with decreased synthesis rate. *Am J Physiol* 276:E205-11.

135. Staal FJ. 1998.Glutathione and HIV infection: reduced reduced, or increased oxidized? *Eur J Clin Invest* 28:194-6.

136. Droge W, Eck HP, Mihm S. 1992. HIV-induced cysteine deficiency and T-cell dysfunction--a rationale for treatment with N-acetylcysteine. *Immunol Today* 13:211-4.

137. Aukrust P, Svardal AM, Muller F, Lunden B, Berge RK, Ueland PM, Froland SS. 1995. Increased levels of oxidized glutathione in CD4+ lymphocytes associated with disturbed intracellular redox balance in human immunodeficiency virus type 1 infection. *Blood* 86:258-67.

138. Mosharov E, Cranford MR, Banerjee R. 2000. The quantitatively important relationship between homocysteine metabolism and glutathione synthesis by the transsulfuration pathway and its regulation by redox changes. *Biochemistry* 39:13005-11.

139. Droge W, Eck HP, Naher H, Pekar U, and Daniel V 1988. Abnormal amino acid concentrations in the blood of patients with acquired immunodeficiency syndrome (AIDS) may contribute to the immunological defect. *Biol.Chem. Hoppe-Seyler* 369:143-148.

140. Muller J, Janz S. 1993. Modulation of the H2O2-induced SOS response in Escherichia coli PQ300 by amino acids, metal chelators, antioxidants, and scavengers of reactive oxygen species. *Environ Mol Mutagen* 22:157-63.

141. Aruoma OK, Halliwell B, Hoey BM, and Butler J. 1989. The antioxidant action of N-acetylcysteine: Its reaction with hydrogen peroxide, hydroxyl radical, superoxide and hypochlorous acid. *Free Rad Biol* 6:593-597.

142. Gutteridge JM, Halliwell B. 1989. Iron toxicity and oxygen radicals. *Baillieres Clin Haematol* 2:195-256.

143. De Rosa SC, Zaretsky MD, Dubs JG, Roederer M, Anderson M, Green A, Mitra D, Watanabe N, Nakamura H, Tjioe I, Deresinski SC, Moore WA, Ela SW, Parks D,

Herzenberg LA, Herzenberg LA. 2000. N-acetylcysteine replenishes glutathione in HIV infection. *Eur J Clin Invest* 30:915-29.

144. Kalebic T, Kinter A, Poli G, Anderson ME, Meister A, Fauci AS. Suppression of human immunodeficiency virus expression in chronically infected monocytic cells by glutathione, glutathione ester, and N-acetylcysteine. Proc Natl Acad Sci U S A 1991 Feb 1;88(3):986-90

145. Miller JW. 2000. Homocysteine, Alzheimer's disease, and cognitive function. *Nutrition* 16:675-7.

146. Kandel ER, Schwartz JH, Jessell TM. 1991. Principles of Neuroscience, 3rd ed. Norwalk, CO: Appleton + Lange.

147. Miller JW. 1999. Homocysteine and Alzheimer's disease. *Nutr Rev* 57:126-9.

148. Hofman A, Ott A, Breteler MMB, *et al.* 1997. Atherosclerosis, apolipoprotein E, and prevalence of dementia and Alzheimer's disease in the Rotterdam Study. *Lancet* 349:151.

149. Roses AD. 1996. Apolipoprotein E alleles as risk factors in Alzheimer's disease. *Annu Rev Med* 47:387.

150. Snowden DA, Greiner LH, Mortimer JA, *et al.* 1997. Brain infarction and the clinical expression of Alzheimer disease: the Nun Study. *JAMA* 277:813.

151. Refsum H, Ueland PM, Nygard O, Vollset SE .1998. Homocysteine and cardiovascular disease. *Annual Rev Med* 49:31-62.

152. Chiu E. What in a name – dementia or dysmentia? (Editorial) Int J Geriatr Psychiatry 9;1-4.

153. Gottfries CG, Lehmann W, Regland B. 1998. Early diagnosis of cognitive impairment in the elderly with the focus on Alzheimer's disease. *J Neural Transm* 105:773-86.

154. Goodwin JS, Goodwin JM, Garry PJ. 1983. Association between nutritional status and cognitive functioning in a healthy elderly population. *JAMA* 249:2917-21.

155. Lindenbaum J, Healton EB, Savage DG, Brust JC, Garrett TJ, Podell ER, Marcell PD, Stabler SP, Allen RH. 1998. Neuropsychiatric disorders caused by cobalamin deficiency in the absence of anemia or macrocytosis. *N Engl J Med* 318:1720-8.

156. Martin DC, Francis J, Protech J, Huff FJ. 1992. Time dependency of cognitive recovery with cobalamin replacement: report of a piolt study. *J Am Geriatr Soc* 40:168-72.

157. Bell IR, Edman JS, Selhub J, Morrow FD, Marby DW, Kayne HL, Cole JO. 1992. Plasma homocysteine in vascular disease and in nonvascular dementia of depressed elderly people. *Acta Psychiatr Scand* 86:386-90.

158. Riggs KM, Spiro A 3rd, Tucker K, Rush D. Relations of vitamin B-12, vitamin B-6, folate, and homocysteine to cognitive performance in the Normative Aging Study. *Am J Clin Nutr* 63:306-14.

159. Lehmann M, Gottfries CG, Regland B. 1999. Identification of cognitive impairment in the elderly: homocysteine is an early marker. Dement Geriatr Cogn Disord 10:12-20.

160. Morris MS, Jacques PF, Rosenberg IH, Selhub J. 2001. Hyperhomocysteinemia associated with poor recall in the third National Health and Nutrition Examination Survey. *Am J Clin Nutr* 73:927-33.

161. Ikeda T, Furukawa Y, Mashimoto S, Takahashi K, Yamada M. 1990. Vitamin B12 levels in serum and cerebrospinal fluid of people with Alzheimer's disease. *Acta Psychiatr Scand* 82:327-9.

162. Levitt AJ, Karlinsky H. 1992. Folate, vitamin B12 and cognitive impairment in patients with Alzheimer's disease. *Acta Psychiatr Scand* 86:301-5.

163. Joosten E, Lesaffre E, Riezler R, Ghekiere V, Dereymaeker L, Pelemans W, Dejaeger E. 1997. Is metabolic evidence for vitamin B-12 and folate deficiency more frequent in elderly patients with Alzheimer's disease? *J Gerontol A Biol Sci Med Sci* 52:M76-9.

164. McCaddon A, Davies G, Hudson P, Tandy S, Cattell H. 1998. Total serum homocysteine in senile dementia of Alzheimer type. *Int J Geriatr Psychiatry* 13:235-9.

165. Breteler MM. Vascular involvement in cognitive decline and dementia. Epidemiologic evidence from the Rotterdam Study and the Rotterdam Scan Study. Ann N Y Acad Sci 2000 Apr;903:457-65.

166. Clarke R, Smith AD, Jobst KA, Refsum H, Sutton L, Ueland PM. 1998. Folate, vitamin B12, and serum total homocysteine levels in confirmed Alzheimer disease. *Arch Neurol* 55:1449-55.

167. Breteler MM, Bots ML, Ott A, Hofman A. 1998. Risk factors for vascular disease and dementia. *Haemostasis* 28:167-73.

168. Diaz-Arrastia R. 1998. Hyperhomocysteinemia: a new risk factor for Alzheimer disease? *Arch Neurol* 55:1407-8.

169. Selhub J, Bagley LC, Miller J, Rosenberg IH. 2000. B vitamins, homocysteine, and neurocognitive function in the elderly. *Am J Clin Nutr* 71:614S-620S.

170. De La Cruz JP, Pavia J, Gonzalez-Correa JA, Ortiz P, Sanchez de la Cuesta F. 2000. Effects of chronic administration of S-adenosyl-L-methionine on brain oxidative stress in rats. Naunyn Schmiedebergs *Arch Pharmacol* 361:47-52.

171. Morrison LD, Smith DD, Kish SJ. 1996. Brain S-adenosylmethionine levels are severely decreased in Alzheimer's disease. *J Neurochem* 67:1328-31.

172. Bottigileri T, Godfrey P, Flynn T, Carney MWP, Toone B.K and Reynolds EH 1990. Cerebrospinal fluid S-adensoylmethionine in depression and dementia: effects of treatment with parenteral and oral S-adensoylmethionine. *J. Neurol. Neurosurg. Psychiatry* 53:1096-1098.

173. Trolin CG, Lofberg C, Trolin G, Oreland L. 1994. Brain ATP:L-methionine S-adenosyltransferase (MAT), S-adenosylmethionine (SAM) and S-adenosylhomocysteine (SAH): regional distribution and age-related changes. *Eur Neuropsychopharmacol* 4:469-77.

174. Bottigileri T, Hyland K, and Reynolds EH. 1994. The clinical potential of adomethionine (S-adenosylmethionine) in neurological disorders. *Drugs* 48:137-152.

175. Fekkes D, van der Cammen TJ, van Loon CP, Verschoor C, van Harskamp F, de Koning I, Schudel WJ, Pepplinkhuizen L. 1998. Abnormal amino acid metabolism in patients with early stage Alzheimer dementia. *J Neural Transm* 105:287-94.

176. Tabet N, Mantle D, Orrell M. 2000. Free radicals as mediators of toxicity in Alzheimer's disease: a review and hypothesis. *Adverse Drug React Toxicol Rev* 19:127-52.

177. Christen Y. 2000. Oxidative stress and Alzheimer disease. *Am J Clin Nutr* 71:621S-629S.

178. Zhou Y, Richardson JS, Mombourquette MJ, Weil JA. 1995. Free radical formation in autopsy samples of Alzheimer and control cortex. *Neurosci Lett* 195:89-92.

179. Lyras L, Cairns NJ, Jenner A, Jenner P, Halliwell B. 1997.An assessment of oxidative damage to proteins, lipids, and DNA in brain from patients with Alzheimer's disease. *J Neurochem* 68:2061-9.

180. Smith CD, Carney JM, Starke-Reed PE, Oliver CN, Stadtman ER, Floyd RA, Markesbery WR. 1991.Excess brain protein oxidation and enzyme dysfunction in normal aging and in Alzheimer disease. Proc Natl Acad Sci U S A 88:10540-3.

181. Markesbery WR, Lovell MA. 1998. Four-hydroxynonenal, a product of lipid peroxidation, is increased in the brain in Alzheimer's disease. *Neurobiol Aging* 19:33-6.

182. Marcus DL, Thomas C, Rodriguez C, Simberkoff K, Tsai JS, Strafaci JA, Freedman ML. 1998. Increased peroxidation and reduced antioxidant enzyme activity in Alzheimer's disease. *Exp Neurol* 150:40-4.

183. Connor JR, Tucker P, Johnson M, Snyder B. 1993. Ceruloplasmin levels in the human superior temporal gyrus in aging and Alzheimer's disease. *Neurosci Lett* 159:88-90.

184. Prohaska JR. 1997. Neurochemical roles of copper as antioxidant or prooxidant. In:Connor JR, ed. *Metals and Oxidative Damage in Neurological Disorders*. New York: Plenum Press 57-75.

185. Mantyh PW, Ghilardi JR, Rogers S, DeMaster E, Allen CJ, Stimson ER, Maggio JE. 1993. Aluminum, iron, and zinc ions promote aggregation of physiological concentrations of beta-amyloid peptide. *J Neurochem* 3:1171-4.

186. Dyrks T, Dyrks E, Hartmann T, Masters C, Beyreuther K. 1992. Amyloidogenicity of beta A4 and beta A4-bearing amyloid protein precursor fragments by metal-catalyzed oxidation. *J Biol Chem* 67:18210-7.

187. Starkebaum G, Harlan JM. Endothelial cell injury due to copper-catalyzed hydrogen peroxide generation from homocysteine. *J Clin Invest* 1986 Apr;77(4):1370-6

188. Zhang F, Slungaard A, Vercellotti GM, Iadecola C. 1998. Superoxide-dependent cerebrovascular effects of homocysteine. *Am J Physiol* 274:R1704-11.

189. White AR, Huang X, Jobling MF, Barrow CJ, Beyreuther K, Masters CL, Bush AI, Cappai R. 2001. Homocysteine potentiates copper- and amyloid beta peptide-mediated toxicity in primary neuronal cultures: possible risk factors in the Alzheimer's-type neurodegenerative pathways. *J Neurochem* 76:1509-20.

190. Anst SD, Morehouse LA, Thomas CE. 1985. Role of metals in oxygen radical reactions. *Journal of Free Radicals in Biology and Medicine* 1:3-25.

191. Sokol RJ, Devereaux M, Mierau GW, Hambidge KM, Shikes RH. 1990. Oxidant injury to hepatic mitochondrial lipids in rats with dietary copper overload. Modification by vitamin E deficiency. *Gastroenterology* 99:1061-71.

192. Huang X, Atwood CS, Hartshorn MA, Multhaup G, Goldstein LE, Scarpa RC, Cuajungco MP, Gray DN, Lim J, Moir RD, Tanzi RE, Bush AI. 1999. The A beta peptide of Alzheimer's disease directly produces hydrogen peroxide through metal ion reduction. *Biochemistry* 38:7609-16.

193. Varadarajan S, Yatin S, Aksenova M, Butterfield DA. 2000. Review: Alzheimer's amyloid beta-peptide-associated free radical oxidative stress and neurotoxicity. J *Struct Biol* 130:184-208.

194. Mutisaya EM. Bowling AC, Beal MF. 1994. Cortical cytochrome oxidase activity is reducedin Alzheimer's disease. *J Neurochem* 63:2179-84.

195. Chandrashekharan K, Giordano T, Brady DR, Stoll J, Martin JL, Rapport SI. 1994. Impairment in mitochondrial cytochrome oxidase gene expressionin Alzheimer's disease. *Mol Brain Res* 24:336-40.

196. Austin RC, Sood SK, Dorward AM, Singh G, Shaughnessy SG, Pamidi S, Outinen PA, Weitz JI. 1998. Homocysteine-dependent alterations in mitochondrial gene expression, function and structure. Homocysteine and H2O2 act synergistically to enhance mitochondrial damage. *J Biol Chem* 273:30808-17.

197. Shoffner JM. 1997. Oxidative phosphorylation defects and Alzheimer's diseae. *Neurogenetics* 1:13-9.

198. Jakubowski H. 1997. Metabolism of homocysteine thiolactone in human cell cultures. Possible mechanism for pathological consequences of elevated homocysteine levels. *J Biol Chem* 272:1935-42.

199. Parsons RB, Waring RH, Ramsden DB, Williams AC. 1998. In vitro effect of the cysteine metabolites homocysteic acid, homocysteine and cysteic acid upon human neuronal cell lines. *Neurotoxicology* 19:599-603.

200. Lipton SA, Kim WK, Choi YB, Kumar S, D'Emilia DM, Rayudu PV, Arnelle DR, Stamler JS. 1997. Neurotoxicity associated with dual actions of homocysteine at the N-methyl-D-aspartate receptor. *Proc Natl Acad Sci U S A* 94:5923-8.

201. Love S, Barber R, Wilcock GK. 1998. Apoptosis and expression of DNA repair proteins in ischaemic brain injury in man. *Neuroreport* 9:955-9.

202. Pieper AA, Verma A, Zhang J, Snyder SH. 1999. Poly (ADP-ribose) polymerase, nitric oxide and cell death. *Trends Pharmacol Sci* 20:171-81.

203. Hughes PE, Alexi T, Schreiber SS. 1997. A role for the tumour suppressor gene p53 in regulating neuronal apoptosis. *Neuroreport* 8:v-xii.

204. Grieve A, Butcher SP, Griffiths R. 1992. Synaptosomal plasma membrane transport of excitatory sulphur amino acid transmitter candidates: kinetic characterization and analysis of carrier specificity. *J Neurosci Res* 32:60-68.

205. Kruman II, Culmsee C, Chan SL, Kruman Y, Guo Z, Penix L, Mattson MP. 2000. Homocysteine elicits a DNA damage response in neurons that promotes apoptosis and hypersensitivity to excitotoxicity. *J Neurosci* 20:6920-6.

206. Hopkins PN, Williams RR. 1981. A survey of 246 suggested coronary risk factors. *Atherosclerosis* 40: 1-52.

207. Graham IM, Daly LE, Refsum HM, Robinson K, Brattstrom LE *et al*. 1997. Plasma homocysteine as a risk factor for vascular disease. The European Concerned Action Project. *JAMA* 277: 1775-1781.

208. Andersson A, Lindgren A, Hultberg B. 1995. Effect of thiol oxidation and thiol export from erythrocytes on determination of redox status of homocysteine and other thiols in plasma from healthy subjects and patients with cerebral infarction. *Clin Chem* 41:361-6.

209. Misra HP. 1974. Generation of superoxide free radical during the autooxidation of thiols. *J. Biol Chem* 249:2151-5.

210. Wall RT, Harlan JM, Harker LA, Striker GE. 1980. Homocysteine-induced endothelial cell injury in vitro: a model for the study of vascular injury. *Thrombo Res* 18:113-21.

211. Starkebaum G, Harlan JM. 1986. Endothelial cell injury due to copper-catalyzed hydrogen peroxide generation from homocysteine. *J Clin Invest* 77:1370-6.

212. Hanaki A, Kamaide H. 1975. Autooxidation of cysteine catalyzed by copper in glycylglycine buffer. *Chem Pharm Bull* 23:1671-1676.

213. Winterbourn CC. 1993. Superoxide as an intracellular radical sink. *Free Radic Biol Med* 14:85-90.

214. Hogg N. 1999. The effect of cyst(e)ine on the auto-oxidation of homocysteine. *Free Radic Biol Med* 27:28-33.

215. Morimoto RI, Tissieres A, and Georgopoulos C. *The Biology of Heat Shock Proteins and Molecular Chaperones*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

216. Watson AJ, Askew JN, Sandle GI. 1994. Characterisation of oxidative injury to an intestinal cell line (HT-29) by hydrogen peroxide. *Gut* 11:1575-81.

217. Upchurch GR Jr, Welch GN, Fabian AJ, Freedman JE, Johnson JL, Keaney JF Jr, Loscalzo J. 1997. Homocyst(e)ine decreases bioavailable nitric oxide by a mechanism involving glutathione peroxidase. *J Biol Chem* 272:17012-7.

218. Outinen PA, Sood SK, Liaw PC, Sarge KD, Maeda N, Hirsh J, Ribau J, Podor TJ, Weitz JI, Austin RC. 1998. Characterization of the stress-inducing effects of homocysteine. *Biochem J* 332:213-21.

219. Loscalzo J. 1996. The oxidant stress of hyperhomocyst(e)inemia. *J Clin Invest* 98:5-7.

220. Heinecke JW, Rosen H, Suzuki LA, Chait A. 1987. The role of sulfur-containing amino acids in superoxide production and modification of low density lipoprotein by arterial smooth muscle cells. *J Biol Chem* 262:10098-103.

221. Hirano K, Ogihara T, Miki M, Yasuda H, Tamai H, Kawamura N, Mino M. 1994. Homocysteine induces iron-catalyzed lipid peroxidation of low-density lipoprotein that is prevented by alpha-tocopherol. *Free Radic Res* 5:267-76.

222. Fohe L. 1989. In *Glutathione: Chemical, Biomedical and Medical Aspects*, eds. Pp. 644-731.

223. Harker LA, Slichter SJ, Scott CR, Ross R. Homocystinemia. 1974. Vascular injury and arterial thrombosis. *N Engl J Med* 291 11:537-43.

224. Simon DI, Stamler JS, Jaraki O, Keaney JF, Osborne JA, Francis SA, Singel DJ, Loscalzo J. 1993. Antiplatelet properties of protein S-nitrosothiols derived from nitric oxide and endothelium-derived relaxing factor. *Arterioscler Thromb* 6:791-9

225. Stamler JS, Osborne JA, Jaraki O, Rabbani LE, Mullins M, Singel D, Loscalzo J. 1993. Adverse vascular effects of homocysteine are modulated by endothelium-derived relaxing factor and related oxides of nitrogen. *J Clin Invest* 1:308-18.

226. Gryglewski RJ, Palmer RM, Moncada S. 1986. Superoxide anion is involved in the breakdown of endothelium-derived vascular relaxing factor. *Nature* 320:454-6.

227. Lentz SR, Sobey CG, Piegors DJ, Bhopatkar MY, Faraci FM, Malinow MR, Heistad DD. 1996. Vascular dysfunction in monkeys with diet-induced hyperhomocyst(e)inemia. *J Clin Invest* 98:24-9.

228. Freedman JE, Loscalzo J, Benoit SE, Valeri CR, Barnard MR, Michelson AD. 1996. Decreased platelet inhibition by nitric oxide in two brothers with a history of arterial thrombosis. *J Clin Invest* 97:979-87.

229. Weiss N, Zhang YY, Heydrick S, Bierl C, Loscalzo J. 2001. Overexpression of cellular glutathione peroxidase rescues homocyst(e)ine-induced endothelial dysfunction. *Proc Natl Acad Sci U S A* 98:12503-8.

230. Welch GN, Upchurch GR, Keany JF, and Loscalzo J. Homocyst(e)ine decreases cell redox potential in vascular smooth muscle cells. *J Am Coll Cardiol* 27:164A

231. Ratnoff OD. 1968. Activation of Hageman factor by L-homocystine. *Science*. 162:1007-9.

232. Rodgers GM, Kane WH. Activation of endogenous factor V by a homocysteine-induced vascular endothelial cell activator. J Clin Invest 6:1909-16, 1986

233. Zoller B, Dahlback B. 1995. Resistance to activated protein C caused by a factor V gene mutation. *Curr Opin Hematol* 2:358-64.

234. Rodgers GM, Conn MT. 1990. Homocysteine, an atherogenic stimulus, reduces protein C activation by the arterial and venous endothelial cells. *Blood* 75:895-901.

235. Lentz SR, Sadler JE. 1991. Inhibition of thrombomodulin surface expression and protein C activation by the thrombogenic agent homocysteine. *J Clin Invest* 88:1906-14.

236. Hagen FS, Gray CL, O'Hara P, Grant FJ, Saari GC, Woodbury RG, Hart CE, Insley M, Kisiel W, Kurachi K, Davie EW. 1986. Characterization of a cDNA coding for human actor VII. *Proc Natl Acad Sci USA* 83:2412- 2416.

237. Fryer RH, Wilson BD, Gubler DB, Fitzgerald LA, Rodgers GM. 1993. Homocysteine, a risk factor for premature vascular disease and thrombosis, induces tissue factor activity in endothelial cells. *Arterioscleor Thromb* 13:1327-33.

238. Nishinaga M, Ozawa T, Shimada K. 1993. Homocysteine, a thrombogenic agent, suppresses anticoagulant heparan sulfate expression in cultured porcine aortic endothelial cells. *J Clin Invest* 92:1381-6.

239. Iiyama K, Hajra L, Iiyama M, Li H, DiChiara M, Medoff BD, Cybulsky MI. 1999. Patterns of vascular cell adhesion molecule-1 and intercellular adhesion molecule-1 expression in rabbit and mouse atherosclerotic lesions and at sites predisposed to lesion formation. *Circ Res* 85:199-207.

240. Cybulsky MI, Gimbrone MA Jr. 1991. Endothelial expression of a mononuclear leukocyte adhesion molecule during atherogenesis. *Science* 251:788-91.

241. Wang G, Siow YL, O K. 2000. Homocysteine stimulates nuclear factor kappaB activity and monocyte chemoattractant protein-1 expression in vascular smooth-muscle cells: a possible role for protein kinase C. *Biochem J* 352:817-26.

242. Welch GN, Loscalzo J. 1998. Homocysteine and atherothrombosis. *N Engl J Med* 338:1042-50.

243. Anderson KM, Castelli WP, Levy D. 1987. Cholesterol and mortality. 30 years of follow-up from the Framingham study. *JAMA* 257:2176-2180.

244. Martin MJ, Hulley SB, Browner WS, Kuller LH, Wentworth D. 1986. Serum cholesterol, blood pressure and mortality: Implications from a cohort of 361, 662 men. *Lancet* 2:933-936.

245. Mayes PA. 1993. *Cholesterol synthesis, Transport and excretion*. In R.K. Murray, D.K. Garner, P.A. Mayes, V.W. Rodwell (eds). Harper's Biochemistry (23rd edition). Appleton & Lange, Norwalk, Connecticut, pp 266-278.

246. Marinetti GV. 1990. *Dietary management of elevated blood lipids. In G.V. Marinetti (ed). Disorders of Lipid Metabolism*. Plenum Press, New York, pp 135-168.

247. Choy PC, Mymin D, Zhu Q, Dakshinamurti K, O K. 2000. Atherosclerosis risk factors: the possible role of homocysteine. *Mol Cell Biochem* 207:143-8.

248. Jayakody L, Senaratne M, Thomson A, Kappagoda T. 1987. Endothelium dependent relaxation in experimental atherosclerosis in the rabbit. *Circ Res* 60: 251-264.

249. Sloop GD. 1996. A Unifying theory of atherogenesis. *Medical Hypotheses* 47: 321-325.

250. Olszewski AJ, Szostak WB, Bialkowska M, Rudnicki S, McCully KS. 1989. Reduction of plasma lipid and homocysteine levels by pyridoxine, folate, cobalamin, choline, riboflavin and troxerutinum in atherosclerosis. *Atherosclerosis* 75: 1-6.

251. Frauscher G, Karnaukhova E, Muehl A, Hoeger H, Lubec B. 1995. Oral administration of homocysteine leads to increased plasma triglycerides and homocystenic acid additional mechanism in homocysteine induced endothelial damage. *Life Sci* 57: 813-817.

252. K, Lynn EG, Chung YH, Siow YL, Man RYK, Choy PC. 1985. Homocysteine stimulates the production and secretion of cholesterol in hepatic cells. *Biochem Biophys Acta* 836: 385-389.

253. Kosykh VA, Preobrazhensky SN, Fuki IV, Zakina OZ, Tsibulksy VP, Repin VS, Smirnov VN. 1985. Cholesterol can stimulate secretion of apolipoprotein B by cultured human hepatocytes. *Biochem Biopys Acta* 836: 385-389.

254. Tonstad S, Joakimsen O, Stensland-Bugge E, Leren TP, Ose L, Russell D, Bonaa KH. 1996. Risk factors related to carotid intima-media thickness and plaque in children with familial hypercholesterolemia and control subjects. *Arterioscler Thromb Vasc Biol* 16:984-91.

255. Tonstad S, Refsum H, Ueland PM. 1997. Association between plasma total homocysteine and parental history of cardiovascular disease in children with hypercholesterolemia. *Circulation* 96:1803-1808.

256. Piyathilake CJ, Macaluso M, Hine RJ, Richards EW, Krumdieck CL. 1994. Local and systemic effects of cigarette smoking on folate and vitamin B-12. *Am J Clin Nutr* 60:559-66.

257. Hoogeveen EK, Kostense PJ, Beks PJ, Mackaay AJ, Jakobs C, Bouter LM, Heine RJ, Stehouwer CD. 1998. Hyperhomocysteinemia is associated with an increased risk of cardiovascular disease, especially in non-insulin-dependent diabetes mellitus: a population-based study. *Arterioscler Thromb Vasc Biol* 18:133-8.

258. Munshi MN, Stone A, Fink L, Fonseca V. 1996. Hyperhomocysteinemia following a methionine load in patients with non-insulin-dependent diabetes mellitus and macrovascular disease. *Metabolism* 45:133-5.

259. Kesavulu MM, Rao BK, Giri R, Vijaya J, Subramanyam G, Apparao C. 2001. Lipid peroxidation and antioxidant enzyme status in Type 2 diabetics with coronary heart disease. *Diabetes Res Clin Pract* 53:33-9.

260. Klag MJ, Whelton PK, Randall BL, Neaton JD, Brancati FL, Ford CE, Shulman NB, Stamler J.1996. Blood pressure and end-stage renal disease in men. *N Engl J Med.* 334:13-8.

261. Brattstrom L, Wilcken DE. 2000. Homocysteine and cardiovascular disease: cause or effect? *Am J Clin Nutr* 72:315-23.

262. Fan B, Bose N, Hargrove JL and Hartle DK. 1999. Computer simulation model for vitamin effects on homocysteine metabolism. In: Anderson JG and Katzper M.eds., Health Sciences Simulation, Society for Computer Simulation International, San Diego, CA, pp.170-174.

263. Chadefaux B, Rethore MO, Raoul O, Ceballos I, Poissonnier M, Gilgenkranz S, Allard D. 1985.Cystathionine beta synthase: gene dosage effect in trisomy 21. *Biochem Biophys Res Commun* 128:40-4.

264. Clarke RL and Daly *et al*. 1991. Hyperhomocysteinemia: an independent risk factor for vascular disease. *N Engl J Med* 324:1149-55.

265. Bukovska GV and Kery *et al*; 1994. Expression of human cystathionine beta-synthase in Escherichia coli: purification and characterization." *Protein Expr Purif* 5:442-8.

266. Janosik M, Kery V, Gaustadnes M, Maclean KN, Kraus JP. 2001. Regulation of human cystathionine β-synthase by S-adensoyl-L-methionine: evidence for two catalytically active conformations involving an autoinhibitory domain in the C-terminal region. *Biochemsitry* 40:10625-10633.

267. Meier M, Janosik M, Kery V, Kraus JP, Burkhard P. 2001. Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *EMBO* J 20:3910-6.

268. Laurent B Coignet J.1973. Cystathionuria and defective enzyme regulation. *Clin Chim Acta* 43:171-82.

269. Tateishi N, Higashi T, Shinya S, Naruse A, Sakamoto Y. 1974. Studies on the regulation of glutathione level in rat liver. *J Biochem* 75:93-103.

270. Steegborn C, Clausen T, Sondermann P, Jacob U, Worbs M, Marnikovic S, Huber R and Wahl M. 1999. Kinetics and inhibition of recombinant human cystathionine γ-lyase. *J Biol Chem* 274:12675-12684.

271. Finkelstein JD and Martin JJ. 1984. Methionine metabolism in mammals. Distribution of homocysteine between competing pathways. J Biol Chem 259:9508-13.

272. Kang SS, Wong PW, Malinow MR. 1992. Hyperhomocysteinemia as a risk factor for occlusive vascular disease. *Annu Rev Nutr* 12:279-98.

273. Millian NS & Garrow TA. 1998. Human betaine-homocysteine methyltransferase is a zinc metalloenzyme. *Arch Biochem Biophys* 356:93-98.

274. Finkelstein JD, Martin JJ. 1984. Inactivation of Betaine-Homocysteine Methyltransferase by Adenosylmethionine and Adenosylethionine. *Biochem. Biophys. Res. Commun.* 118, 14-19.

CHAPTER I

EXPRESSION OF RECOMBINANT HUMAN BETAINE: HOMOCYSTEINE S-METHYLTRANSFERASE FOR X-RAY CRYSTALLOGRAPHIC STUDIES AND FURTHER CHARACTERIZATION OF INTERACTION WITH S-ADENOSYLMETHIONINE[1]

ABSTRACT

Elevated homocysteine as a result of dysfunctional metabolic enzymes is an independent risk factor for arteriosclerosis. Betaine:homocysteine S-methyltransferase (BHMT) (EC 2.1.1.5) is an important enzyme in the pathway of homocysteine metabolism in that it recycles methionine from homocysteine and non-folate methyl donors. To initiate X-ray crystallographic structural studies, we created a BHMT expression construct for use in *E. coli* that has a polyhistidine purification tag with no extraneous protein, usually found in commercial vectors, between the tag and protein sequence. The extra amino acids can hinder the crystallization process. A modified pET28b vector was designed to produce N-terminal poly-histidine tagged proteins with a simple construction scheme having broad applicability because of the use of rare *Sap* I cloning sites. BHMT expressed using this vector could be rapidly purified using metal chelate chromatography. Gel exclusion chromatography analysis showed that recombinant polyhistidine-tagged human BHMT is a tetramer. S-adenosylmethionine (SAMe) has no effect on the recombinant BHMT's ability to methylate homocysteine nor does the enzyme appear to bind SAMe when examined by microcalorimetry.

KEYWORDS

ABBREVIATIONS

BHMT, betaine:homocysteine methyltransferase; MS, methionine synthase (vitamin B_{12} dependent); SAMe, S-adenosylmethionine; DL-Hcy, DL-homocysteine; BME, β-mercaptoethanol; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl β-D-thiogalactopyranoside; PCR, polymerase chain reaction; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis.

INTRODUCTION

Homocysteine, an intermediate in the active methyl cycle, is an independent risk factor for the development of arteriosclerotic heart disease (1). Elevated homocysteine levels can result from deficiencies in any one of the enzymes used to process homocysteine in the body. The enzyme betaine-homocysteine S-methyltransferase (BHMT) catalyzes the conversion of betaine and homocysteine to dimethylglycine and methionine, respectively. The better-known folate/vitamin B-12 dependent methionine synthase (EC 2.1.1.13) is the only other enzyme known to methylate homocysteine in mammalian cells. BHMT plays a key role in the regulation of methionine metabolism by maintaining hepatic concentrations of methionine during periods of inadequate intake of this amino acid (2,3) and by removing excessive homocysteine (4). In addition to the role of BHMT in methionine conservation, the utilization of betaine by BHMT is an obligatory reaction in the catabolism of choline in mammalian tissues (5).

Earlier *in vitro* studies have indicated that the recycling of homocysteine to methionine by remethylation is shared equally between BHMT and methionine synthase (1). Oral doses of betaine have been used therapeutically to reduce homocysteine (6,7,8)

in patients suffering from homocystinuria due to cystathionine-beta-synthase deficiency and non-responsive to vitamin B$_6$ therapy. The efficacy of betaine treatment, which could be partly due to increased methylation of homocysteine by the BHMT-catalyzed reaction, makes BHMT a pharmacologically interesting target for the treatment of homocystinuria.

BHMT is a zinc metalloenzyme (9) that shares limited homology to the N-terminal region of *E. coli* methionine synthase (10) and also *E. coli* S-methylmethionine:homocysteine methyltransferase (11). Human BHMT (subunit Mw 45,000) isolated from liver has a reported M$_r$ of 270,000 making it a hexamer in solution. Studies on the enzyme isolated from rat liver have demonstrated that S-adenosylmethionine (SAMe) inhibits BHMT (12). However, a lack of sequence similarity between BHMT and other SAMe binding proteins suggests that BHMT does not have a SAMe binding motif. The recent discovery of a new BHMT gene in humans and mice, betaine-homocysteine methyltransferase-2 (BHMT-2) that has 73% amino acid identity to BHMT (13) may clarify some of these discrepancies. Since BHMT preferentially utilizes betaine as the methyl donor and simultaneously reduces the choline levels, BHMT-2 might interact with SAMe as an alternative methyl donor. All these biochemical aspects of the enzyme could be better understood if the atomic structure of the enzymes were known. Structural information obtained by X-ray crystallographic studies could be helpful to identify the critical residues involved in its mechanism, predict mutations in BHMT that could make individuals pre-disposed to heart disease, and discover better methyl donors since they could be more potent than betaine in lowering plasma homocysteine (10).

Here we report the over-expression of recombinant human liver BHMT in *E. coli* using a modified pET28b vector designed by us. The modified vector uses *Sap* I restriction sites, which because of its 7 base recognition sequence, occur rarely in particular in human genes. We produced amino terminal histidine tagged BHMT, which enabled us to purify the enzyme by a quick, two-step process using metal affinity and ion exchange chromatography. The molecular weight determination of the purified BHMT using size exclusion chromatography supports our crystallographically derived suggestion that the recombinant enzyme is a tetramer (14). We also studied the effect of S-adenosylmethionine (SAMe) on the activity of BHMT to compare the human enzyme to the one in the rat using the same strategy, a radiochemical assay, performed in the rat enzyme study. Isothermal titration calorimetry (ITC) of BHMT was done to confirm our finding that SAMe does not directly interact with BHMT.

MATERIALS AND METHODS

Materials: A sample of an expanded human liver cDNA library (Stratagene) was provided by Dr. Harry Dailey's lab at the University of Georgia (UGA). "PCR-Script", "Seamless", and "Quick-Change" kits and BL21 (DE3) RIL cells for expression were purchased from Stratagene. The expression vector, pET28b, was from Novagen. Betaine and Dowex – OH were purchased from Sigma Chemical Company (St. Louis, MO). American Radiolabeled Chemicals, Inc. (St. Louis, MO) was the source of [^{14}C]betaine labeled in the methyl groups. The Molecular Genetics Instrumentation Facility at the University of Georgia performed DNA sequencing and synthesis of oligonucleotides. Oligonucleotide sequences are shown in Table1. Restriction enzymes were obtained from

Promega, Boehringer Mannheim and New England Biolabs. *Sap* I (2000 U/ml) came

from New England Biolabs (NEB). All other reagents were of the highest analytical or

molecular grade available from commercial vendors.

Creation of Sap I based expression vector for crystallization purposes: A T7

promoter based vector, pET28b, was used as the template vector for further modification.

The single *Sap* I site of the vector at position 3108 was knocked out by site-directed

mutagenesis (Primers: Sap-knockout Prim1 and Sap-knockout Prim2) using the Quick-

Change protocol (Stratagene) which entails PCR amplification of the template with *Pfu*

polymerase using mutation primers, followed by *Dpn* I digestion of the original template.

The vector with the correct mutation was identified by restriction analysis using *Sap* I.

Utilizing the *Nco* I and *Eco*R I sites present in the original pET28b vector's multiple

cloning site (MCS), DNA encoding a polyhistidine tag (His$_5$) was introduced after the

initiating ATG with a double overlapping *Sap* I sites placed between the initiating ATG

and a double stop codon. (Primers: NHSap-Forward and NHSap-Reverse). The resulting

vector called pET28b-NHSap can incorporate any DNA of interest with *Sap* I sites

flanking at both the ends of the DNA (Figure 1). Digestion with *Sap* I detaches the

recognition site from both the vector and the desired insert which allows the sticky ends

flanking the desired sequences to ligate and form a seamless junction. Despite the

removal of the *Sap* I restriction site near the origin of replication, the vector appears

stable.

Cloning of BHMT into the pET28b-NHSap construct: The DNA encoding

BHMT was amplified by PCR from a human liver cDNA library (Stratagene). The PCR

reactions were performed using *Pfu* Turbo DNA polymerase (Stratagene) with the

oligonucleotide primers, BHMT forward and BHMT reverse (Table 1) and analyzed by agarose gel electrophoresis. The resulting 1220 bp PCR product was then ligated into the PCR product compatible vector pCRScript, and the desired clone was identified as white colonies on IPTG/X-GAL plates. The DNA sequence of one clone containing the BHMT insert was confirmed by restriction analysis and sequencing. The PCR vector containing the protein coding sequence of BHMT was further PCR amplified with Sap I restriction sites engineered upstream of the first ATG of the cDNA and with a C-terminal stop codon so that the gene was in the proper reading frame with an N-terminal purification tag (Primers: NHSap BHMT-Forward and NHSap BHMT-Reverse). Synthesis of the desired PCR product was again verified by agarose electrophoresis, and the product purified using a PCR Quick Spin Kit (Qiagen). This PCR product was then inserted into the expression construct pET28b-NHSap using the "Seamless Strategy" of Stratagene, but using Sap I rather than Eam1104 as the restriction enzyme. 75 ng of the modified pET28b vector and 50 ng PCR product (1:2 molar ratio) were digested together with 2 units of Sap I in a 10 μl reaction overnight at 37°C (NEB-Buffer 4). To the digested mix, 1 unit T_4 DNA ligase (Gibco), 4 μl of 5X ligase buffer, 6 additional units of Sap I restriction enzyme (3 μl), and 2 μl of water were added to get a 20 μl reaction volume. After overnight ligation, 1 μl of the ligation reaction was transformed into 40 μl $E.\ coli$ strain XL1- Blue (Stratagene) by electroporation. The colonies were screened for the vector containing the insert by single colony screening and restriction analysis. The resulting expression construct was then verified by DNA sequencing and named pET28b-NHBhmt.

Expression of N-terminal poly-histidine tagged BHMT: The expression

construct, pET28b-NHBhmt was then transformed into electrocompetent BL21 Codon

Plus (DE3) RIL cells (Stratagene) by electroporation. Two liters of LB media (1 L per

Fernbauch flask), each containing 50 μg/ml kanamycin and 250 μM zinc chloride were

inoculated with a 2 ml overnight culture. The zinc chloride is added to ensure adequate

occupancy of the Zn in this metalloenzyme. After growing the culture to A_{600} = 0.6 at

37°C, the culture was transferred to room temperature and then induced with 0.5 mM

IPTG for 4 hrs and maintained at room temperature. Following the induction period, the

cells were collected by centrifugation at 13,000 x g for 20 min (4°C) and resuspended in

40 ml of extraction buffer (20mM sodium phosphate, 0.5M NaCl, pH 7.4) containing a

EDTA-free protease tablet (Boehringer-Mannheim) and either used immediately, or

frozen at -20°C for later use. The cell suspension was cooled in ice to 10°C and sonicated

for 3 min in 1-min increments with a Fisher Scientific 550 sonic dismembrator set at 50%

duty cycle while maintaining the temperature below 15°C. Cellular debris was removed

by centrifugation at 60,000 x g for 15 min (4°C), and the clarified supernatant was

purified by chromatography. A small aliquot was frozen at -20°C for further assay.

   Purification of recombinant BHMT: The entire purification process was done

on a Pharmacia ÄKTA HPLC system at room temperature. The first step of purification

was by Ni^{2+}-metal chelate affinity chromatography with a Pharmacia NTA Hi-Trap

Chelating 5ml column. The clarified lysate was applied (1 ml/min) to the nickel-charged

NTA column that had been equilibrated with 20 mM sodium phosphate, 0.5 M NaCl, pH

7.4. After washing unbound protein off, the protein was then eluted using a linear

gradient of imidazole (0-0.5 M imidazole, 0.5 M NaCl, 20 mM sodium phosphate, pH

7.4) over 20 column volumes. The protein eluted out as two closely associated peaks. The fractions corresponding to the two peaks were pooled separately and the purity was checked by SDS-PAGE gel electrophoresis using a Pharmacia PHAST system. Even though SDS-PAGE analysis showed a single band after the metal-chelate column, a second purification step, utilizing anion exchange chromatography was done to improve the prospects of obtaining better quality crystals. After the metal-chelate chromatography, the major protein peak was dialyzed into 50 mM tris HCl buffer (pH 8.0), 5 mM beta-mercaptoethanol (BME) and 0.5 mM EDTA using 2 x 1 L changes. The dialyzed protein was then loaded onto the Pharmacia Hi-Trap Q column (5ml) equilibrated with 50mM tris HCl, (pH 8), 5mM BME and 0.5 mM EDTA. The protein was eluted using a linear gradient of 0-1.0 M NaCl over 20 column volumes. The purity was again checked by SDS-PAGE analysis. At all stages of the purification, small aliquots of protein were saved at -20°C. Later, the protein concentrations of the frozen aliquots were determined by the Lowry method (15) using bovine serum albumin as a standard and then assayed for BHMT activity.

In vitro assay for BHMT activity: The standard radiochemical assay for BHMT activity that has been previously described (16, 10) was used with some modifications. *DL*-Hcy was prepared fresh as described by Garrow (10). All the reaction tubes were capped with rubber stoppers and kept on ice until the assay was started. The final concentrations of the assay components were 10 mM *DL*-homocysteine, 2.5 mM betaine (1 µCi/ml), 5mM BME, and 50 mM Tris (pH 8.0). The reaction was initiated with the addition of protein sample to give a total volume of 0.5 ml and the reactants were incubated at 37°C for one hour. Following the incubation, the samples were chilled in ice

water and 2.5 ml of ice cold water was added. The samples were applied to a Dowex 1-X4 (OH-; 200-400 mesh), column (5ml), which was previously equilibrated with the 50 mM Tris HCl buffer, pH 8.0. The unreacted betaine was washed from the column with water (4 x 5 ml). Dimethylglycine and methionine were eluted into scintillation vials with 6 ml of 1.5 N HCl. The exact volume of the elution for each sample was noted. To 1 ml of the eluted sample, 750 µl of 2 N NaOH and 12 ml of scintillation mixture (Ecolume, ICN) were added and then counted. Of the counts measured, methionine contributed one-third and dimethylglycine, two-thirds. For each assay, blanks were prepared from the same buffer conditions as the sample to be measured. All samples were assayed in triplicate.

Molecular weight determination: The molecular weight of the native BHMT was determined by gel filtration on a Sephacryl S-300 High Resolution column (26 X 60 cm) (Pharmacia Biotech). Calibration was done with the following protein standards: Thyroglobulin (669,000 Da), Apoferritin (443,000 Da), β-Amylase (200,000 Da), Alcohol Dehydrogenase (150,000 Da), Albumin (66,000 Da) and Carbonic Anhydrase (20,000 Da). Aliquots of the post-Q column BHMT sample (1 ml of 1 mg/ml) were applied to the column that had been previously equilibrated with 50mM Tris HCl, 5mM BME and 0.5mM EDTA, pH 8.

SAMe inhibition study using the radiochemical assay: SAMe was added to the previously described radiochemical assay conditions to achieve three concentrations of 100, 200 and 400 nM. This study was done in two parts. In the first part the enzyme was pre-incubated with the above concentrations of SAMe at 37°C for 30 mins. The pre-incubated enzyme was then assayed by the normal radioactive assay. In the second part

of the study, the same concentrations of SAMe were added to the enzyme assay mixture immediately before starting the assay. The specific activities of the enzyme samples from both parts of the study were compared to look for SAMe inhibition of BHMT. All the samples of enzyme were assayed twice. The purity and the stability of SAMe were verified by electrospray mass spectroscopy.

Qualitative Isothermal Titration Calorimetry (ITC): ITC experiments were performed using an isothermal titration calorimeter of the MCS system from Microcal, Inc. (Northhampton, MA). The reference cell was filled with water. All solutions were degassed for 10 min with gentle stirring under vacuum. Solutions of the protein were filled in the sample cell (cell volume = 1.3424 ml) and titrated with *DL*-homocysteine and SAMe. Ligand solutions were prepared in the same buffer as the protein sample. The ligand concentration in the injection syringe was usually 25 times higher than the concentration of protein binding sites. In individual titrations, injections of 5 μl of 1 mM SAMe were added from the computer controlled 250 μl microsyringe into the protein solution (2mg/ml) at 4 min intervals with stirring at 350 rpm. A positive control experiment was performed by injecting DL-homocysteine (1mM) into the protein solution. Control experiments to nullify the effects of dilution were performed by making identical injections of homocysteine and SAMe into a cell containing buffer with no protein.

RESULTS

Cloning of BHMT and preparation of Expression Construct pET28b-NHBhmt:
The PCR amplified BHMT from the human cDNA liver library was cloned into pCRScript cloning vector using the vector's *Srf I* site. Successful selection of the right

85

clone containing the gene in the desired orientation was achieved by blue-white selection and restriction analysis. The sequence analysis of the clone matched with the published sequence (10). The T7 based expression vector pET28b was modified so that it contains two *Sap* I sites and a polyhistidine tag oriented on the amino terminus of the protein sequence (pET28b-NHSap). The sequenced BHMT gene within the PCR vector was then PCR re-amplified with *Sap* I sites flanking the gene. Digestion of both the vector and the PCR product together with *Sap* I was followed directly by ligation. After transformation into electrocompetent *E. coli (*XL1-Blue) cells, kanamycin resistant colonies were screened for proper insertion using restriction analysis. The correct clone (pET28b-NHBhmt), re-checked by restriction analysis and sequencing was used for further transformation into the expression strain.

  Expression and purification of BHMT: The expression of the N-terminal histidine tagged BHMT was carried out in BL21 Codon Plus (DE3) RIL cells, a strain that increases the concentrations of several rare (but common for eukaryotes) *E. coli* codons. Initially some solubility problems at 37°C were encountered, but they were overcome by growing the cells at room temperature. BHMT was then purified by metal-chelate chromatography. BHMT eluted out as two closely associated peaks. The SDS-PAGE analysis of the two peaks could not differentiate between them (Figure 2). It could be because of some proteolytic degradation presumably caused by bacterial exopeptidase activity as shown in earlier work (9). The two peaks were assayed and were found to have approximately the same specific activity. So the main peak of the protein was considered for further analysis. This single step of purification reported here enriched the activity about 15 fold with a final yield of 40% and the preparation was judged to be

homogenous after SDS-PAGE yielding a single band of Mr 45,000 as shown in Figure 2. The preparation shown in Table 2 had a specific activity of 1432 U/mg with a unit defined as nanomoles of methionine formed per hour. The protein purification took a total of 4 hours up to this step. The single most important factor in the success of a crystallization experiment is the quality of the macromolecule. Thus our major emphasis was to ensure that the protein is of the highest purity. So we decided to add an additional purification step using anion exchange chromatography. Total purification time including dialysis was eight hours. This final purification step yielded a homogenous protein of specific activity of 1531 U/mg. This is in the similar range of the activity found for protein isolated from human liver (1100 U/mg) or recombinant enzyme (1957 U/mg) (9).

Size exclusion chromatography: Protein standards were used to generate a calibration curve of log molecular weight versus elution volume (Figure 3). BHMT eluted as a single symmetric peak between alcohol dehydrogenase (150,000 Da) and β-amylase (200,000 Da). When calculated, the BHMT peak was found to elute at a volume consistent with a molecular mass near 180,000 Da. This result when combined with the fact that BHMT is composed of 45,000 Da subunits, suggests that the recombinant his-tagged human BHMT is a tetramer.

SAMe Inhibition Study Using Radiolabeled Assay: As indicated by the Table 3, the activity of BHMT was found to be unaffected by the incorporation of SAMe in the assay. When the enzyme was incubated with the given concentrations of SAMe, the specific activity of the enzyme remained the same. Contrary to the results obtained from previous rat liver study, no inhibition of human BHMT by SAMe was observed (12).

87

Qualitative Isothermal Titration Calorimetry: In order to confirm BHMT's

possible affinity for SAMe, binding studies were conducted employing Isothermal

Titration Calorimetry. During the course of these experiments, *DL*-homocysteine and

SAMe were titrated individually into a solution of BHMT at an experimental temperature

of 30°C. For each injection of the ligands, the heat change inside the protein was

recorded. For each ligand, a control was performed by titrating the ligand into a buffer

solution with no protein. If binding between the protein and the ligand occurs, a clearly

measurable positive or negative heat change, which is considerably different from the

control should be observed and, as the experiment proceeds and binding is saturated, this

heat change should decrease in value finally falling to zero. When *DL*-homocysteine was

titrated into the buffer with no protein, no significant heats of dilution were observed. But

the titration of BHMT against *DL*-homocysteine was characterized by a significant

exothermic heat change. The binding isotherm of *DL*- Hcy showed negative heat change

at the beginning of the experiment and then it gradually decreased as binding sites were

saturated. Thus DL-Hcy binds to BHMT. However, titration with SAMe showed a

different behavior. The control titration of SAMe into the buffer solution showed a

negative heat change. Also the titration of the SAMe into the protein showed the same

heat signals as with the buffer but no heat effect typical of a binding reaction was

observed. When the heat change per injection of SAMe was plotted versus the injection

number, no titration curve could be obtained (Results not shown). This supports the

radiochemical assay data suggesting that BHMT does not bind SAMe. Due to the

commercial unavailability of *L*-Homocysteine, no exact determination of a binding

constant was possible. Only a qualitative statement about BHMT's ability to bind homocysteine and SAMe could thus be obtained.

DISCUSSION

Commercial expression vectors tend to introduce additional amino acids, beyond a simple purification tag, that confer undesired conformational flexibility and thus decrease the likelihood of crystallization. The earlier methods of expressing BHMT as a fusion protein with beta-galactosidase was not considered appropriate for crystallization studies (10, 17). So a short purification tag like the polyhistidine tag that does not require protease cleavage was preferred. The "seamless strategy" used to make the expression construct simplifies cloning and also decreases the number of superfluous amino acids on the amino or carboxyl ends. Further, the use of *Sap* I ensures that most genes can be inserted into this vector because of the rare occurrence of the 7 base restriction recognition sequence.

Introduction of the his-tagged protein allowed us to adopt a simple one-step purification protocol. A complete purification could be performed within four hours on our system as compared to the earlier purification procedures, which took 16 to 18 hours (10, 17). But for crystallization purposes we did include one additional step of purification by anion exchange chromatography that added an extra four hours including dialysis time.

As published in our preliminary X-ray crystallography studies of the recombinant human BHMT crystals, the non-crystallographic two-fold symmetry and the crystallographic C2 (monoclinic) symmetry supported the likelihood that the protein is a

tetramer with molecular 222 symmetry (14), but this finding was inconsistent with the earlier reports that human BHMT is a hexamer (19). So to support the crystallographic observations, a gel filtration study of the recombinant enzyme was performed. The elution profile of the recombinant BHMT proved to be consistent with the native BHMT being a homotetramer with a molecular mass of 180,000 Da. This observation is contrary to the results obtained by Skiba *et al.,* wherein the BHMT isolated from human liver was shown to be a hexamer with molecular mass around 270,000 Da. This discrepancy could be a result of either our recombinant protein having a his-tag at the N-terminus or some post-translational modification, both of which could confer different oligomerization properties to the protein. Alternatively, the presence of stabilizing agents like N,N-dimethylglycine and *DL-* homocysteine used by Skiba *et al*. may change the physical behavior of BHMT (18).

The Isothermal Titration Calorimetry and SAMe inhibition studies showed that SAMe does not inhibit or interact with BHMT. BHMT has significant similarity with *E. coli* methionine synthase, but only to the N-terminal part of methionine synthase (residues 2 to 353), which is the homocysteine and zinc-binding site (10, 19). The SAMe binding domain of methionine synthase is at the C-terminal end (10). So it seems illogical that BHMT should have a SAMe binding site based on sequence similarity arguments. The rat liver BHMT study showed that rat BHMT is inhibited by SAMe suggesting a negative regulatory aspect of SAMe on BHMT. Down regulation of BHMT by SAMe is metabolically sensible as high SAMe would indicate adequate methionine and less need for the remethylation pathway. This action of SAMe on BHMT is in contrast to its action on cystathionine-beta-synthase (CBS), the enzyme in the transsulfuration pathway, where

SAMe activates the enzyme. One consequence of the possible opposing actions of SAMe on CBS and BHMT is that any therapeutic approach that utilizes a SAMe like molecule will in turn activate CBS while inhibiting BHMT. Since the active methyl cycle is involved in many important metabolic pathways, it seems important not to inhibit methionine synthase or BHMT. Activation of BHMT alone may be adequate to significantly reduce homocysteine levels since this will increase SAMe levels, and thus in turn activate CBS. Recently the discovery of a second BHMT homologue (BHMT-2) in humans and mice, which has 73% amino acid identity to BHMT, was reported. BHMT and BHMT-2 have sequence similarity to S-methylmethionine-homocysteine S-methyltransferases found in *E. coli* and *Astragalus bisculatus,* which interact with both S-methylmethionine and SAMe (11, 20). There is thus a possibility that BHMT-2 preferentially uses one of the alternative substrates like SAMe as a methyl donor (9). The results of the rat liver BHMT showing that SAMe inhibits BHMT have to be scrutinized more closely because the assay was done on a partially purified enzyme. It could contain a mixture of BHMT and BHMT-2. If so, BHMT-2 might utilize SAMe preferentially as a methyl donor over the ^{14}C labeled betaine. Since SAMe is not radiolabeled in the assay, non-radiolabeled methionine would be produced giving a false interpretation of SAMe inhibition of BHMT. Alternatively, it is also important to concede that the lack of SAMe interaction with our recombinant BHMT could be a result of the N-terminal poly-histidine residues that might be sterically hindering the larger ligand SAMe, while allowing *L*-homocysteine to bind unhindered. Structural studies of BHMT and BHMT-2 could help to unravel this discrepancy.

ACKNOWLEDGEMENTS

Table1: Oligonucleotides used for the creation of the Sap I knockout vector and the cloning of BHMT

Primer	Sequence
BHMT forward	5' TGTCTGGACACCACAAAGATGCC 3'
BHMT reverse	5' TTGTAGGAACTGTGACCCAAACACC 3'
Sap-knockout Prim1	5' GAAAATACCGCATCAGGCWCTWTTCCGCTTCCTCGCTCAC 3' [†]
Sap-knockout Prim2	5' GTGAGCGAGGAAGCGGAAWAGWGCCTGATGCGGTATTTTC 3' [†]
NHSap-Forward	5' CATGCACCATCATCATCACTGAAGAGCTCTTCATAAT 3'
NHSap-Reverse	5' GTGGTAGTAGTAGTGACTTCTCGAGAAGTATTATTAA 3'
NHSap BHMT-Forward	5' CGCTCTTCTCACATGCCACCCGTTG 3'
NHSap BHMT-Reverse	5' TGCTCTTCGTTACTGTGATTTGAATTTTTG 3'

[†]W = A or T

Table 2: **Purification table of recombinant human liver BHMT**

Fraction	Volume	Activity[a]	Protein[b]	Sp.Act.	Purification	Yield
	(ml)	(U)	(mg/ml)	(U/mg)	(x-fold)	(%)
1. Crude	81	767	8	96	1	100
2. Nickel Affinity						
Peak I	29	859	0.6	1432	14.9	40.1
Peak II	1	958	0.6	1558	16.2	1.5
3. Anion Exchange	18	1074	0.7	1531	15.9	31.1

[a]Units are nmol methionine formed per hour
[b]Protein determined by Lowry Assay

Table 3: *Effect of S-Adenosylmethionine on betaine:homocysteine methyltransferase*

Study	Concentration of SAMe in 0.5 ml reaction (nM)	Specific Activity of BHMT (U/mg) *
Control: No addition of SAMe	0	1536
Preincubation of BHMT with SAMe at 37°C for 30 mins	100	1469
	200	1520
	400	1450
Addition of SAMe immediately before the assay reaction	100	1442
	200	1400
	400	1500

*All values indicate the average of three readings taken for each concentration of SAMe

Sap-knockout Prim1   5 ——✕—— 3
Sap-knockout Prim2   3 ——✕—— 5

Quick-Change Mutagenesis →

MCS

pET28b

Sap I

Xho I
Not I
Eag I
Hind III
Sal I
Sac I
EcoR I
BamH I
Nhe I
Nde I
Nco I

PET28b-SapKnockOut

Nco I and *EcoR* I digestion
+
NHSap-Forward
NHSap-Reverse

Methionine   **Histag**   *Sap* I   **Stop**

ATGCACCATCATCAT CACTGAAGA*GCTCTTC*A TAATAAT
GTGGTAGTAGTAGTGA*CTTCTCG*AGAAGTATT ATTAA
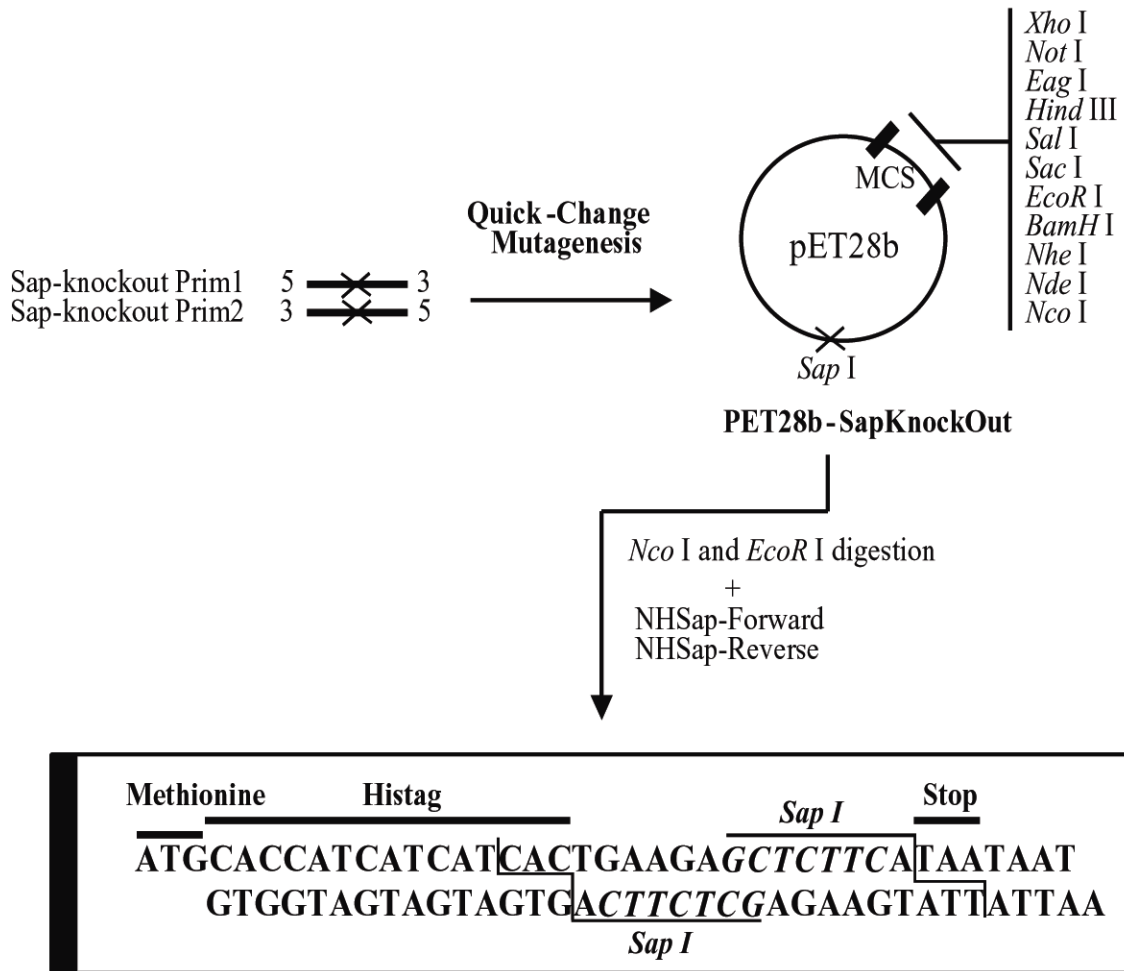
Sap I

Figure 4: Scheme for constructing the Sap I based expression vector

96
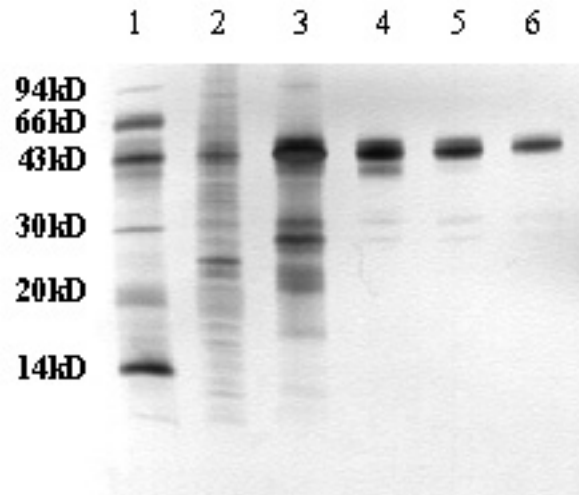
Figure 5: SDS-PAGE analysis of BHMT purification.

BHMT samples after each purification step were treated with SDS-PAGE sample buffer and loaded (0.2 μl) onto a Pharmacia PHAST homogeneous 20% gel and then silver stained. Lane1 – Molecular weight markers with the masses marked along the side of the lane; Lane 2 – Crude extract (1:25 dilution); Lane 3 - Crude extract (1:10 dilution); Lane 4 – First (major) peak of BHMT after affinity chromatography; Lane 5 – Second (minor) peak of BHMT after affinity chromatography; Lane 6 –BHMT after ion exchange column.
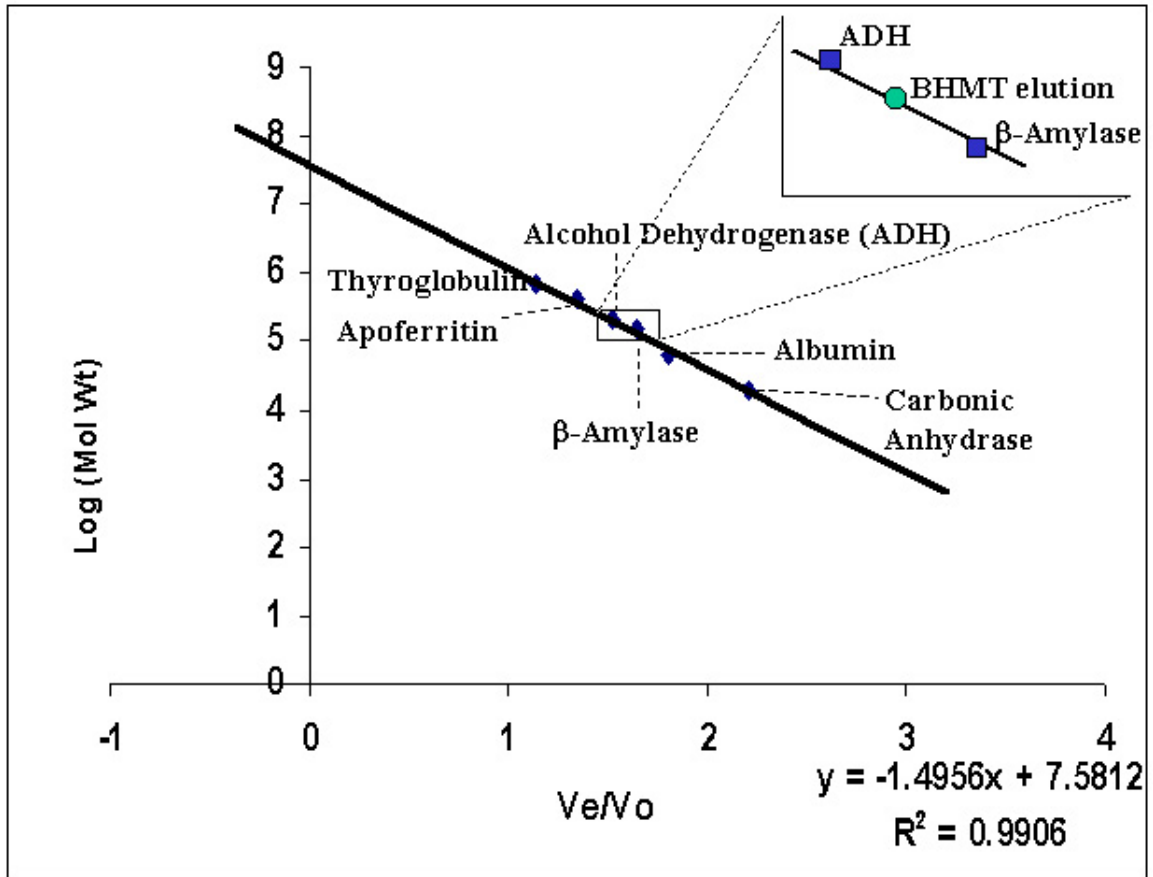
Figure 6: Standard curve of log molecular weight versus elution volume of the protein standards and BHMT

REFERENCES

1. Finkelstein, J. D. & Martin, J. J. (1984b). Methionine metabolism in mammals. Distribution of homocysteine between competing pathways. *J. Biol. Chem.* 259, 9508-9513.

2. Finkelstein, J. D., Martin, J. J., Harris, B. J. & Kyle, W. E. (1982). Regulation of the betaine content of rat liver. *Arch. Biochem. Biophys.* 218, 169-173.

3. Finkelstein, J. D., Harris, B.J., Martin, J. J., Kyle, W.E. (1982). Regulation of hepatic betaine-homocysteine methyltransferase by dietary methionine. *Biochem. Biophys. Res. Commun.* 108(1), 344-348.

4. Finkelstein, J. D., Martin, J. J., Kyle, W. E. & Harris, B. J. (1978). Methionine metabolism in mammals: regulation of methylenetetrahydrofolate reductase content of rat tissues. *Arch. Biochem. Biophys.* 191, 53-160.

5. Finkelstein, J. D., Harris, B.J., Martin, J. J., Kyle, W.E (1983). Regulation of hepatic betaine-homocysteine methyltransferase by dietary betaine. *J.Nutr.* 113, 519-521.

6. Smolin, L.A., Benevenga, J.N., and Berlow, S. (1981). The use of betaine for the treatment of homocystinuria. *J.Pediatr.* 99 (3), 467-472.

7. (1984). Betaine Therapy for Homocystimuria. *Nutrition Reviews* 42 (5), 180-181.

8. Mudd, S. H., Levy, H. L. & Skovby, F. (1995). The metabolic and molecular bases of inherited disease (Edited by Scriver, C. R., Beaudet, A. L., Sly, W. S., and Vale, D., ed.). McGraw-Hill, New York. pp. 1279-1327.

9. Millian, N. S. & Garrow, T. A. (1998). Human betaine-homocysteine methyltransferase is a zinc metalloenzyme. *Arch. Biochem. Biophys.* 356, 93-98.

10. Garrow, T. A. (1996). Purification, kinetic properties, and cDNA cloning of mamalian betaine-homocysteine methyltransferase. *J. Biol. Chem.* 271, 22831-22838.

11. Thanbichler, M., Neuhieri, B., and Bock, A. (1999) S-Methylmethionine metabolism in Escherichia coli. *J. Bacteriol.* 181(2), 662-665.

12. Finkelstein, J.D., Martin, J. J. (1984a). Inactivation of betaine-homocysteine methyltransferase by adenosylmethionine and adenosylethionine. *Biochem. Biophys. Res. Commun.* 118, 14-19.

13. Chadwick, L.H, McCandless, S.E, et.al. (2000). Betaine-homocysteine methyltransferase- 2:cDNA cloning, gene sequence, physical mapping and expression of the human and mouse genes. *Genomics.* 70, 66-73.

14. Bose, N., Momany, C. (2001) Crystallization and preliminary X-Ray crystallographic studies of recombinant human betaine:homocysteine S-methyltransferase. *Acta Cryst.* D57, 431-433.

15. Lowry,O.H., Rosenbrough, N.J., Farr, A.L., and Randall, R.J.(1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193:265.

16. Skiba, W.E., Wells, M.S., and *et.al.* (1987). Betaine-Homocysteine S-Methyltransferase (Human). *Methods Enzymol.* 143, 384-390.

17. Breska, A.P & Garrow, T. A. (1999). Recombinant human liver Betaine-homocysteine S- methyltransferase: identification of three cysteine residues critical for zinc binding. *Biochemistry.* 38, 13991-13998.

18. Skiba, W.E., Taylor, P.M., *et al.* (1982) Human hepatic methionine biosynthesis, purification and characterization of Betaine:Homocysteine S-Methyltransferase. *J. Biol. Chem.* 257(24), 14944-14948.

19. Goulding, C.W. and Matthews, R.G. (1997). Cobalamin-dependent methionine synthase from Escherichia coli: involvement of zinc in homocysteine activation. *Biochemistry.* 36, 15749-15757.

20. Neuhierl, B., Thanbichler, M and *et.al.* (1999). A family of S-methylmethionine-dependent thiol/selenol methyltransferases. Role in selenium tolerance and evolutionary relation. *J. Biol. Chem.* 274(9), 5407-5414.

CHAPTER II

CRYSTALLIZATION AND PRELIMINARY X-RAY CRYSTALLOGRAPHIC STUDIES OF RECOMBINANT HUMAN BETAINE-HOMOCYSTEINE S-METHYLTRANSFERASE[1]

Synopsis

Human betaine:homocysteine S-methyltransferase was cloned, expressed in *E. coli*, purified and crystallized. The crystals belong to the monoclinic space group C2 with cell dimensions *a* = 109.190 Å, *b* = 91.319 Å, *c* = 88.661 Å and β = 122.044° and diffract to 2.9 Å.

Abstract

Betaine:homocysteine S-methyltransferase (BHMT) catalyzes a reaction essential for regulation of methionine and homocysteine metabolism and the catabolism of choline in mammalian tissues. Human recombinant BHMT (M_r 45kDa) has been crystallized by the hanging drop vapor diffusion method at 294°K using ethylene glycol as the precipitant. The crystals belong to the monoclinic space group C2 with cell dimensions *a* = 109.190 Å, *b* = 91.319 Å, *c* = 88.661 Å and β = 122.044° and diffract up to 2.9 Å resolution on a local rotating- anode X-ray source. Rotation function analysis and the Matthews coefficient, V_m=2.46 Å^3 Da^{-1}, are consistent with a dimer in the asymmetric unit, suggesting that the active enzyme is a tetramer with 222 symmetry.

Introduction

Regulation of homocysteine metabolism has gained recent attention since epidemiological studies have established that homocysteine is an independent risk factor, even with only moderate elevations in the plasma, for the development of arteriosclerotic vascular disease (Finkelstein & Martin, 1984b). The enzyme betaine-homocysteine S-

methyltransferase (BHMT; EC 2.1.1.5) has an important role in the modulation of plasma homocysteine in that it catalyzes the the reaction,

$$\text{L - homocysteine} \ + \ \text{L - betaine} \ \xrightarrow{\ \ BHMT\ \ } \text{L - methionine} \ + \ \text{L - dimethylglycine}$$
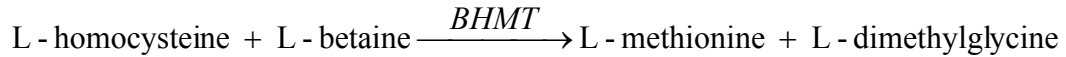
The only other enzyme known to methylate homocysteine in mammalian cells is the folate/vitamin B-12 dependent methionine synthase (EC 2.1.1.13). The reaction catalyzed by BHMT is significant in the regulation of methionine metabolism both as a means for the maintenance of hepatic concentrations of methionine during periods of inadequate intake of this amino acid (Finkelstein *et al.*, 1982) and for the removal of excessive homocysteine (Finkelstein *et al.*, 1978). In addition to the key role of BHMT in methionine metabolism, the utilization of betaine by BHMT is an obligatory reaction in the catabolism of choline in mammalian tissues.

In vitro studies indicate that homocysteine remethylation is shared equally between BHMT and methionine synthase. Large oral doses of betaine have been used therapeutically to reduce homocysteine (Mudd *et al.*, 1995) in patients suffering from homocystinuria. The efficacy of betaine treatment is due, at least in part, to increased methylation of homocysteine by the BHMT-catalyzed reaction. BHMT is therefore a pharmacologically interesting target for the treatment of homocystinuria since a better methyl donor than betaine may have more potent plasma homocysteine lowering effects.

BHMT is a zinc metalloenzyme (Millian & Garrow, 1998) that shares limited similarity to the amino-terminal region of bacterial and eukaryotic methionine synthases (Garrow, 1996). Previous studies on the liver enzyme isolated from rats have shown that it is inhibited by S-adenosylmethionine (SAM) (Finkelstein & Martin, 1984a). However, BHMT does not appear to contain a SAM binding motif recognizable by sequence

analysis. All these biochemical aspects of the enzyme could be better understood if the atomic structure of the enzyme were known. Structural information obtained by X-ray crystallographic studies could be helpful in identifying the critical residues involved in its mechanism and also predict mutations in BHMT that could make individuals pre-disposed to heart disease. Further, drugs that act as more efficient methyl donors could be useful therapeutically. Here we report the crystallization and preliminary diffraction analysis of human recombinant BHMT.

Materials and Methods

Cloning, Expression and Purification of BHMT

The DNA encoding BHMT was amplified by PCR from a human cDNA library (Stratagene) and cloned into an expression vector, a modified pET28b, to contain a polyhistidine purification tag at N-terminus of the protein. Details concerning the construction of the expression plasmid will be reported in detail elsewhere. Expression was performed at room temperature using BL21 Codon Plus (DE3) cells (Stratagene) in LB media supplemented with 250 μM zinc chloride. Cells were sonicated in 0.5M NaCl, 20mM sodium phosphate, pH 7.4, centrifuged at 60000 x g, and the clarified supernatant purified by Ni^{2+}-metal chelate affinity chromatography with a Pharmacia NTA Hi-Trap column using an imidazole gradient for elution (0-0.5 M imidazole, 0.5 M NaCl, 20 mM sodium phosphate, pH 7.4). After purification the protein was dialyzed into 50 mM tris HCl buffer, pH 8.0. A second purification step, using a Pharmacia Hi-Trap Q column, was introduced (elution: 0-1.0 M NaCl, 20 mM tris HCl, pH 8.0) to improve crystal diffraction quality.

Crystallization

A broad range of crystallization conditions following the sparse-matrix sampling method (Jancarik & Kim, 1991) were examined by vapor diffusion (McPherson, 1976) trials using 2 μl drops of protein incubated with equal volumes of well solutions from Crystal Screens 1 and 2 and PEG, and PEG/lithium chloride kits of Hampton Research.

Data Collection and Processing

Immediately before data collection, a single plate-like crystal was quickly transferred into a cryoprotectant solution containing 30% v/v 2,3-butanediol and 70% v/v artificial mother liquor (13% PEG 20000, 8% ethylene glycol, 0.1M Hepes, pH 7.5). The crystal was flash cooled to 100°K in a nitrogen stream produced by a MSC X-STREAM cooling system. X-ray data were collected using a Rigaku R-Axis IIC image plate detector mounted on a Rigaku RU −200 rotating anode X-ray (CuKα) generator equipped with Osmic mirrors and operating at 50kV and 100mA. The data was processed using the DENZO and SCALEPACK programs in the HKL2000 suite (Otwinowski & Minor, 1997).

Results and Discussion

Initial crystallization conditions were obtained with 10 mg/ml protein and 5% PEG 6000 as precipitant with protein purified in the single metal-chelate column step. Tiny crystals appeared the day following setup, but they did not increase in size after the initial growth. Unsuccessful optimization trials included the addition of additives like salts (sodium chloride, lithium chloride, ammonium chloride and ammonium sulfate) and organics (dioxane and glycerol). Since optimization of the initial conditions failed to produce large crystals, an additional purification step, ion-exchange chromatography

using a Pharmacia Hi-Trap Q column, was introduced. Crystal Screens 1 and 2 were

again screened with the higher quality protein. Crystals with a hexagonal morphology,

which unfortunately failed to diffract, were obtained from 10 mg/ml protein in 20 mM

tris HCl, pH 8.0 with well solutions containing 10% w/v PEG 8000, 8% w/v ethylene

glycol, 0.1M Hepes Na, pH 7.5. Ethylene glycol at 8% w/v was absolutely essential for

crystal formation. Additional optimization included screening with PEG 400, 600, 2000,

3000, 5000 and 20,000. Two different types of crystals were obtained from similar

conditions. Condition 1 (Figure 1A) consisted of 2 μl of 7 mg/ml protein, 3% PEG

20,000 and 3% ethylene glycol crystallized with 2 μl of the well solution containing only

8% ethylene glycol. These crystals appeared in a week. Condition 2 (Figure 1B)

consisted of 2 μl of 4mg/ml of protein with the same conditions as condition 1. These

clear, plate-like crystals appeared in 4 weeks. Crystals obtained from condition 1 failed

to diffract while the crystals obtained by condition 2 diffracted up to 2.9 Å resolution.

The diffraction data collected on an area detector (Figure 2) were consistent with

the space group C2. Diffraction spots were observable at 2.9Å resolution, but the data set

was processed to 3.0 Å with an overall R_{merge} of 8.5 %. Diffraction is anisotropic,

consistent with a plate-like crystal. Refined cell parameters were $a = 109.190$ Å, $b =$

91.319 Å, $c = 88.661$ Å and $\beta = 122.044°$. Data collection statistics are summarized in

Table 1. The molecular weight of the monomer is 45 kDa, so assuming two molecules

per asymmetric unit, the Matthews coefficient, V_m is 2.46 $\text{Å}^3\,\text{Da}^{-1}$ corresponding to a

solvent volume fraction of 49.5% (Matthews, 1968). The unit cell volume is 884,041 Å^3.

Using the area detector data, self-on-self rotation functions were calculated using the

program GLRF (Tong & Rossmann, 1997) to identify likely non-crystallographic

symmetry. The rotation function analyses, calculated with 30.0 - 4.0 Å resolution data and radius of integration of 20 Å, showed only strong peaks at $\kappa = 180°$ (Figure 4) that would be consistent with non-crystallographic two fold axes perpendicular to the crystallographic b axis ($\psi = 90°$ and $\phi = 55°$) rotated 55° from the *a* axis. The non-crystallographic two-fold symmetry and the crystallographic C2 symmetry support the likelihood that the protein is a tetramer with molecular 222 symmetry, but this is inconsistent with earlier assumptions that it BHMT is a hexamer (Lee *et al.*, 1992). This discrepancy requires closer scrutiny since our recombinant protein contains a polyhistidine tag at the N-terminus that could interfere with proper oligomerization, but the prediction that the protein is a hexamer is based on studies on the rat liver enzyme, and so they may not hold in the case of the human enzyme.
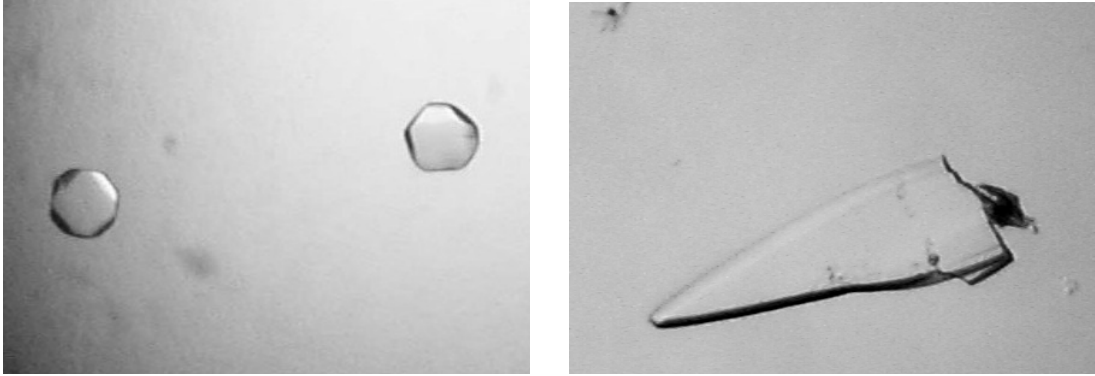
Figure 7. Crystals of human BHMT using PEG 20,000 and ethylene glycol.

Panel A, Crystals of human BHMT using PEG 20,000 and ethylene glycol (protein concentration, 7 mg/ml). Panel B, crystal of different morphology obtained using a lower protein concentration (4 mg/ml), but otherwise same conditions.
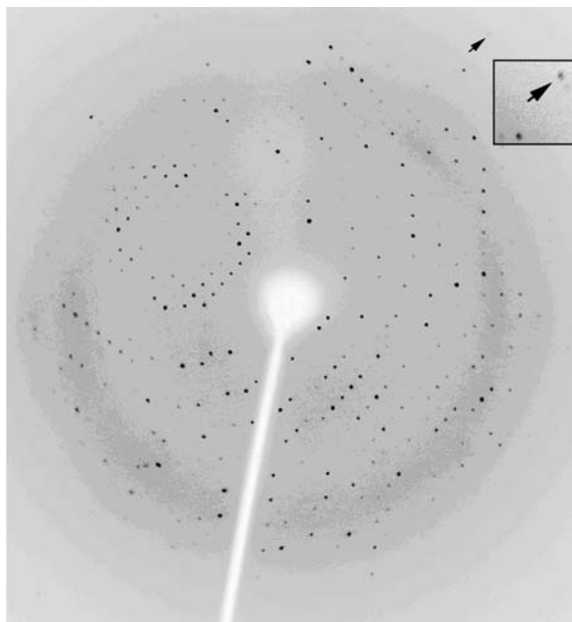
Figure 8: A typical diffraction pattern of a crystal of recombinant human BHMT.

Shown is an image as represented by the HKL 2000 data processing package. Data collection settings were: crystal to detector distance, 200 mm; oscillation range, 1.0º; exposure time, 10 min; total frames collected, 180. The arrow at the top right of the figure points to a reflection at 2.9 Å resolution. The inset is a magnified window around the arrow.
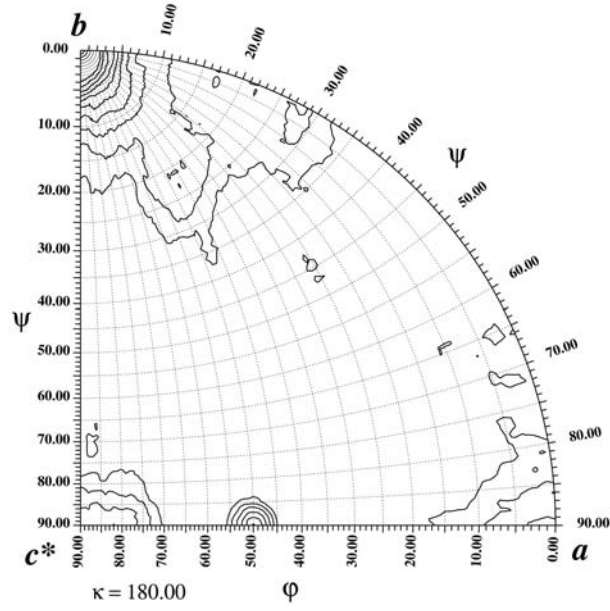
Figure 9. Self- rotation function of a crystal BHMT calculated using GLRF (Tong & Rossmann, 1997).

Table 4: Data Collection Statistics

Wavelength (Å)	1.5418	
Space group	C2	
Unit cell (Å)	$a = 109.190$, $b = 91.319$, $c = 88.661$	
	$\beta = 122.044°$	
	$V = 884{,}041$ Å^3	
Number of observed reflections (unique)	51,853	(14,797)
Complete resolution range (Å) (outer shell)	30.0 – 3.0	(3.05 – 3.0)
Completeness overall (%) (outer shell)	94.2	(70.1)
Overall I/σ (I) > 3 (%) (outer shell)	75.3	(60.0)
Overall R$_{merge}$† (%) (outer shell)	8.5	(36.1)

$$† \ R_{merge} = \sum_h \sum_i I(h)_i - <I(h)> / \sum_h \sum_i I(h)_i$$

References

Finkelstein, J. D. & Martin, J. J. (1984a). Biochem Biophys Res Commun 118, 14-19.

Finkelstein, J. D. & Martin, J. J. (1984b). J Biol Chem 259, 9508-9513.

Finkelstein, J. D., Martin, J. J., Harris, B. J. & Kyle, W. E. (1982). Arch Biochem Biophys 218, 169-173.

Finkelstein, J. D., Martin, J. J., Kyle, W. E. & Harris, B. J. (1978). Arch Biochem Biophys 191, 153-160.

Garrow, T. A. (1996). J Biol Chem 271, 22831-22838.

Jancarik, J. & Kim, S. H. (1991). J. Appl. Cryst. 24, 409-411.

Lee, K. H., Cava, M., Amiri, P., Ottoboni, T. & Lindquist, R. N. (1992). Arch Biochem Biophys 292, 77-86.

Matthews, B. W. (1968). J. Mol. Biol. 33, 491-497.

McPherson, A. (1976) J. Biol. Chem. 251, 6300-6303.

Millian, N. S. & Garrow, T. A. (1998). Arch Biochem Biophys 356, 93-98.

Mudd, S. H., Levy, H. L. & Skovby, F. (1995). (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Vale, D., ed.), pp. 1279-1327. McGraw-Hill, New York.

Otwinowski, Z. & Minor, W. (1997). Methods Enzymol 276, 307-326.

Tong, L. & Rossmann, M. G. (1997). Methods Enzymol 276, 594-611.

CHAPTER III

BETAINE: HOMOCYSTEINE S-METHYLTRANSFERASE AND

CYSTATHIONINE γ-LYASE: CRYSTALLIZATION TRIALS

Betaine: homocysteine S-methyltransferase. BHMT was previously described in Chapters I and II.

Cystathionine γ-lyase. The metabolic interconversion of sulfur-containing amino acids cysteine and methionine mainly consists of two pathways, transsulfuration and reverse transsulfuration. The transsulfuration pathway is mainly found in bacteria, fungi and plants. This pathway involves the transformation of cysteine into homocysteine via *L*-cystathionine with the participation of two enzymes, cystathionine γ-synthase (CGS) and cystathionine β-lyase (CBL). In contrast, the reverse transsulfuration pathway, which involves the conversion of homocysteine to cystathionine by cystathionine β-synthase (CBS) and cystathionine to cysteine by cystathionine γ-lyase (CGL), is found only in fungi and mammals (1, 2). The only exception to this rule is that the reverse transsulfuration enzyme, CGL was found in filamentous bacteria, *Actinomycetaceae* (1). CBL, CGS and CBS, all are very extensively studied enzymes as far as kinetic and high resolution structural studies are concerned (3, 4, Chapter IV and the references therein). CGL has been well characterized in rat and *Streptomyces phaeochromogenes* (5), but studies of the human CGL has been marred by solubility and low yield problems (6, 7).

114

Cystathionine γ-lyase, also called cystathionase (EC 4.4.1.1), the second enzyme in the reverse transsulfuration pathway catalyzes the irreversible conversion of *L*-cystathionine to *L*-cysteine, α-ketobutyrate and ammonia. The product cysteine, other than being incorporated into protein, either leads to the formation of the tripeptide, glutathione or is catabolized via cysteine-sulfinate dependent pathways. In humans, the enzyme is associated with a broad spectrum of diseases. Cystathionuria is an autosomal recessive disorder due to cystathionine γ-lyase deficiency, the clinical presentation ranging from asymptomatic to mental retardation (8). Cystinosis is another disease that has been linked with CGL (9). The rate of hepatic glutathione synthesis is limited by the availability of cysteine, which in turn makes cystathionase the primary rate limiting factor in the synthesis of glutathione (10; 11). Thus, in humans a low hepatic cystathionase activity is associated with various pathological and toxicological conditions in which low levels of glutathione are encountered. The most researched disease in association with glutathione deficiency is AIDS. Martin *et al*. reported that hepatic CGL activity was substantially reduced in the liver samples of AIDS patients and therefore suggested that low CGL activity could be associated with the systemic deficiency of glutathione and cysteine seen in HIV infected patients (12). An interesting observation that rats are more resistant to the development of alcoholic liver injury as compared to humans was attributed to the markedly reduced human hepatic CGL activity relative to rat's CGL activity. Humans with a defective CGL had low glutathione levels and therefore were more susceptible to alcohol liver damage (13, 14, 15). The enzyme has been linked to homocystinuria and atherosclerosis, but less frequently than CBS. But the enzyme's role in maintaining glutathione levels could be indirectly related to

glutathione's ability to prevent the oxidation of circulating fats in the bloodstream, including cholesterol, retarding the process of plaque formation in the arteries - the underlying cause for most heart disease and stroke (6). CGL has also been linked to malignancy of lymphoid cells. Normal lymphoid cells are capable of converting homocysteine to cysteine by cystathionase action whereas malignant cells have markedly reduced levels of CGL and do not grow in media devoid of L-cysteine (6, 16, 17). Thus CGL activity can serve as an important marker for cellular differentiation.

The gene for human CGL (1204 bp) was cloned and sequenced by Lu *et al*. (18) with the first expression system for recombinant human CGL developed in 1999 (7). Two isoforms of human CGL exist that are splice variants. A shorter isoform had a deletion of 132 bp as compared to a longer form. Strong expression of both mRNA isoforms was detected only after 19th gestational week onwards, the longer form being dominant throughout the development (19).

CGL is a pyridoxal 5'-phosphate (PLP)-dependent enzyme and shares extensive sequence homology to CBL and CGS, which are all members of the PLP γ-family (20). By virtue of its homology, CGL has also been predicted to be a homotetramer composed of 45 kDa subunits, having an aspartate aminotransferase fold and carrying one PLP cofactor per monomer covalently bound via a Schiff base to an active site lysine (21). Even though human CGL is structurally very homologous to CBL and CGS, there are several reasons for a detailed structural study of CGL. CBL and CGS have been structurally studied as antibacterial targets. Since the active sites of CGS and CBL are very homologous to that of human CGL, it is a challenge to design antibacterials that can distinguish between the active sites and selectively inhibit only bacterial CGS or CBL.

116

Aminoethoxyvinylglycine (AVG), an inhibitor known for its selectivity for CBL was recently shown to have a comparable effect on human CGL thus rendering it unusable (7). CBL and CGL from yeast and rat liver have shown dual specificities, *i.e.*, they are able to cleave both C-γ-S and C-β-S bonds of *L*-cystathionine. In contrast, human CGL is very specific for C-γ-S bond cleavage indicating that there must be some subtle, but specific, structural constraints at the active site responsible for its regiospecificity. Propargylglycine and trifluoroalanine are much stronger inhibitors of CGL than CBL, again supporting the fact that there must be some subtle differences in the active site residues or structural orientation of some conserved residues that lead to the observed effect (7). CGS can be differentiated from CGL on the basis of their first substrate, but an understanding of the structure of CGL at an atomic level could lead to the development of more potent and specific inhibitors of CGS and CBL.

Considering the importance of the enzymes betaine: homocysteine methyltransferase and cystathionine γ-lyase in a broad range of pathological conditions, the need for a detailed structural study of these enzymes for structure-based drug design can be greatly appreciated. The goal of this chapter is to describe in detail the attempts to crystallize BHMT and CGL and the problems associated with it.

Theory of crystallization of macromolecules. Crystallization is a process by which a metastable supersaturated solution experiences a thermodynamic driving force to reestablish a state of maximum probability for the system, its energy minimum by reduction of solute concentration (22). At saturated conditions, the solution is at equilibrium and there is transfer of neither the solute to the soluble state or from the solution to the solid state. But in the supersaturated solution there is a force that promotes

the exclusion of molecules from the solution to accumulate in the solid state and this is what drives crystallization. A certain amount of activation energy is required for the formation of crystals. The higher the degree of the supersaturation of protein solution, the lower is the activation energy required to initiate the formation of the solid state. The three steps in the crystallization process are i) nucleation, ii) growth of crystals and iii) termination of growth. An understanding of the phase diagram pertaining to crystallization is necessary to explain the three stages of crystallization (Figure 10) (23). A 'labile' region in the phase diagram corresponds to the state of extreme supersaturation. This stage corresponds to a state that has high energy and wants to reach its energy minimum. Therefore there is a high probability that nucleation will occur spontaneously at this stage. Nucleation is the process by which molecules in the supersaturated solution, collide with each other, overcome the activation barrier, establish weak interactions and form a thermodynamically stable aggregate with a repeating lattice. The next step, crystal growth, generally starts at solute concentrations sufficient for nucleation to occur, and continues at concentrations beneath the nucleation threshold. This is termed the "metastable" region of supersaturation. In this region, no new nuclei are formed but the stable nuclei continue to grow. The rate of growth is determined by a combination of the nature of the growing crystal surface and the diffusion rate (24, 25). Cessation of growth occurs due to various reasons: the attainment of the equilibrium state, the deprivation of crystallizing solute for further growth, lattice strain which effectively prevents addition of molecules to the surface once a certain critical volume is reached (25), accumulation of defects by the crystal or by the buildup of impurities by the growing crystal surfaces.
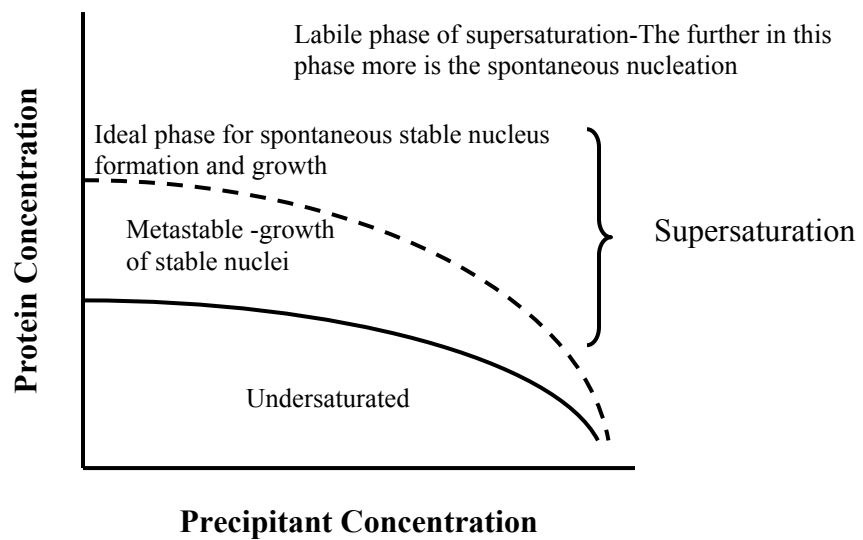
Figure 10: Phase separation diagram for crystallization of

macromolecules (23 & 26)

Methods for creating supersaturation. To make a supersaturated solution of a macromolecule, it is necessary to start with an undersaturated solution. Then gradually the chemical and physical properties of the system are modified so that it enters the supersaturated region without a sudden change in the energy of the system. There are several ways of creating supersaturated solutions of macromolecules: i) concentration of macromolecule, ii) addition of precipitating agents like salts, polymeric precipitants, organic solvents and non-volatile alcohols (27, 28, 29, 30, 31), iii) alteration of temperature, iv) alteration of pH, v) addition of ligand (32, 33), vi) addition of additives (34, 35, 36) and vii) dialysis against deionized water.

Methods of crystallizations. There are several devices and methods available for crystallization purposes. The common methods are: i) bulk crystallization (37, 38), ii) batch crystallization (39), iii) vapor diffusion (40, 41), iv) microdialysis (42) and v) seeding (30). Figure 11 shows some of the most common methods of crystallization.

Strategies for screening crystallization conditions. Various strategies have been set forth to initially screen conditions for crystallization. These vary from somewhat rational approaches (screening at the isoelectric point, pI) to highly regimented approaches (successive grid screening) to analytical approaches (incomplete factorials, solubility assays, perturbation) to randomized approaches (sparse matrices) (43). The principle of the sparse matrix approach, which is the most common method, is to provide by randomization of various conditions, a broad sampling space for obtaining initial crystals. The 'hit condition' is then used as a starting point for crystallization, which is then improved upon.

Initial condition   **Final condition**

cover
slip
sealed
with

$$[ppt]_{drop} = \frac{[ppt]_{reservoir}}{2}$$ $$[ppt]_{drop} = [ppt]_{reservoir}$$
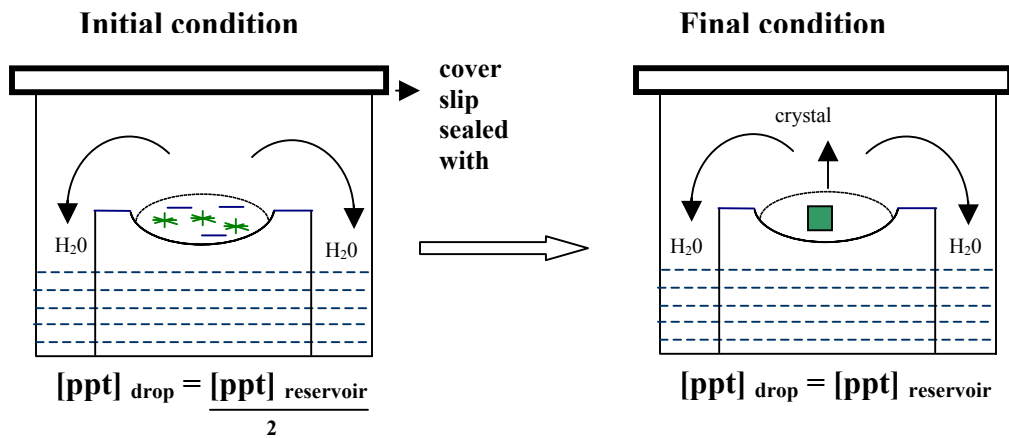
A) VAPOR DIFFUSION: SITTING DROP TECHNIQUE

$$[ppt]_{drop} = \frac{[ppt]_{reservoir}}{2}$$ $$[ppt]_{drop} = [ppt]_{reservoir}$$
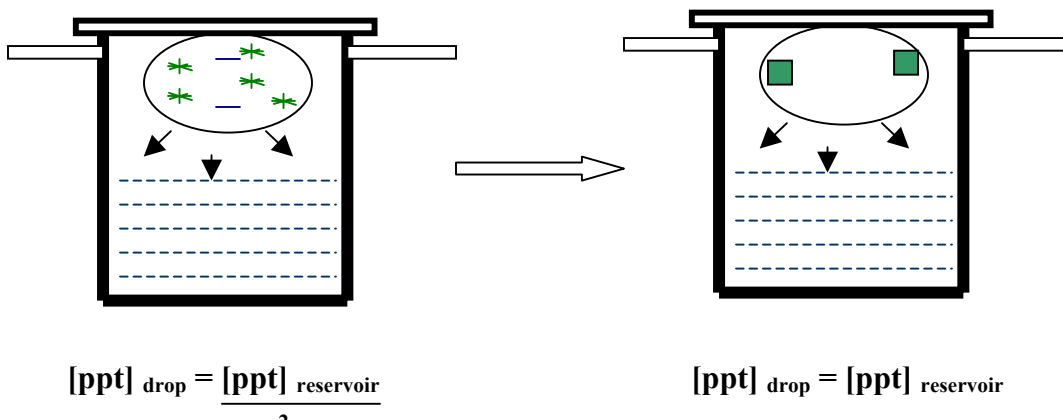
B) VAPOR DIFFUSION: HANGING DROP TECHNIQUE
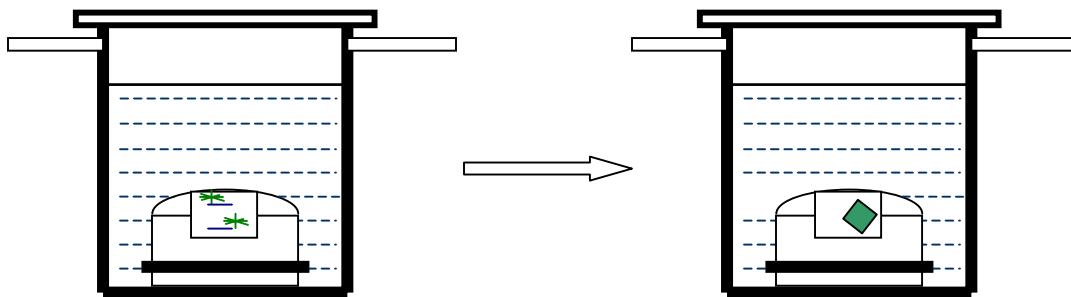
C) MICRODIALYSIS CRYSTALLIZATION

Figure 11: Common methods of crystallization used in the crystallization of BHMT and CGL (26)

MATERIALS AND METHODS:

Cloning, expression and isolation of BHMT. As described in chapter I (Figure 13).

Crystallization attempts. As described in chapter II, the condition that produced a crystal that diffracted up to 2.9 Å consisted of 2 μl of 4 mg/ml protein, 3% PEG 20,000 and 3% ethylene glycol crystallized with 2 μl of the well solution containing only 8% ethylene glycol in 0.1 M Hepes buffer, pH 7.5. Attempts were made to reproduce the crystals in identical conditions for two different batches of isolation of BHMT. On failure of reproducibility of crystals, a whole new approach to produce BHMT crystals began.

A new batch of BHMT was isolated with the elution gradient of the anion exchange chromatography spread over 40 column volumes instead of 20. Also the purification included 5 mM BME (β-mercaptoethanol) as reducing agent. This resulted in three poorly resolved peaks that were collected separately and set up for crystallizations using the previously defined crystallization condition. The three peaks were analyzed by SDS-PAGE analysis and then sent for mass spectroscopy analysis.

 After seeing symptoms of proteolysis in SDS-PAGE analysis, a fresh batch of BHMT was isolated. The changes that were incorporated in the expression and isolation were that after induction for 4 hrs at room temperature, the cells instead of being frozen at -20°C, were immediately processed. Before sonicating, 3 tablets of EDTA-free protease inhibitor were added to the cells to inhibit proteolysis. The protein obtained after purification was checked for proteolysis by SDS-PAGE and then set up for crystallizations. The previously defined crystallization parameters were used, but three additional test conditions were include: one with 1 mM EDTA in the well solution, a

second with 5 mM BME in the well solution, and a third with both 1 mM EDTA and 5 mM BME.

Since 5 mM BME was found to be critical for crystallization, efforts were channeled to obtain a single large crystal for X-ray diffraction purposes. The crystals obtained were amidst a fair amount of precipitate. Therefore the approach taken to avoid the precipitation was to first make a stock solution of the protein drop mixed with the reservoir solution and additives. This stock solution was centrifuged at 9,000 x g for 1 min and then only used for crystallization.

A temperature change was introduced in which the crystallization plate was first incubated at 4°C for 2 days followed by incubation in the styrofoam box for 2 days. It was then transferred for further incubation in the 25°C incubator.

Another approach was to screen within a narrow pH range. The concentration of the precipitant ethylene glycol was changed in the range of 8 to 30% along one of the axes of the grid. A pH range of 7 to 8 was tested along the second axis (Figure 12).

Use of organics like dimethyl sulfoxide (DMSO) and dioxane was considered as the next step. The screening grid consisted of ethylene glycol in the concentration range of 8 to 30 % on one axis and DMSO in the range of 2 to 12% on the other axis. This set up was also repeated with dioxane replacing DMSO.

Crystallization in the presence of substrates, products and metal ions were also set up. One of the setups included one substrate and one product, homocysteine and dimethylglycine. For this purpose, freshly isolated BHMT was dialyzed against 1 mM dimethylglycine and 1 mM *DL*-homocysteine before the crystallization trials using similar conditions. In another setup, methionine was used in the concentration range of 2

to 10 mM in the protein drop. In the third setup, zinc chloride was used as additive in the concentration range of 0.1 to 2 mM.

Dialysis crystallization was also performed in the hope of getting diffraction quality crystals. In the method, a dialysis button contained a total volume of 20 μl that included 10 μl protein (4 mg/ml), 5 mM BME, 1 mM EDTA and 4% ethylene glycol. It was covered by a dialysis membrane and sealed tightly by an O-ring. This button was then introduced into a vial containing 10 ml of reservoir solution (8% ethylene glycol, 10 mM BME and 2 mM EDTA). The vial was closed tightly and then placed in the 25°C incubator. The button was checked intermittently for crystals under the microscope. At no signs of nucleation, the concentration of precipitant in the reservoir solution in the vial was increased in steps up to 30%.

As a final resort, microseeding and macroseeding of crystals were also tried. For microseeding, a drop containing tiny crystals (drop conditions-4 mg/ml protein, 3% PEG-20,000 and 3.4% ethylene glycol; well solution- 8% ethylene glycol) was pipetted, crushed by vortexing and then serially diluted in the well solution (1-50 dilutions). A crystallization plate was setup wherein the drops contained ethylene glycol in the concentration range of 3.5 to 7.5%. To these drops, the various dilutions of the seed stock (1 μl) was added and then allowed to equilibrate at 25°C. For macroseeding, single tiny crystal from the same old drop was transferred by a loop to each of the fresh protein drops containing ethylene glycol in the concentration range of 3.5 to 7.5%. This setup was also kept in the incubator at 25°C.
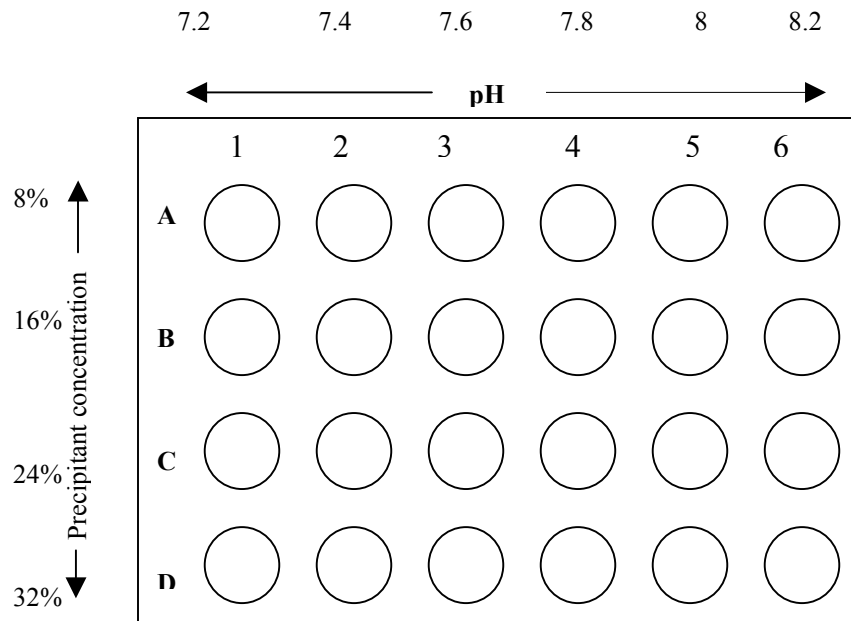
Figure 12: Example of a grid using the Limbro plate where the precipitant concentration and the pH are varied on each of the axes

High throughput crystallization screening at Hauptman-Woodward

Institute, New York. After all attempts failed to generate large crystals of BHMT, the protein was sent to the Hauptman-Woodward institute for high throughput crystallization trials. Two samples of BHMT were sent: BHMT (12.5 mg/ml) in 50 mM Tris, pH 8, 1 mM EDTA and 5 mM BME, and BHMT-G (15 mg/ml) in 50 mM Tris, pH 8, 1 mM EDTA, 5 mM BME and 15% glycerol. At this institute, they provide a service of screening the samples against 1536 crystallization conditions. All the crystallizations are performed by the microbatch method under oil. The drops are very small in volume, containing 0.2 μl of protein solution combined with 0.2 μl of cocktail solution underneath a layer of paraffin oil. The experimental plate is set up and stored at 25°C. Multiple photographic recordings of the plate are made: just prior to the addition of macromolecule, immediately after the addition of the macromolecule solution, 24 hours after setup, 72 hours after setup, 1 week after setup and 2 weeks after setup. All these images are compressed and sent to the investigator via a compact disc. MACROSCOPE software is provided that allows the investigator to view the results of the crystallization experiments. The results of the high throughput crystallization trials of BHMT and BHMT-G were reviewed using the software. The successful 'hits' were carefully studied and the conditions were replicated (In the jargon of high throughput screening, 'hits' are any observable crystal form including needles and salt crystals).

Cloning, expression and isolation of CGL. The method used was identical to that used for the preparation of expression constructs of BHMT and CBS described in Chapters I and IV. The pET28bNH construct was used to create the expression construct of CGL, pET28b-NHCgl, with a his-tag at the N-terminus (All the primers used for

cloning of CGL are listed in Table 5). All the expression parameters were also similar to that of BHMT and CBS. The purification strategies for this enzyme also consisted of metal chelate chromatography (Pharmacia HiTrap NTA charged with Ni^{2+}) in phosphate-NaCl buffer with imidazole gradients, followed by anion exchange chromatography (Pharmacia Hi Trap Q) in tris buffer (Figure 13). An additional purification step was introduced after the first crystallization screens. An amino-hexyl column was used for the final purification step with the buffer conditions similar to that used for the Q column. The purity of the protein was checked using SDS-PAGE analysis.

Enzymatic assay of CGL. A colorimetric ninhydrin based assay was identical to that used to assay CBS except that the pink colored complex between cysteine and ninhydrin was monitored at 560 nm (44).

Crystallization attempts. A broad range of crystallization conditions following the sparse-matrix sampling method (43) were examined by vapor diffusion trials (40) using 2 μl drops of protein (CGL in 50 mM Tris, pH 8) incubated with equal volumes of well solutions from Hampton Screens I and II. This was followed by a more focused screens using PEG, PEG/lithium chloride, ammonium sulfate and methyl-pentane-diol (MPD) kits of Hampton Research. After the additional purification step, all of the above screens were repeated with the protein containing 5 mm BME. The screens were performed thrice so that each screen condition was incubated at three different temperatures: 4°C, 15°C and 25°C. Additional screens were performed using different combinations of the precipitants and additives. Three screens using combination of PEGs and three different salts were prepared. PEGs of different molecular weights were used

Table 5: Oligonucleotides used for building the expression construct of CGL (pET28b-NHcgl)

Primer	Sequence
CGLforward	5'-TTCGCGGTTCAGCATGCAGG-3'
CGLreverse	5'-CTAGCTGTGAATTCCACTTGGAGG-3'
NHSapCGL-Forward	5'-CGCTCTTCTCACATGCAGGAAAAAGACG-3'
NHSapCGL-Reverse	5'-TGCTCTTCGTTAGCTGTGAATTCCACTTGG-3'

A) BHMT purification using metal chelate affinity chromatography
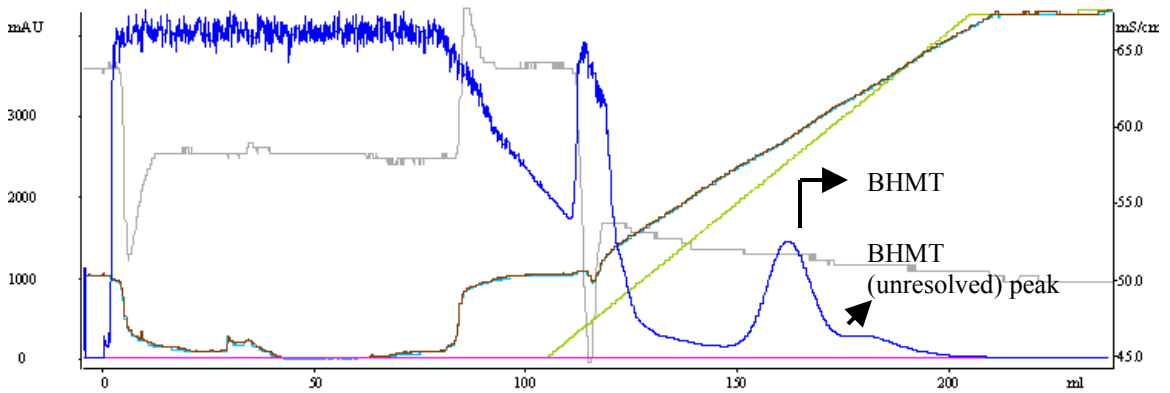
B) BHMT purification using anion exchange chromatography
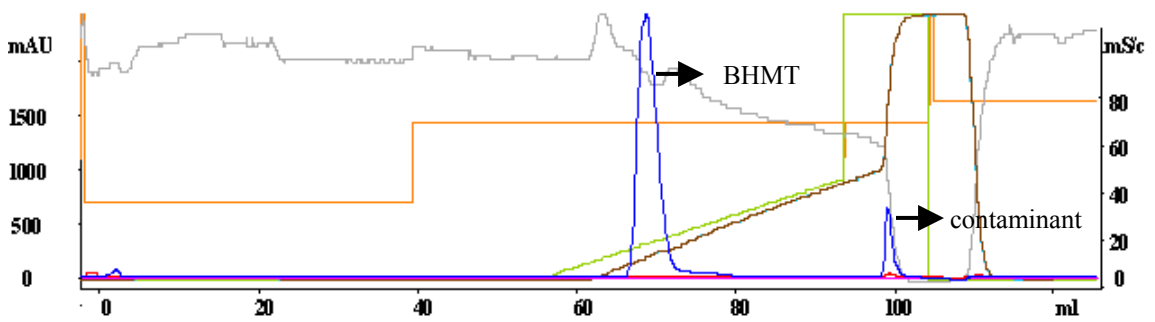
Figure 13: Chromatograms of betaine:homocysteine methyltransferase

absorbance at 280 nm

absorbance at 412 nm (PLP proteins)

A: BHMT eluted off the nickel column as closely associated peaks, both of which had similar activities.

B: The major protein peak of BHMT, after loading onto anion-exchange column was eluted using a linear gradient of 0.1-1.0M NaCl over 20 column volumes

C) CGL purification using metal chelate affinity chromatography
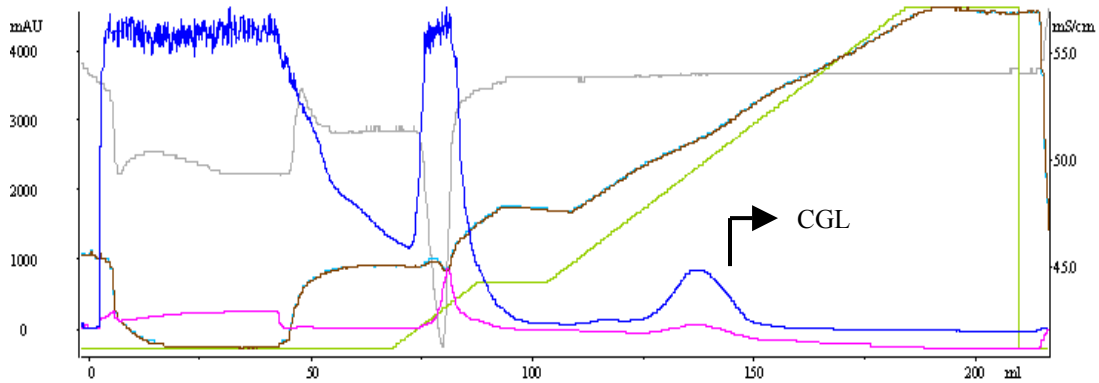
D) CGL purification using anion exchange chromatography
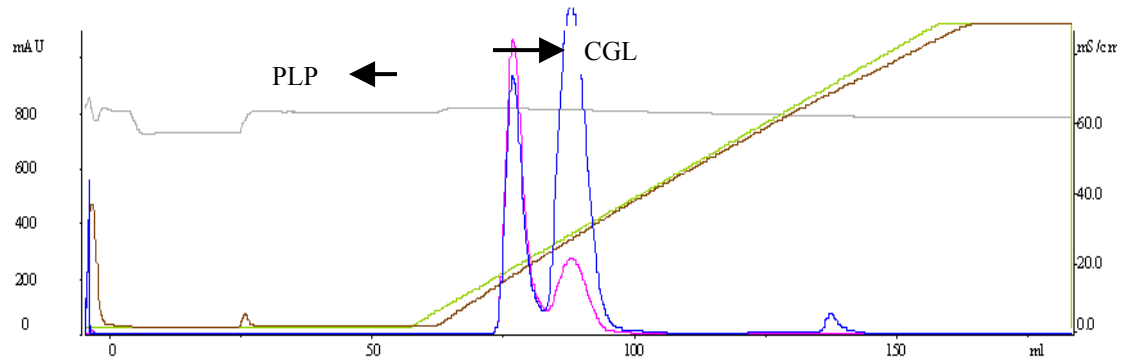
Figure 13(cont'd): Chromatograms of cystathionine γ-lyase purification.

——————   absorbance at 280 nm

——————   absorbance at 412 nm (PLP proteins)

C: CGL eluted off the nickel column as a single homogenous peak

D: The anion-exchange chromatogram of CGL shows that CGL eluted off

   as a single peak in close association with the PLP peak.

as precipitants in the concentration ranges of 5-30%. These were mixed with different

salts like lithium sulfate, sodium chloride and ammonium sulfate in the concentration

ranges of 0.2 to 3 M. Another screen that was manually designed was with the precipitant

ammonium sulfate in the concentration range of 0.5 to 1.4 M and MPD as an additive in

the concentration range of 1 to 6% MPD. A similar screen was prepared where MPD was

replaced with dioxane. These two organics were also used in combination with the

different PEGs in the concentrations mentioned above. Finally the protein was also

screened using the detergent screen kit from Hampton Research. For all the mentioned

screens, the concentration of CGL used was 20 mg/ml.

When all of the above attempts to crystallize CGL failed, it also was sent for a

high throughput crystallization screening (HTCS) at Hauptman-Woodward Institute. The

CGL that was sent to this institute was purified by using metal chelate affinity, anion-

exchange column and gel-filtration chromatography. One half of the CGL isolated was

dialyzed against 50 mM Tris, pH 8, 0.2 mM PLP and 5 mM BME and 1mM EDTA. It

was then concentrated to 16.5 mg/ml (CGL). The other half of CGL was dialyzed against

the same buffer with the inclusion of 15% glycerol. This protein was concentrated to 10

mg/ml (CGL-G).

The results of crystallization screen of CGL with the 1536 conditions were then

viewed using MACROSCOPE. The hits obtained by the HTCS screening were attempted

to be reproduced by both microbatch under oil and hanging drop method. The HTCS

conditions of crystallizations that showed crystals and were subsequently tried to

reproduce in house are listed in Table 8. The conditions were tried to be reproduced using

131

the same concentration of the protein as well as reducing the protein concentration to half. The oils used for microbatch were paraffin oil as well as Al's oil (mixture of paraffin as well as silicon oil). In the hanging drop vapor diffusion technique, the protein concentration as well as precipitant concentration was reduced to half of that used for microbatch crystallization. The crystallization setups were incubated at 25°C and then checked intermittently for crystals for 3 weeks.

RESULTS

    BHMT crystallization attempts. As mentioned in chapter II, all the attempts to crystallize BHMT, either resulted in crystals that did not diffract or crystals that are thin plates. The crystal on which the preliminary X-ray data was collected was also a thin plate. Therefore the main goal was to induce crystal growth in the third dimension. But the next two batches of BHMT failed to produce any crystals at all. When the protein was isolated freshly and eluted using a shallow gradient during anion exchange chromatography, symptoms of proteolysis were seen. The chromatogram, SDS-PAGE analysis and the mass spectroscopy results confirmed the possibility of proteolytic cleavage (Figure 14). The main peak in the chromatogram was mainly the full-length BHMT, the second peak showed an approximately equal percentage of both the full-length and the cleaved protein and the third peak was mostly cleaved BHMT. The mass spectroscopic results of the main and the third peak showed that the molecular weight of the main peak was 46,246 and that of the third peak was 42,925, clearly indicating proteolysis. All the individual peaks were set up for crystallization trials, but only the main peak crystallized (Figure 14). Thus rationalizing that proteolytic cleavage of the

protein was severely affecting crystallization reproducibility, changes in the isolation protocol were incorporated. The BHMT isolation protocol was modified to avoid any freezing of the cells and to include excess of protease inhibitors. Crystallization of the freshly isolated protein was then performed in identical conditions with the addition of reducing agent, BME. It was found that in the presence of BME, BHMT crystallized, albeit small in size. The small size of the crystals was then attributed to the precipitate from which the crystals grew. To obviate this problem, the drop of protein and the precipitant was premixed in a microfuge tube, centrifuged and then used for crystallization. Although this removed the precipitate that was thought to hinder the crystal growth, the ultimate result was showers of tiny crystals.

After finalizing the conditions for isolation of BHMT, various methods to obtain large diffraction quality crystals were attempted using BHMT. Variation of the temperature, which is known to affect crystallization by virtue of its affect on the solubility of protein, did not produce any appreciable change in the size of the crystals of BHMT. Change of pH within a narrow range also did not seem to affect the crystal size.

The discrete use of organics as additives in crystallization conditions of certain macromolecules has resulted in prevention of aggregation of proteins and better crystals in the case of nucleic acids. But in the case of BHMT, the use of organic solvents like DMSO and dioxane resulted in denaturation of the protein.

The inclusion of substrates, products, inhibitors or cofactors during protein purification usually helps in shielding the active sites and maintaining structural integrity.

Anion exchange chromatography of BHMT on a shallow gradient

Second peak

Main peak

Third peak
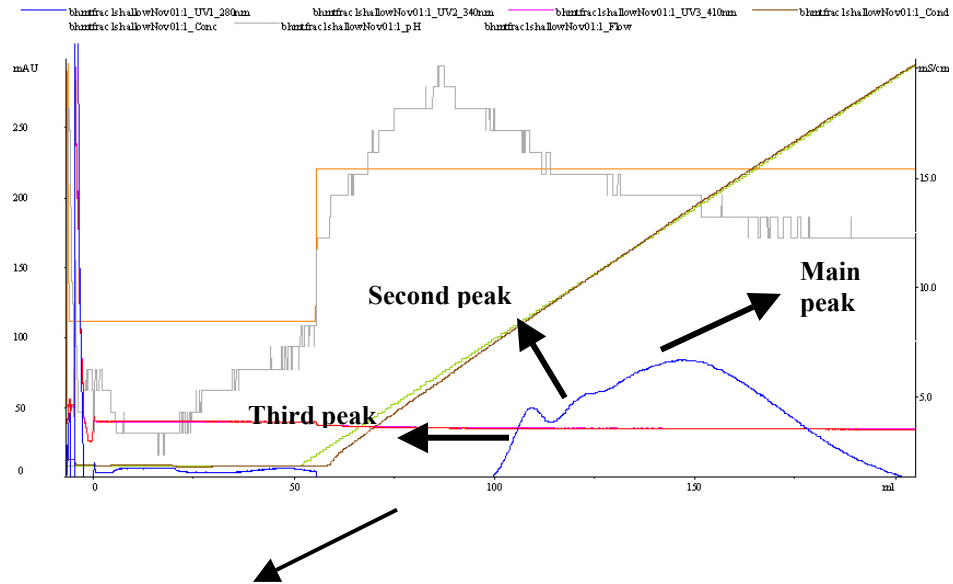
Second Main Third MW

43 kDa
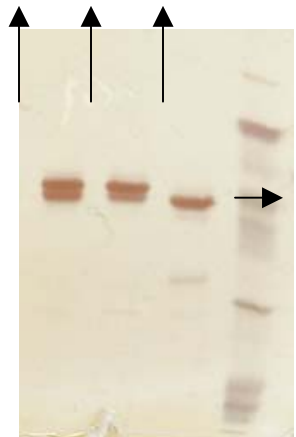
SDS-PAGE analysis of all the three peaks. Lane MW contains the molecular mass markers

Crystallization of the main peak showing precipitate along with the crystal

Figure 14: Evidence of proteolytic cleavage of BHMT given by the chromatogram, the SDS-PAGE analysis and the crystallization of the main peak of the protein

Addition of the substrate or the coenzyme induces conformational change of the protein and can yield a more stable and compact protein, provided that active metabolism of the substrate is not occurring. It can also have deleterious effects on protein crystallization. In the case of BHMT, addition of its product, methionine, in the protein drop resulted in complete dense precipitation of the protein. In contrast, the addition of its substrate *DL*-homocysteine and one of its products, N, N-dimethylglycine did produce crystals, but of insignificant size. Zinc, which plays a role in the catalytic mechanism of BHMT by activating homocysteine, was then tried as an additive in the protein drop. This also resulted in precipitation of the protein.

In the microdialysis crystallization trial, the solution in the dialysis button remained clear initially. But as the precipitant concentration in the reservoir solution in the vial was increased, the protein solution in the button precipitated heavily.

The microseeding technique did result in fewer crystals but not of large size. When single crystals of BHMT were macroseeded, no evidence of crystal growth was seen.

All the attempts of obtaining bigger crystals of BHMT revolved around the same precipitant and buffer. To search for completely new conditions of BHMT crystallization, the samples of BHMT were shipped for high throughput crystallization screening. After reviewing the crystallization results of both the samples (BHMT and BHMT-G), the successful conditions (Table 6 & 7) were then pursued. According to the high throughput result, the interesting observation was that BHMT-G, *i.e.*, BHMT dialyzed in glycerol had better 'hits'. The 'hits' were obtained in the pH range of 7-9, which was consistent to the pH range already established for BHMT. The additives were different from what was

previously tested on BHMT. The additives used included ammonium bromide, sodium bromide, sodium nitrate, potassium bromide etc. When BHMT crystallizations were set up using these 'hit' conditions by microbatch method under oil, heavy precipitation of the protein was observed. This was very consistent with the precipitant seen in the background of crystals in all the figures. Therefore it was anticipated that the crystals would eventually grow from the precipitate, but no crystals were obtained from any of the conditions. The same conditions were set up again by reducing the precipitant concentration to half or reducing the concentration of protein to half. All the attempts failed to reproduce any of the crystals seen in the high-throughput results.

Expression, isolation and crystallization attempts of CGL. The expression of CGL at 37°C produced insoluble protein. The expression temperature was lowered to room temperature and tested for expression and solubility of CGL after 6, 12 and 18 and 24 hrs. The SDS-PAGE gel showed that most of the protein is insoluble and remains in the pellet until 12 hrs. The crude extract of the 18 hr and the 24 hr time point contained some soluble protein. Therefore 24 hrs was decided as the optimum time of induction for CGL. After metal chelate affinity and anion exchange chromatography, approximately 9 mg of CGL was obtained from one liter of *E.coli* grown in LB media. The CGL was active using a ninhydrin-based assay and had a specific activity of 10 μmole/min/mg. This method of expression and purification of CGL proved to be very efficient as compared to the earlier attempt of isolating human recombinant CGL (7) wherein only 5 mg of protein was obtained after a long purification protocol, with a specific activity of 2.5 μmole/min/mg. Thus the pET28b-NH based CGL expression and isolation proved to

136

Table 6: The successful conditions obtained for BHMT without glycerol by the high throughput crystallization screening method
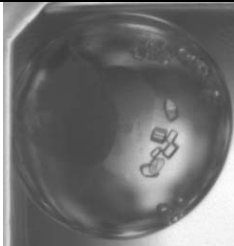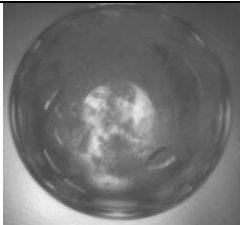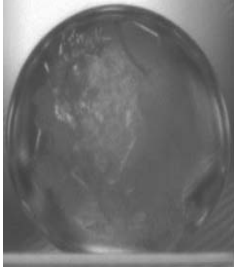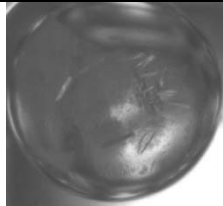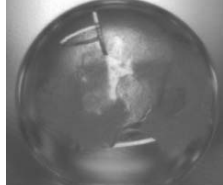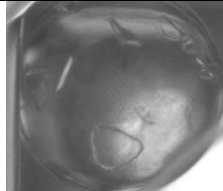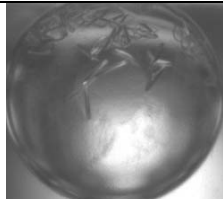
	Chemical Additive	Buffer	pH	PEG % (w/v)	Figure
1	0.1M Cobalt sulfate heptahydrate	0.1M Tris	8	20% 20000	
2	0.1 M Ammonium bromide	0.1 M CAPS	10	20% 4000	
3	0.1 M Potassium nitrate	0.1 M Tris	8	20% 4000	

Table 7: The successful conditions obtained for BHMT with glycerol by the high throughput crystallization screening method

	Chemical Additive	Buffer	pH	PEG % (w/v)	Figure
5	0.1 M Lithium chloride	0.1 M HEPES	7.5	20% 4000	
6	0.1 M Sodium nitrate	0.1 M Tris	8	40% 4000	
7	0.1 M Potassium bromide	0.1 M TAPS	9	20% 4000	
8	0.1 M Potassium chloride	0.1 M TAPS	9	40% 4000	

be not only a fast method but also efficient as far as yield of active protein was concerned.

After concentrating CGL to 20 mg/ml, a full battery of crystallization attempts was made using the sparse matrix screens, Crystal Screen I & II. This was followed by crystallization setups by screens using ammonium sulfate, PEG, PEG/LiCl and MPD. The protein seemed to precipitate under many screening conditions. Since no initial lead was obtained by these screens, it was decided to purify the protein further. After further purification on a PLP specific amino hexyl column and addition of 5 mM BME to the protein solution, crystallization using all the above screens was repeated. In all the screens, at all the three temperatures, either the protein precipitated, although to a lesser degree, or the drops remained clear. Also in all the combination screens of PEG with different salts, ammonium sulfate with MPD and ammonium sulfate with dioxane, the protein showed no signs of crystallization. Even the screen of CGL using various types of detergents failed to show any hints towards any viable direction. Thus, having no lead to be examined further, CGL was sent to Hauptmann-Woodward institute as two samples (CGL and CGL-G). The overview of the results obtained by this high throughput screen was that PEG-20,000 might be essential for crystallization of CGL because most of the positive conditions had PEG-20,000 as the precipitant. CGL showed needle-like crystals in the pH range of 7-10. It was surprising that even though CGL has aggregation problems, CGL-G (with glycerol), which tends to prevent aggregation, had no positive hits. All the positive hit conditions (Table 8) were attempted to be reproduced using the same batch of protein that was kept at 4°C. Most of the conditions using the microbatch method under oil remained clear for about 3 weeks and then precipitated. The hanging

drop method also resulted mostly in precipitation. Thus the high throughput results were not reproducible for CGL also.

DISCUSSION

Crystallization problems of BHMT. The discovery of crystallization conditions appropriate for specific macromolecules is a completely empirical process. The overall goal of crystallization is to obtain crystals of some kind, grow them in an acceptable form and grow them to a required size. Crystallization can be hampered by problems at three stages, initiation of nucleation, rate and nature of crystal growth and termination of crystal growth.

BHMT with its lack of growth in the third dimension and inability to grow to a bigger size could be experiencing problems in the phases of crystal growth and cessation of growth. Several reasons could be attributed to BHMT's lack of reproducibility and inability of crystal growth. The prime reason could be sample heterogeneity. During metal chelate chromatography, the protein eluted off the nickel column as two closely associated peaks (Chapter I). Also during elution with a shallow gradient off an anion exchange column, there were three unresolved peaks. The SDS-PAGE analysis and the mass spectroscopy analysis confirmed the speculation that BHMT was being cleaved by proteolysis. The evidence of proteolytic cleavage in recombinant human BHMT was first provided by Millan *et al*. where in the mass spectroscopy analysis and amino acid sequencing of recombinant preparations suggested three forms of BHMT, one was the full-length and the other two were proteolytically cleaved (46). An earlier study on

140

Table 8: The successful conditions obtained for CGL without glycerol by the high throughput crystallization screening method
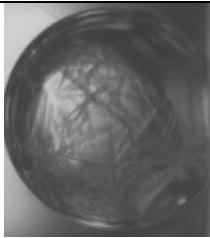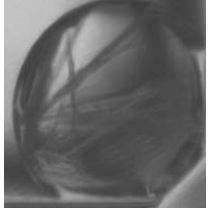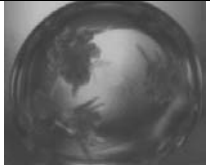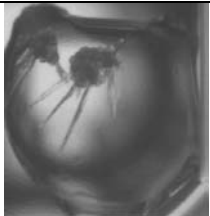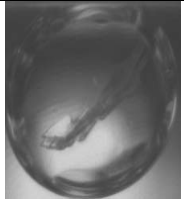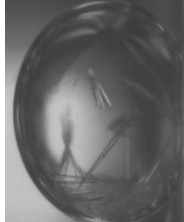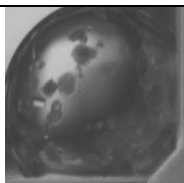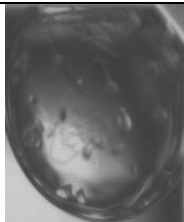
	Chemical Additive	Buffer	pH	PEG % (w/v)	Figure
1	0.1 M Lithium chloride	0.1 M TAPS	9	20% 20000	
2	0.1 M Lithium chloride	0.1 M CAPS	10	20% 20000	
3	0.1 M Potassium acetate	0.1 M HEPES	7.5	20% 20000	
4	0.1 M Potassium chloride	0.1 M HEPES	7.5	20% 20000	

Table 8: Continuation of the successful conditions obtained for CGL without glycerol by the high throughput crystallization screening method

	Chemical Additive	Buffer	pH	PEG % (w/v)	Figure
5	0.1 M Rubidium chloride	0.1 M TAPS	9	20% 20000	
6	0.1 M Potassium thiocyanate	0.1 M CAPS	10	20% 20000	
7	0.1 M Magnesium sulfate heptahydrate	0.1 M MOPS	7	20% 8000	
8	0.1 M Magnesium sulfate heptahydrate	0.1 M MOPS	7	20% 4000	

BHMT isolated from rat liver demonstrated by chromatofocusing, the presence of three active isoforms of BHMT with very closely related pI values (45). Thus taking this problem into consideration, the BHMT isolation procedure was modified to include a high dosage of protease inhibitor at every step of the purification. Even though the SDS-PAGE and IEF (isoelectric focusing) analysis showed a single band of protein, contamination with trace amounts of the isoforms cannot be ruled out. This is further accentuated by the fact that most of the crystals obtained are mixed with a fair amount of precipitate. This could be rationalized as one form of the enzyme precipitates out and the other form crystallizes. It seems logical that the proteolytically cleaved protein can get incorporated onto the face of an otherwise perfect protein crystal and can cause poisoning of the growth surface of crystal. The knicked protein could be unable to form the same lattice contacts with the growing crystal as the other newly added molecules. As growth slowly progresses, the population of the cleaved protein increases because a growing crystal selectively incorporates perfect over damaged molecules. This increases the probability of defective protein getting incorporated into the crystal, leading to crystal growth defects and premature termination of growth. The solution to this problem could be to prepare samples at 4°C to keep the protease activity low, increase the load of protease inhibitors or try to resolve the different isoforms of BHMT using isoelectric focusing chromatography.

Another interesting observation that could provide some insight into BHMT's crystallization problem is that the presence of reducing agent, BME, was critical to control the non-specific precipitation. Reducing agents have been one of the important variables in crystallization of those proteins that have sensitive cysteine residues that

need to be protected in order to maintain the activity and structural integrity. BHMT has three cysteine residues (Cys-299, Cys-300, third-unidentified) that are absolutely essential for zinc binding and mutation of these residues abolished zinc binding and therefore the activity of the enzyme (46). Besides these cysteine residues, there are six other cysteine residues in BHMT. If any of the cysteine residues are at the surface or a sensitive position, they could make BHMT very oxidation sensitive. Oxidation of these cysteine residues could alter the conformation of the protein thereby diminishing its prospects of crystallizability. This fact could also explain the failure of being able to reproduce the high throughput screening conditions. Even though BME was included in the protein during purification, the protein stock after HTCS was stored at 4°C for at least a month. This allows enough time for BME to oxidize. Thus when crystallizations were performed using the protein stock, it may have already undergone oxidative damage. Thus the oxidation problem can be avoided by freshly adding the volatile reducing agent during crystallization and/or carrying out the entire procedure of purification and crystallization under anaerobic chambers.

Another speculation can be made based on sequence similarity between BHMT and methionine synthase. BHMT has significant similarity with *E. coli* methionine synthase, but only to the N-terminal part of methionine synthase (residues 2 to 353), which is the homocysteine and zinc-binding site (47). Methionine synthase being a modular enzyme has been broken down into domains and then studied structurally. It has been very difficult to structurally study the enzyme as a whole. All the domains have been studied except the homocysteine binding domain to which BHMT is homologous. The fact that this domain hasn't been crystallized yet could suggest some potential

problems either in isolation or crystallizability. BHMT by virtue of its sequence similarity to methionine synthase could also have some conformational disposition that could pose crystallization problems.

Crystallization problems of CGL. A highly desired feature of crystallization is that the protein molecules exist as single oligomeric species. Earlier study on recombinant human CGL had demonstrated that about 90% of the heterologous gene product was insoluble, and renaturation experiments from purified inclusion bodies met with limited success (7). The his-tagged CGL isolated in the pET28b-NH vector also had solubility problems although some soluble protein could be attained after a prolonged induction period. Even though the SDS-PAGE analysis and gel filtration chromatography of a freshly isolated sample indicated a homogenous preparation of CGL, this observation cannot be extrapolated to the stability of a protein over time. The fact that the fresh protein could be crystallized by high throughput screening methods, but after a month of storage at 4°C it failed to crystallize could suggest that the protein must have undergone time dependent insults like deamidation of side chains or random oxidation events. CGL may be undergoing some conformational change over the time that could lead to aggregation. Aggregates have the property of increasing the heterogeneity of the macromolecular population, competing with the molecules for normal ordering in the crystal and thus creating a major hindrance to the nucleation process and crystal growth.

Aggregation is a difficult problem to tackle, but there are some approaches that can be utilized for success in crystallization. Detergents and surfactants cover the hydrophobic regions of the macromolecule's surface with their hydrophobic tails, while solubilizing them by virtue of their hydrophilic (ionic or non-ionic) head. Even though

CGL was screened using the detergent kit from Hampton Research, which was unsuccessful, there are more variables to be considered. The amount of detergent really varies according to the precipitant present in the protein, for instance crystals grown from PEG were most affected at higher detergent concentrations, in some cases exceeding the critical micelle concentration (48). The most informative technique for monitoring the formation, existence or alleviation of aggregation is light scattering. Protein Solutions, Inc. manufactures analyzers based on light scattering for the specific purpose of evaluating the extent of aggregation of macromolecules in solutions that could potentially become mother liquors. Therefore if one has knowledge about the protein's tendency towards aggregation, it is ideal to screen the protein against various precipitants, additives and concentration of detergents. This approach with appropriate controls can give a fairly good idea about which conditions are conducive to aggregation of proteins or which detergent/combination of detergents can maintain the protein as a homogenous species.

Another possible reason for CGL's instability could be its sensitivity towards oxidation. CGL could also have some sensitive surface cysteines, which can be a major problem in crystallization. These residues can be identified by performing a cursory homology modeling of the protein since it is very homologous to prokaryotic cystathionine β-lyase and cystathionine γ-synthase (Blast analysis). If such residues are identified, they could be either protected by incorporation of fresh reducing agents or mutated to some other residue (genetic surface engineering), *i.e.*, if they are not involved in any important catalytic or structural role.

REFERENCES

1. Nagasawa T, Kanzaki H, Yamada H. 1984. Cystathionine gamma-lyase of Streptomyces phaeochromogenes. The occurrence of cystathionine gamma-lyase in filamentous bacteria and its purification and characterization. *J Biol Chem* 25:10393-403.

2. Flavin M. 1971. Methods Enzymol 17B:416-417.

3. Clausen T, Huber R, Laber B, Pohlenz HD, Messerschmidt A. 1996. Crystal structure of the pyridoxal-5'-phosphate dependent cystathionine beta-lyase from Escherichia coli at 1.83 A. *J Mol Biol* 262:202-24.

4. Clausen T, Huber R, Prade L, Wahl MC, Messerschmidt A. 1998. Crystal structure of Escherichia coli cystathionine gamma-synthase at 1.5 A resolution. *EMBO J* 17:6827-38.

5. Kashiwamata S, Greenberg DM. 1970. Studies on cystathionine synthase of rat liver. Properties of the highly purified enzyme. *Biochim Biophys Acta* 212:488-500.

6. Uren JR, Ragin R, Chaykovsky M. 1978. Modulation of cysteine metabolism in mice--effects of propargylglycine and L-cyst(e)ine-degrading enzymes. *Biochem Pharmacol* 27:2807-14.

7. Steegborn C, Clausen T, Sondermann P, Jacob U, Worbs M, Marnikovic S, Huber R and Wahl M. 1999. Kinetics and inhibition of recombinant human cystathionine γ-lyase. *J Biol Chem* 274:12675-12684.

8. Laurent B Coignet J.1973. Cystathionuria and defective enzyme regulation. *Clin Chim Acta* 43:171-82.

9. Allsop J, Watts RW. 1975. Methionine adenosyltransferase, cystathionine beta-synthase and cystathionine gamma-lyase activity of rat liver subcellular particles, human blood cells and mixed white cells from rat bone marrow. *Clin Sci Mol Med Suppl* 48:509-13.

10. Tateishi N, Higashi T, Shinya S, Naruse A, Sakamoto Y. 1974. Studies on the regulation of glutathione level in rat liver. *J Biochem* 75:93-103.

11. Reed DJ, Orrenius S. 1977. The role of methionine in glutathione biosynthesis by isolated hepatocytes. *Biochem Biophys Res Commun* 7:1257-64.

12. Martin JA, Sastre J, de la Asuncion JG, Pallardo FV, Vina J. 2001. Hepatic gamma-cystathionase deficiency in patients with AIDS. *JAMA* 285:1444-5.

13. Videla LA, Iturriaga H, Pino ME, Bunout D, Valenzuela A, Ugarte G. 1984. Content of hepatic reduced glutathione in chronic alcoholic patients: influence of the length of abstinence and liver necrosis. *Clin Sci (Lond)* 66:283-90.

14. Poulsen HE, Ranek L, Andreasen PB. 1981. The hepatic glutathione content in liver diseases. *Scand J Clin Lab Invest* 4:573-6.

15. Look MP, Riezler R, Reichel C, Brensing KA, Rockstroh JK, Stabler SP, Spengler U, Berthold HK, Sauerbruch T. 2000. Is the increase in serum cystathionine levels in patients with liver cirrhosis a consequence of impaired homocysteine transsulfuration at the level of gamma-cystathionase? *Scand J Gastroenterol* 35:866-72.

16. Sturman JA, Beratis NG, Guarini L, Gaull GE. 1980. Transsulfuration by human long term lymphoid lines. Normal and cystathionase-deficient cells. *J Biol Chem* 255:4763-5.

17. Glode LM, Epstein A, Smith CG. 1981. Reduced gamma-cystathionase protein content in human malignant leukemia cell lines as measured by immunoassay with monoclonal antibody. *Cancer Res* 41:2249-54.

18. Lu Y, O'Dowd BF, Orrego H, Israel Y. 1992. Cloning and nucleotide sequence of human liver cDNA encoding for cystathionine gamma-lyase. *Biochem Biophys Res Commun* 189:749-58.

19. Levonen AL, Lapatto R, Saksela M, Raivio KO. 2000. Human cystathionine gamma-lyase: developmental and in vitro expression of two isoforms. *Biochem J* 347:291-5.

20. Alexander FW, Sandmeier E, Mehta PK, Christen P. 1994. Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. Regio-specific alpha, beta and gamma families. *Eur J Biochem* 219:953-60.

21. Yamagata S, D'Andrea RJ, Fujisaki S, Isaji M, Nakamura K. 1993. Cloning and bacterial expression of the CYS3 gene encoding cystathionine gamma-lyase of Saccharomyces cerevisiae and the physicochemical and enzymatic properties of the protein. *J Bacteriol* 175:4800-8.

22. Weber PC. 1991. Physical principles of protein crystallization. *Adv Protein Chem* 41:1-36.

23. Miers HA and Isaac F. 1907. The spontaneous crysallization of binary mixtures: Experiments on salol and betol. *Proc Roy Soc Lond A* 79:332-338.

24. Boistelle R and Astier J P. 1988. Crystallization mechanisms in solution. *J of Cryst Growth* 90:14-30.

25. Feher G and Kam Z. 1985. Nucleation and growth of protein crystals: general principles and assays. *Meth Enz* 114:77-111.

26. McPherson A.1999. *Crystallization of Biological macromolecules*. Cold Spring Harbor Laboratory Press.

27. McPherson 1991. A. A brief history of protein crystal growth. *J Cryst Growth* 110:1-10.

28. Polson A, Potgieter GM, Largier JF, Mears GEF, Joubert FJ. 1964. The fractionation of protein mixtures by linear polymers of high molecular weight. *Biochim Biophys Acta* 82:463-475.

29. McPherson A. 1976. Crystallization of proteins from polyethylene glycol. *J Biol Chem* 251:6300.

30. McPherson A and Schlichta P. 1988. The use of heterogeneous and epitaxial nucleants to promote the growth of protein crystals. *J Cryst Growth* 90:47-50.

31. Lietzke SE, Barnes CL and Kundrot CE. 1995. Crystallization and structure determination of RNA. *Curr Opin Sruct Biol* 5:645.

32. Liljas A and Rossman MG. 1974. X-ray studies of protein interactions. *Anny Rev Biochem* 43:475

33. McPherson A. 1987. Interactions of biological macromolecules visualized by X-ray crystallography. *Crystall Rev* 1:91-95.

34. Trakhanov S, Kreimer DI, Parkin S, Ferro-Luzzi G, Rupp B. 1998. Calcium-induced crystallization of proteins. II. Crystallization of the *Salmonella typhimurium* histidine-binding protein in complex with L-histidine, L-arginine, or L-lysine. *Protein Sci* 7:600-10.

35. Abel JJ. Ceiling EMK. Rouiller CA. Bell FK. Wintersteiner O. 1927. Crystalline insulin. *J Pharmacol Exp Ther* 31:65-70.

36. Harding MM, Hodgkin DC, Kennedy AF. O'Connor A and Weitzmann PDJ. 1966. The crystal structure of inslulin. II. An investigation of rhombohedral zinc insulin crystals and a report of other crystalline forms. *J Mol Biol*. 16:212.

37. Northrop JH, Kunitz M and Herriott RM. 1948. *Crystalline Enzymes*. Columbia University Press, New York, New York.

38. Summer JB and Sommers GF. 1943. *The Enzymes*, 1st edition. Academic Press, New York.

39. http://www.hamptonresearch.com/hrproducts/3419.html

40. McPherson A. 1976. The growth and preliminary investigation of protein and nucleic acid crystals for X-ray diffraction techniques. Methods Biochem Anal 23:249-57.

41. McPherson A. 1982. *The Preparation and Analysis of Protein Crystals*. Wiley, New York.

42. Carter CW, Jr. Baldwin ET, and Frick L. 1988. Statistical design of experiments for protein crystal growth and the use of a precrystallization assay. *J of Cryst. Growth* 90:60-73.

43. Jancarik J and Kim S-H.J. 1991. Sparse matrix sampling: a screening method for crystallization of proteins. *Appl. Cryst* 24:409-411.

44. Gaitonde MK. 1967. A spectrophotometric method for the direct determination of cysteine in the presence of other naturally occurring amino acids. *Biochem J* 104:627-33.

45. Lee KH, Cava M, Amiri P, Ottoboni T, Lindquist RN. 1992. Betaine:Homocysteine Methyltransferase from rat liver: Purification and inhibition by a boronic acid substrate analog. *Arch Biochem Biophys* 292:77-86.

46. Millian NS & Garrow TA.(1998). Human Betaine-Homocysteine Methyltransferase is a Zinc Metalloenzyme. *Arch Biochem Biophys* 356:93-98.

47. Garrow TA. 1996. Purification, Kinetic Properties, and cDNA Cloning of Mamalian Betaine-Homocysteine Methyltransferase. *J Biol Chem* 271:22831-22838.

48. Shick MJ, Ed. 1966. *Nonionic Surfactants*. Marcel Dekker, New York.

CHAPTER IV

CYSTATHIONINE β - SYNTHASE: FROM INSOLUBILITY TO SOLUBILITY

INTRODUCTION

Cystathionine β-synthase (CBS) catalyzes the first irreversible step of the transsulfuration reaction of homocysteine metabolism in eukaryotes. CBS, a hepatic enzyme (L-serine hydro-lyase, EC 4.2.1.22), conjugates *L*-serine and homocysteine to form cystathionine in the presence of coenzyme pyridoxal 5'- phosphate (PLP), an allosteric activator S-adenosylmethionine (SAMe) and heme whose function is not clear.

Homocystinuria caused by deficiency of cystathionine β-synthase is a well-known, autosomal recessively inherited inborn error of sulfur amino acid metabolism. The prominent clinical manifestations of this disorder include dislocated optic lenses, thrombotic vascular disease, and mental retardation. The major biochemical changes in affected patients are elevated plasma and urinary concentration of methionine and homocysteine and a much reduced plasma cysteine (1). The human CBS gene has been localized to chromosome 21 and this locus is associated with the Down syndrome phenotype. Overexpression of CBS may be involved in some aspects of the generation of this syndrome (2).

CBS deficiency as the cause of homocystinuria, which, in turn, is strongly related to premature vascular disease has been the subject of numerous studies. Epidemiology studies have shown that one-third of patients with premature arteriosclerosis carry a

heterozygous deficiency of the gene encoding CBS (3). According to an update by Kraus, an analysis of 205 homocystinuric alleles has been performed with 64 mutations found (4). Animal models for mild and severe hyperhomocysteinemia demonstrated that homozygous mutants completely lacking CBS suffered from severe growth retardation and a majority of the mice died within 5 weeks after birth (5).

Since the enzyme requires PLP as coenzyme, a 43% decrease in CBS activity has been observed in rats deficient in vitamin B_6 (6). Further evidence for the role of vitamin B_6 was given by Smolin *et al*. by showing that the bound homocysteine and cysteine and plasma free homocysteine concentrations in rats deficient in vitamin B_6 were in the similar concentration range as those found in patients with CBS deficiency (7). PLP can be given in pharmacological doses as a treatment for CBS deficiency. Two clinical forms can be distinguished on the basis of patient's responsiveness to the treatment with the cofactor, PLP responsive and PLP non-responsive patients (8, 9, 10, 11).

Characteristics of CBS. The CBS gene is 23-30 kbp long and encodes a subunit of 63 kDa (12). It assembles as a homotetramer and probably larger oligomers. It contains one heme and one PLP per each subunit. It also binds to S-adenosylmethionine with a 1:1 stoichiometry (12).

CBS is an unusual PLP enzyme in that it is associated with heme. The observation that the nuclei acid sequence of rat CBS was identical to that of the rat hemoprotein H-450 led to the suspicion that the rat CBS had heme (13). Characterization of recombinant human CBS showed a Soret absorbance band (λ_{max} = 428 nm), which shifted to 448 nm after reduction with sodium dithionite, revealing the presence of heme in the enzyme (14). The speculations that heme plays a role in the catalytic mechanism of the enzyme

154

and is absolutely essential for the proper functioning of the CBS enzyme were disproved by the sequencing and expression of yeast CBS, which contained only PLP and no heme. Even though the deduced sequences of human rat and *Saccharomyces cereveisiae* CBS are similar, and there is functional conservation of human and yeast CBS, the absence of heme in the functional yeast enzyme suggests that heme does not play an essential catalytic role in the rat and human enzymes (15). The presence of heme in CBS was also supported by a study wherein the supplementation of cells with δ-aminolevulinate (precursor of heme) during expression of recombinant human CBS increased the yield, heme saturation of CBS and specific activity of CBS (16). There are controversial results regarding the number of heme bound to a fully saturated enzyme. According to Kery *et al*. the enzyme binds 1 mol of heme and 1 mol of PLP per mol of subunit, while Taoka *et al*. found one equivalent of heme to be bound per dimer (16, 17). Several functions have been assigned to the heme moiety: a) heme is functionally incorporated into CBS during protein folding and it is a prerequisite for PLP binding (18) which explains why mutations affecting heme binding influences the responsiveness and unresponsiveness of certain patients towards pyridoxine therapy, b) heme cofactor is involved in the redox regulation of the enzyme with the ferrous CBS being approximately 2-fold less active than the ferric enzyme (17), c) heme, by displacing its cysteine ligand and directly coordinating with homocysteine could activate homocysteine to generate a thiolate for nucleophilic attack on the aminoacrylate species (19). The role of heme being involved in catalysis by activating homocysteine was disproved by the polarized absorption microspectrophotometry of the recombinant CBS crystal. In this study, when heme-free enzyme crystals were exposed to increasing concentrations of natural substrate *L*-serine,

155

the key catalytic intermediate α-aminoacrylate was formed. Finally on exposure to homocysteine, the α-aminoacrylate absorption band disappeared with the concomitant appearance of absorption band of internal aldimine. This demonstrated that heme is not directly involved in the catalytic mechanism of CBS (20).

Based on ligand binding studies and homologous searches with protein sequences in the database, enzymatic functions were assigned to specific regions of CBS. Residues 68-209 are important for PLP binding, residues 241-341 for heme binding and residues 421-469 for SAMe binding (21). The C-terminal domain called the 'CBS domain' consists of residues (414-551) that are not only necessary for SAMe binding but also essential for maintaining the tetrameric structure (22). Trypsin digestion of the enzyme separates this domain and forms a 45 kDa monomeric core. This truncated core forms a dimer in contrast to the tetrameric organization of the native enzyme, is twice as active as the full-length enzyme, and does not respond to SAMe activation (22). The first evidence of SAMe activation of the full length CBS was provided by Finkelstein *et al*. wherein the preincubation of partially-purified CBS (from liver extracts) with SAMe resulted in a marked increase of the activity of CBS (23). SAMe was found to be an allosteric activator of rat CBS, reducing the K_m for homocysteine by eight-fold (24). This was later confirmed by the kinetic studies performed on human recombinant CBS, which also showed that the enzyme subunits exhibit positive cooperativity for SAMe binding (14). Another study also confirmed the allosteric activation of SAMe, but it assigned SAMe as a V-type allosteric effector that increases the V_{max} 2-3 fold but has no effect on the K_m for either substrate (17). SAMe has also been proposed to have a role in protecting the enzyme against oxidizing conditions. CBS has several cysteine residues exposed to the

protein surface and hence the redox status of these residues greatly influences the size and activity of the enzyme. The redox potential of the environment plays a vital role in the oligomerization of CBS units, enzyme activity and SAMe activation. SAMe binds to the two exposed cysteine residues out of total 11 cysteine residues per CBS subunit and thus protects the sulfhydryl group against oxidation (4). By virtue of SAMe activation of CBS and inhibition of synthesis of methyltetrahydrofolate (25), SAMe can act as a switch that controls the various pathways of homocysteine metabolism.

CBS as a PLP enzyme belongs to the tryptophan β-synthase family (26). Prior to the structural study of the truncated CBS, sequence homology studies of human CBS to other PLP dependent proteins like, O-acetylserine sulfhydrylase (OASS), O-acetyl-L-serine (thiol) lyase (OASTL) and threonine deaminase (26, 27), indicated that the PLP binding site is located in the N-terminal third of the protein. Kinetic and spectroscopic studies revealed that serine binds in the PLP pocket and homocysteine binds in the proximity of heme (19). Heme binds tightly to the enzyme in an oxidizing environment and can be reversibly released from CBS under reducing conditions only. Displacement from the axial heme ligands causes local unfolding of the N-terminal residues and release of prosthetic group. The redox state of heme has no impact on the CBS activity and heme is not required for catalysis (28). A CBS enzyme with first 70 residues truncated still maintained 25 % activity and this further supported the finding that heme moiety essentially has no role in catalysis and is involved only in proper protein folding and cofactor binding of the protein. The regulation of the enzyme activity by redox changes was attributed to the spatially adjacent oxidoreductase active site motif (28). No structural information about the C-terminal, SAMe binding, CBS-domain is available

although there is a high degree of conservation of this domain with functionally and evolutionarily distinct proteins such as the human voltage-gated chloride channel, ABC transporters, human inosine-5'-monophosphate and several other proteins from *Methanococcus jannaschii* (29). A recent study of the mutants of CBS having mutations in the SAMe binding domain put forth a very interesting and convincing mode of regulation of CBS by SAMe. According to this study, the C-terminal domain of CBS contains an autoinhibitory region that covers the active site. This region gets displaced from the active site upon binding of the allosteric activator SAMe. The interesting result is that it is not just the SAMe that can cause this conformational change but also thermal denaturation or some specific point mutations can also cause this change. This could also explain the activity of the proteolytically cleaved enzyme by rationalizing that the regulatory domain is cleaved from the enzyme and is therefore in the constitutively activated conformation (30).

Thus even though the CBS core has been extremely well characterized structurally and kinetically, it is very doubtful whether all the evaluated kinetic parameters and structural conformations can be extended to the full-length enzyme. Also to understand the complicated regulatory mechanism of CBS via SAMe, it becomes extremely critical to study the protein in its full length with its C-terminal. Therefore the objective of this project was to obtain full-length, soluble and active human CBS for structural studies. The focus is to address the aggregation problems of the full-length CBS that is a major hindrance to the crystallization and X-ray diffraction studies of this enzyme.

MATERIALS

A sample of an expanded human liver cDNA library (Stratagene) was provided by Dr. Harry Dailey's laboratory at the University of Georgia (UGA). "PCR-Script", "Seamless", and "Quick-Change" kits and BL21 (DE3) RIL cells for expression was purchased from Stratagene. The expression vector, pET28b, was from Novagen. The IMPACT T7 system was from New England Biolabs. The Molecular Genetics Instrumentation Facility at the University of Georgia performed DNA sequencing and synthesis of oligonucleotides. Oligonucleotide sequences are shown in Table 9. Restriction enzymes were obtained from Promega, Boehringer Mannheim and New England Biolabs. *Sap* I (2000 U/ml) came from New England Biolabs (NEB). Tryptone-phosphate medium: 2% bacto-tryptone, 0.2 % Na_2HPO_4, 0.1 % KH_2PO_4, 0.8 % NaCl, and 1.5 % yeast extract. All other reagents were of the highest analytical or molecular grade available from commercial vendors.

METHODS

Creation of Sap I based expression vectors. A T7 promoter based vector, pET28b, was used as the template vector for further modification. The single *Sap* I site of the vector at position 3108 was knocked out by site-directed mutagenesis (Primers: Sap-knockout Prim1 and Sap-knockout Prim2) using the Quick-Change protocol (Stratagene) which entails PCR amplification of the template with *Pfu* polymerase using mutation primers, followed by *Dpn* I digestion of the original template (All the primers used in the experiments will be listed in Table 9).

Table 9: Oligonucleotides used for the cloning of CBS

Primer	Sequence
CHSap-Forward	5'-CATGGGAAGAGCTCTTCACACCATCATCATCACTAAT-3'
CHSap-Reverse	5'-AATTATTAGTGATGATGATGGTGTGAAGAGCTCTTCC-3'
WHSap-Forward	5'-CATGGGAAGAGCTCTTCATAAT-3'
WHSap-Reverse	5'-AATTATTATGAAGAGCTCTTCC-3'
CBSforward	5'-AGCATGCCTTCTGAGACCC-3'
CBSreverse	5'-GCTCCGGACTTCACTTCTGG-3'
NHSapCBS-Forward	5'-CGCTCTTCGCACATGCCTTCTGAGACCCCC-3'
NHSapCBS-Reverse	5'-CGCTCTTCATTACTTCTGGTCCCGCTCC-3'
CHSapCBS-Forward	5'-CGCTCTTCTATGCCTTCTGAGACCCCCC-3'
CHSapCBS-Reverse	5'-CGCTCTTCAGTGCTTCTGGTCCCGCTCC-3'
WHSapCBS-Forward	5'-CGCTCTTCTATGCCTTCTGAGACCCCCC-3'
WHSapCBS-Reverse	5'-CGCTCTTCATTACTTCTGGTCCCGCTCC-3'
CBScys1	5'-GGAGGGCCAGCGCTGCGTGGTCATTC-3'
CBScys2	5'-GAATGACCACGCAGCGCTGGCCCTCC-3'
CBSforwardSap	5'-GGTAGATGCTCTTCCAACATGCCTTCTGAGACCCC-3'
CBSreverseXho	5'-GGTGGTCTCGAGTTATCACTTCTGGTCCCGCTCCTG-3'

The vector with the correct mutation was identified by restriction analysis using *Sap* I. Utilizing the *Nco* I and *Eco*R I sites present in the original pET28b vector's multiple cloning site (MCS), three expression constructs were designed: DNA encoding a polyhistidine tag (His$_5$) was introduced after the initiating ATG with a double overlapping *Sap* I sites placed between the initiating ATG and a double stop codon. (Primers: NHSap-Forward and NHSap-Reverse); in the second type, a polyhistidine tag was introduced at the C-terminus just before the stop codon with the *Sap* I sites placed between the ATG codon and CAC (coding for histidine) (Primers: CHSap-Forward and CHSap-Reverse); the third type of vector was designed with no histags (WHSap-Forward and WHSap-Reverse). The resulting vectors called pET28b-NHSap, pET28b-CHSap and pET28b-WHSap, can incorporate any DNA of interest with *Sap* I sites flanking at both the ends of the DNA. Digestion with *Sap* I detaches the recognition site from both the vector and the desired insert which allows the sticky ends flanking the desired sequences to ligate and form a seamless junction. Despite the knockout of the *Sap* I restriction site near the origin of replication, the vector appears stable.

Cloning of CBS into the pET28b-NHSap construct. The DNA encoding CBS was amplified by PCR from a human liver cDNA library (Stratagene). The PCR reactions were performed using *Pfu* Turbo DNA polymerase (Stratagene) with the oligonucleotide primers, CBSforward and CBSreverse and analyzed by agarose gel electrophoresis. The resulting 1650 bp PCR product was then ligated into the PCR product compatible vector pCRScript, and the desired clone was identified as white colonies on IPTG/X-GAL plates. The DNA sequence of one clone containing the CBS insert was confirmed

161

by restriction analysis and sequencing. The PCR vector containing the protein coding sequence of CBS was further PCR amplified with *Sap* I restriction sites engineered upstream of the first ATG of the cDNA and with a C-terminal stop codon so that the gene was in the proper reading frame with an N-terminal purification tag (Primers: NHSap CBS-Forward and NHSap CBS-Reverse). Synthesis of the desired PCR product was again verified by agarose electrophoresis, and the product was purified using a PCR Quick Spin Kit (Qiagen). This PCR product was then inserted into the expression construct pET28b-NHSap using the "Seamless Strategy" of Stratagene, but using *Sap* I rather than *Eam*1104 as the restriction enzyme. 75 ng of the modified pET28b vector and 50 ng PCR product (1:2 molar ratio) were digested together with 2 units of *Sap* I in a 10 μl reaction overnight at 37°C (NEB-Buffer 4). To the digested mix, 1 unit T_4 DNA ligase (Gibco), 4 μl of 5X ligase buffer, 6 additional units of *Sap* I restriction enzyme (3 μl), and 2 μl of water were added to get a 20 μl reaction volume. After overnight ligation, 1 μl of the ligation reaction was transformed into 40 μl *E. coli* strain XL1- Blue (Stratagene) by electroporation. The colonies were screened for the vector containing the insert by single colony screening and restriction analysis. The resulting expression construct was then verified by DNA sequencing and named pET28b-NHCbs.

$\qquad$ ***Test expression of N-terminal poly-histidine tagged CBS***. The expression construct, pET28b-NHCbs was then transformed into electrocompetent BL21 (DE3) cells (Stratagene) by electroporation. 100 ml of LB media, containing 50 μg/ml kanamycin and 0.3 mM δ-aminolevulinate was inoculated with a 2 ml overnight culture. After growing the culture to A_{600} = 0.6 at 37°C, the culture was then induced with 1 mM

IPTG. After induction, aliquots of 1 ml were taken and frozen for expression testing at time interval of 2 hrs, 4hrs, 6hrs and 24 hrs.

Expression and solubility testing. 300 μl of each aliquot at different time points were microfuged at 14,000 x g for 15 mins. The whole cell pellet was boiled directly in reducing SDS/PAGE buffer and then analyzed by SDS-PAGE. For solubility testing, another 300 μl of each aliquot was microfuged at 14,000 x g for 15 mins. The pellet was then resuspended in 60 μl of phosphate buffer and 3 μl of 6 mg/ml of lysozyme. The samples were then kept on ice and then sonicated using a bench-top sonicator (Power level 10). The samples were then centrifuged at 14,000 x g for 10 mins. The supernatant and the cell debris were treated separately: 10 μl of supernatant was boiled with 10 μl of SDS-PAGE buffer for 5 mins; the cell debris pellet was also mixed with 10 μl of SDS-PAGE buffer and then boiled for 5 mins. All the samples (the supernatant and cell pellet for each time point were then analyzed by SDS-PAGE.

CBS solubilization efforts.

a) Extraction of protein from the pellet. All the sample pellets were extracted in buffer with a detergent. The buffer composition used was 20mM Sodium phosphate, pH 7.4, 0.5M NaCl, 50 μM PLP, 5mM BME and 0.1% Triton. They were then analyzed by SDS-PAGE analysis.

b) Preparation of C-terminal and Wild-type (No His-tag) CBS expression constructs. The PCR vector containing the protein coding sequence of CBS was PCR amplified with *Sap* I restriction sites engineered such that the gene was in the proper reading frame with an C-terminal purification tag (Primers: CHSapCBS-Forward and

CHSapCBS-Reverse). The PCR vector containing CBS was again PCR amplified with

Sap I restriction sites engineered such that the gene was inserted in the proper reading

frame with the construct with no histag (Primers: WHSap CBS-Forward and WHSap

CBS-Reverse). Synthesis of the desired PCR product was again verified by agarose

electrophoresis, and the product purified using a PCR Quick Spin Kit (Qiagen). The PCR

products, CHSap-CBS and WHSap-CBS were then inserted into the expression

constructs pET28b-CHSap and pET28b-WHSap respectively, using the "Seamless

Strategy" of Stratagene described earlier. These two expression constructs, named

pET28b-CHCbs and pET28b-WHCbs were then tested for expression and solubility of

CBS as described earlier.

c) Usage of BL21 DE3 RIL cells as expression strain. All the three expression

constructs were transformed into electrocompetent BL21 Codon Plus (DE3) RIL cells

(Stratagene) by electroporation. They were all then tested for expression and solubility of

CBS.

d) Expression using varied temperature and IPTG concentrations. The

expression of the pET28b-NHCbs was tested by lowering the temperature conditions to

room temperature and 16°C. Induction with lower amounts of IPTG at 0.5 mM, 0.1 mM,

and no IPTG for 6, 12, 24, 30, 36, 48 hrs were also tried.

e) Change of growth media. The pET28b-NHCbs was expressed using a change

in the composition of media, for instance instead of LB, the cells were grown in enriched

media like superbroth and tryptone-phosphate buffered media.

f) Mutation of cysteine to arginine (AGTGCT to AGCGCT). This single base

mutation was carried out by site-directed mutagenesis (Primers: CBScys1 and CBScys2)

using Quick-change protocol (Stratagene). pET28b-NHSap construct was used for correction of the mutation. The vector with the correct mutation was identified by sequencing and restriction analysis using *Afe* restriction enzyme. This construct was checked for expression and solubility of CBS.

g) **Denaturation and renaturation of CBS.** Expression conditions optimized for maximal CBS production in the insoluble pellet as evaluated by SDS-PAGE were used to produce substantial quantities of CBS in the cellular pellet after sonication. 20 mM sodium phosphate buffer, pH 7.4 with 0.5 M urea was used to pre-extract the cell wall and outer membrane components from the aggregates and produce washed pellets. The aggregated protein was then solubilized in a strong protein denaturant, 6 M guanidine hydrochloride (20 mM sodium phosphate buffer with 6M guanidine hydrochloride, 0.5 M NaCl, 1 mM EDTA and 5 mM DTT (dithiothreitol)). This solubilization was carried out at room temperature for 2 hrs. The denatured his-tagged protein was then partially purified by using metal chelate affinity chromatography. Then the protein was renatured by 10-fold dilution into a renaturation buffer, 20mM sodium phosphate containing 0.5 M NaCl, 1 M guandine hydrochloride, 1mM glutathione (GSH-reduced form), 0.2 mM glutathione disulfide (GSSG-oxidized form), 0.5 M *L*-arginine, 100 µM heme (dissolved in 100% dimethyl sulfoxide) and 50 µM PLP. This renaturation was carried out at 4°C for 24 hrs. The protein was then dialyzed back into dialysis buffer (20 mM sodium phosphate buffer with 50 µM PLP and 2 mM DTT) with two changes with each change being 6 hrs long. The dialyzed protein was then checked for purity on SDS-PAGE. It was then repurified by metal chelate affinity chromatography. The protein was then checked for activity by ninhydrin based colorimetric assay.

165

In the second attempt at denaturation and renaturation of CBS, after the initial step of denaturation, renaturation was done by including S-adenosylmethionine (30 μM) in the renaturation buffer. This renaturation was done by step-wise dilution. Each step of dilution was done with the renaturation buffer, which included everything but GSH and GSSG. Each step was allowed to remain at room temperature for 30 mins and then the next dilution was performed. Before the final dilution, the protein solution was divided into two parts, one part was diluted to the final concentration with the renaturation buffer without GSH/GSSG and the other half was diluted to the final concentration with GSH/GSSG included in the renaturation buffer. Both these solutions were then allowed to remain at 4°C for 24 hrs. The protein solutions were dialyzed back into the dialysis buffer to remove the remnants of guanidine hydrochloride and then repurified by metal chelate affinity followed by anion exchange chromatography (Q column). The buffer used for loading onto the Q column was 50 mM Tris buffer, pH 8, 5 mM BME, 1 mM EDTA and 50 μM PLP. The protein was then assayed by colorimetric assay.

h) Expression in a different system (IMPACT T7 - <u>I</u>ntein <u>M</u>ediated <u>P</u>urification with an <u>A</u>ffinity <u>C</u>hitin-binding <u>T</u>ag).

Preparation of CBS-Intein/Chitin-binding fusion construct (pTYB11-CBS). pTYB11 is a N-terminal fusion vector in which the N-terminus of the target protein is fused to the intein tag. The PCR vector containing the protein coding sequence of CBS was PCR amplified with *Sap* I restriction sites engineered upstream of the first ATG of the cDNA and *Xho* I site at the C-terminus (Primers: CBSforwardSap and CBSreverseXho). Synthesis of the desired PCR product was again verified by agarose electrophoresis, and the product purified using a PCR Quick Spin Kit (Qiagen). This PCR

product and the pTYB11 were then digested with *Sap* I and *Xho* I. Following the

purification of the DNA fragments with the PCR Quick Spin kit, T4 DNA ligase (Gibco)

was used to ligate the PCR product into the vector. The ligase reaction product was

transformed into XL-1 Blue (Stratagene). The resulting expression construct was verified

by DNA sequencing and named pTYB11-CBS. The expression construct was then

transformed into electrocompetent BL21 (DE3) RIL cells (Stratagene) by electroporation.

Expression of pTYB11-CBS. Small-scale expression testing was performed at

three temperatures, 37°C, RT and 15°C. 250 ml cultures (containing 100 μg/ml

carbenicillin and 0.3 mM ALA) were inoculated with 3 ml overnight cultures and were

allowed to grow at 37°C until OD 0.6. They were then transferred to respective

temperatures and induced with 0.3 mM IPTG. 1 ml aliquots at 0, 6, 18, 24 hrs were taken

and frozen. They were then analyzed for expression and solubility. Two types of lysis

buffer were tested for extraction of protein from the cell pellet: 20 mM Tris pH8, 500

mM NaCl, 1 mM EDTA, 0.2% Triton, 0.2 mM PLP and 1 mM TCEP (reducing agent);

the other type of buffer had similar composition but without Triton.

Large scale expression and isolation of CBS (optimized). 10 ml of overnight

culture of pTYB11 was inoculated into 1 L of LB with 100 μg/ml carbenicillin. The

mixture was allowed to grow until OD 0.6 and then was supplemented with 0.3 mM ALA

(aminolevulinic acid) and 0.3 mM IPTG. The culture was induced at 15°C for 10 hrs.

Following the induction period, the cells were collected by centrifugation at 13,000 x g

for 15 min (4°C) and resuspended in 40 ml of lysis buffer (20 mM Tris pH8, 500 mM

NaCl, 1 mM EDTA, 0.2% Triton, 0.2 mM PLP, 1 mM TCEP) containing an EDTA-free

protease tablet (Boehringer-Mannheim) and used immediately. The cell suspension was

167

cooled in ice to 10°C and sonicated for 3 min in 1-min increments with a Fisher Scientific 550 sonic dismembrator set at 50% duty cycle while maintaining the temperature below 15°C. Cellular debris was removed by centrifugation at 14,000 x g for 30 mins (4°C), and the clarified supernatant was purified by chromatography. The clarified lysate was applied (1 ml/min) to a chitin affinity column (60 ml) that had been previously equilibrated with the lysis buffer. The flow through from the column was collected and loaded back onto the column. The column was washed with 5 column columns of the same buffer. The column was then rapidly washed with 2 column volumes of cleavage buffer (20 mM Tris pH8, 500 mM NaCl, 1 mM EDTA, 0.2% Triton, 0.2 mM PLP and 100 mM BME (5 ml/min). The column was then capped and allowed to stand at room temperature for 48 hours. Then the protein was eluted from the column using cleavage buffer without the BME. The fractions were pooled and then analyzed by SDS-PAGE. The protein solution was then dialyzed into 50 mM Tris buffer, pH 8, 1 mM EDTA, 5 mM BME and 0.2 mM PLP. The dialysis was done with three changes to ensure that the Triton was completely dialyzed out. The dialyzed protein was then loaded onto the Pharmacia Hi-Trap Q column (5ml) equilibrated with 50mM tris HCl, (pH 8), 5mM BME and 1 mM EDTA. The protein was eluted using a linear gradient of 0-1.0 M NaCl over 20 column volumes. The purity was again checked by SDS-PAGE analysis. The same procedure was repeated for testing the expression of pTYB11-CBS in a defined media.

i) Expression of pET28b-NHCbs in defined media. The expression conditions were similar to that of pTYB11-Cbs except that the media was the defined media instead of LB. For purification, metal affinity chromatography was used as the first step, the

buffer being the same. The protein was eluted using a linear gradient of imidazole (0 –

0.5 M imidazole). The second step was anion exchange column chromatography with

conditions used for pTYB11-Cbs. An extra column of purification was used which is

very specific for PLP proteins. The protein was loaded to an amino hexyl column with

the loading and elution buffer similar to that of anion exchange column.

Enzymatic assay. The assay volume of 0.2 ml contained different enzyme

concentrations (buffer: 50 mM Tris, pH 8, 5 mM BME, 0.2 mM PLP) 10 mM L-

homocysteine (prepared just before use by heating homocysteine thiolactone

hydrochloride for 3 min with 2 equivalent of NaOH), 10 mM serine, 0.2 mM PLP. The

enzyme without the substrates is preincubated at 37°C for 10 mins. The substrates were

then added to the reaction mixture and then incubated in a shaker incubator at 37°C for

45 mins. At the end of the incubation period, 0.1 ml of 50 % TCA was added, precipitate

was removed by centrifugation and the clear supernatant fluid was mixed with 3.2 ml of

ninhydrin reagent (0.5 g of ninhydrin dissolved in 50 ml of glacial acetic acid and 17 ml

of glacial phosphoric acid). After boiling the ninhydrin treated samples for 5 mins, cooled

the mixture in ice for 2 mins and then allowed the color to develop at room temperature

for 20 mins. The absorbance at 453 nm was recorded and compared to standard curve

prepared using cystathionine.

RESULTS

The CBS protein sequence as initially obtained from the cDNA PCR

amplification contained a missense mutation with a cysteine substituting for arginine. The

sequencing result showed that what should be AGCGCT was AGTGCT in our sequence.

An *Afe I* restriction site should be present at the mutation site but restriction analysis of the original PCR product and several isolated clones do not show any evidence for this site. Therefore it is likely that the mutation was present in the individual from which the original cDNA library was constructed.

Of all the expression constructs: pET28b-NHCbs, pET28b-CHCbs and pET28b-WHCbs, only pET28b-NHCbs showed expression. But after examining the levels of soluble protein recoverable in cell extracts after sonication and centrifugation, it was found that the protein is insoluble (Figure 15). Various solubilization efforts were carried out to solubilize CBS, but all of them proved to be futile. With the corrected protein, attempts were made to renature CBS from the insoluble cellular debris after cell disruption. After the final dialysis against phosphate buffer, the renatured protein appeared soluble in the SDS-PAGE analysis. When the protein was tested for its activity, it was found to be inactive. In the next attempt at renaturation, S-adenosylmethionine was included in the renaturation buffer with the rationale that it could help in the proper folding of the C-terminal of the protein. This effort also yielded soluble but inactive protein (Figure 16). An interesting observation in this experiment was that when half the amount of protein was renatured in the renaturation buffer without the reduced and oxidized glutathione, the protein precipitated completely. This suggested the presence of disulfide bonds in the proteins that require the presence of redox shuttling system.

The expression testing in the IMPACT T7 system on a small scale demonstrated that the protein formed inclusion bodies at 37°C. According to SDS-PAGE analysis, room temperature induction for 6 hrs and 15°C induction for 10 hrs showed evidence of soluble protein in the crude extract. The protein was detected by a strong band at
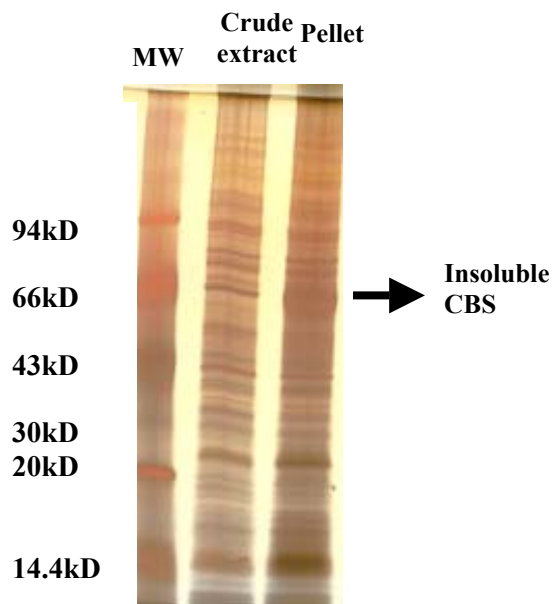
Figure 15: SDS-PAGE analysis of N-terminal his-tagged CBS

The solubility testing of the CBS showed that all of the protein formed inclusion bodies in the cell (pellet) and the crude extract had no traces of the soluble protein. Lane MW contains molecular mass markers.

molecular weight of 118,000 (fusion of intein- 55,000 and CBS- 63,000). When the expression of the protein was scaled up to 1 L in similar temperature conditions, the protein formed aggregates at the step of high-speed centrifugation (60,000 x g for 20 mins) and was not recoverable in the supernatant. Therefore isolation of CBS was carried out at lower centrifugation speeds. The results of SDS-PAGE also led to the conclusion that inclusion of detergent (0.2 % Triton) gave better yields of soluble protein than that obtained in lysis buffer without detergent. Therefore the final optimized condition accepted for CBS expression and extraction was induction at room temperature for 6 hrs and extraction with lysis buffer that included detergent (20 mM Tris buffer, pH 8, 5 mM BME, 1 mM EDTA, 0.5 M NaCl, 0.2% Triton and 0.2 mM PLP). The purity of the protein obtained after two steps of purification, the chitin affinity column and the anion exchange column (Q column) was about 95%. SDS-PAGE analysis showed that along with the 63,000 molecular weight band of soluble CBS, there was a contaminant band at 45 kDa. This protein band was attributed to either the cleaved intein or the proteolytically cleaved core of CBS (Figure 17). The final yield of the protein obtained by this method was very low, approximately 1 mg of protein from 1L. The same procedure of pTYB11-Cbs expression and isolation was repeated by using the defined medium. The defined medium is a synthetic medium containing a cocktail of salts, 20 amino acids and vitamins (30). The soluble protein yield increased to approximately 4 mg per liter. The specific activity of the protein when assayed by colorimetric method was found to be 4.2 μmole/min/mg, which was very comparable to the earlier reported value, 4 μmole/min/mg, for CBS purified as a fusion protein (14).
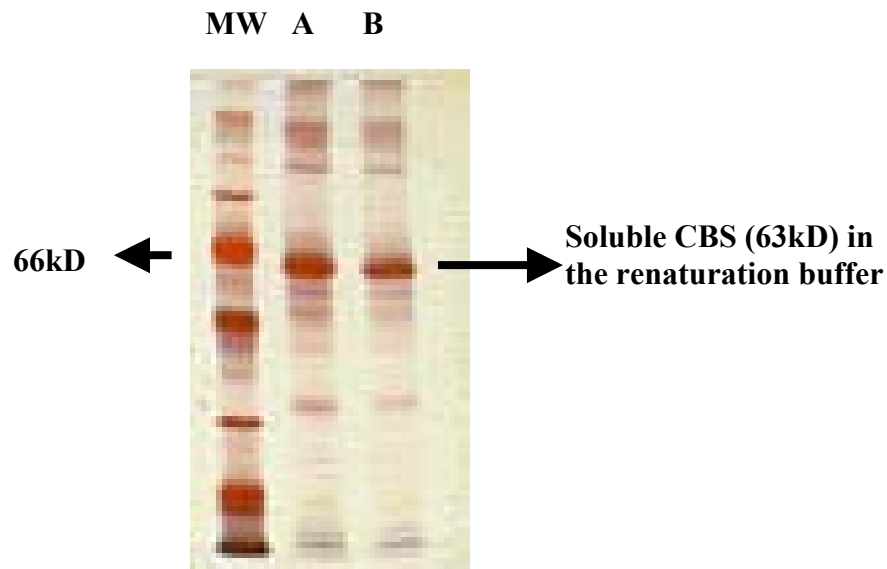
Figure 16: SDS-PAGE gel showing the renatured CBS as soluble protein in the renaturation buffer

The renaturation buffer with the soluble CBS showed many contaminant bands indicating that the metal chelate affinity chromatography was not successful in isolating just the CBS. Lane MW contains molecular mass markers. Lane A shows 1:5 dilution of the protein and lane B shows 1:10 diluted protein.
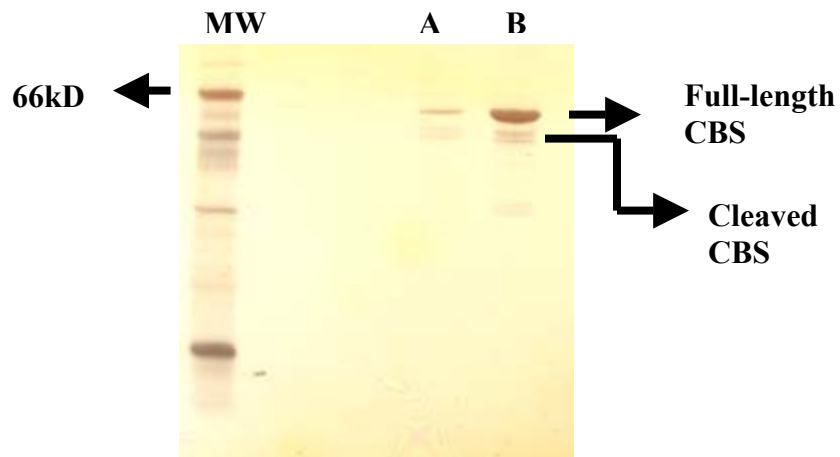
Figure 17: SDS-PAGE analysis of the soluble CBS obtained by using the IMPACT T7 system

The soluble CBS after chitin affinity and anion exchange chromatography was not a homogenous species. According to the gel, the full-length CBS is susceptible to proteolysis that separates the core of CBS (45kDa) and is seen as the extra band along with the full-length CBS band at 63kDa. Lane MW contains the molecular mass markers. Lane A shows 1:25 dilution of the protein. Lane B contains the undiluted protein.

Having obtained soluble and active CBS after various efforts, the next step was to find alternatives to the problems associated with it. First and foremost, the CBS obtained in this method is expressed as a fusion protein with intein. This method defeats our purpose to express full length CBS without any fusion protein attached to it. The protein band observed at molecular weight 45,000 along with the full-length protein band at 63,000 confirmed the doubts of proteolytic degradation. The possibility of the degraded product (45,000 band) being the cleaved intein was ruled out because according to the product literature, intein binds very tightly to the chitin column and it requires 0.3 M NaOH as the stripping solution. Besides, the molecular weight of intein is 55,000. The factors that could be considered as causes of proteolytic degradation of CBS is that it is being degraded either in the cells during long induction period of 10 hrs or during the 48 hrs incubation for cleavage on the column. This proteolysis problem could also be one of the reasons for low protein yield.

With the optimized parameters for expression of pTYB11-Cbs in the IMPACT T7 system, it was decided to again try the expression of pET28b-NHCbs. The parameters of expression and isolation used were: defined medium as the cell growth medium, induction period for only 6hrs at room temperature, low speed spin for pelleting the cells, composition of extraction buffer similar to that of the pTYB11-Cbs (50 mM Tris buffer, pH 8, 0.5 M NaCl, 5 mM BME, 0.2 mM PLP and 0.2 % Triton, 2 tablets of EDTA-free protease inhibitor) and immediate loading on the metal affinity column without freezing the cells. This was followed by purification using anion exchange column and amino-hexyl column chromatography. The SDS-PAGE analysis of the pooled fractions after each chromatography still showed the cleaved 45,000 protein band along with the full

length CBS band. The yield of soluble protein was 6 mg per liter and the specific activity

measured was 3.64 μmoles of cystathionine/min/mg. Thus this method of expression and

purification of full length CBS did not prevent the proteolytic degradation of CBS

although it did have advantages of improved yield and rapid purification.

DISCUSSION

For crystallization purposes, it is advantageous to have the protein in its native

state with no fusion to any other protein or with no additional amino acids added onto the

protein. Commercial vectors tend to introduce these additions to the protein of interest,

beyond a simple purification tag, that confer undesired conformational flexibility and

thus decrease the likelihood of crystallization. Therefore pET28b-NHSap, the vector that

was modified to introduce only a short purification tag and no protease cleavage, was the

initial method of choice for expression of pET28b- NHSap. The histidine tag at the N-

terminus has been the preferred choice because it should have the least interference with

the S-adenosylmethionine binding to the C-terminal and oligomerization of the protein.

Therefore the entire focus was largely on the pET28b-NH construct for solubilization of

CBS.

CBS has been reported to have significant solubility problems and so it was not

surprising that various attempts to obtain full length, soluble CBS did not succeed. After

trying various strategies of solubilizing CBS like reducing the rate of protein synthesis,

supplementing the medium with δ-ALA to increase cellular heme concentrations,

correcting the R368C mutation, the last attempt of solubilizing NHCbs was the *in vitro*

denaturation and refolding of CBS. This procedure did give soluble protein but the activity was lost. This method has been effective for a very limited number of proteins.

It was also not very surprising that the IMPACT T7 system gave soluble CBS because in earlier work, CBS had been expressed in fusion with β-galactosidase (14). When the successful conditions for expression and isolation in this system were applied to the pET28b-NHCbs vector, the goal of obtaining full length, soluble, active CBS without any fusion protein attached was met. The combinations of the synthetic defined media, low IPTG concentrations, shorter times of induction and change in the lysis buffer composition were the factors that interplayed to give active CBS.

The next big hurdle was the purity of the CBS obtained, since it plays a very important role in crystallization of proteins. The SDS-PAGE analysis of the purified CBS shows that proteolysis of the protein gives a 45,000 molecular weight band. This band was assigned to be the stable active core of CBS because it also behaves like the full-length CBS. Metal chelate affinity, anion exchange, amino hexyl and gel filtration chromatography, all failed to separate the core protein (Figure 18). This could be explained by the existence of a non-homogenous mixture of the core, which is known to be a dimer and the full length CBS, which is a tetramer. The dimer could associate with another dimer and this dimer of dimer could make the resolution between the core and the full length difficult.

The main problem of aggregation of CBS can be tackled at different levels. First, at the expression level, the protein should be screened with different detergents and then look for prevention of aggregation. A preliminary screen was done by running a native gel on CBS incubated with different detergents. Among all the detergents, zwitter ion

177
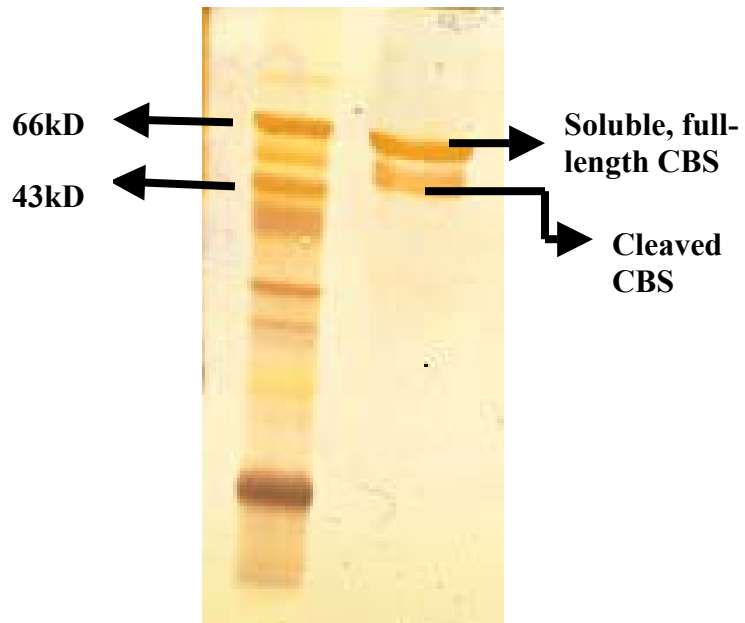
Figure 18: SDS-PAGE gel of the soluble, active CBS obtained from the defined media

Even after metal chelate affinity, anion exchange and gel-filtration chromatography, the

cleaved CBS (45kD) could not be separated from the full-length CBS (63kD)

gave the appearance of a tight band on the gel indicating the ability of that detergent to prevent the protein aggregation. This could be an interesting lead as a relatively new zwitterion that is being widely used for insoluble proteins is sulfobetaine. Non-detergent sulfobetaines (NDSB) are particularly suited for protein work because they are non denaturing (most enzymes remain active in the presence of NDSB) (32). Expression in yeast (*Pichia pastoris*) can be one of the alternative approaches for obtaining soluble and homogenous CBS. Pichia pastoris has a strong, inducible promoter that can be used for protein production. It is capable of generating post-translational modifications that are more similar to human protein modifications than what *S. cerevisiae* was capable of doing (33).

Attempts can also be made to separate the core and the full-length CBS. Wild-type CBS contains a trypsin site that removes a fragment of the C-terminus to form the stable, active core. Therefore the core cannot bind S-adenosylmethionine but the full-length enzyme has the SAMe binding domain intact. A method, which utilizes this property of both the proteins, is the affinity chromatography using immobilized S-adenosylhomocysteine. This method has been used for separation of SAMe-dependent methyltransferases (34). This method involves immobilizing SAH on the gel, binding of the SAMe dependent enzyme on the column using higher pH and then eluting the protein by lowering the pH.

Thus even though full-length, soluble, active human CBS was obtained, more efforts have to be put in before obtaining crytsallizable protein for the achievement of the long term goal of structural study of SAMe binding domain of CBS.

REFERENCES

1. Scriver CR and Rosenberg LE. 1973. Sulfur amino acids. In Amino Acid Metabolism and Its Disorders. WB Saunders Company, Philadelphia. 207-233.

2. Chadefaux B, Rethore MO, Raoul O, Ceballos I, Poissonnier M, Gilgenkranz S, Allard D. 1985.Cystathionine beta synthase: gene dosage effect in trisomy 21. *Biochem Biophys Res Commun* 128:40-4

3. Clarke RL and Daly *et al*. 1991. Hyperhomocysteinemia: an independent risk factor for vascular disease. *N Engl J Med* 324:1149-55.

4. Kraus JP. 1998. Biochemistry and molecular genetics of cystathionine β-synthase deficiency. *Eur J Peidatr* 157:S50-S53.

5. Watanabe M, Osada J, Aratani Y, Kluckman K, Reddick R, Malinow MR, Maeda N. Mice deficient in cystathionine beta-synthase: animal models for mild and severe homocyst(e)inemia.1995. *Proc Natl Acad Sci U S A* 92:1585-9.

6. Binkley F. Christensen GM and Jensen WN. 1952. Pyridoxine and the transfer of sulfur. *J. Biol. Chem* 194:109-113.

7. Smolin LA and Benevenga NJ. 1982. Accumulation of homocysteine in vitamin B-6 deficiency: a model for the study of cystathionine β-synthase deficiency. *J Nutr* 112:1264-1272.

8. Barber GW and Spaeth GL. 1969. The successful treatment of homocystinuria with pyridoxine. *J.Pediat* 75:463-478.

9. Mudd SH, Levy HL, Skovby F. 1989. Disorders of transsulfuration. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic basis of inherited disease, 6th ed. McGraw-Hill, New York, pp 693-734.

10. Lipson MH, Kraus JP and Rosenberg EL. 1980. Affinity of cystathionine β-synthase for pyridoxal 5'-phosphate in cultured cells. *J. Clin.Invest*. 188-193.

11. Kim JY, Rosenberg EL. 1974. On the mechanism of pyridine responsive homocystinuria. II. Properties of normal and mutant cystathionine β-synthase from cultured fibroblasts. *Proc Nat Acad Sci. USA*. 71:4821-4825.

12. Kraus JP, Le K, Swaroop M, Ohura T, Tahara T, Rosenberg LE, Roper MD, Kozich V. 1993. Human cystathionine beta-synthase cDNA: sequence, alternative splicing and expression in cultured cells. *Hum Mol Genet* 2:1633-8.

13. Ishihara S, Moroshashi KI, Sadano H, Kawabata SI, Gotoh O and Omura T. 1990. Molecular cloning and sequence analysis of cDNA coding for rat liver hemoprotein H-450. *J.Biol Chem* 108:899-902.

14. Bukovska GV and Kery *et al*; 1994. Expression of human cystathionine beta-synthase in Escherichia coli: purification and characterization." *Protein Expr Purif* 5:442-8.

15. Jhee KH, McPhie P, Miles EW. 2000. Yeast cystathionine beta-synthase is a pyridoxal phosphate enzyme but, unlike the human enzyme, is not a heme protein. *J Biol Chem* 275:11541-4.

16. Kery V, Elleder D and Kraus JP. 1995. Delta-aminolevulinate increases heme saturation and yield of human cystathionine beta-synthase expressed in Escherichia coli. *Arch Biochem Biophys* 316:24-9.

17. Taoka S, Ohja S, Shan X, Kruger WD, Banerjee R. 1998. Evidence for heme-mediated redox regulation of human cystathionine beta-synthase activity. *J Biol Chem* 273:25179-84.

18. Kery V, Bukovska G, Kraus JP. 1994. Transsulfuration depends on heme in addition to pyridoxal 5'-phosphate. Cystathionine beta-synthase is a heme protein. *J Biol Chem* 269:25283-8.

19. Taoka S, West M, Banerjee R. 1999. Characterization of the heme and pyridoxal phosphate cofactors of human cystathionine beta-synthase reveals nonequivalent active sites. *Biochemistry* 38:7406.

20. Bruno S, Schiaretti F, Burkhard P, Kraus JP, Janosik M, Mozzarelli A. 2001. Functional properties of the active core of human cystathionine beta-synthase crystals. *J Biol Chem* 276:16-9.

21. Taoka S, Widjaja L, Banerjee R. 1999. Assignment of enzymatic functions to specific regions of the PLP-dependent heme protein cystathionine beta-synthase. *Biochemistry* 38:13155-61.

22. Kery V, Poneleit L, Kraus JP. 1999. Trypsin cleavage of human cystathionine beta-synthase into an evolutionarily conserved active core: structural and functional consequences. *Arch Biochem Biophys* 355:222-32.

23. Finkelstein JD, Kyle WE, Martin JL, Pick AM. 1975. Activation of cystathionine synthase by adenosylmethionine and adenosylethionine. *Biochem Biophys Res Commun* 66:81-7.

24. Roper MD, Kraus JP. 1992. Rat cystathionine beta-synthase: expression of four alternatively spliced isoforms in transfected cultured cells. *Arch Biochem Biophys* 298:514-21.

25. Kutzbach C, Stokstad EL. 1971. Mammalian methylenetetrahydrofolate reductase. Partial purification, properties, and inhibition by S-adenosylmethionine. *Biochim Biophys Acta* 250:459-77.

26. Swaroop M, Bradley K, Ohura T, Tahara T, Roper MD, Rosenberg LE, Kraus JP. 1992. Rat cystathionine beta-synthase. Gene organization and alternative splicing. J Biol Chem 267:11455-61.

27. Kraus JP. 1994. Komrower Lecture. Molecular basis of phenotype expression in homocystinuria. *J Inherit Metab Dis* 17:383-390.

28. Meier M, Janosik M, Kery V, Kraus JP, Burkhard P. 2001. Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *EMBO* J 20:3910-6.

29. Bateman A. 1997. The structure of a domain common to archaebacteria and the homocystinuric disease protein. *Trends Biochem Sci* 22:12-13.

30. Janosik M, Kery V, Gaustadnes M, Maclean KN, Kraus JP. 2001. Regulation of human cystathionine β-synthase by S-adensoyl-L-methionine: evidence for two catalytically active conformations involving an autoinhibitory domain in the C-terminal region. *Biochemsitry* 40:10625-10633.

31. Ramakrishna V, Finch JT, Graziano V, Lee PL & Sweet RM. 1993. Crystal structure of globular domainof histone H5 and its implications for nucleosome binding. *Nature* 362:219-23.

32. Vuillard L, Rabilloud T, Goldberg ME. 1998. Interactions of non-detergent sulfobetaines with early folding intermediates facilitate in vitro protein renaturation. *Eur J Biochem* 256:128-135.

33. Cregg, JM, Barringer KJ, Hessler AY and Madden KR.1985. Pichia pastoris as a host for transformations. *Mol Cell Biol* 5:3376-3385.

34. Mack JP, Slaytor MB. 1978. Affinity chromatography of an S-adenosylmethionine-dependent methyltransferase using immobilized S-adenosylhomocysteine. Purification of the indolethylamine N-methyltransferases of phalaris tuberosa. *J Chromatogr* 157:153-9.

CHAPTER V

CONCLUSIONS

BETAINE: HOMOCYSTEINE METHYLTRANSFERASE

Specific aim 1: Design a simple expression vector for BHMT, using which, the protein can be expressed with few superfluous amino acids attached and, can be purified rapidly. The expression vector used to express all the three proteins, BHMT, CGL and CBS, was designed with the aim of broad applicability. The pET28b vector was designed to produce N-terminal, C-terminal or No-histagged proteins with a simple construction scheme having broad applicability because of the use of rare *Sap* I cloning sites. Thus, being ensured that most of the genes of interest will not have the 7 base restriction recognition sequence, they can be PCR modified to include *Sap* I restriction sites and then directly cloned into any of the modified constructs depending on the position of his-tag. The design of this construct not only simplifies cloning but also decreases the number of additional amino acids appended on the amino and carboxyl ends commonly found in commercial expression vectors. Introduction of the his-tagged protein obviates the use of proteases and allows adoption of a simple one-step purification protocol. BHMT expressed using this vector could be rapidly purified using metal chelate chromatography.

Specific aim 2: Perform crystallization trials on BHMT. The initial conditions that resulted in crystals of BHMT that diffracted up to 2.9 Å had ethylene glycol as precipitant and PEG-20,000 as additive. But later, all the attempts to crystallize BHMT were seriously hampered with uncontrolled nucleation and termination of growth. The crystals of BHMT were either too small to diffract or diffracted to a low resolution. The problems of crystallization of BHMT could be associated to its susceptibility to proteolysis. The proteolytically cleaved protein can get incorporated onto the face of an otherwise perfect protein crystal and can cause poisoning of the growth surface of the crystal. Some solutions to this problem include preparing samples at lower temperature to keep the protease activity under check, increasing the load of protease inhibitors or trying to resolve the different isoforms of BHMT using isoelectric focusing chromatography. BHMT has several cysteine residues, some of which are involved in catalytic activity, and they likely make the protein oxidation sensitive. Oxidation of these cysteine residues could alter the conformation of the protein thereby diminishing its prospects of crystallizability. The oxidation problem can be avoided by freshly adding reducing agents just before crystallization or carrying out the entire procedure of purification and crystallization under aerobic chambers.

Specific aim 3: Clarify the ambiguity about the oligomerization property of BHMT. A preliminary X-ray crystallography analysis of these recombinant BHMT crystals indicated a C2 (monoclinic) space group with a non-crystallographic two-fold symmetry axis. The crystallographic studies suggested the likelihood that the protein is a tetramer with molecular 222 symmetry. This was the first contrary evidence to the earlier reports of human BHMT being a hexamer. The follow up gel-filtration study of the

recombinant enzyme gave consistent results of native BHMT being a homotetramer with a molecular mass of 180,000 Da. The discrepancy of BHMT being a hexamer or a tetramer could be a result of either our recombinant protein having a his-tag at the N-terminus or some post-translational modification, both of which could confer different oligomerization properties to the protein. Alternatively, the presence of stabilizing agents like N,N-dimethylglycine and *DL-* homocysteine used by earlier investigators may change the physical behavior of BHMT.

Specific aim 4: *Study the effect of SAMe on human recombinant BHMT.* The regulatory aspect of BHMT by SAMe was investigated by qualitative isothermal titration calorimetry and SAMe inhibition studies. Both methods proved that SAMe does not interact or inhibit BHMT. An alternative to the observed opposing result was attributed to the discovery of a second BHMT homologue (BHMT-2) in humans and mice, which has 73% amino acid identity to BHMT. BHMT and BHMT-2 have sequence similarity to S-methylmethionine-homocysteine S-methyltransferases found in *E. coli* and *Astragalus bisculatus,* which interact with both S-methylmethionine and SAMe. There is thus a possibility that BHMT-2 preferentially uses one of the alternative substrates like SAMe as a methyl donor. Also there is a possibility that the lack of SAMe interaction with the recombinant BHMT could be a result of the poly-histidine tag that might offer steric hindrance to the binding of ligand SAMe. Structural investigations of BHMT and BHMT-2 could help to sort the differences. BHMT-2 has already been cloned in pET28b-NH vector but its expression procedure is associated with solubility problems (Data not shown).

In conclusion, recombinant human BHMT is a tetrameric enzyme that is not affected by S-adenosylmethionine levels in the body. Structural studies of BHMT, which could give further insights on this aspect, are hampered by crystallization problems like irreproducibility and oxidation sensitivity. Successful crystallization trials of BHMT may require fresh isolations and crystallization under anaerobic conditions.

CYSTATHIONINE γ-LYASE

Specific aim 1: Develop an efficient expression system for human cystathionine γ-lyase. The pET28b-NH based CGL expression and isolation proved to be not only a fast method, but also efficient as far as yield of the active protein was concerned. Compared to the earlier week-long purification protocol of human CGL, the designed expression system enabled the purification of recombinant CGL to be completed in 4 hrs. Also the yield and activity of CGL was significantly higher as compared to that of the earlier approaches.

Specific aim 2: Perform various crystallization trials of CGL in order to find a suitable condition that could give large, diffraction-quality and reproducible crystals Crystallization trials of CGL failed to give any successful leads because the protein seemed to precipitate non-specifically over time. CGL has been reported to have aggregation problems. Even though the CGL isolation from pET28b-NH vector had some initial solubility problems, the SDS-PAGE analysis and gel filtration chromatography indicated a homogenous preparation of CGL. Also CGL could be crystallized from freshly isolated sample by high-throughput crystallization screening method, but not after storage for a month at 4°C. This led to the conclusion that CGL must be undergoing some

conformational change over the time that could lead to aggregation. Aggregates have the property of increasing the heterogeneity of the macromolecular population, competing with the molecules for normal ordering in the crystal and thus creating a major hindrance to the nucleation process and crystal growth. Another possible reason for CGL's instability could be its sensitivity towards oxidation. CGL could also have some sensitive surface cysteines that can be a major problem in crystallization. Alternatively, it is also possible that the appended his-tag is contributing some conformational change to the protein or the tag instead of folding compactly with the protein stays as a separate floppy arm taking serious tolls on crystallization. Taking into consideration the problems associated with CGL, the ideal way to approach them would be to perform a very detailed study of the effect of various precipitants, additives, individual detergents and mixture of detergents by light scattering. This method could give a good understanding about the conditions that are conducive to aggregation of proteins or which detergent/combination of detergents can maintain the protein as homogenous species. The oxidation sensitivity problem can be addressed either by including fresh reducing agents in the buffers or by mutating specific surface residues that could make crystallization amenable. The final approach would be to remove the his-tag attached to the protein and then perform crystallization trials on the native protein.

CYSTATHIONINE β-SYNTHASE

     Specific aim: *Obtain human, full-length, soluble and active CBS with no fusion protein attached.* The initial strategies applied for solubilizing CBS were reducing the rate of protein synthesis, supplementing the medium with δ-ALA to increase cellular

heme concentrations, and correcting an R368C mutation. The last attempt to solubilize NHCbs was the *in vitro* denaturation and refolding of CBS. This procedure gave soluble protein albeit *sans* the activity. Finally the goal of obtaining full length, soluble, active CBS without any fusion protein attached was met by combination of synthetic defined media, low IPTG concentrations, shorter induction times and by changing the lysis buffer.

Proteolysis was the next major hurdle associated with CBS. Metal chelate affinity, anion exchange, amino hexyl column and gel filtration chromatography, all failed to separate the cleaved, core protein. It was finally concluded that the major problem with CBS is proteolytic cleavage followed by the association of the partially cleaved dimer into higher oligomers, making the resolution of the aggregated mass of the cleaved and the full-length enzyme difficult. Thus, even though full-length, soluble, active human CBS was obtained, prevention of proteolysis will be the next major struggle. The next step would be to either try to prevent proteolysis and aggregation by the appropriate combinations of protease inhibitor, temperature control, use of detergents, chaotropes or non-detergent sulfobetaines or design a HPLC based protocol to separate the cleaved protein from the full-length on the basis of their differential SAMe binding characteristics.

APPENDIX A

ANALYSIS OF FUNCTIONALLY AND STRUCTURALLY IMPORTANT AMINO ACID RESIDUES OF DIAMINOPIMELATE DECARBOXYLASES BASED ON THE CRYSTAL STRUCTURE OF THE *E.COLI* DIAMINOPIMELATE DECARBOXYLASE[1]

---

[1]Bose, N. and Momany, C. To be submitted to *Protein Science*

ABSTRACT

Diaminopimelate decarboxylase is an enzyme found in bacteria and plants involved in the biosynthesis of lysine from diaminopimelic acid. Based on the three dimensional structure of *E. coli* diaminopimelate decarboxylase, a complete structural and functional analysis of the amino acid residues of the enzyme is reported. We performed two multiple sequence alignments: one, of all the related amino acid decarboxylases for which sequences are available and second, of all the available diaminopimelate decarboxylases which are homologous to that of the *E.coli*. The invariant or conservatively substituted amino acid residues were identified and assigned specific roles as residues involved in cofactor binding, selective substrate binding or stereospecific decarboxylation and the residues that are important for structural integrity. In this process, various amino acid residues came into light, which on substitution could yield only subtle changes in the structure but dramatically affect the overall mechanism or reaction rates.

KEYWORDS

There are two distinct, mutually exclusive pathways for lysine biosynthesis: the diaminopimelate pathway in bacteria and plants and the α-aminoadipate pathway in euglenoid algae and higher fungi (Vogel., 1965; Vogel., 1961). In the diaminopimelate pathway, lysine synthesis occurs in the final step of a series of six highly regulated reactions from aspartic semialdehyde (Cohen., 1985). In this last step, diaminopimelate decarboxylase (DapDC) (meso-2,6-diaminopimelate carboxy-lyase, EC 4.1.1.20) catalyzes the conversion of a D- amino acid substrate, *meso*-diaminopimelic acid (DAP) to *L*-lysine (Figure A1). DAP is an important component of the cell wall and disruption of its biosynthesis results in cell death due to the absence of cross-linking of peptidoglycan polymers. As evidence of the importance of DAP in bacterial cell wall stability, most diaminopimelate auxotrophs undergo lysis in a complex medium unless diaminopimelate is present at low concentrations (Meadow., 1957). Lysine, besides being important in cell wall synthesis of Gram positive cocci, is mainly important for protein synthesis. Since inhibitors of amino acid biosynthesis are potentially useful therapeutically, inhibition of lysine biosynthesis is a prime site for the design of rationally based antibiotics. The idea of inhibition of lysine biosynthesis as antibacterial target is further accentuated by the fact that mammals lack this metabolic pathway and require *L*-lysine as an essential dietary supplement (Jouanneau., 1985). Hence inhibitors of DapDC could prove to be selectively lethal to bacteria only. In spite of renewed interest in the potential antibacterial properties of inhibitors of this enzyme, lack of structure based mechanistic studies has been a stumbling block in the design of compounds with strong antibacterial activity. The developed inhibitors of DapDC as antibacterial compounds such as β,γ-unsaturated meso-diamimopimelate, α-difluoromethyl derivative of DAP,

193

hydrazine and phosphonate derivatives of diaminopimelate, all exhibit only a modest antimicrobial action (Girodeau, 1985; Kelland, 1985; Song, 1994). In this respect, our structure of *E.coli* DapDC (Levdikov et al., 2002, submitted) will broaden the scope of design of very potent inhibitors of DapDC.

DapDC, requiring pyridoxal phosphate (PLP) as a coenzyme, is the only known amino acid decarboxylase that acts on the carbon atom with *D* configuration in the substrate. As per the mechanism of all the PLP-dependent decarboxylases, the first step in DapDC catalyzed reaction is the transimination reaction leading to the displacement of the internal aldimine linkage between PLP and Lys-54 by the external aldimine linkage between the substrate, DAP and PLP. Decarboxylation occurs only when the carboxyl-α-carbon bond is perpendicular to the plane of the conjugated π system. The carbanion generated following decarboxylation is stabilized by the formation of the quininoid tautomer. The charge neutralization is assisted by an ionic interaction between the pyridinium nitrogen and Glu-268. In the final step of the reaction, the quininoid tautomer is protonated with overall stereochemical inversion, leading to the formation of the *L*-lysine-PLP imine, which is transiminated before the product release (Dunathan, 1974). Thus the reaction mechanism of DapDC is very similar to the other group members of group III decarboxylases like eukaryotic ornithine decarboxylases (ODC) and arginine decarboxylases (ADC) (Grishin et al., 1995). ODCs which have been the target of cancer therapeutics, catalyze the conversion of ornithine to putrescine, the first committed step in the biosynthesis of polyamines (Tabor & Tabor, 1984). Many structures of ODCs like the human ODC, truncated mouse ODC (Almrud et al., 2000 and Kern et al., 1999), the *Trypanosoma bruceii bruceii* ODC with and without the drug difluoromethylornithine

194

(Grishin et al., 1999) as well as the mutant structure (Jackson et al., 2000) have been published. ADCs are alternative sources of putrescine (Tabor & Tabor, 1984 and Slocum et al., 1984) and no structural information is available about this protein although it has been crystallized (Rodriguez et al., 1994). DapDC has high structural and sequence similarities to these *L*-amino acid decarboxylases and therefore it is only logical to deduce that the difference in the reaction stereospecificity between DapDC and other group III decarboxylases must result from subtle changes in the residues involved in catalysis. These residues can be identified and assigned a definite functional role only after comparison with the corresponding amino acid residues in other DapDCs, ODCs and ADCs.

Here we report a complete analysis of the conserved amino acid residues in all the available sequences of diaminopimelate decarboxylases that have some sequence identity to *E.coli* DapDC. We have performed a multiple alignment of the DapDC sequences and tried to assign a particular functional or structural role to all the conserved amino acids after comparing and analyzing its corresponding amino acids position and role in *E.coli* DapDC. We have also compared the sequences of DapDCs with other group III *L*-amino acid decarboxylases like ODCs and ADCs so as to provide a logical explanation for the differences in the substrate selectivity and stereospecificity. Our primary goal is to provide a lucid picture of all the critical residues playing an important role in the active site and spatial orientation of the DapDCs, which could provide a new insight in the prevalent subject of numerous mechanism-based inhibitor studies or open up a new realm of structure-based inhibition studies.

RESULTS

DapDCs, ODCs and ADCs have been reported to be members of the superfamily of alanine racemase (Grishin et al., 1995) and the structure of DapDC confirmed the fold (Levdikov et al., 2002, submitted). Like the previously determined structures of alanine racemase and ODC, DapDC consists of an N-terminal β/α TIM like barrel containing the PLP binding site and a C-terminal Greek key β-barrel that provides the active site for catalysis and interactions for oligomerization (Figure A2) (Levdikov et al., 2002, submitted). Keeping this structure in mind, we performed a multiple alignment of the representative DapDCs, ODCs and ADCs (Figure A3). The ADCs exist as active tetramers (Wu et al., 1973) and the DapDCs and the ODCs are both dimers (Grishin et al., 1999; Levdikov et al, 2002 submitted). The ADCs have long insertions, which might play a role in tetramerization of the protein. These long insertions cause sequence alignment program to introduce gaps in the middle of secondary structure elements. Therefore the multiple alignment obtained from CLUSTALW had to be adjusted "by eye" using the computer program 'Bioedit' to force gap positions outside a secondary structure element. Sequence similarity breaks down at the C-termini and the N-termini so the residues at both the ends of all the sequences are removed for clarity.

All the DapDC sequences obtained from multiple bacterial species from an initial blast search of *E.coli* DapDC were multiple aligned (Figure A4). When the conserved residues from this alignment were compared with those of the global alignment of the DapDCs with the ODCs and the ADCs, a very clear picture arose as to what structural or functional roles do the conserved residues play, which residues could have a role in the stereochemical selection of the substrate and which residues may need to be considered

196

in the design of broad-specificity inhibitors of DapDC. The conserved residues were divided according to their functional role: residues involved in cofactor binding, substrate binding, and maintenance of structural integrity.

Residues involved in cofactor binding. Several residues characteristic of a PLP binding enzyme are conserved (Figure A5). The lysine involved in Schiff base formation is Lys-54. This lysine is conserved in all the DapDCs, ODCs and the ADCs. Mutation of this residue compromises the positioning of the carboxyl group of the substrate for efficient catalysis and thereby decreases the rate of decarboxylation step by greater than 10^4 fold in the ODCs (Osterman et al., 1999).

The main chain nitrogen of Gly-227, a part of the absolutely conserved GGG sequence found in all the group III decarboxylases, hydrogen bonds to the phosphate of the cofactor. While one of the glycine residues of the GGG sequence interacts with the cofactor, the tertiary structure is locally confined and requires an extended structure with no side chains, explaining the sequence conservation. Gly-270 and Tyr-378, both of which are conserved in all the DapDCs, ODCs and ADCs contribute a hydrogen bond to the phosphate of the cofactor. Arg-271, again conserved in all three decarboxylases, forms an ionic bond with the phosphate of PLP. The corresponding arginine in the _Trypanosoma bruceii bruceii_ ODC (tbODC) is Arg-277 and this is also observed to have charge interaction with the cofactor's phosphate. But the Arg-277 is at a distance of 2.8 Å from Asp-332 (conserved in all ODCs) and thus is involved in a salt bridge with the aspartic acid (Grishin et al., 1999). Indirectly it thus is involved in orientation of the substrate rather than simply interaction with the cofactor. DapDC has no acidic residue corresponding to ODC's Asp-332; instead it has Arg-307, which is involved in substrate

binding (discussed later). Proline-269, which is very conserved in all the DapDCs and the ODCs, orients Gly-270 and Arg-271 for interaction with the cofactor. Gly-376 and Ala-377, conserved in all the group III decarboxylases, bridge the loop between the β- strands S20 and S21 and orient Tyr-378 for its interaction with the phosphate of the cofactor.

The role of His-191, which is again absolutely conserved in all three decarboxylases, is not well defined. Mutation studies in ODC indicate that this conserved histidine plays an important role in catalysis (Tsirka and Coffino, 1992; Tsirka et al., 1993). In the structural study of tbODC, as far as its role in the cofactor binding site is concerned, it has been speculated to function as a "lid" to the PLP binding pocket or it may contribute to the tautomeric state of PLP (Grishin et al., 1999). In the DapDCs, His-191 has been speculated to have an important role in the substrate interaction and substrate orientation favorable for decarboxylation and this is discussed in the substrate-binding section.

Arg-142, very conserved in the DapDCs, the ODCs and the ADCs, interacts with the phenolic oxygen of the cofactor through a water molecule. It also hydrogen bonds to His-191 mainchain carbonyl, thereby helping to orient the histidine residue. In tbODC, it has been suggested that this residue may have an influence on the tautomeric state of the PLP cofactor (Grishin et al., 1999). Glu-268 in the *E.coli* DapDC and also conserved in all the DapDCs has a very well defined role. Like all the other PLP-dependent enzymes that have an acidic residue interacting with the pyridine nitrogen, this glutamic acid residue interacts with the pyridinium nitrogen, lowers the pKa of the internal aldimine, and stabilizes the protonated state of the cofactor as well as the carbanion intermediate produced after decarboxylation. This glutamic acid, which is well conserved in all the

ODCs too, has been mutated and was found to lower the k_{cat} by 50-fold (Osterman et al., 1995). Our sequence alignment shows that the corresponding glutamic acid residue in the ADCs (Glu-360 in ADC_ECOLI and Glu-374 in ADC_ARATH) could have an identical role since it is conserved.

Asp-73 and Glu-79 is a highly conserved pair of residues in most of the DapDCs. They are referenced as a pair because they are absolutely conserved in the ODCs and have been found to form interactions with the conserved lysine when it swings away from the C4 atom of PLP breaking the Schiff base bond, giving access to the substrate (Jackson et al., 2000). But according to our speculation, for lysine-54 in the DapDC to interact with these residues, it has to swing a substantial distance. Instead our analysis shows that Asp-73 is more appropriate as a contributor to electronegative sink associated with the pyridinium nitrogen of the PLP cofactor. Glu-79 forms a hydrogen bond with Val-75 and water molecules and this hydrogen bond could help in the packing of helix H4 against beta strand S3. In the ADCs Glu-79 is absolutely conserved and the residue corresponding to Asp-73 is a glutamic acid (Glu-148 in ADC_ECOLI and Glu-159 in ADC_ARATH)

Phe-51, Ala-52, Ala-55 and Ile-80 are not conserved but are replaced by similar hydrophobic residues in all of the DapDCs. They help in maintaining the hydrophobic environment for the positioning of the pyridine ring as suggested by the placement of similar residues on one side of the pyridine ring in alanine racemase (Shaw et al., 1997).

Residues involved in substrate binding. The substrate is bound at the dimer interface with residues from both the subunits contributing to the enzyme-substrate interactions (Figure A6). Cys-342'in DapDC is at the tip of a hairpin loop formed by the

199

residues 340-346 (prime will be used to denote residues from the two-fold related subunit with respect to the subunit containing the Lys-54 to which the PLP cofactor is bound). This positioning is very similar to the Cys-360' of tbODC wherein it is placed at the tip of the hairpin formed by the residues (358-364) (Grishin *et al.*, 1999). But there is a difference in the orientation of the side chain of both the cysteine residues. This difference may be due to the movement of the cysteine residue in tbODC while interacting with the inhibitor, α-difluoromethylornithine (Grishin *et al.*, 1999). Various functions have been assigned to the Cys-360' in tbODC; it could influence the orientation of the substrate's leaving carboxylate group or it could actually act as a general acid to protonate the C_α carbon after decarboxylation (Grishin *et al.*, 1999). Cys-360' has also been demonstrated to be important in preventing the side reaction of decarboxylation-dependent transamination reaction by preventing the reprotonation at C4' of PLP (Jackson *et al.*, 2000). In DapDC, for the Cys-342' to play these roles, it has to move significantly from its position. It definitely interacts with the phenolic oxygen of PLP through hydrogen bonding. Thus even though the exact role of Cys-342' is not known in the catalytic mechanism of DapDC, it definitely has a very important role to play because it is not only conserved in all the DapDCs, but also in the ODCs and the ADCs. The biochemical evidence supporting the role of cysteine in the catalysis was given by the mutagenesis studies of Cys to Ala, which decreased the k_{cat} by 50-fold (Swanson *et al.*, 1998).

Glu-343' located on the hairpin loop is conserved in all the DapDCs and forms a hydrogen bond with the distal amino group of DAP. Tyr-311, again conserved in all the DapDCs, ODCs and ADCs, forms a hydrogen bond with the DAP through a water

molecule. Arg-271 and Arg-307 form a positively charged area for the carboxylate group of the substrate. Thus the substrate-binding pocket formed by these four residues is very specific for DapDC and are critical for substrate discrimination. The substrate-binding pocket in the tbODC contains of two acidic residues Asp-361 and Asp-332, which again are conserved only in the ODCs. The sequences of all the three decarboxylases in this substrate binding pocket does not align well and so it is difficult to accurately predict the substrate binding residues of the ADCs, but they do have conserved arginines (Arg-363 in *E. coli*-ADC and Arg-377 in Arab-ADC), aspartates (Asp-529 in *E. coli* ADC and Asp-525 in Arab-ADC) and tyrosines (Tyr-486 in *E. coli*-ADC and Tyr-481 in Arab-ADC) in this region.

The residues that are not directly involved in substrate interaction, but help in orienting the substrate binding residues, are Asp-346', Asn-387' and Tyr-232. The buried charge of Asp-346' probably plays a structural role by positioning the active site residues Cys-342' and Glu-343' that are placed at the tip of the hairpin loop (Grishin *et al*., 1999). The corresponding residue in the tbODC, Asp-364 is said to form a salt bridge with Lys-169 and stabilize the dimer interface. In the *E.coli* DapDC, Lys-163 is at a distance of 4.14 Å from the Asp-346' and so it is unlikely that this interaction can occur. Asp-346'is very critical in holding the hairpin loop in position and it does so by hydrogen bonding to Pro-340', Leu-341' and Val-347'. It also bridges to the loop connecting H10 and S13 by hydrogen bonding to Phe-304' (Figure A7). To further strengthen the structural role of the Asp-346' residue, mutation of this aspartate residue in the tbODC to Ala resulted in severe loss of activity (Osterman *et al*., 1995). Asn-387', by hydrogen bonding with the mainchain nitrogens of Cys-342', helps in the orientation of the cysteine residue for

interaction with the substrate. Asn-387' and Asp-346' are very good candidates for mutagenesis study in DapDC as they are absolutely conserved in all the three decarboxylases. Tyr-232, conserved in most of the DapDCs, is very specific for the DapDCs because it hydrogen bonds with Arg-271 and probably helps to orient the arginine residue for interaction with the carboxylate of the substrate.

His-191 in DapDC is speculated to play a role in the catalytic mechanism based on the studies performed in dopaDC and orotidine 5'-monophosphate (OMP) decarboxylase. In the protonated form, it may interact via a hydrogen bond with the leaving carboxyl group of the substrate and thus stabilize the transition state. This hypothesis is strengthened by the findings that the corresponding histidine residue in the dopaDC (His-192) forms a hydrogen bond to the carboxylate group of carbidopa (Burkhard *et al*., 2001) and the histidine to alanine mutant of dopaDC loses catalytic activity (Ishi *et al*., 1996), whereas the histidine to glutamine mutant does not (Tramonti *et al*., 1998). This residue was also speculated to ensure protonation of the quinonoid Cα (Jansonius, 1998) but this was disputed by the mutation study carried out in glutamate decarboxylase where the analogous histidine (His-167) to aspartate mutant was found to be catalytically active (Tramonti *et al*., 1998). Another possible mechanism is that the unprotonated histidine may repel the carboxyl group and assist it to leave as observed in the OMP (Appleby *et al*., 2000). These roles could be better established by determining the pKa of His-191.

According to our *E.coli* DapDC structure, Tyr-386' is a very interesting residue from the standpoint of stereochemical selectivity. The X-ray structure shows that the hydroxyl group projects towards the *D*-amino acid-binding site and thus it is

hypothesized that it probably prevents *L*-carboxyl groups from interacting with the cofactor. The ODCs have a phenylalanine and the ADCs have a histidine (His-575 in ADC_ECOLI and His-573 in ADC_ARATH) in the corresponding position and thus leave adequate space for the *L*-carboxylate binding. Mutating Tyr to a Phe and then looking for a change in substrate specificity (*L,L*-DAP as substrate) might corroborate this hypothesis. Another role that could be assigned to Tyr-386' is that of a catalytic acid for the protonation reaction. Being in close proximity to a protonated Lys-54, its pKa might be shifted adequately enough to donate its proton (Levdikov *et al.*, 2002, submitted).

Another subtle, but very important structural feature that can affect stereochemical selectivity, is the loop formed by the residues His-191 through Asp-197. The third and the fourth residues in the loop *i.e.* Gly-193 and Ser-194, are very conserved in all the three decarboxylases. In tbODC, they correspond to Gly-199 and Ser-200. This flexible Gly-199 in the tbODC takes a turn and positions Ser-200 in such a way that the hydroxyl group may sterically prevent D-amino acid carboxyl groups from orienting properly or prevent catalysis by stabilizing the leaving group via hydrogen bond interactions. In contrast, Gly-193 in DapDC orients Ser-194 away from the substrate and bridges the loop between the β-strand S8 and α-helix H8 by allowing the formation of a hydrogen bond between Ser-194 and His-200. Thus in DapDC, Ser-194 causes no steric block to the carboxylate group of the D-amino substrate and hence the selectivity.

Residues involved in the structural orientation. Several residues are involved in stabilizing the dimeric complex of DapDC. Ser-76 hydrogen bonds with the Asn-387' mainchain oxygen and helps in the monomer/monomer interface stabilization. This serine

is largely conserved in all the three decarboxylases. Another serine that plays a similar role is Ser-123. Placed on the turn between β-strand S5 and α-helix H6, Ser-123 hydrogen bonds to Lys-288' and the mainchain carbonyl of Asp-102. In some DapDCs, this serine is replace by glutamine, another hydrogen bond donor. Lys-163 contributes hydrogen bonds between the monomers by interacting with the mainchain oxygens of Leu-341' and Gly-339'. It is also conserved in a majority of the DapDCs except that some have another basic residue, arginine instead. The corresponding lysine in the tbODC, Lys-169 is predicted to form a salt bridge with Asp-364 (Asp-346' of DapDC) of the other subunit. But the distance between them in the DapDC is too large for this interaction to occur. This lysine is very conserved in the ODCs and the ADCs. Another lysine stabilizing subunit interactions is Lys-288. It hydrogen bonds with carbonyls of Lys-163', Gly-165' Gly-122' and Ser-123'. The Lys-294 in tbODC forms a salt bridge with Asp-134 from the other subunit. Surprisingly there is no acidic residue corresponding to the Asp-134 in the *E.coli* DapDC, whereas all the other DapDCs have an aspartate or a glutamate in this position. Gly-165, which is absolutely conserved in all the three decarboxylases is placed on a turn before β-strand S7 and hydrogen bonds with Lys-288'. While the tbODC has an aromatic acid zipper passing through the two-fold axis (Grishin *et al*., 1999), DapDC has several methionine residues at the interface (Met-306, Met-310 and Met-382) forming a hydrophobic cluster. Of these methionines, only Met-382 is conserved in the DapDCs.

A number of glycines that are conserved are present at the turns. Gly-187 is present in the sharp bend that the β-strand S8 takes to be tightly packed against strands S9 & S6. This glycine is present in the ODCs but it is substituted by serine in some of the

DapDCs. Gly-193, absolutely conserved in all the three decarboxylases, is placed in the

turn between the β-strand S8 and α-helix H8. The loop between the strand S7 and S8

contains Thr-157 and Gly-158 that are fairly conserved in all the DapDCs. Thr-157 holds

the loop against the TIM barrel through hydrogen bonding and Gly-158 stabilizes the turn

between S7 and S8. Gly-367, very conserved in all the three decarboxyalses, stabilizes

the turn between β-strands S19 and S20. Asp-299, again conserved in all the DapDCs,

the ODCs and the ADCs, stabilizes the turn between β-strands S11 and S12 by hydrogen

bonding to Thr-282 hydroxyl group and Gln-283 backbone nitrogen. Asp-268, substituted

by glutamates in some DapDCs, bridges the loop between S19 and S20 by hydrogen

bonding to Lys-365 mainchain nitrogen. It also interacts with the residues on beta strand

S14 via a water network. Arg-408, very conserved in all the DapDCs locks down the N-

terminus to the C-terminus by hydrogen bonding with the oxygen of Pro-27, which again

is also very conserved.

DISCUSSION

When the *E.coli* DapDC sequence was analyzed by BLAST against the

SWISSPROT database of proteins, a number of sequences of pathogenic microorganisms

were found including those of *Haemophilus influenzae (*Fleischmann *et al*., 1995) *,*

Pseudomonas aeruginosa (Stover *et al*., 2000)*, Vibrio cholerae* (Heidelberg *et al*., 2000)*,*

Neisseria meningitides (Parkhill *et al*., 2000)*, Campylobacter jejuni* (Parkhill *et al*.,

2000), *Helicobacter pylori* (Alm *et al*., 1999), *Mycobacterium tuberculosis* (Cole *et al*.,

1998)*, and Mycobacterium leprae* (Cole *et al*., 2001). These sequences have a high

homology to the *E. coli* DapDC and so its complete sequence and structure-function

analysis could throw light on the structural features of the other DapDCs from a therapeutic standpoint.

The cofactor-binding site of DapDC is similar to the other PLP binding sites in the alanine racemase superfamily. The only substantial difference between the alanine racemase and DapDC is the residue that is involved in interaction with the pyridinium nitrogen. DapDC being involved in releasing a keto group, has a negatively charged residue (Glu-268) near the pyridinium nitrogen, whereas the alanine racemase has a positively charged arginine residue, Arg-219 (Shaw *et al*., 1997). An interesting study carried out by Ophem *et al*., demonstrated that when the Glu-177 interacting with the pyridinium nitrogen in D-Amino acid transaminase was mutated to a positively charged lysine, its reaction specificity changed to that of a racemase (Ophem *et al*., 1999). This fact can be very important in the case of DapDC because it was found in the DapDC/*L*-lysine complex that *L*-lysine had racemised to *D*-lysine (Levdikov *et al*., 2002, submitted). So if DapDC can already play the role of a racemase, albeit inefficiently, what will be the effect of mutating Glu-268 to a positively charged residue? Will its racemase action be accentuated? This mutation will be interesting to study because very few mutagenesis experiments that alter the reaction specificity as opposed to substrate specificity have been reported.

In the substrate binding pocket, there appears to be potentially significant local differences among the DapDCs and the ODCs as well as among the DapDCs themselves. An interesting difference lies in the conserved sequence H[191]IGS of the DapDCs and the ODCs. In the ODCs the serine residue interacts with the PLP cofactor whereas in the *E.coli* DapDC this serine (Ser-194) is steered away from the active site to play a

206

structural role. However, in other bacterial DapDCs, this serine could have the same orientation as that of the ODCs and thus complicate the development of broad-spectrum antibiotics.

Some residues are not essential for the catalysis, but are considered to contribute to the activity through conformational or other effects. An interesting mutation of this type that has been identified in the *E.coli* DapDC is the Ser-384-Phe mutation. This mutant was found to require betaine for proper folding and enzymatic activity (Bourot *et al.*, 2000). The S384F mutation is rather remarkable since both DDC and the ODC have hydrophobic residues in this region and it is not obvious why a serine to phenylalanine mutation should show this phenotype. A phenylalanine should be stable at the monomer/monomer interface and in fact increase association of monomers to form the active dimer. However this would increase the hydrophobicity of the local environment that could result in incorrect folding that is corrected by betaine. Some additional volume would be required to accept a phenylalanine in that position, and it would be difficult to imagine betaine also interacting directly at the subunit interface. This serine is fairly conserved between the DapDCs and the ODCs but not the ADCs and so it must be definitely playing an important role. It could either help in the oligomerization of the protein or in orienting the Tyr-386 residue for substrate interaction.

In the structure of the *E. coli* DapDC, several water molecules are found in the cofactor-binding site. These water molecules form a pocket around the cofactor binding sites and interact with many residues, which are very conserved. The important residues around this pocket that could play a role in drug interaction are Asp-73, Val-75, Glu-79, Arg-142, His-191, Gly-225, Gly-226, Gly-277, Pro-269, Gly-270, Arg-271 and Tyr-378.

Besides these residues which could be considered in the design of mechanism based inhibitors, there are other residues like Asp-346, which could be very critical in maintaining structural integrity and hence important for consideration in the study of structure based inhibitors.

In conclusion, the roles of the catalytically and the structurally important amino acid residues probed in this study by the sequence alignment provides a molecular basis for the development of potent inhibitors of diaminopimelate decarboxyalses with better pharmacological characteristics.

METHOD

Accession codes: The PDB accession codes of the crystal structure of E.coli Diaminopimelate decarboxylase with and without D/L-lysine are 1KNW and 1KO0 respectively.

Multiple sequence alignment: Homology search of the SwissProt and PDB databanks using the BLAST algorithm (Altschul *et al*., 1990) was carried out using the server: http://www.ncbi.nlm.nih.gov/BLAST. All the homologous sequences were compiled together in the Fasta format. These sequences were then aligned using the computer program CLUSTALW (Higgins & Sharp, 1988) and PILEUP (GCG package; Devereux *et al*., 1984). The alignment was then adjusted 'by eye" using the alignment editor 'BioEdit'.

HOOC—*H / R / NH$_2$ ... H / S / NH$_2$—COOH → **DDC** → *H / NH$_2$ ... H / S / NH$_2$—COOH

meso-Diaminopimelic acid   **L-Lysine**

Figure A1: The reaction scheme of meso-diaminopimelate decarboxylase

Figure A2: Schematic representation of the secondary structure element in the three-dimensional structure of E.coli diaminopimelate decarboxylase

```
Sec_Struct       sshhhhhhhh~~~~~~~~~~~~~~~~~ssssss~~~~~hhhhhhhhhh~~~~~~ssss~~~hhhhhhhhhh~~~~~~~~~~~~~~~~~~~ssssss~~~~
DapDC_ECOL   32  DAQIIRRQIAALKQF~~~~~~~~~~~~DVVRFAQKACSNIHILRLMREQG----VKVDSVSLGEIERALAAGYN--------P-QTHPDDIVFTADVID 105
DapDC_BACS   39  DVALIRERAKSFKQAFISAGLKA------QVAYASKAFSSVAMIQLAEEEG----LSLDVVSGGELYTAVAAGFP--------AERIHFHGNNKSREELR 120
DapDC_MYCT   46  DEDDFRSRCRETAAAFGSG---A------NVHYAAKAFLCSEVARWISEEG----LCLDVCTGGELAVALHASFP--------PERITLHGNNKSVSELT 124
DapDC_METJ   40  SEEQIKINYNRYIEAFKRWEEET--GKEFIVAYAYKANANLAITRLLAKLG----CGADVVSGGELYIAKLSNVP--------SKKIVFNGNCKTKEEII 125
ODC_HUMAN    44  DLGDILKKHLRWLKALPRV---------TPFYAVKCNDSKAIVKTLAATG----TGFDCASKTEIQLVQSLGVP--------PERIIYANPCKQVSQIK 121
ODC_TRYPB    42  DLGDIVRKHETWKKCLPRV---------TPFYAVKCNDDWRVLGTLAALG----TGFDCASNTEIQRVRGIGVP--------PEKIIYANPCKQISHIR 119
ODC_MOUSE    44  DLGDILKKHLRWLKALPRV---------TPFYAVKCNDSRAIVSTLAAIG----TGFDCASKTEIQLVQGLGVP--------AERVIYANPCKQVSQIK 121
ADC_ECOLI    93  PQILQHRLRSINAAFKRARESYGYNGDYF~LVYPIKVNQHRRVIESLIHSGEP--LGLEAGSKAELMAVLAHAGM--------TRSVIVCNGYKDREYIR 181
ADC_SYNY     84  FSDILADRLERLNSCFAKAIARYNYPNTYQAVYPVKCNQQRHLVEALVRFGQTSQCGLEAGSKPELMIALATLPPPLDRQDKHTKPLIICNGYKDQDYLE 183
ADC_ARATH   101  FPDVLKNRLECLQSAFDYAIQSQGYDSHYQGVYPVKCNQDRFIIEDIVEFGSGFRFGLEAGSKPEILLAMSCLCKGN------PEAFLVCNGFKDSEYIS 194

Consensus                                  :  K                G              :  :  E                                    K
                                            54               73 76 79
```

```
Sec_Struct       hhhhhhhhhh~~sssss~hhhhhhhhhh~~~~~~ssssssss~~~~~~~~~~~~~~~~~~~~~sss~~~~hhhhhhhhhh~~~~~~sssssss~~~~~~
DapDC_ECOL  106  QATLERVSELQIPVNAGSVDMLDQLGQV~~~~SPGHRVWLRVNPGFGHGHSQKTNTG-GENSKHGIWYTDLPAALDVIQRH~~~~HLQLVGIHMHIGSG- 195
DapDC_BACS  121  MALEHRIGCIV~~~~VDNFYEIALLEDLCKETGHSIDVLLRITPGVEAHTHDYIITG-QEDSKFGFDLHNGQTERAIEQVLQSE~HIQLLGVHCHIGSQI 214
DapDC_MYCT  125  AAVKAGVGHIV~~~~VDSMTELERLDAIAGEAGIVQDLVRLTVGVEAHTHEFISTA-HEDQKFGLSVASGAAMAAVRRVFATD~HLRLVGLHSHIGSQI 218
DapDC_METJ  126  MGIEANIRAFN~~~~VDSISELILINETAKELGETANVAFRINPNVNPKTHPKISTG-LKKNKFGLDVESGIAMKAIKMALEME~YVNVVGVHCHIGSQL 219
ODC_HUMAN   122  YAANNGVQMMT~~~~FDSEVELMKVARA~~~~~PKAKLVLRI~~~~~~~~ATDDSKAVCRLSVKFGAT-LRTSRLLLERAK~~ELNIDVVGVSFHVGSGC 201
ODC_Tryp    120  YARDSGVDVMT~~~~FDCVDELEKVAKT~~~~HPKAKMVLRI~~~~~~~~STDDSLARCRLSVKFGAK--VEDCRFILEQAK~-KLNIDVTGVSFHVGSGS 200
ODC_MOUSE   122  YAASNGVQMMT~~~~FDSEIELMKVARA~~~~HPKAKLVLRI~~~~~~~~ATDDSKAVCRLSVKFGAT--LKTSRLLLERAK~ELNIDVIGVSFHVGSGC 202
ADC_ECOLI   182  LALIGEKMGHKVYLVIEKMSEIAIVLDEAERLNVVPRLGVR-ARLRSQGSGKWQSSG-GEKSKFGLAATQVLQLVETLR-~EAGRLDSLQLLHFHLGSQM 277
ADC_SYNY    184  TALLAKRLGHRPIIIIEQLRELEWVLHISQQLNIKPMLGVRARLSCQSLKSSEISSGNGDRAKLGLTMPDIVTVIHRLE-~ENNCLDCLKMLHFHLGTQV 281
ADC_ARATH   195  LALFGRKLELNTVIVLEQEELDLVIDLSQKMNVRPVIGLR-AKLRTKHSGHFGSTS-GEKGKFGLTTVQILR-VVRKLS~QVGMLDCLQLLHFHIGSQI 290

Consensus        A                E                   R                  KFG                    G H H:GS
                                                        142              163-165                      191-194
```

```
Sec_Struct       ~~hhhhhhhhhhhhhhhhhh~~~~~~~~~~sssss~~~~~~~~~~~~~~~~~~~~~hhhhhhhhhhhhhhhhhh~~~~~ssssss~~~~~~~ssss
DapDC_ECOL  196  VDYAHLEQVCGAMVRQVIEFG-~~~~---QDLQAISAGGGLSVPYQQGEE-------AVDTEHYYGLWNAAREQIARHLG-HPVKLEIEPGRFLVAQSGVL 280
DapDC_BACS  215  FDTAGFVLAAEKIFKLDEWRDSYS----FVSKVLNLGGGFGIRYTEDDE-------PLHATEYVEKIEAVKENASRYGFDIPEIWIEPGRSLVDGAGTT 303
DapDC_MYCT  219  FDVDGFELAAHRVILLRDVVGEFGPEKTAQIATVDLGGGLGISYLPSDD-------PPPIAELAAKLGTIVSDESTAVGLPTPKLVVEPGRAIAGPGTIT 311
DapDC_METJ  220  TDISPFIEETRKVMFVVELKEEG-----IEIEDVNLGGGLGIPYYKDKQ-------IPTQKDLADATINTMLKYKDKV--EMPNLILEPGRSLVATAGYL 305
ODC_HUMAN   202  TDPETFVQAISDARCVFDMGAEVG---FSMYLLDIGGGFPGS---EDV-------KLKFEEITGVINPALDKYFPSD--SGVRIIAEPGRYYVASAFTL 285
ODC_Tryp    201  TDASTFAQAISDSRFVFDMGTELG----FNMHILDIGGGFPGT--RDA-------PLKFEEIAGVINNALEKHFPPD--LKLTIVAEPGRYYVASAFTL 284
ODC_MOUSE   203  TDPDTFVQAVSDARCVFDMATEVG----FSMHLLDIGGGFPGS---EDT-------KLKFEEITSVINPALDKYFPSD--SGVRIIAEPGRYYVASAFTL 286
ADC_ECOLI   278  ANIRDIATGVRESARFYVELHKLG----VNIQCFDVGGGLGVDYEGTRSQSDCSVNYGLNEYANNIIWAIGDACEENGL-PHPTVITESGRAVTAHHTVL 372
ADC_SYNY    282  SDIALIKEAMREASQLYVELVKLG----AKMRYLNVGGGLAVDYDGSKTN-YPASKNYNMQNYANDIVAAIQDACELGQVSPPILVSESGRAIMAHQSVL 376
ADC_ARATH   291  PSTALLSDGVAEAAQLYCELVRLG----AHMKVIDIGGGLGIDYDGSKSGESDLSVAYSLEEYAAAVVASVRFVCDQKSVKHPVICSESGRAIVSHHSVL 386

Consensus                             GGG     Y                          :                    EPGR   :
                                     227-229  232                                        268-271
```

Figure A3: Sequence alignment of DapDCs, ODCs and ADCs related to the sequence of E.coli DapDC

211

```
Sec_Struct       ssssssssss~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ssssssss~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
DapDC_ECOL   281 ITQVRSVKQMGS-~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~-RHFVLVDAGFN---------------------------~~~~~~~~~~~~~~~ 303
DapDC_BACS   304 LYTVGSQKEVPG~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~-VRQYVAVDGGMN---------------------------~~~~~~~~~~~~~~~ 327
DapDC_MYCT   312 LYEVGTVKDVDVSAT~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~-AHRRYVSVDGGMS---------------------------~~~~~~~~~~~~~~~ 339
DapDC_METJ   306 LGKVHHIKETPVT~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~-KWVMIDAGMN---------------------------~~~~~~~~~~~~~~~ 328
ODC_HUMAN    286 AVNIIAKKIV~~~~~~~~~~~~~~~~~~~~~~~~~~~~~LKEQTGSDDEDESSEQTFMYYVNDGVYG---------------------------~~~~~~~~~~~~~~~ 323
ODC_Tryp     285 AVNVIAKKVT~~~~~~~~~~~~~~~~~~~~~~~~~~~~~PGVQTDVGAHAESNAQSFMYYVNDGVYG---------------------------~~~~~~~~~~~~~~~ 322
ODC_MOUSE    287 AVNIIAKKTVW~~~~~~~~~~~~~~~~~~~~~~~~~~~~~KEQPGSDDEDESNEQTFMYYVNDGVYG---------------------------~~~~~~~~~~~~~~~ 324
ADC_ECOLI    373 VSNIIGVERNEYTVPTAPAEDAPRALQSMWETWQEMHEPGTRRSLREWLHDSQMDLHDIHIGYSSGIFSLQERAWAEQLYLSMCHEVQKQLDPQNRAHRP 472
ADC_SYNY     377 VFDVLGSNQTGFSEPHPPDENAHPLLKNLWECYETITAEQYQEQYHDALQLKTEASSLFNFGYLSLTERGQAEQIHWACCRKIFEITRQLEYIPE----- 471
ADC_ARATH    387 IFEAVSAGQQHETPTDHQ-------FMLEGYSEEVRGDYENLYGAAMRGDRESCLLYVDQLKQRCVE-GFKEGSLGIEQLAGVDGLCEWVIK------- 470

Consensus             K                                                 D
                      288                                               299


Sec_Struct       ~~~~~~~~~~hhhh~~~~~sssss~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ssssss~~~~~~~~~ssssss~~~sss~sss~~~~~~~~~~~~sssss
DapDC_ECOL   303 ~~~~~~DLMRPAMYGSYHHISALAADGR---------------SLEHAPTVETVVAGPLCES---GDVFTQQEGGNVETRALP--EVK----AGDYLVL 372
DapDC_BACS   327 ~~~~~~DNIRPALYQAKYEAAAANR------------------IGEAHDKTVSIAGKCCES---GDMLIWD-------IDLP--EVK----EGDLLAV 385
DapDC_MYCT   339 ~~~~~~DNIRTALYGAQYDVRLVSR------------------VSDAPPVPARLVGKHCES---GDIIVRD-------TWVPD-DIR----PGDLVAV 398
DapDC_METJ   328 ~~~~~~DMMRPAMYEAYHHIINCKVK------------------NE-KEVVSIAGGLCES---SDVFGRD-------RELD--KVE----VGDVLAI 384
ODC_HUMAN    323 ~~~~~~SFNCILYDHAHVKPLLQKRP----------------KPDEKYYSSSIWGPTCDG--LDRIVER------CDLP--EMH----VGDWMLF 382
ODC_Tryp     322 ~~~~~~SFNCILYDHAVVRPLPQREP----------------IPNEKLYPSSVWGPTCDG--LDQIVER------YYLP--EMQ----VGEWLLF 381
ODC_MOUSE    324 ~~~~~~SFNCILYDHAHVKALLQKRP----------------KPDEKYYSSSIWGPTCDG--LDRIVER------CNLP--EMH----VGDWMLF 383
ADC_ECOLI    473 IIDELQERMADKMYVNFSLFQSMPDAWG-----IDQLFPVLPLEGLDQVPERRAVLLDITCDSDGAIDHYIDGDIATTMPMPEYD--PENPPMLGF~~~~ 561
ADC_SYNY     472 DFQALDKIMTDIYYVNLSVFQSAPESWS-----LDQLFPILPIHHLNEKPSQRVILADLTCDSDGKIDRFIDLWDVSYLEVHPLEN-DGNPYYLGM~~~~ 561
ADC_ARATH    470 -~-AIGASDPVLTYHVNLSVFTSIPDFWG----IDQLFPIVPIHKLDQRPAARGILSDLTCDSDGKINKFIGGESSLPLHEMDNNGCSGGRYYLGM~~~~ 559

Consensus                  Y                                         C::     D                          G
                         311                                       339------346                    367
```

Figure A3: Sequence alignment of DapDCs, ODCs and ADCs related to the sequence of E.coli DapDC

212

```
Sec_Struct        ss~~~~~~~~~~~~~~~~~~sssss~~~~
DapDC_ECOL   373  HDTGAYGASMSSNYNSRPLLPEVLFD----
DapDC_BACS   386  FCTGAYGYSMANNYNRIPREAVVFVE----
DapDC_MYCT   399  AATGAYCYSLSSRYNMVGRPAVVAVH----
DapDC_METJ   385  FDVGAYGISMANNYNARGRPRMVLTS----
ODC_HUMAN    383  ENMGAYTVAAASTFNGFQRPTIYYVM----
ODC_Tryp     382  EDMGAYTVVGTSSFNGFQSPTIYYVV----
ODC_MOUSE    384  ENMGAYTVAAASTFNGFQRPNIYYVM----
ADC_ECOLI    562  FMVGAYQEILGNMHNLFGDTEAVDVFVFP-
ADC_SYNY     562  FLVGAYQEIMGNLHNLFGDINVVHIATTPQ
ADC_ARATH    560  FLGGAYEEALGGVHNLFGGPSVVRVLQS--

Consensus            GAY        YN
                   376-378    386-387
```

Figure A3: Sequence alignment of DapDCs, ODCs and ADCs related to the sequence of E.coli DapDC

The organisms and their respective accession numbers are ECOL – E.coli (sp-P00861), BACS - Bacillus subtilis (sp-P23630),

MYCT - Mycobacterium tuberculosis (sp-P31848), METJ - Methanothermobacter thermautotrophicus (sp-Q58497),

ODC_Human (sp-P11926) , ODC_TRYPB – Trypanosoma bruceii bruceii (tr-Q9TZZ6), ODC_Mouse (sp-P00860),

ADC_ECOLI – E.coli (sp-P21170), ADC_SYNECH - Synechocystis sp. (strain PCC 6803) (sp-P72587), ADC_ARATH –

Arabidopsis Thaliana (sp-O23141).

The important amino acids discussed in the paper are bold in the consensus and numbered.

```
Sec_Struct        ~~~sssss~~~~~~hhhhhhhhh~~ssss~~hhhhhhhhh~~~~~~~~~~~~~~sssssss~~~~hhhhhhhhhh~~~sssss~hhhhhhhhhh~~~~~~
DapDC_ECOL   47   ~~~VVRFAQKACSNIHILRLMREQGVKVDSVSLGEIERALAAG~~~~YNPQTHPDDIVFTADVIDQATLERVSEL~QIPVNAGSVDML~DQLGQVSP~~~  135
DapDC_ARCF   50   ~~~~LLYAVKANNNLALMRIIASHGFGADVFSDGELYLASLAG~~~~~~~~FRKDMVLFNGNSKSRKEIEMGVTA~GVKFSVDSLDEL~RTISKIAKEVG  136
DapDC_METJ   64   KEFIVAYAYKANANLAITRLLAKLGCGADVVSGELYIAKLSN~~~~~~~~VPSKKIVFNGNCKTKEEIIMGIEANIRAFNVDSISEL~ILINETAKELG  154
DapDC_HELP   39   ~~~LICYALKANSNLSILSLLAHLESGADCVSIGEIYRALKAG~~~~~~~~IKPYRIVFSGVGKSGFEIEQALKLNILFLNVESFMEL~TTIETIAQSLG  127
DapDC_HELP   39   ~~~LICYALKANSNLSILSLLANLGSGADCVSIGEIQRALKAG~~~~~~~~IKPYKIVFSGVGKSAFEIEQALKLNILFLNVESFMEL~KTIETIAQSLG  127
DapDC_CAMJ   38   ~~~QIFYAVKANSNLSLLQMLANLDSGFDCVSIGEVKRALKAG~~~~~~~~AKAYKIIFSGVGKTKBELRQALEYDILYINLESEAEM~MLLESVAKELN  126
DapDC_NEIM   45   ~~~LVCYAVKANGNLSIIKHFASLGSGFDIVSGGELARVLAAG~~~~~~~~GDAAKTIFSGVGKSEAEIEFALNAGVKCFNMESIPEI~DRIQKVAARLG  133
DapDC_NEIM   45   ~~~LVCYAVKANGNLSIIKHFASLGSGFDIVSGGELARVLAAG~~~~~~~~GDAAKTIFSGVGKSEAEIEFALNAGVKCFNMESIPEI~DRIQKVAARLG  133
DapDC_PSEF   53   ~~~LVCFAVKANSNLGVLNVLARLGAGFDIVSGGELERVLAAG~~~~~~~~GSADKIVFSGVGKTREDMRRALEVGVHCFNIESTDELGSRLQIVAAELG  142
DapDC_PSEA   53   ~~~LVCFAVKANSNLGVLNVLARLGAGFDIVSRGELERVLAAG~~~~~~~~GDPAKVVFSGVGKTRDDMRRALEVGVHCFNVESGEEL~ERLQRVAAELG  141
DapDC_VIBC   54   ~~~LICYAVKANSNLGVLNTLARLGSGFDIVSGELERVLAAG~~~~~~~~GDPSKVVFSGVGKTEAEMKRALQLKIKCFNVESEPEL~QRLNKVAGELG  142
DapDC_HAEI   53   ~~~LICFAVKSCSNIGVLNIMAKLGSGFDIVSQGELERVLAAG~~~~~~~~GDASKVVFSGVAKSREEIMRALEVRIRCFNVESVSEL~KHINQIAGEMG  141
DapDC_ZYMM   56   ~~~LVAFAVKANPSQAILASFAKEGLGADVVSAGEIRRAVHAG~~~~~~~~IPPERIVFSGVGKTABEMRYALEIGIGQFNIESVSEI~EMLAEVATSLG  144
DapDC_AQUA   58   ~~~LICYAVKANFNPHLVKLLGELGAGADIVSGGELYLAKKAG~~~~~~~~IPPERIVYAGVGKTEKELTDAVDSEILMFNVESRQEL~DVLNEIAGKLG  146
DapDC_TM     38   ~~LLPTFAVKANNNPVLLKILREEGFGMDVVTKGELLAAKLAG~~~~~~~~VPSHTVVWNGNGKSRDQMEHFLREDVRIVNVDSFEEM~EIWRELNPEG~  126
DapDC_BACS   61   ~~~QVAYASKAFSSVAMIQLAEEEGLSLDVVSGGELYTAVAAG~~~~~~~~FPAERIHFHGNNKTEBEELRMALEHRIGCIVVDNFYEI~ALLEDLCKETG  149
DapDC_BACAN   1   ~~~~~~~~~~~~~~~~~~RVAREENMSLDVVSGGELYTALQAGF~~~~~~~~PASRIHFHGNNKTEBEIVMALQANIGCFVVDNFLEL~EILHDLAVQHG   73
DapDC_BACM   59   ~~~QVAYASKAFSTVAMIQLAEEEGLSLDVVSGGELYTALVAG~~~~~~~~FPVHKIHFHGNNKSRABELEMALEHQIGCIVVDNFHEL~DLIDSICSEKN  147
DapDC_BACH   59   ~~~QVAYASKAFSCIAMFQLAEEELGLSLDVVSGGELYTAQQAG~~~~~~~~VSMERIHFHGNNKSRABEIEMAVEAGIGCFVVDNFYEL~DLLADICEQKR  147
DapDC_SYNY   86   ~~~QVIYASKAWSCLAVVAIAAQEGLGFDVVSGGELFTTVSALKQLGWDEAEIAEKIYFHGNNKSVQELQEAIAINC~TIIVDNWLEL~ETLTKLAADSG  181
DapDC_MYCT   65   ~~~NVHYAAKAFLCSEVARWISEEGLCLDVCTGGELAVALHAS~~~~~~~~FPPERITLHGNNKSVSELTAAVKAGVGHIVVDSMTEI~ERLDAIAGEAG  153
DapDC_MYCL   90   ~~~NVHYAAKAFLCTEIARWIDEEGLSLDVCSGGELAVALHAG~~~~~~~~FPPERISLHGNNKSVAELKDAVKAGVGIVLDSTTEI~ERLDAIAGEAG  178
DapDC_MYCS   90   ~~~VVHYAAKAFLCSEIARWVDEEGLSLDVATGGELAVALHAG~~~~~~~~FPASRITVHGNNKSVAELTAAVQAGVGHVVLDSEIEI~ERLDQIAGAAG  178
DapDC_CORG   68   ~~~NVHYASKAFLTKTIARWVDEEGLALDIASINELGIALAAG~~~~~~~~FPASRITAHGNNKGVEFLRALVQNGVGHVVLDSAQEL~ELLDYVAAGEG  156
DapDC_ACTP   74   ~~~EVVFAGKALPCREVYRWVADEGLSLDVCSAGELAIARSVG~~~~~~~~FPAERVLLHGNVKTPEDLKAALGYGVGRVVVDSFDEI~EQLGALAEGP~  161
DapDC_STRC   54   ~~~VIRYAQKACSNLHILRLMREEGVHVDAVSEGEIERALAAG~~~~YRVGGDDEPIVFTADLLNRSTLRRVVEL~GIPVNAGSPQML~DQVGRAAPGHP  145
DapDC_BUCA   47   ~~~VIRFAQKSCSNINILRLMKNKNVKIDAVSLGEIERALLSG~~~~FKPSTN~~EIIFTADILDEETLLKVAKY~KIPVNAGSLDML~KQLGKISPGHH  136
DapDC_DEIR   41   ~~~RIFYAMKANPNLNLLRRYAAAGVGFECVSLGELLRAEAAG~~~~AGGER~~~~MILNGPAKSDAFYAAAAR~LGATIVVDRE~~~~EEVVLLPPGSR  125

Consensus         :A KA              G   D VS GE:  :: AG          : G  K   E   A            :S  E:           A
                  51---55                73  76 79                                                 123
```

Figure A4: Sequence alignment of DapDCs related to E.coli DapDC

214

```
Sec_Struct        ~~ssssssss~~~~~~~~~~~~~~~~~~~~~sss~~~~~~hhhhhhhhhh~~sssssss~~~~~~~hhhhhhhhhhhhhhhhhhhh~~~~~~~~~~~~~~sss
DapDC_ECOL   135  ~GHRVWLRVNPGFGHGHSQKTNTGGENSKHGIWYTD~~~LPAALDVIQRHHLQLVGIHMHIGSGVDYAHLEQVCGAMVRQVIEF~~~~~~~~~GQDLQAI 222
DapDC_ARCF   137  KEVEIAFRVNPDVDPKTHPKIATGLRESKFGIPHE~~MVREAYEMALKLDGVVPVGIHCHIGSQILDLSPFVHALNKVMDIAVDIEKL~~~~~GVELSFV 229
DapDC_METJ   155  ETANVAFRINPNVNPKTHPKISTGLKKNKFGLDVESGIAMKAIKMALEMEYVNVVGVHCHIGSQLTDISPFIEETRKVMDFVVELKEE~~~~~GIEIEDV 249
DapDC_HELP   128  IKARISIRINPNIDAKTHPYISTGLKENKFGVGEK~~EALEMFLWAKKSAFLEPVSVHFHIGSQLSDLEPIIEASQKVAKIAKSLIAL~~~~~GIDLRFF 220
DapDC_HELP   128  IKARISIRINPNIDAKTHPYISTGLKENKFGVEEK~~EALEMFLWAKKSAFLEPVSVHFHIGSQLLDLDPITEASQKVAKIAKSLIAL~~~~~GIDLRFF 220
DapDC_CAMJ   127  LKARISIRVNPNVDAKTHPYISTGLNENKFGVEID~~IARKMYLYAKNSSFLEPVGVHFHIGSQLLDISPIHEAAAIVAKLVRELKAL~~~~~QIDLKFF 219
DapDC_NEIM   134  KTASVSLRVNPDVDAKTHPYISTGLKANKFGIAYA~~DALEAYRHAAQQPNLKIIGIDCHIGSQLTDLSPLVEACERILILVDALAAE~~~~~GIVLEHL 226
DapDC_NEIM   134  KTAPVSLRVNPDVDAKTHPYISTGLKANKFGIAYA~~DALEAYHYAAQQPNLKIIGIDCHIGSQLTDLSPLVEACERILILVDALAAE~~~~~GIVLEHL 226
DapDC_PSEF   143  VRAPISLRVNPDVDAGTHPYISTGLKENKFGIAIA~~DAEDVYVRAAQLPNLEVLGVDCHIGSQLTTLEPFIDALDRLLALVDRLGDC~~~~~GIYLRHI 235
DapDC_PSEA   142  VKAPVSLRVNPDVDAQTHPYISTGLKENKFGIAID~~EAEAVYARAAELDHLEVIGVDCHIGSQLTQLEPFLDALERLLGLVDRLAGK~~~~~GIGIRHL 234
DapDC_VIBC   143  VKAPISLRINPDVDAKTHPYISTGLRDNKFGITFD~~RAAQVYRLAHSLPNLDVHGIDCHIGSQLTALAPFIDATDRLLALIDSLKAE~~~~~GIHIRHL 235
DapDC_HAEI   142  KIAPISLRVNPDVDAHTHPYISTGLKENKFGVSVN~~EAREVYKLASTLPNIKITGMDCHIGSQLTELQPFLDATDRLIVLMEQLKED~~~~~GITLKHL 234
DapDC_ZYMM   145  KKAAVALRINPDVDPHTHAKIATGKADTKFGIAAE~~DALSAYEKLASYPSLKIQGIASHIGSQITDLAPFEAAAERIYEIITALEKA~~~~~GHAIETA 237
DapDC_AQUA   147  KKARIAIRVNPDVDPKTHPYIATGMQKSKFGVDIR~~EAQKEYEYASKLENLEIVGIHCHIGSQILDISPYREAVEKVVSLYESLTQK~~~~~GFDIKYL 239
DapDC_TM     126  ~~VEYFIRVNPEVDAKTHPHISTGLKKHKFGIPLE~~DLDSFMERFRS~~~MNIRGLHVHIGSQITRVEPFVEAFSKVVRASERYG~~~~~~~~~~FEEI 209
DapDC_BACS   150  HSIDVLLRITPGVEAHTHDYITTGQEDSKFGFDLHNGQTERAIEQVLQSEHIQLLGVHCHIGSQIFDTAGFVLAAEKIFKKLDEWRDS~~~~~YSFVSKVL 245
DapDC_BACA    74  KFVNILIRVTPGVEAHTHEYITTGQDDSKFGFGVSNGQAMQAIELALQKSNYNVLGIHSHIGSQIFETAGFVRAIEVLHQFLEEVRERTN~~~~~YVVKVL 169
DapDC_BACM   148  VKTNILLRVTPGIEAHTHDYILTGQEDSKFGFDLQNGQAEKALQIALNSNFVEVLGVHCHIGSQIFDTTGFVLAARKIFEKLKEWKDR~~~~~LSYEPKVL 243
DapDC_BACH   148  ATANILLRITPGVEAHTHDYISTGQEDSKFGFDLVSGQATLAVKKALAVDSIHLLGVHSHIGSQIFETTGFVMAVEKIFEHLVTWRED~~~~~HGFEPSVL 243
DapDC_SYNY   182  APVKIMLRLTPGIECHTHEYIKTGHLDSKFGFDP~~NQLEAVFTYIAQQPSLHCLGLHAHIGSQIFERQ~~~~PHKDLGEVLVQWFTKGLT~YGLPLTEL 274
DapDC_MYCT   154  IVQDVLVRLTVGVEAHTHEFISTAHEDQKFGLSVASGAAMAAVRRVFATDHLRLVGLHSHIGSQIFDVDGFELAAHRVIGLLRDVVGEFGPEKTAQIATV 253
DapDC_MYCL   179  IVQDVLVRLTVGVEAHTHEFIATAHEDQKFGLSVASGAAMAAVRRVFATDNLRLVGLHSHIGSQIFDVAGFELAAHRVIGLLCDIVGEFDPEKTAQLSIV 278
DapDC_MYCS   179  VVQDVLVRVTVGVEAHTHEFISTAHEDQKFGLSLATGAAMNAIRRVFATDNLRLVGLHSHIGSQIFDVAGFEIAAHRVIGLLRDVVAEFGVEKTSQMSIV 278
DapDC_CORG   157  KIQDVLIRVKPGIEAHTHEFIATSHEDQKFGFSLASGSAFEAAKAANNAENLNLVGLHCHVGSQVFDAEGFKLAAERVLGLYSQIHSELG~~~~VALPEL 252
DapDC_ACTP   161  ~~QDLVLRVTPGVDPRTHRAVATGVEDQKFGFSLAGGDALEAVLRVVEQPSLRLVGLHCHVGSQVRHVAVYEEAARRVVGLIAS~CG~~~~~~~VRIEQL 251
DapDC_STRC   145  ~~~~VWIRINPGFGHGHSRKTNTGGEHSKHGIWHEH~~~LEESLALVDRHGLDLVGLHMHIGSGVDYGHLESVCETMVKQVRMAGRD~~~~~~~~~~IRAI 229
DapDC_BUCA   136  ~~~~VWLRINPRFGYGHSKKTNTGGENSKHGIWHP~~~~~EQAIPIIKKYKLKLIGLHMHIGSGVNYLHLKKVSQAMIKYAFQLNQK~~~~~~~~~~ISSI 218
DapDC_DEIR   125  ~~~~VLVRVNPAMTVSTHEHLATGTARSKFGLTPEQ~~~VPGTLAELRDAGHEVLGLHMHIGSAIEQAEDFTAAFARVTELRAHIGG~~~~~~~~~~LSVL 209

Consensus              R  P      TH  I TG   KFG                    G  H HIGSQ          A
                       142           157-158 163-165              187   191-194
```

Figure A4: Sequence alignment of DapDCs related to E.coli DapDC

```
Sec_Struct      sss~~~~~~~~~~~~~~hhhhhhhhhhhhhhhhhhhhh~~~~~sssssss~~~~~~~sssssssssssssss~~~~~~~sssssss~~~~~~~hhhh~~~~~s 187
DapDC_ECOL  223 SAGGGLSVPYQQGEE~AVDTEHYYGLWNAAREQIARHLGHPVKLEIEPGRFLVAQSGVLITQVRSVKQMGS~~~~~RHFVLVDAGFNDLMRPAMYGSYHH 316
DapDC_ARCF  230 DMGGGLGIDYE~GKGAPTPKDLASAILPEFEGRKA~DLTSDPQLWLEPGRSIVGNTTVLITRVNAVKKG~Y~~~~~KNFVAVDAGFNVLIRPAMYGSYHR 321
DapDC_METJ  250 NLGGGLGIPYYKDKQIPTQKDLADAIINTMLKYK~~DKVEMPNLILEPGRSLVATAGYLLGKVHHIKETPV~~~~~TKWVMIDAGMNDMMRPAMYEAYHH 342
DapDC_HELP  221 DVGGGIGVSYE~NEETIKLYDYAQGILNSLQ~~~~~~~GLDLTIICEPGRSIVAESGELITQVLYEKKAQN~~~~~KRFVVVDAGMNDFLRPSLYHAKHA 307
DapDC_HELP  221 DVGGGIGVSYE~NEETIKLYDYAQGILNSLQ~~~~~~~GLDLTIICEPGRSIVAESGELITQVLYEKKAQN~~~~~KRFVIVDAGMNDFLRPSLYHAKHA 307
DapDC_CAMJ  220 DIGGGLGVAYEKNECEPDLYDYAQGILAQLH~~~~~~~GLDLTIGMEPGRYLVAKSGEFVCSVLYEKQNKT~~~~~KRFVVVDAGAMNDLIRPSLYEAYHE 307
DapDC_NEIM  227 DLGGGVGIVYQDEDV~PDLGAYAQAVQKLIG~~~~~~~TRRLKLIILEPGRSLVGNAGALLTRVEFVKYGEE~~~~~KNFVMVDAAMNDLMRPALYDAYHH 313
DapDC_NEIM  227 DLGGGVGIVYQDENV~PDLGAYAQAVQKLIG~~~~~~~TRRLKLIILEPGRSLVGNAGSLLTRVEFVKYGEE~~~~~KNFVMVDAAMNDLMRPALYDAYHH 313
DapDC_PSEF  236 DLGGGVGVRYRDEEP~PLVADYIKAVRERLD~~~~~~~GRDLALMFEPGRYIVANAGVLLTQVEYLKHTEH~~~~~KDFAIVDAAMNDLIRPALYQAWMD 322
DapDC_PSEA  235 DLGGGLGVRYRDEQP~PLAGDYIRAIRERLH~~~~~~~GRDLTLVFEPGRSIVANAGVLLTRVEYLKHTEH~~~~~KDFAIVDAAMNDLIRPALYQAWMD 321
DapDC_VIBC  236 DVGGGLGVVYRDELP~PQPSEYAKALLDRLER~~~~~~HRDLELIFEPGRAIAANAGVLVTKVEFLKHTEH~~~~~KNFAIIDAAMNDLIRPALYQAWQD 323
DapDC_HAEI  235 DLGGGLGVTYDETP~PHPSDYANALLEKLKN~~~~~~~YPELEIIILEPGRAISANAGILVAKVQYLKSNES~~~~~RNFAITDTGMNDMIRPALYEAYMN 322
DapDC_ZYMM  238 DLGGGLGVRYKDDQPEPPSVEAYGEMIKRVTK~~~~~~GWNCRLIFEPGRSLIANAGVLLSKVIRIKESKT~~~~~ARFVILDAAMNDLVRPTLYDAYHE 326
DapDC_AQUA  240 DIGGGLGIKYKPEDKEPAPQDLADLLKDLLE~~~~~~~NVKAKIILEPGRSIMGNAGILITQVQFLKDKGS~~~~~KHFIVDAGMNDLIRPSIYNAYHH 327
DapDC_TM    210 NIGGGWGINYSGEE~~LDLSSYREKVVPDLK~~~~~~~~RFKRVIVEIGRYIVAPSGYLLLRVVLVKRRHN~~~~~KAFVVVDGGMNVLIRPALYSAYHH 294
DapDC_BACS  246 NLGGGFGIRYTEDDEPLHATEYVEKIIEAVKENASRYGFDIPEIWIEPGRSLVGDAGTTLYTVGSQKEVPG~~~~VRQYVAVDGGMNDNIRPALYQAKYE 341
DapDC_BACAN 170 NVGGGFGIRYTESDTPLTLETYVRAVTSAVREQFTVCNYPLPEIWVEPGRSIVGDAGTTLYTVGSVKDIPG~~~~IRKYVSVDGGMTDNLRPALYNARYE 265
DapDC_BACM  244 NLGGGFGIRYTEEDDPIPASQVVKEIINEVKKQVSAYSMKMPEIWIEPGRSLVGDAGTTLYQIGSRKDVPN~~~~VRHYVAVDGGMSDNLRPALYNAKYE 339
DapDC_BACH  244 NLGGGFGIRYIEGDTPRPVGDYVKEMIRAVKQQIAEHEMSMPEIWIEPGRSLVGDAGTTLYTIGSRKEIPN~~~~VRHYLSVDGGMSDNLRPALYQAEYE 339
DapDC_SYNY  275 NIGGGLGICYTESDDPPSIEEWAQVAAISVAKACDRQNIPYPKLIAEPGRSLVGSACVTAYRVGGRKVVPN~~~~IRTYISVDGGMSDNPRPITYQSVYR 370
DapDC_MYCT  254 DLGGGLGISYLPSDDPPPIAELAAKLGTIVSDESTAVGLPTPKLVVEPGRAIAGPGTITLYEVGTVKDVDVSATAHRRYVSVDGGMSDNIRTALYGAQYD 353
DapDC_MYCL  279 DLGGGLGISYLPDDDPPPIFELAAKLGAIVSNESAAVGLPVPKLMVEPGRAIAGPGTITLYEVGTIKDVDVSATAHRRYVSIDGGMSDNIRTALYDAQYD 378
DapDC_MYCS  279 DLGGGLGISYLPHDDPPPMKELADKLLEIVRTESAAVGLPTPKLVVEPGRAIAGPGTITLYEVGTVKDVAVSQTAHRRYVSVDGGMSDNIRTSLYAAQYD 378
DapDC_CORG  253 DLGGGYGIAYTAAEEPLNVAEVASDLLTAVGKMAAELGIDAPTVLVEPGRAIAGPSTVTIYEVGTTKDVHVDDDKTRRYIAVDGGMSDNIRPALYGSEYD 352
DapDC_ACTP  252 DLGGGFAVPYLPGEGEFDLGGFAHRVRVALSHECALRRVPVPRLSIEPGRAVVARAGVTLYRVAAVK~~~R~~~GVRRVFVAVDGGMSDNPRPALYGSRYA 346
DapDC_STRC  230 SAGGGLSVPYTPGDP~EIDTDRYFELWDAARRELVSELGHPVRLEIEPGRFLVAGAGVLAAEVRAQKPVGS~~~~~NYFVLVDAGFNDLMRPAMYGSNHR 323
DapDC_BUCA  219 SVGGGLPIPYKFNDQ~PIDIKKYFMLWDIARKKISKFLGEKIELEIEPGRFLVAESGILISKVWATKKMGN~~~~~KNFVLVDVGFNDLMRPTMYGSYHH 312
DapDC_DEIR  210 NVGGGWSLNADLEG~~~~~~~~~~~~~~IAYEAHEAARVFGAELWVEPGRYLVASAGWLLTRVVGTKRTG~~~~~~RNFCLVDAGMTEFLRPMLYGASHP 289

Consensus       GGG G  Y                            EPGR  V  G     V     K            ::  VD GM  D  RPALY
                227-229 232                          268-271                288             299    307 311

Sec_Struct      ssss~~~~~~~~~~~~~ssssssss~~~~~sssssss~~~sss~sss~~~~~sssssss~~~~~~~~~~~~~~~~~~sssss~ss~ssss~~~~~~~hhh
DapDC_ECOL  317 ISALAADGRSLEHAPTVETVVAGPLCESGDVFTQQEGGNVETRALP~EVKAGDYLVLHDTGAYGASMSSNYNSRPLLPEVLF~DNGQARLIRRRQTIEEL 414
DapDC_ARCF  322 VAVANK~~~~MDAEPEEVYTVVGPICESGDVLARD~~~~~~~~RKLP~KVEVGDLIAVFDAGAYGFVMSSQYNGRPRCAEVLV~SGDRWDVIREKESYGDL 408
DapDC_METJ  343 IINCKV~~~~KNEK~~EVVSIAGGLCESSDVFGRD~~~~~~~~RELD~KVEVGDVLAIFDVGAYGISMANNYNARGRPRMVLT~SKKGVFLIRERETYADL 427
DapDC_HELP  308 IRVITP~~~~SKGREISPCDVVGPVCESSDTFLKD~~~~~~~~AHLP~ELEPGDKLAIEKVGAYGSSMASQYNSRPKLLELAL~EDHKIRVIRKREALEDL 394
DapDC_HELP  308 IRVITP~~~~CGGREISPCDVVGPVCESSDTFLKD~~~~~~~~ANLP~ELEPGDKLAIEKVGAYGSSMASQYNSRPKLLELAL~EDHKIRVIRKREALEDL 394
DapDC_CAMJ  308 I~~ILP~~~~YNQGEESLCDVVGGICESGDFFAKA~~~~~~~~RSLP~STQSDDIMVIKNTGAYGFSMSSNYNTRNKVCELAL~EEGQVRLIRQRENFEDQ 392
DapDC_NEIM  314 IEAV~E~~~~PKNIAPLTANIVGPICETGDFLGKD~~~~~~~~RTI~~ACEEGDLLLIRSAGAYGASMASNYNARNRAAEVLV~DGNEYRLIRRRETLEQQ 398
DapDC_NEIM  314 IEAV~E~~~~TKDIATLTANIVGPICETGDFLGKD~~~~~~~~RTI~~ACEEGDLLLIRSAGAYGASMASNYNARNRAAEVLV~DGNEYRLIRRRETLEQQ 398
DapDC_PSEF  323 VTAV~R~~~~PRDTAARSYDIVGPICETGDFLAKG~~~~~~~~REL~ALEEGDLLAVHSAGAYGFVMSSNYNTRGRCAEVLV~DGDQAFEVRRRETVAEL 407
DapDC_PSEA  322 VQAV~R~~~~PRDAAPRRYDLVGPICETGDFLAKD~~~~~~~~RDL~ALAEGDLLAVRSAGAYGFVMSSNYNTRGRAAEVLV~DGEQTHEVRRRETVQEL 406
DapDC_VIBC  324 IIPL~R~~~~PRQGEAQTYDLVGPVCETSDFLGKD~~~~~~~~RDL~VLQEGDLLAVRSSGAYGFTMSSNYNTRPRVAEVMV~DGNKTYLVRQREELSSL 408
DapDC_HAEI  323 IVEIDR~~~~TLEREKAIYDVVGPVCETSDFLGKQ~~~~~~~~REL~SIAEGDYIAQCSAGAYGASMSSNYNSRARTAEVLV~DGDQSYLIRRRETLQEL 408
DapDC_ZYMM  327 IKAV~~~~~~TPSAQTYQADIVGPVCETGDIFARN~~~~~~~~RSIS~AVKADDLMAIMSAGAYGATMASAYNSRPLVAEVMV~SGNKSALIRKRQSVEDL 411
DapDC_AQUA  328 IIPV~E~~~~TKERKKVVADIVGPICETGDFLALD~~~~~~~~REIE~EVQRGEYLAVLSAGAYGFAMSSHYNMRPRAAEVLV~ENGSVKLIRKRENYDYI 413
DapDC_TM    295 IFVLG~~~~~~KQGKEMRADVVGPLCESGDVIAYD~~~~~~~~RELP~EVEPGDIIAVENAGAYGYTMSNNYNSTTRPAEVLVRENGRISLIRRRETEMDI 380
```

Figure A4: Sequence alignment of DapDCs related to E.coli DapDC

216

```
DapDC_BACS   342 AAAANR~~~~IGEAHDKTVSIAGKCCESGDMLIWD~~~~~~~IDLP~EVKEGDLLAVFCTGAYGYSMANNYNRIPRPAVVFV~ENGEAHLVVKRETYEDI 428
DapDC_BACAN  266 AMLANR~~~~GNDENEETVSIAGKCCESGDMLIWD~~~~~~~ITLP~KVASADLLAISCTGAYGYSMANNYNRIRRPAVVFV~KGGTSQIVVERETYENI 352
DapDC_BACM   340 AVLANK~~~~PLAKADETVSIAGKCCESGDMLIWD~~~~~~~LPLP~KADSDDILAVFCTGAYGYSMANNYNRIPRPAVVFV~ENGESMLVVKRKHMRTS 426
DapDC_BACH   340 GALANR~~~~VNEQPVGMFSVAGKCCESGDMLIWD~~~~~~~LPLP~EANHEDILAVFCTGAYGYSMANNYNRIPRPPVVFV~EDGDAQLVIQRERYEDL 426
DapDC_SYNY   371 VALANR~~~~MNDEITETVTVAGKHCESGDILVKD~~~~~~~VALP~AAEPGDIMVVAATGAYNHSMASNYNRLGRPAAVLV~NQGQANLILQRETYTDL 457
DapDC_MYCT   354 VRLVSR~~~~VSDAPPVPARLVGKHCESGDIIVRD~~~~~~~TWVPDDIRPGDLVAVAATGAYCYSLSSRYNMVGRPAVVAV~HAGNARLVLRRETVDDL 441
DapDC_MYCL   379 VRLVSR~~~~TSDAPAAPASIVGKHCESGDIVVRD~~~~~~~TWVPDDLKPGDLVGVAATGAYCYSLSSRYNMLGRPAVVAV~CAGQARLILRRETVDDL 466
DapDC_MYCS   379 ARLVSR~~~~VSDAPPALARIVGKHCESGDIVVRD~~~~~~~TWVSDDIAPGDLIGVAATGAYCYSMSSRYNLLCRPAVVAV~ADGQARLVLRRETVDDL 466
DapDC_CORG   353 ARVVSR~~~~FAEGDPVSTRIVGSHCESGDILIND~~~~~~~EIYPSDITSGDFLALAATGAYCYAMSSRYNAFTRPAVVSV~RAGSSRLMLRRETLDDI 440
DapDC_ACTP   347 VRLVRR~~~~~~GGRRAPVTVVGRHCEAGDVLAED~~~~~~~VPLPEDVRAGDLLAVPVTGAYHHALASNYNAVGRPPVVGV~RDGVARVLVRRETEEDL 432
DapDC_STRC   324 VSVLDADG~APRASDARDTVLAGPLCESGDVFTQVEGGDVEPVPVP~RTDVGDLVVFHDTGAYGASMSSTYNSRPLIPEVLV~DGAETRLIRRRQTVAEL 420
DapDC_BUCA   313 ISVIYGDDRKMNEKETIDTVVAGPLCESGDIFTQKEGGTVQTRKLP~TIKVGDYLIFHDTGAYGAAMSSNYNTRPLIPEILL~KNNDSIVIRRRQTIEEI 410
DapDC_DEIR   290 LYPMWD~~~~~~~ALATEVWDVAGPACESGDLIARG~~~~~~~VPLP~TPQRGHLLLLIGEAGAYGASMSSTYLSRPRPAEVLW~TGHDWQLLRRRETPQDI 374

Consensus            :G   CESGD     D              GD        GAYG   M  S  YN    R     :::          :   RRET
                    339----346             367        376-378 382—387                    408
```

Figure A4: Sequence alignment of DapDCs related to E.coli DapDC

The organisms listed here are: ECOL-E.coli (sp-P00861), ARCF - Archaeoglobus fulgidus (sp-O29458), METJ - Methanothermobacter thermautotrophicus (sp-Q58497), HELP - Helicobacter pylori (sp-P56129), HELPJ - Helicobacter pylori J99 (sp-Q9ZME5), CAMJ - Campylobacter jejuni (sp-Q9PII5), NEIMA - Neisseria meningitides serogroupA (sp-Q9JWA6), NEIMB - Neisseria meningitides serogroupB (sp-Q9JXM2), PSEF - Pseudomonas fluorescens (sp-O05321), PSEA - Pseudomonas aeruginosa (sp-P19572), VIBC - Vibrio cholerae (sp-Q9KVL7), HAEI - Haemophilus influenzae (sp-P44316), ZYMM - Zymomonas mobilis (sp-Q9Z661), AQUA - Aquifex aeolicus (sp-O67262), TM - Thermotoga maritime (sp-Q9X1K5), BACS - Bacillus subtilis (sp-P23630), BACN - Bacillus anthracis (TIGR-unfinished genome project), BACM - Bacillus methanolicus (sp-P41023), BACH - Bacillus halodurans (sp-Q9KCM5), SYNY - Synechocystis sp. PCC 6803 (sp-Q55484), MYCT - Mycobacterium tuberculosis (sp-P31848), MYCL - Mycobacterium leprae (sp-Q50140), MYCS - Mycobacterium smegmatis (sp-Q9X5M1), CORG - Corynaebacterium glutamicum (sp-P09890), ACTP Actinosynnema pretiosum subsp. Auranticum (sp-O69203), STRC - Streptomyces coelicolor (sp-Q9ZBH5), BUCA - Buchnera aphidicola (sp-P57513), DEIR - Deinococcus radiodurans (sp-Q9RTK2).
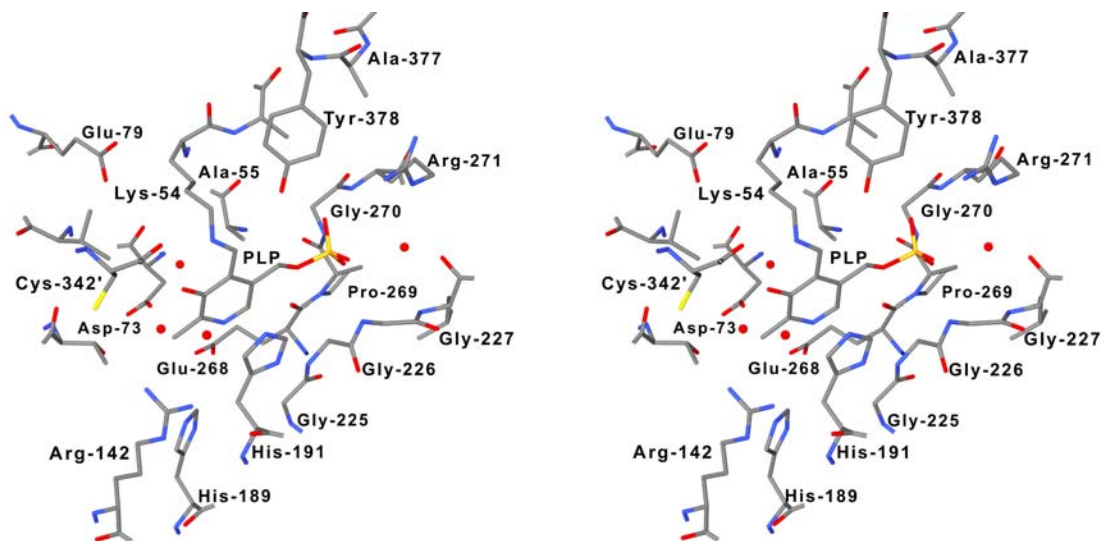
217

Figure A5: Stereo view of the conserved amino acid residues around the cofactor,
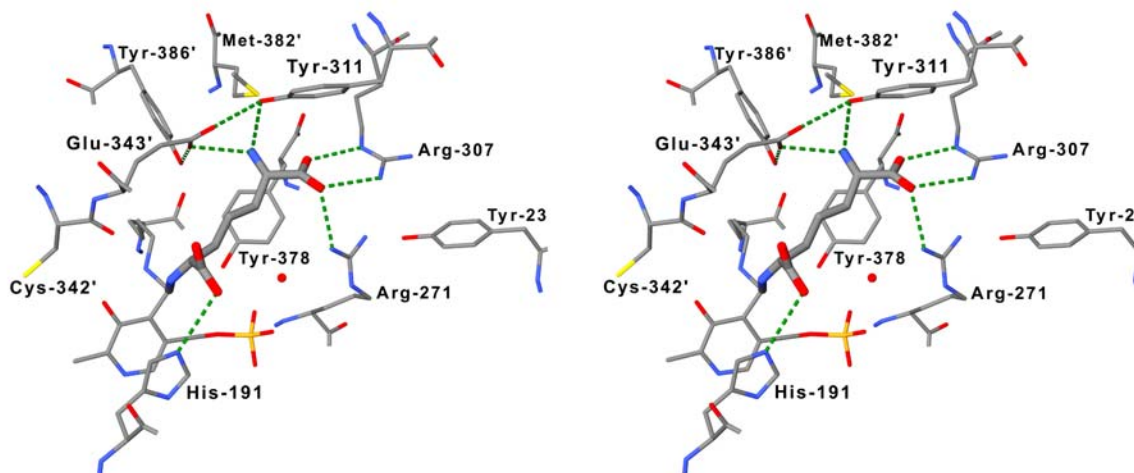
pyridoxal-5'-phosphate.

Figure A6: Stereo view of the conserved amino acid residues around the substrate binding site

In this figure, the substrate, meso-diaminopimelate is obtained by merging the D-lysine and L-lysine. The bond and atom radii of DAP is increased in the figure for clarity purposes.
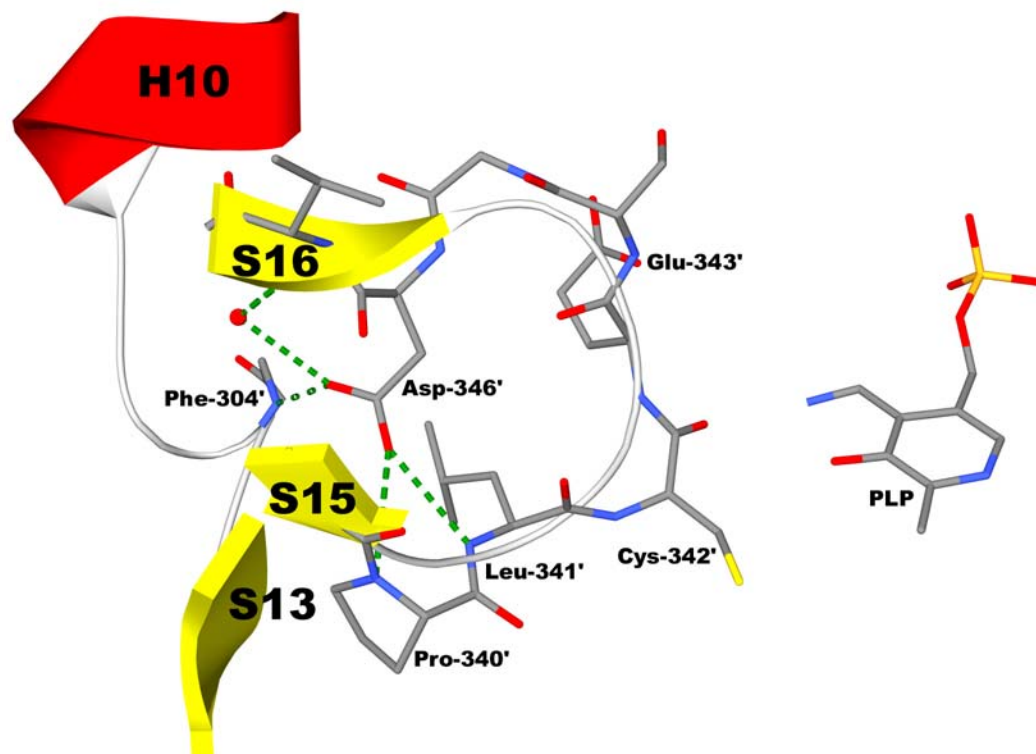
Figure A7: Asp-346' – the critical residue in maintaining the structural integrity

Asp-346' holds the hairpin loop in position to interact with the substrate by

hydrogen bond interactions.

REFERENCES

1) H.J. Vogel in V.Bryson and H.J. Vogel, Evolving Genes and Proteins, Academic Press, New York, 1965, p.25.

2) H.J. Vogel , Symp. Evolutionary Biochem., Proc. 5th, Intern. Congr. Biochem., Moscow, 1961, Pergamon Press, 1963, p.341.

3) Cohen GN. 1985. Aspartate kinases I, II, and III from *Escherichia coli*. *Methods Enzymol 113:*596-9.

4) Meadow, P., Hoare, D.S., Work, E. Biochem.J. 1957, 66, 270.

5) Jouanneau J, Stragier P, Bouvier J, Patte JC, and Yaniv M. 1985. Expression in mammalian cells of the diaminopimelic acid decarboxylase of Escherichia coli permits cell growth in lysine-free medium. *Eur.J.Biochem. 146*:173-178

6) Girodeau JM, Agouridas C, Masson M, Pineau R. and Le Goffic F. 1985. The lysine pathway as a target for a new genera of synthetic antibacterial antibiotics? *J Med Chem 29*:1023-30.

7) Kelland JG, Palcic MM, Pickard MA. and Vederas JC. 1985. Stereochemistry of lysine formation by meso-diaminopimelate decarboxylase from wheat germ: use of 1H-13C NMR shift correlation to detect stereospecific deuterium labeling. *Biochemistry 24*:3263-7.

8) Song Y *et al*.1994. Stereospecific synthesis of phosphonate analogues of diaminopimelic acid (DAP), their interaction with DAP enzymes, and antibacterial activity of peptide derivatives. *J. Org. Chem.* 59:5784-5793.

9) Dunathan HC and Voet JG. 1974. Stereochemical evidence for the evolution of pyridoxal-phosphate enzymes of various function from a common ancestor. *Proc Natl Acad Sci U S A 71*:3888-91.

10) Tabor CW, Tabor H. 1984. Polyamines. *Annu. Rev. Biochem. 53:*749-790.

11) Almrud JJ. *et al*. 2000. Crystal structure of human ornithine decarboxylase at 2.1 A resolution: structural insights to antizyme binding. *J Mol Biol 295:*7-16.

12) Kern AD, Oliveira MA, Coffino P & Hackert ML. 1999. Structure of mammalian ornithine decarboxylase at 1.6 A resolution: stereochemical implications of PLP-dependent amino acid decarboxylases. *Structure Fold Des 7:*567-81.

13) Grishin NV, Osterman AL, Brooks HB, Phillips MA and Goldsmith EJ. 1999. X-ray structure of ornithine decarboxylase from Trypanosoma brucei: the native structure and the structure in complex with alpha-difluoromethylornithine. *Biochemistry 38:*15174-84.

14) Jackson LK, Brooks HB, Osterman AL, Goldsmith EJ. and Phillips MA. 2000. Altering the reaction specificity of eukaryotic ornithine decarboxylase. *Biochemistry 39*:11247-57.

15) Slocum RD, Kaur-Sawhney R, Galston AW. 1984. The physiology and biochemistry of polyamines in plants. *Arch Biochem Biophys 235(2):*283-303.

16) Rodriguez BR, Carroll DW, Mitchell D, Momany C and Hackert ML. 1994. Crystallization of biosynthetic arginine decarboxylase from Escherichia coli. *Acta. Cryst. D50:*175-177.

17) Grishin NV, Phillips MA and Goldsmith EJ. 1995 Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci 4*:1291-304.

18) Wu WH and Morris DR. 1973. Biosynthetic arginine decarboxylase from Echerichia coli. Subunit interactions and the role of magnesium ion. *J. Biol. Chem. 248(5) :*1696-9.

19) Osterman AL, Brooks HB, Jackson L, Abott JJ and Phillips MA. 1999. *Biochemistry 38*:11814-11826.

20) Tsirka S, and Coffino P. 1992. Dominant negative mutants of ornithine decarboxylase. *J.Biol.Chem. 267:*23057-23062.

21) Tsirka S, Truck CW. and Coffino P. 1993. Multiple active conformers of mouse ornithine decarboxylase. *Biochem. J. 293*:289-295

22) Jansonius JN. 1998. Structure, evolution and action of vitamin B6-dependent Curr Opin Struct Biol. *8(6):*759-69.

23) Osterman AL, Kinch LN, Grishin NV and Phillips MA. 1995. Acidic residues important for substrate binding and cofactor reactivity in eukaryotic ornithine decarboxylase identified by alanine scanning mutagenesis. *J Biol Chem 270:* 11797-802

24) Burkhard P, Dominici P, Borri-Voltattorni C, Jansonius JN. & Malashkevich VN. 2001. Structural insight into Parkinson's disease treatment from drug-inhibited DOPA decarboxylase. *Nat Struct Biol 8:*963-7.

25) Ishii S., Mizuguchi H, Nishino J, Hayashi H. & Kagamiyama H. 1996. Functionally important residues of aromatic L-amino acid decarboxylase probed by sequence alignment and site-directed mutagenesis. *J Biochem (Tokyo) 120:* 369-76.

26) Tramonti A, De Biase D, Giartosio A, Bossa F. & John RA. 1998. The roles of His-167 and His-275 in the reaction catalyzed by glutamate decarboxylase from Escherichia coli. *J Biol Chem 273:*1939-45

27) Appleby TC, Kinsland C, Begley TP & Ealick SE. 2000. The crystal structure and mechanism of orotidine 5'-monophosphate decarboxylase. *Proc Natl Acad Sci U S A 97:*2005-10.

28) Shaw JP, Petsko GA.and Ringe D. 1997. Determination of the structure of alanine racemase from Bacillus stearothermophilus at 1.9-A resolution. *Biochemistry 36:*1329-42.

29) Swanson T, Brooks HB, Osterman AL, O'Leary M, and Phillips MA. 1998. *Biochemistry 37*:14943-14947

30) van Ophem PW, Peisach D, Erickson SD, Soda K, Ringe D, Manning JM. 1999. Effects of the E177K mutation in D-amino acid transaminase. Studies on an essential coenzyme anchoring group that contributes to stereochemical fidelity. *Biochemistry. 38(4):*1323-31

31) Fleischmann RD *et al*.1995. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science:269(5223):496-512.

32) Stove CK *et al*. 2000. Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen. *Nature:*406(6799):959-64.

33) Heidelberg JF *et al*. 2000. DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae. *Nature:*406(6795):477-83.

34) Parkhill J *et al*. 2000. Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491. *Nature:*404(6777):502-6.

35) Parkhill J *et al*.2000. The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. *Nature:*403(6770):665-8.

36) Alm RA *et al*. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. *Nature:*397(6715):176-80.

37) Cole ST e al.1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature:*393(6685):537-44.

38) Cole ST *et al*. 2001. Massive gene decay in the leprosy bacillus. *Nature*:409(6823):1007-11.

39) van Ophem PW, Peisach D, Erickson SD, Soda K, Ringe D, Manning JM. 1999. Effects of the E177K mutation in D-amino acid transaminase. Studies on an essential coenzyme anchoring group that contributes to stereochemical fidelity. *Biochemistry. 38(4):*1323-31

40) Bourot S *et al*. 2000. Glycine betaine-assisted protein folding in a lysA mutant of *Escherichia coli*. *J Biol Chem 275*:1050-6

41) Altschul SF, Gish W, Miller, W, Myers, EW, Lipman, DJ. 1990. Basic local alignment search tool. *J Mol Biol, 215*(3):403-10

42) Higgins DG, Sharp PM. 1988. CLUSTAL: A package for performing multiple sequence alignment on a minicomputer. *Gene 73:237-244*

43) Devereux J, Haeberli P, Smithies O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res 12*:387-395.

APPENDIX B

BIOCHEMICAL CHARACTERIZATION OF THE *E.COLI*

DIAMINOPIMELATE DECARBOXYLASE

EXPRESSION, ISOLATION AND DYNAMIC LIGHT SCATTERING ANALYSIS OF

THE WILD-TYPE AND MUTANTS OF E.COLI DIAMINOPIMELATE

DECARBOXYLASE

INTRODUCTION

Previous size analysis experiments performed on the wild-type DDC generated

contradictory results about the oligomeric state of the active enzyme. White *et al*.

demonstrated by ultracentrifugation studies that DapDC has an apparent molecular

weight of 200,000. Since DapDC has a monomeric molecular weight of 46,099, this

study suggested that the active enzyme is a tetramer (1). Sizing analysis by gel filtration

experiments showed that DapDC is a monomer (2, 3). These results were in contrast to

the dimeric structure of ornithine decarboxylases to which DapDC has high sequence and

structural similarity (4, 5). Also the structural study of *E.coli* DapDC showed that the

active site of DDC sits at the interface of two monomers related by two-fold. Hence to

clear the discrepancy of DapDC being a monomer or a dimer, a dynamic light scattering

(DLS) experiment was performed.

Based on the structure-oriented sequence analysis presented in Appendix A, several mutant enzymes of DapDC in addition to the wild-type enzyme were studied. These mutants were prepared because they appeared to have unique relevance in DapDC and they have not been studied in other systems. The set of mutants consisted of, i) S384F, a folding mutant that requires the presence of osmoprotectant like glycine betaine for proper folding and activity (3), ii) Y386F, which could play a vital role in stereochemical specificity, iii) Y378F, which contributes to the hydrogen bonding required for cofactor binding and also forms the back support of the substrate-binding pocket, and iv) H191F, which plays an important role in catalysis since it is conserved in all the PLP binding decarboxylases. Light scattering experiments were performed on these sets of mutants to look for unusual folding properties that could explain some of the properties of the mutants.

MATERIALS AND METHODS

Expression of C-terminal poly-histidine tagged wild type and mutants of DapDC. The appropriate clones of the wild-type and mutants of DapDC transformed into the *E.coli* expression strain BL21(DE3) was prepared in Dr. Momany's laboratory by Karino Resto and Geoff Smith (2). A common expression strategy was used for the wild-type as well as the mutants.

A 100 ml culture of *E.coli* expression strain BL21(DE3) transformed with the expression vector, pET28b containing the appropriate gene was grown for 15-17 hrs at 37°C in LB medium containing 50 μg/ml kanamycin from a single colony. After inoculation into 1L of the same medium, the cells were grown for 4-5 hrs, then 0.1 mM

meso-DAP, 10 µM pyridoxine and 1mM IPTG (isopropyl-β-D-thiogalactoside) were
added and the cells were allowed to grow for 6 hrs. The cells were collected by
centrifugation at 4°C for 15 min at 10,000 x g and suspended in 30 ml extraction buffer
(0.02 M sodium phosphate, 0.5 M NaCl, 0.2 mM pyridoxal-5'-phosphate, 25 mM
imidazole, pH 7.4) and either used immediately, or frozen at -20°C for later use. The cell
suspension was cooled in ice to 10°C and sonicated for 3 min in 1-min increments with a
Fisher Scientific 550 sonic dismembrator set at 50 % duty cycle while maintaining the
temperature below 15°C. Cellular debris was removed by centrifugation for 20 min at
25,000 X g at 4°C followed by centrifugation at 60,000 x g for 20 min. The clarified
supernatant was then used for chromatography.

Purification of the wild type and mutants of DapDC. The entire purification
process was done on a Pharmacia ÄKTA HPLC system at room temperature. The first
step of purification was by Ni^{2+}-metal chelate affinity chromatography with a Pharmacia
NTA Hi-Trap Chelating 5ml column. The clarified lysate was applied (1 ml/min) to the
nickel-charged NTA column that had been equilibrated with 20 mM sodium phosphate,
0.5 M NaCl, pH 7.4. After washing unbound protein off, the protein was then eluted
using a linear gradient of imidazole (0.025-0.5 M imidazole, 0.5 M NaCl, 20 mM sodium
phosphate, pH 7.4) over 20 column volumes. The fractions containing DapDC (yellow
colored) were combined and dialyzed against 20 mM tris HCl buffer, pH 8.0, containing
1 mM BME (β - mercaptoethanol) and 0.1 mM PLP using 2 x 1 L changes. The dialyzed
protein was then loaded onto a Pharmacia Hi-Trap Q column (5ml) (anion-exchange
column) equilibrated with Buffer T (20 mM tris HCl, pH 8 and 1mM BME). The protein
was eluted using a linear gradient of 0-1.0 M NaCl over 20 column volumes. This second

228

step of purification was tried using different buffer conditions also. As a second

approach, the DapDC eluted off the nickel affinity column was dialyzed against 50 mM

MES buffer, pH 6.0, containing 0.2 mM PLP. The resulting solution was then applied to

5 ml Pharmacia HiTrap SP column (cation-exchange column) pre-equilibrated with 50

mM MES buffer, pH 6.0 lacking PLP at a rate of 2 ml/min (Buffer M). The enzyme was

washed with the same buffer and then eluted using a linear gradient of 0-0.5 M NaCl in

Buffer M over 20 column volumes. The brightly yellow fractions containing DapDC

were combined and then dialyzed against 50 mM MES buffer, pH 6.0, containing 0.2

mM PLP (Buffer C). For crystallization trials, dialysis of the protein into Buffer C with 1

mM BME was also tried.

   Dynamic light scattering experiments of the wild type and mutants of DapDC.

Dynamic light scattering measurements were made using a 90Plus particle-sizing

instrument from Brookhaven Instruments Corporation. DapDC (wild-type and mutants)

was dialyzed into 50 mM MES, 0.2 mM PLP, 1mM BME, pH 6.0 and diluted to 5 mg/ml

in the same buffer. Measurements below this protein concentration did not have adequate

counting statistics. All the protein solutions were filtered through a Millipore Millex-FG,

0.2 μM filter unit. Two samples of similar concentration were measured for consistency.

Each sample was read ten times, each reading being 3 min duration. The samples were

illuminated at 659 nm using a solid-state laser. The MAS OPTION particle-sizing

software of the instrument was used to analyze the results.

RESULTS

Expression and isolation of the wild-type and the mutants of DapDC. From the

metal-chelate column, the wild-type as well as the mutants of the DapDC eluted as a

single peak. The purified proteins appeared homogenous (>95% pure) when checked by

SDS-PAGE (Pharmacia PHAST system). The wild type, the Y386F mutant, the S384F

mutant, the H191F mutant, all appeared as a characteristic PLP-dependent protein, a

yellow color with absorption maximum around 426 nm. The Y378F mutant was

completely devoid of the yellow color. After dialysis into Buffer T (tris-HCl buffer) and

loading onto the anion-exchange column (Q column), all the mutants eluted out as

multiple-species. When all the proteins were dialyzed into Buffer M (MES buffer) and

then loaded onto the cation-exchange column (SP column), all of them except the S384F

mutant eluted as a single peak. Two closely associated peaks of S384F were eluted out

(S384F-I and S384-II). Also previous results from our laboratory had demonstrated that

crystals of wild-type DapDC were colorless if BME was included in the buffer *i.e.*, the

protein was stripped off from its cofactor, PLP. This could also be the reason that DapDC

was reported to have 43% of the activity in BME versus BAL (2,3-dimercaptopropan-1-

ol) (White & Kelley, 1965). Taking this action of BME into account, the isolation of the

wild type and the mutants of DapDC were done with buffers, which excluded BME.

Dynamic light scattering results. The MAS OPTION particle sizing software of

the instrument determines the particle size by calculating the diffusion coefficient (D) and

then calculating the particle diameter from the Stokes-Einstein equation, $D = kT/3\pi\eta d$,

where η, k, T and d are viscosity of the liquid in which the particle is moving, the

Boltzmann constant, absolute temperature and particle diameter respectively.

The average diameter of wild-type DapDC was found to be 9.3 nm with a standard deviation of 0.1 nm (Figure B1). Also the sample was monodisperse giving no evidence that the his-tagged DapDC oligomerizes into a tetrameric complex. Among the mutants, S384F-I had the similar radius as that of wild-type. S384F-II showed very high effective stokes diameter indicating the formation of very large aggregates (Figure B2).

DISCUSSION

The elution of the wild-type and most of the mutants of DapDC as single peaks when dialyzed into MES buffer and loaded onto cation-exchange column as against the multiple peaks observed in tris-HCl buffer, indicated that MES has a stabilizing effect on the protein. This was further supported by the structure of DapDC wherein a molecule of MES from the crystallization was seen to be bound near the active site (6).

The dimensions of the wild-type DapDC dimer were very consistent with the structural data. From the structural studies, the dimer appears like a rhombohedral plate with dimensions 9.6 nm by 7.4 nm by 4.7 nm (6). The shape of DapDC is elongated with short dimensions along the minor axes. The protein folds itself in such a way that there is not a significant difference in the sizes of the monomer (maximally 7.5 nm long) and dimer of DDC, which makes it difficult to resolve on a gel filtration column chromatography. The gel filtration size analysis can be explained by the elongated shape with the short dimensions along the minor axes that could give it a smaller apparent molecular weight. Thus the results of the light scattering experiment clearly showed that the size analysis by previously reported gel filtration is in error and that the active DapDC is not a monomer but a dimer in solution.

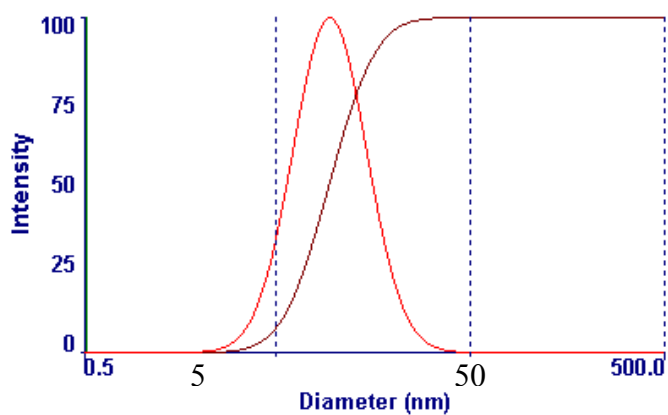| Run	Eff. Diam. (nm)
1	9 . 8
2	9 . 2
3	9 . 4
4	9 . 2
5	8 . 8
6	9 . 1
7	10 . 5
8	9 . 5
9	9 . 1
10	9 . 1
Mean	9 . 3
Std. Error	0 . 1
Combined	9 . 4

Figure B1: The effective diameter and the monodispersity of the wild-type DapDC as calculated by the MAS OPTION software of the particle sizing instrument

Run	Eff. Diam. (nm)
1	148.1
2	148.5
3	151.1
4	147.2
5	151.3
6	147.2
7	146.8
8	149.3
9	150.2
10	150.9
Mean	149.1
Std. Error	0.5
Combined	149.3
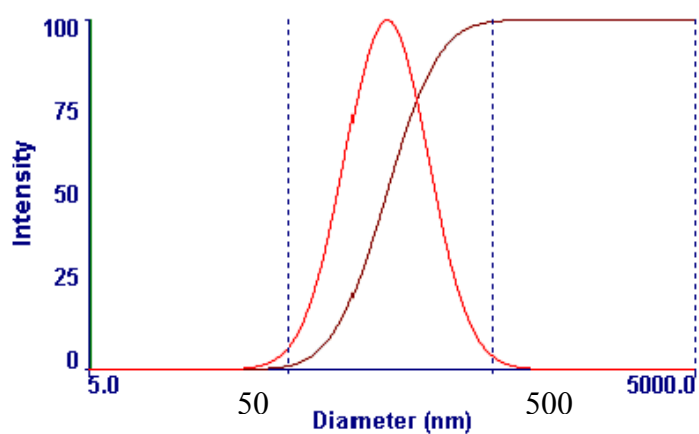
Figure B2: The average diameter calculated for the S384F-II shows that very large aggregates of the protein are being formed.

The S384F mutant is very interesting because it appears to be a population of species. One of the species likely forms improperly folded large aggregates while the other species behaves like wild-type in terms of size and thus in proper folding. When the S384F-II mutant was analyzed by DLS in the presence of betaine, no evidence of the protein folding back to its original shape was seen. Therefore the results of Bourot *et al*. wherein the S384F mutant was found to regain the activity in the presence of glycine betaine have to be carefully analyzed (3). If the mutant is a mixture of species, the activity values of the enzyme could lead to misinterpretations. A structural study of both the species with and without the presence of betaine could shed some light on this aspect.

HPLC ASSAY FOR CHIRAL SEPARATION OF D- AND L-LYSINE

INTRODUCTION

In the structural study of *E.coli* DapDC, the characterization of the substrate binding pocket was done by preparing a complex of the enzyme with its product inhibitor, *L*- lysine (6). Over the course of three days, some of the *L*- lysine appeared to have racemized to *D*-lysine since the substrate binding site contained an apparent mixture of both *D*-and *L*-lysine. This observation was supported by the optical rotation study of the solution of *L*-lysine incubated with DapDC where the rotation shifted towards zero. Alternative explanations for these two observations were: the *D*-lysine seen in the structure could be a misinterpretation of significantly rotated *L*-lysine; the optical rotation shift could be the result of formation of by-products like decarboxylation of *L*-lysine to 1,5-diaminopentane or other compounds (6). It was therefore necessary to develop a

sensitive assay system, which could resolve and detect *D*-lysine, *L*-lysine, DAP isomers, reaction intermediates and by-products of the reaction.

Reversed-phase high-performance liquid chromatographic resolution of amino acid enantiomers by the formation of diastereomers using a chiral reagent, 2,3,4,6-tetra-O-acetyl-β-D-glucopyransoyl isothiocyanate (GITC) was the method of choice. This method has been used in the resolution of common amino acid enantiomers (7, 8), α-methyl amino acid enantiomers (9), antibiotics (10) etc.

Liquid chromatographic resolution of optical isomers can be done by either direct resolution of enantiomers on chiral stationary phases, using chiral eluents or using pre-column derivatizing reagent followed by liquid chromatography of the diastereomers on conventional columns (11, 12, 13). Chiral derivatization agents, which convert the enantiomers into their corresponding diastereomers, are compounds that selectively and readily react with the enantiomers, form appropriate chromophores or fluorophores that may be then detected after chromatographic resolution. GITC is one such chiral derivatizing agent, which reacts with the amino functional group under alkaline conditions to form a thiourea derivative that can be detected using UV with high sensitivity (11) (Figure B3). GITC, when used to derivatize the enantiomers of amino acids, introduces a bulky tetraacetylglucosyl moiety into the amino acids, which increases their conformational rigidity (7). The degree of separation of diastereomers is dependent on the rigidity of conformation (11, 15). The differences in the conformational rigidity and the hydrophobic surfaces of both the amino acid derivatives that are assumed to interact with the hydrophobic ODS residues of the column can affect the order of elution
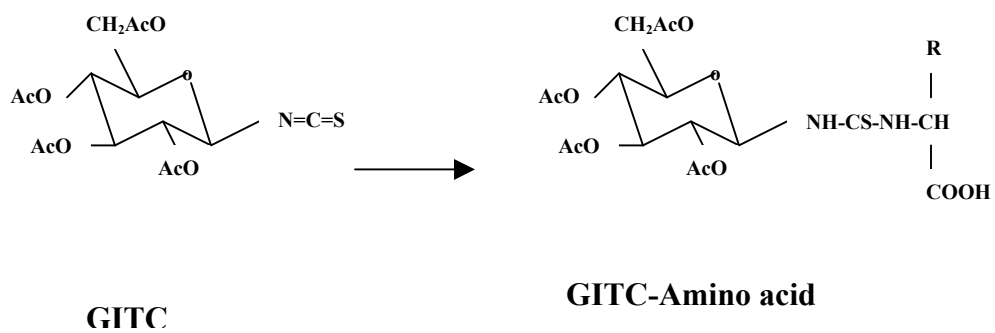
GITC

GITC-Amino acid

Figure B3: Reaction of GITC with an amino acid (8)

from the column (8). The advantages of using GITC as chiral derivatizing agent are: simple and rapid separation of enantiomeric amino acids without the protection of free carboxyl residue (8), the derivatized mixture can be directly injected into the HPLC system, both the GITC and the thiourea derivative absorb at 250 nm although the reagent has significantly lower absorbance (molar extinction coefficient – 1000) than those of the diastereomers (molar extinction coefficient – 12,000), and the excess reagent used for derivatization does not interfere in the analysis of the diastereomer and gives a single peak (7).

The discontinuous chiral separation HPLC assay was developed to accomplish the following goals: 1) to qualitatively monitor the racemization of *L*-lysine to *D*-lysine by DapDC, 2) if racemization occurs or lysine is decarboxylated to diaminopentane, demonstrate the ability to separate the enantiomers as well as the reaction substrate and by-products, 3) to show that this method can be used to achieve the long term goal of using this method as a standard assay for DapDC with the ability to estimate kinetic parameters of the wild-type and the mutants of DapDC and also to determine the stereochemical selectivity of Tyr-386 mutant and other potential engineered mutants towards *L,L*-diaminopimelate.

MATERIALS AND METHODS:

Materials. *DL*-diaminopimelic acid, *L*-lysine, *D*-lysine, diaminopentane and GITC were purchased from Sigma-Aldrich (St. Louis, USA). All other reagents were of the highest analytical or molecular grade available from commercial vendors.

Apparatus. The HPLC system comprised a Gilson liquid chromatographic system (Model 306 Pump with 10 ml/min head, 231 Autosampler, 401 Dilutor, and 811 Dynamic Mixer) and an Applied Biosystems 759A UV/VIS detector. The detector wavelength that was examined is 250 nm.

Crystallization of* meso-*DAP. *Meso*-DAP was separated by fractional crystallization. (1,16). An aqueous solution of the *DL*-DAP (Sigma) was made to a final concentration of 15 mg/ml. Ethanol was added until there was a slight permanent turbidity at room temperature. The precipitate was allowed to stand at room temperature for 24 hrs. The precipitate was collected, dissolved in water and reprecipitated with ethanol. Two more precipitations gave *meso*-isomer as determined chromatographically.

Preparation of standards. Initially the standards, *meso*-DAP, *D* and *L*-lysine and diaminopentane were prepared in deionised water. Solutions of *D*-lysine, *L*-lysine, *meso*-DAP and diaminopentane were prepared in the concentration ranges of 800 μg/ml, 400 μg/ml, 200 μg/ml, 100 μg/ml and 50 μg/ml. These concentrations were chosen as initial concentrations to start the experiment. Depending upon the sensitivity of the instrument and limit of detection of these samples, the concentration range for the standards were changed to 5μg to 25 μg/ml. The same standards in the same concentration ranges were then prepared in the assay buffer (100 mM sodium phosphate buffer, pH 6.8, 50 μM PLP) to simulate the enzyme assay conditions. For this purpose, a stock solution of 5 mg/ml of all the four compounds was first prepared in deionized water. Then the appropriate dilutions were made with the assay buffer to get the final volume of 1 ml and the desired concentrations. (For example: 25 μg/ml standard solution – 5 μl of each of the stock solutions + 880 μl of assay buffer).

Enzyme Assay. A 0.5 mg/ml stock solution of *meso*-DAP was made in the assay buffer (100 mM sodium phosphate buffer, pH 6.8, 50 μM PLP). 0.8 ml of this stock solution was mixed with enzyme (enzyme is in 50 mM MES and 0.2 mM PLP) and the final volume was adjusted to 1 ml with the assay buffer. This assay mixture was incubated at 37°C in a shaker incubator. An aliquot of 50 μl from this reaction mixture was taken at time interval of 30 mins, 1, 3, 6, 12, 24, 48 and 72 hrs**.** To test the racemization reaction, the assay procedure was repeated with *D*-lysine and *L*-lysine as the substrate. The enzyme assay was repeated for all the mutants and a qualitative assessment about their activity was made by comparing the respective area of the peaks at a particular time point and protein concentration with that of the wild type.

Pre-column derivatization. At every time point, 10 μl of the reaction mixture was diluted ten times with water (90 μl of water) and then 50 μl of this diluted solution was mixed with 50 μl of 100% acetonitrile containing 0.8% (w/v) triethylamine to give a final volume of 100 μl. To this mixture, 50 μl of 0.2% (w/v) chiral reagent, GITC in acetonitrile was added. The resulting mixture was incubated at room temperature for 60 mins with vortexing at every 20 mins. To study the stability of the GITC derivative of the compounds, one of the compound's derivative (*L*-lysine and GITC complex) was kept at room temperature and checked for degradation of the complex by comparing the area of the peaks at different time intervals.

Chromatography procedure. A Gilson (Madison, Wis.) HPLC unit with an Applied Biosystems absorbance detector (model 759A; San Jose, Calif) was used. A Whatman RP-18 5 micron 250 X 4.6 mm column (Applied Biosystems, San Jose, Calif.) fitted with a 15 X 3.2 mm 7.0 micron guard column (RP-18 Newgard; Applied

Biosystems) was used for separation. The column temperature was maintained at room temperature. Eluents were: Buffer A-10 mM phosphate buffer, prepared from potassium phosphate monobasic and the pH was adjusted to 2.8 by addition of phosphoric acid, Buffer B-100% methanol. The derivatized reaction mixture was then injected manually through a 20 µl sample loop. The column was eluted isocratically at a flow rate of 0.9 ml/min with a mobile phase prepared by mixing methanol and 10 mM phosphate buffer in the ratio of 50:50. Different ratios of the eluents A and B were also tried to optimize resolution. The eluted compounds were detected at 250 nm, sensitivity set at 0.01 a.u.f.s (absorbance units full scale). The data were collected, recorded and integrated using a personal computer (Gilson 718 sample manager software). Peak assignments were carried out by comparing retention times with that of the standards.

RESULTS

The standard solutions in water when injected onto the column and eluted isocratically (50:50) gave the best resolution as compared to that obtained from 60:40 (methanol:phosphate buffer), 40:60 and a gradient of increasing methanol concentration. Most of the reagent peaks were well separated from the derivatized products and did not interfere with the detection. Triethylamine (TEA), which was added to provide the necessary basicity for the derivatization reaction and also to stop the enzymatic reaction. The reaction products of TEA with the GITC reagent eluted much faster than any of the amino acid derivatives and hence did not affect the detection of other derivatized products.

The retention time ranges of each of the derivatized product were (Figure B4): *meso*-DAP, 18 to 18.5 mins; the unreacted excess GITC, 20 to 22 mins; *L*-lysine eluted out immediately after the GITC peak, 23 to 24 mins; *D*-lysine peak followed the *L*-lysine, 24.5 to 25.5 mins and diaminopentane eluted out after all the peaks, 35 to 35.5 mins. In the stability study carried out to find how long the GITC complex remains intact without degradation, the *L*-lysine-GITC complex was loaded after 1 hr, 4hrs, 10 hrs and 19 hrs. When the area of the peaks at the different time points was compared, it was found that the area declines progressively with time of incubation (Figure B5).

After seeing adequate resolution of the substrate, products and by-products as standards, an actual enzymatic reaction was performed. The area of the peaks of the product, *L*-lysine increased between time points 30 mins and 1.5 hrs and then it remained constant until 24 hrs. There was a reduction in the peak areas of *L*-lysine at time points 48 and 72 hrs. A very tiny peak of *D*-lysine was seen at 6 hrs but it could not be integrated. This *D*-lysine peak increased in area at 24 hrs time point but then remained almost constant until 72 hrs (Figure B6).

After seeing a very slow racemization of *L*-lysine to *D*-lysine, the enzymatic reaction was carried out with *D*- and *L*-lysine as substrate because if DapDC is a true racemase then the reaction should work in both the directions. However, in both the reactions, the area of the peaks of the *D*-lysine and *L*-lysine were almost constant at all time points (4hrs, 24 hrs, 48hrs) and no interconversion of *L* to *D* or *D* to *L* was observed.

When the assay was used to qualitatively quantitate the enzyme activities of the mutants, it was found that the mutants Y378F and Y386F had comparable percent activities to that of the wild type. H191F, the mutant that is speculated to interact with the

cofactor as well as play a role in the substrate interaction, showed 46% activity as compared to that of the wild type. S384F-II, the folding mutant exhibited 47% activity (Figure B7).

DISCUSSION

GITC has previously been used to derivatize compounds with a single amino functional group. In the enzymatic assay of DapDC, the end product, lysine has two functional amino groups and so it was doubtful whether GITC could react with both the enantiomers of lysine and provide differential conformational rigidity and hydrophobicity to resolve the diastereomers. But GITC proved to be a suitable chiral agent for derivatization and chiral separation of lysine.

The observation that racemization of *L*-lysine to *D*-lysine occurs in the DapDC active site was further supported by the evidence of *D*-lysine being produced after *L*-lysine accumulates from DAP decarboxylation. The speculation that the shift in optical rotation of the solution of *L*-lysine and DapDC could be due to the decarboxylation of *L*-lysine to diaminopentane was negated because there were no signs of diaminopentane being eluted out. Putting these results together, it can be concluded that DapDC is a very slow racemase and there could be some chemical modifications of the enzyme that could be tapped to make it a better racemase. But the results of the enzyme reaction with *D*-lysine and *L*-lysine as substrates, which did not show any interconverison of the enantiomers of lysine, created some ambiguity about the racemization potential of DapDC. Another possibility is that the peak, which is being interpreted as *D*-lysine could be some reaction by-product that has similar retention time as that of *D*-
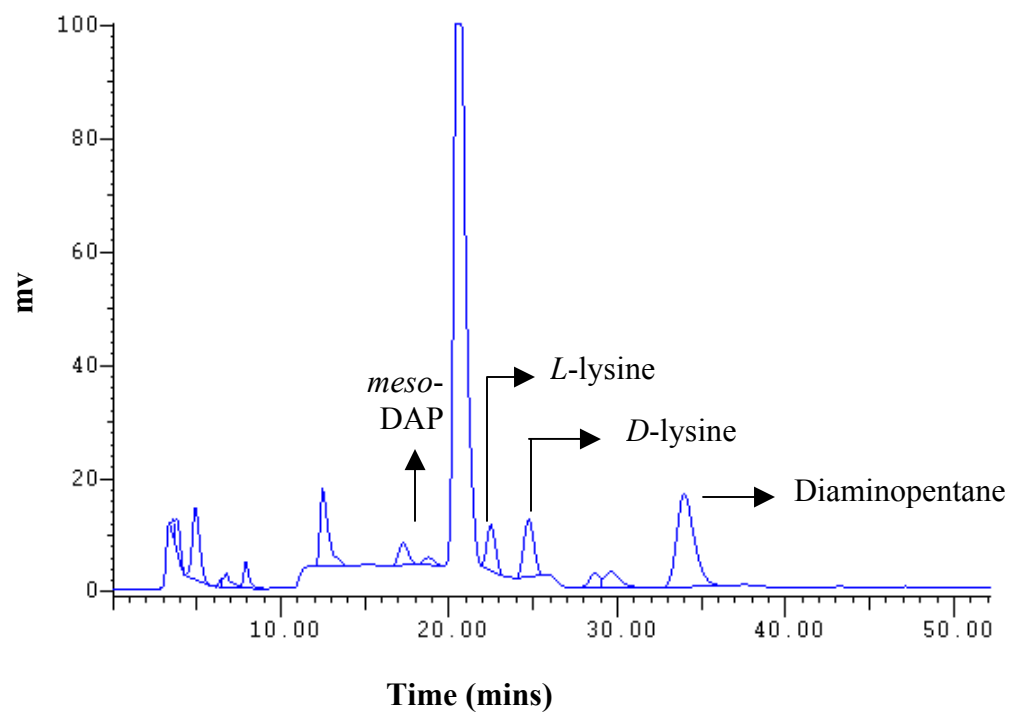
Figure B4: Chromatogram showing separation of meso-DAP, D-lysine, L-lysine and diaminopentane.

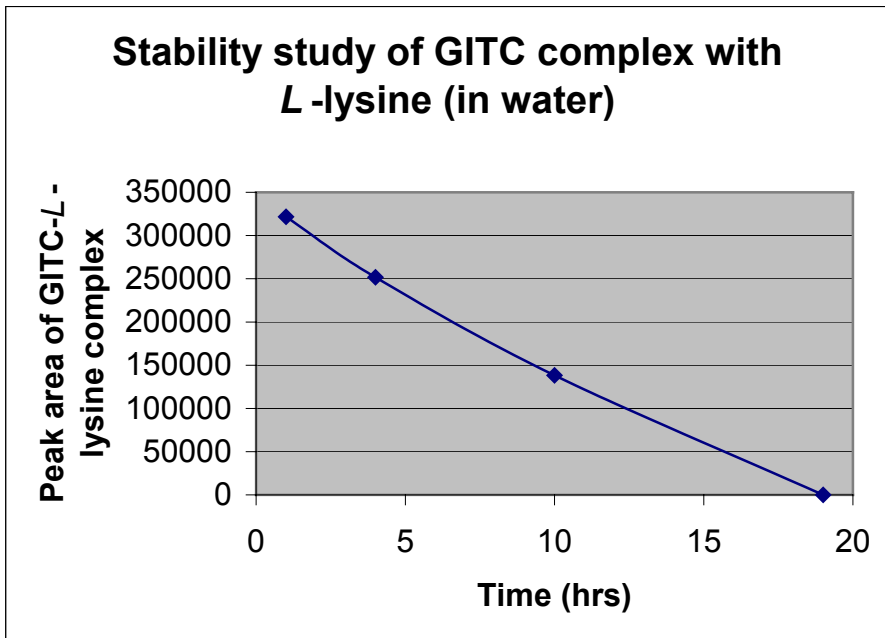**Stability study of GITC complex with *L*-lysine (in water)**

Figure B5: Stability curve of GITC derivatized L-lysine

The graph shows that the GITC complex is not a very stable complex at room temperature. The area under the peak of the complex progressively decreases over the time. Hence one hour was decided as the optimum incubation time for the precolumn derivatization.
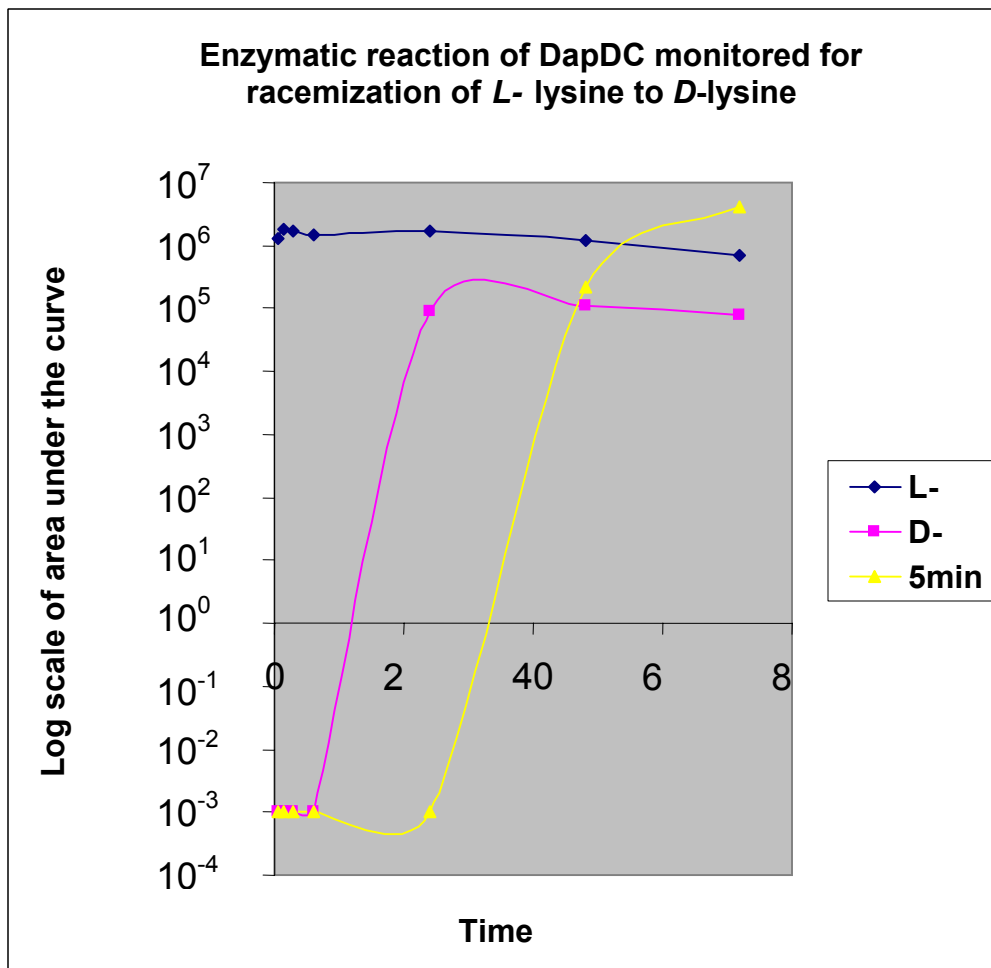
Figure B6: Graph of the enzymatic reaction of DapDC using meso-DAP as the substrate as monitored by the discontinuous HPLC based assay.

The GITC-derivatized samples of the DapDC enzyme reaction at time points 0, 0.5, 1.5, 3, 6, 12, 24, 48 and 72 hrs when eluted from the column showed that a small amount of *L*-lysine was racemized to *D*-lysine after 24 hrs. The area under the peak of *L*-lysine also showed gradual reduction over the time period of 72 hrs. An unidentified peak appeared after 48 hrs that had the retention time of 5 mins.

Qualitative assesment of the activities of mutants of DapDC using discontinous HPLC assay method
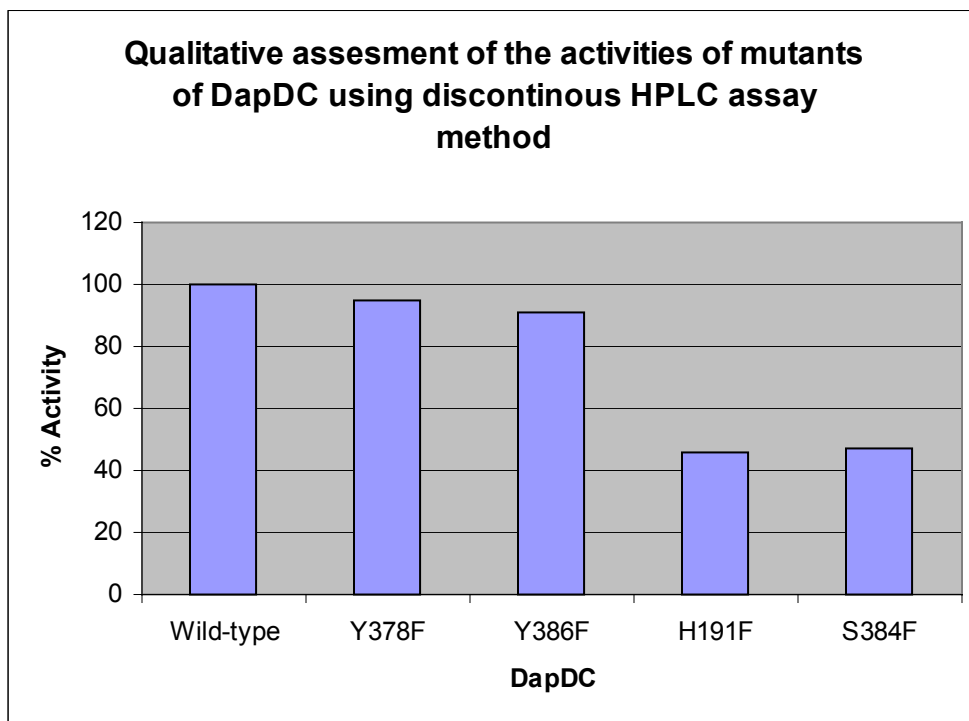
Figure B7: A qualitative measurement of the enzymatic activities of the mutants of DapDC using the HPLC assay method

The area under the peak of all the mutants of DapDC (at the same time point and similar protein concentration) was compared to that of the wild-type and then converted to percent activity with the wild-type DapDC being considered as 100% active.

lysine. This could be clarified by carefully analyzing PLP dependent reactions and considering the various reaction intermediates or by-products that could be formed. The identity of the peak resembling *D*-lysine could also be characterized by mass-spectroscopic analysis.

It is definitely possible to achieve the long-term goal of using this assay as a standard assay to get quantitative data and kinetic parameters. This method is very sensitive and can detect very low quantities of products of the reaction. But before trying to get quantitative results from this method, certain problems have to be tackled. When the standards are prepared in the assay buffer, the peak areas over time became significantly lower as compared to that of the standards prepared in water. Also the reproducibility of the chromatography was decreased because the complex of GITC and the compound probably became more prone to degradation. One explanation to this problem could be that PLP in the phosphate buffer could be interacting with the compounds and forming a complex, which either does not interact with GITC, chemically modifies the *D*- or *L*- lysine, or the newly formed complex does not absorb at 250 nm. Also the MES and PLP in the enzyme form some complex with the compounds themselves or GITC, or the derivatized product, which coelutes with the derivatized compounds. This hurdle towards making the assay method as a standard procedure for derivation of catalytic constants of wild-type as well as the mutants, could require a thorough search of buffer conditions, which do not interfere with the detection and quantification of the products and by-products formed in the enzyme reaction.

DOCKING STUDIES OF DIAMINOPIMELATE DECARBOXYLASE

INTRODUCTION

The structural study of *E.coli* DapDC revealed certain structural constraints in the active site that could be the cause of poor binding properties of the currently available inhibitors of DapDC. Most of the drugs synthesized for DapDC have been designed with the structure of DapDC homologue, ODC as reference. Even though the overall structural fold of DapDC and ODCs are very similar there are subtle changes in the active site residues that have to be considered in the design of inhibitors specific to DapDC. The most appropriate example to support the structural differences in the DapDC and the ODCs is the failure of α-difluoromethyl-DAP analogue to act as a potent inhibitor of DapDC as opposed to α-difluoromethyl-ornithine analogue, which is a very potent inhibitor of ODCs. The rationale that can be attributed to this difference is that in DapDC, the Cys-342 residue of the second subunit is oriented in such a position that it can sterically block the difluoromethyl group of the analogue and thus prevent proper binding (6). Another residue, which is present in both the enzymes, but plays different roles, is His-191. This residue, in ODC, is mainly involved in the maintenance of electronic properties of PLP whereas in DapDC it might have an additional role of hydrogen bonding with the substrate (if protonated) or providing the repelling push to the leaving group (if unprotonated) (6, Appendix A). Another important structural difference in both the enzymes is the placement of residues on the hairpin loop formed by the residues His-191 through Asp-197. This hairpin loop is structurally very critical in terms of placing residues for interaction with the substrate. In DapDC, two residues, Gly-193

and Ser-194 are oriented very differently than the corresponding residues in ODCs, which could seem subtle but it could potentially lead to a change in stereochemical selectivity (Appendix A). These changes are small, but can have significant implications on drug-design and therefore need to be incorporated in the strategies of designing a very selective inhibitor of DapDC. Thus it becomes important to find compounds that can selectively interact with the active site residues of DapDC. Docking of compounds from ligand libraries to the active site of DapDC is definitely the first step towards the achievement of this goal.

DOCK, a program developed by a group headed by Professor Kuntz at University of California, San Francisco, was used to perform docking studies on DapDC (17). Docking methods are computational algorithms that, besides predicting the three-dimensional structures of ligand-receptor complexes, also evaluate the relative affinity or free energy of binding for these bound ligands or drugs to the receptor molecule. The various applications of DOCK are: 1) search databases of chemical compounds which can act as enzyme inhibitors or bind to target receptors, 2) screen databases for DNA binding compounds, 3) examine the various possibilities of binding orientations or conformations of protein-protein and protein-DNA complexes, 4) identify inhibitors which are species selective and 5) design diverse combinatorial drug libraries aimed at multiple but unknown targets or directed ligand libraries aimed at optimizing hits against individual targets (18, 19).

To begin with, DOCK requires the structure of the macromolecule of interest. The procedure is divided into several steps. The first step is to generate a molecular surface for the receptor by using the AutoMS program. Only the surface around the residues at

the active site needs to be generated. For this step, a particular portion of the macromolecule is chosen (for e.g., cofactor) is chosen as ligand and all the residues or receptor atoms within the specific radius of that known ligand (excluding the ligand) is chosen as the surface for MS calculation. This program creates a set up file for the next program, SPHGN by creating the INSPH file. SPHGN defines the docking region by generating spheres that fill the active site. The shape of the cavities in the receptor is used to define a set of spheres. A very dense representation of spheres is created over the entire surface, which is then filtered to remove all the small spheres not specifically associated at the ligand site. The filtered set is then clustered and the largest cluster which contains the ligand binding site of the receptor molecule is chosen to dock the ligand. The program GRID constructs "energy" grid in the local environment around the spheres that describes the electrostatic interactions. It also computes a bump grid, which prevents a ligand atom from sterically overlapping with a receptor atom. Sphere centers are then matched to the ligand atoms, to determine possible orientations for the ligand. Each oriented molecule is then scored for its fit. Energy scoring is one of the scoring schemes, which utilizes the implementation of a force field scoring. A number of random orientations can be tried and refined for each ligand in the library. The best orientation for each ligand associated with an energy score is given as output. These ligands can then be ranked according to their energy scores. The lowest energy ligands can then be picked and further analyzed.

METHOD

In the structure of *E.coli* DapDC, several water molecules are found in the cofactor-binding site such as WAT-10, WAT-27, WAT-40, WAT-78, WAT-124, WAT-196, and WAT-328 AND WAT-345. For the generation of active site, the initial ligand file included the cofactor PLP, the product L-lysine and the water molecules around that site. This ligand file was used to identify the residues or receptor atoms within 10 Å of the ligand atoms. This portion of the receptor file, which was used for MS calculation, was called the 'keep_exclude.pdb' file. This keep_exclude file was then used in program autoMS for the generation of a setup file for the program SPHGN. The output of SPHGN, 'keep_exclude.sph', was used by SHOWSPHERE to write out the sphere center coordinates in PDB format and is helpful for visualizing the clusters. The clusters were spread over the entire surface and so a filter was applied to select the clusters localized exactly at the ligand region. This was performed by using the script pick_sphere.csh, which picks clusters within a certain radii of the ligand. The output file of the script was named as 'limited_clusters.pdb' (Figure B8). Using the files generated until now as input files, the program GRID was used to generate the grid files. As a prelude to grid, the program SHOWBOX was run that allowed the visualization of the location and size of the grids that will be calculated by the program GRID. The output, 'showbox_output,' was seen by using one of the graphics program that can display the PDB format. Then in the final step, the site points and the precomputed grids were used to run DOCK and generate ligand orientations. The Cambridge database was used as the ligand library (approximately 170,000 compounds) in DOCK to screen for potential compounds that
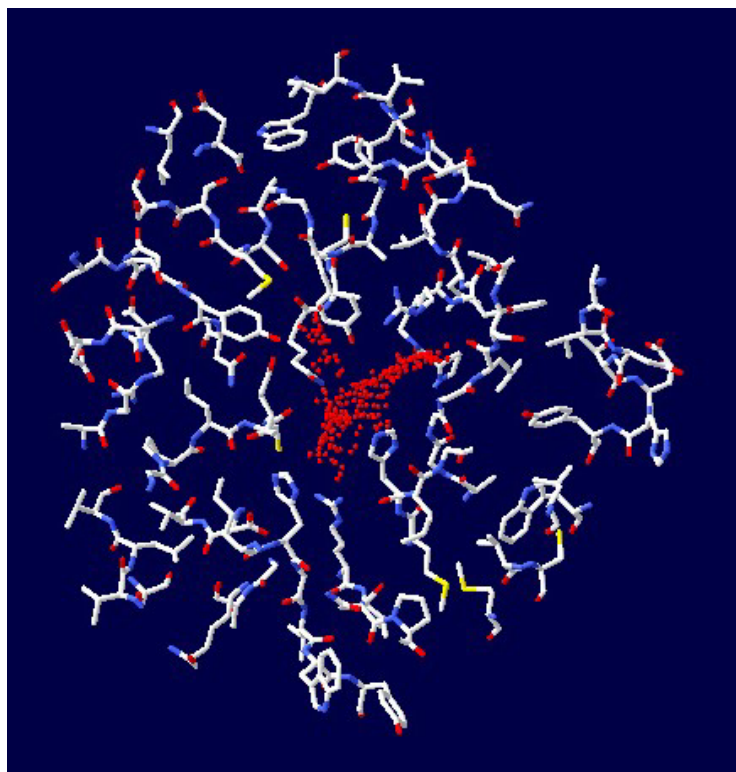
251

Figure B8: The keep_exclude.pdb file along with the limited clusters of spheres within a certain radii of the ligand

The ligand file containing the cofactor, the product lysine and some water molecules around the active site was used to identify the residues or receptor atoms within 10 Å of the ligand atoms. This portion of the receptor file, which was used for MS calculation, was called the 'keep_exclude.pdb' file. This keep_exclude.file was then used in program autoMS for the generation of a setup file for the program SPHGN. The clusters of spheres were spread over the entire surface and thus had to be filtered out to choose the clusters localized exactly at the ligand region.

can interact with the active-site residues of DapDC. DOCK was executed in batch mode with all the relevant parameters entered in an input file 'dock.inp'. After the program was executed, two important output files were generated: a coordinate file that contained docked and minimized molecules and a '*.info' file that contained a summary information about the current ligand rankings. The best scoring molecules or orientations were then viewed by SYBYL, a molecular graphics program and sifted for only those molecules that had meaningful interactions at the active site. The details of the 'hits' were viewed using the program 'CONQUEST'. The entire virtual drug screening was performed on a SGI dual-processor octane for one week.

RESULTS AND DISCUSSION

The crystallization of *E.coli* DapDC requires MES (morpholino-ethyl-sulfonate) as buffer. The structure when refined revealed the presence of a molecule of MES in the active site (2, 6). MES is located right next to the cofactor molecule and is held in position by hydrogen bonds and ionic interactions. This was the beginning of the search of lead compounds for DapDC. Interestingly this condition was very similar to that of alanine racemase where the crystallized protein had an acetate molecule, a component of the crystallization buffer, in the active site and this lead to the screening of other small chain carboxylic acids as inhibitors of alanine racemase (20).

Taking the MES molecule as a pioneer in the goal of docking compounds in the active site of DapDC, it was decided to begin the search in the direction of finding similar molecules. The docking site of DapDC was therefore defined by removing the cofactor,

PLP with the rationale that the compounds that mimic PLP i.e. compounds with aromatic character had a higher probability of being selected and would have drug-like properties.

True to the expectation, an interesting molecule that was docked right at the active site was a moiety with phenolsulfonate functional group.

DATBASE ENTRY NAME: AZPOSM (21)

Formula: $C_6H_4N_2O_4S_1$, H_2O

Compound Name: 2-Diazonium-4-phenolsulfonate monohydrate

The compound had an energy score of –23.39 and ranked 15th in the list. It was docked in the cofactor-binding site. When compared with the structure with PLP and MES, AZPOSM is placed right next to PLP at and below MES with the sulfonate groups of both MES and AZPOSM intersecting (Figure B9). The positioning of this molecule is very close to the hairpin loop structure of the second subunit. Even though the location of the molecule is at the critical place, it falls short of good solid interactions with the protein. The hydroxyl functional group of the molecule hydrogen bonds with the Cys-142 residue located at the tip of the hairpin loop. Arg-142, which is known to interact with the phenolic oxygen of the cofactor, hydrogen bonds to the phenolic group of AZPOSM. Unlike the PLP molecule whose phosphate moiety is held by various hydrogen bond and ionic interactions, AZPOSM's sulfonate moiety does not reach upto Tyr-378, Arg-271 and the Gly-Gly-Gly loop for hydrogen bond interactions. The only residue that can interact with the sulfonate group of AZPOSM is Lys-54. There is also a water molecule at the vicinity of the sulfonate group that could be another source of a hydrogen bond. If His-191 is protonated, it could provide a suitable ionic interaction with the sulfonate group. The diazonium group is again not within the distance of interaction with the
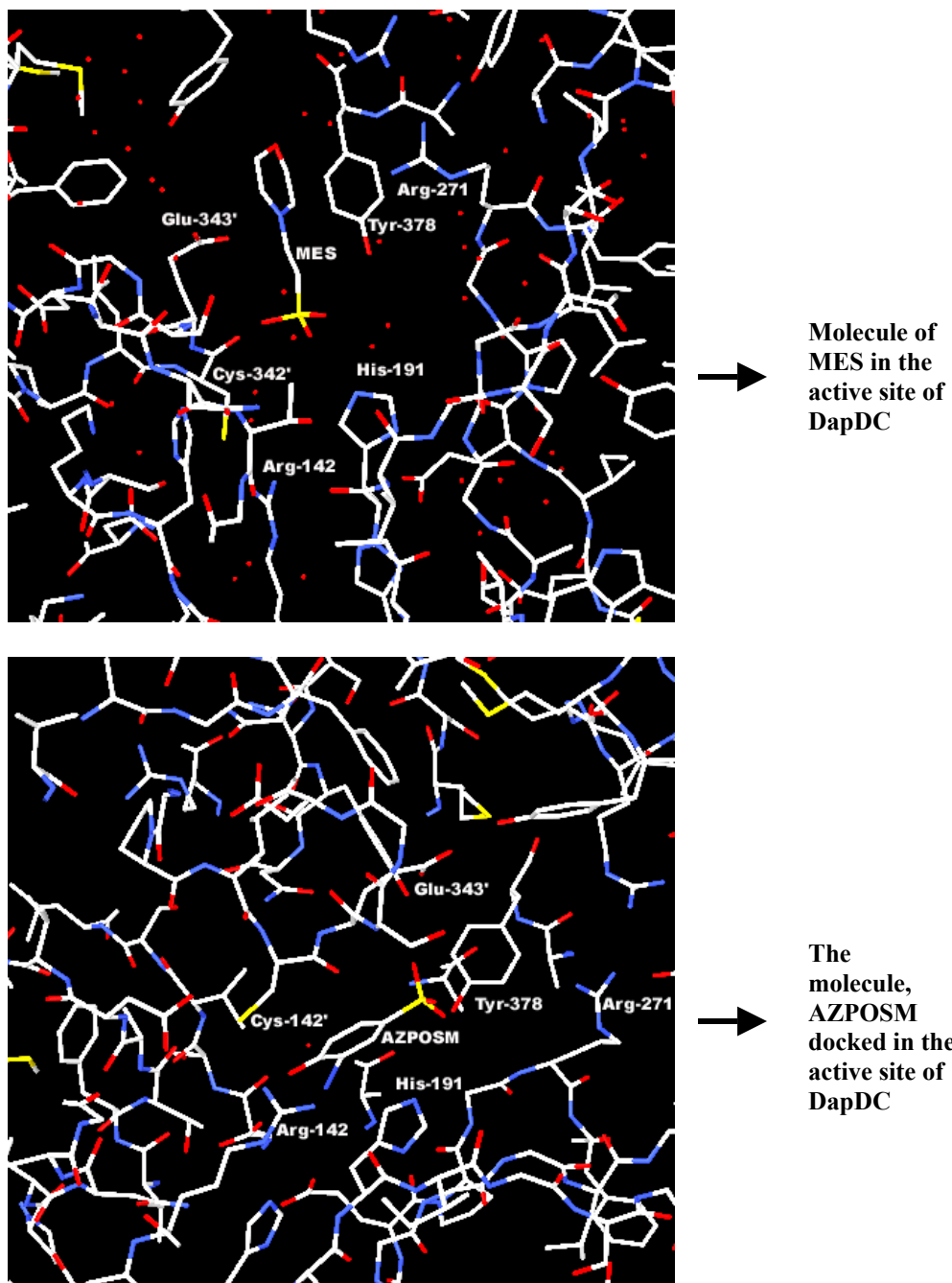
Molecule of MES in the active site of DapDC

The molecule, AZPOSM docked in the active site of DapDC

Figure B9: AZPOSM (2-Diazonium-4-phenolsulfonate monohydrate), structurally similar to MES is docked at the cofactor-binding site.

Even though the location of the molecule is at the critical place, it falls short of good solid interactions with the protein.

surrounding residues. Glu-268 is at a distance of 3.7 Å from the diazonium group. There is a water molecule situated in a bridging position between the diazonium group and Glu-268. His-191 is not exactly parallel to the face of the molecule, but it could provide some stabilization to the aromatic ring of AZPOSM. Thus AZPOSM, the molecule that has the closest resemblance to the MES molecule, can also provide some new direction to the virtual drug screening of DapDC. Several other aromatic compounds docked into the cofactor binding site, but all of them failed to utilize the available residues at the site for hydrogen bonding and ionic interactions. Apparently docking was only successful in filling up the space of the protein molecule with aromatic compounds depending on their shape. Aromatic compounds with flat rings were placed in an orientation that would occupy the vacant space created by the absence of PLP molecule (Figure B10), but they were too close to some of the residues or too far from the residues that could provide strong interaction potentials.

In the docking results, among the non-aromatic compounds that were docked at the active site, HIBZUN, a dipeptide could lead to some new thinking.

DATBASE ENTRY NAME: HIBZUN

Formula: $C_9H_{12}N_3O_5^{1-}H_4N_1^{1+}$

Compound Name: Ammonium N-γ-L-glutamyl-β-cyano-L-alaninate

HIBZUN is a dipeptide of glutamic acid and alanine (22). Although this compound also showed fewer interactions with the amino acid residues, it lead to the thought that if one of the amino acids of the dipeptide could be *D*-lysine residue, it can not only emulate the substrate binding but could also occupy the entire pocket area thereby locking up more of the protein residues in strong binding interactions (Figure B11).
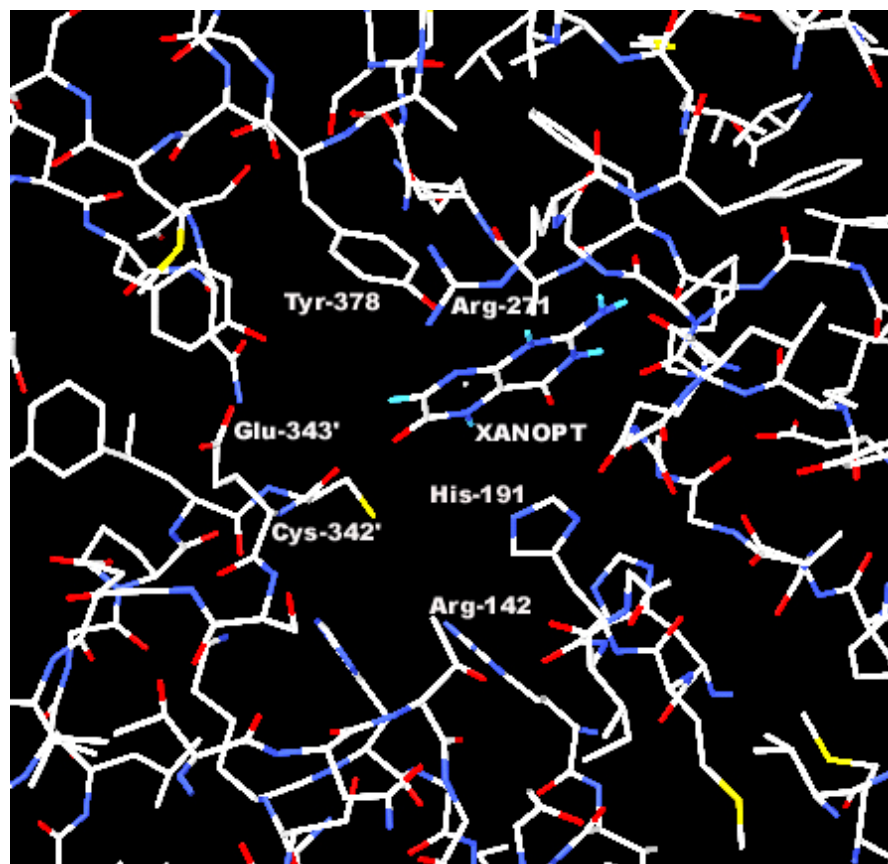
Figure B10: XANOPT (Xanthopterine hydrochloride), another aromatic compound docked in the active site of DapDC (23)

Aromatic compounds with flat rings were placed in an orientation that would occupy the vacant space created by the absence of PLP molecule, but they were either too close to some of the residues or too far from the residues that could provide strong interaction potentials. Docking was only successful in filling up the space of the protein molecule with aromatic compounds depending on their shape.

Another approach to perform docking studies of DapDC is to make DOCK select more substrate analogues that would have high binding affinity to the cofactor to make a weak covalent Schiff base type of interaction. This can be done by manually modifying the active site so that the Lys-54 of DapDC is removed from the way and a functional group that would attract the amino group of the substrate analogues is introduced at the C4' position of the cofactor.

The application of docking studies of *E.coli* DapDC for the development of broad-spectrum antibiotics could be limited. The complication could be due to the subtle but significant differences in the active site of DapDC group members as well the ODCs. Structural and sequence analysis (chapter V) had revealed the difference in the orientation of the conserved sequence H[191]IGS between the *E.coli* DapDC and the ODCs. Ser-194 in *E.coli* DapDC is steered away from interaction with the phosphate group of the cofactor whereas the conserved serine clearly interacts with the cofactor in the ODCs. The consequence of the different orientations of the serine residue is that in the DapDC, it prevents steric hindrance to the carboxylate group of the *D*-amino acid substrate. Analysis of the sequences around this residue in other DapDCs revealed the possibility of the conserved serine being oriented similar to that of the serine in the ODCs. If that is the case, the serine interaction with the cofactor in other DapDCs might sterically prevent some compounds from interacting at the active site that could otherwise bind to the *E.coli* enzyme. Thus the discovery of lead compounds for *E.coli* DapDC through docking procedures could prove to be very organism-specific. Structural studies of DapDCs from additional species would really be beneficial to put this doubt at rest.
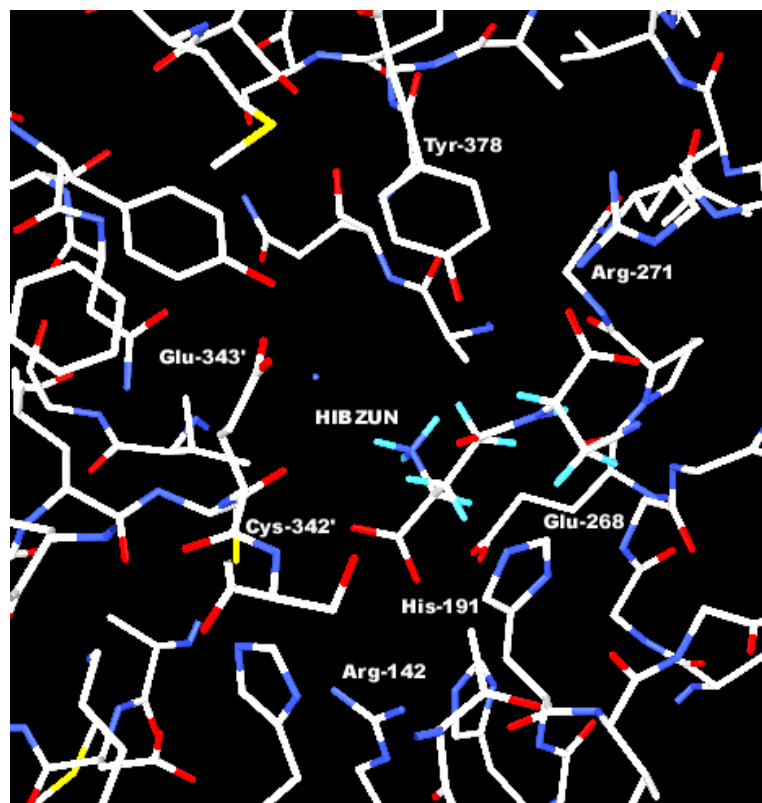
Figure B11: HIBZUM (Ammonium N-γ-L-glutamyl-β-cyano-L-alaninate), a dipeptide docked at the active site

Although this compound also showed few interactions with the protein, it lead to the thought that if one of the amino acids of the dipeptide could be *D*-lysine residue or DAP, it can not only maintain the critical chirality but also emulate the substrate binding. It could also occupy the entire pocket area thereby involving more of the protein residues in strong binding interactions.

REFERENCES

1. White P & Kelley B. Purification and properties of diaminopimelate decarboxylase from Escherichia coli. 1965. *Biochem. J.* 96:75-84.

2. Momany C, Levdikov V, Blagova L & Crews K. 2002. Crystallization of diaminopimelate decarboxylase from Escherichia coli- a stereospecific D-amino acid decarboxylase. *Acta Crystallographica D* (In press).

3. Bourot S et al. 2000. Glycine betaine-assisted protein folding in a lysA mutant of Escherichia coli. *J Biol Chem* 275:1050-6.

4. Grishin NV, Phillips MA & Goldsmith EJ. 1995. Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci* 4:1291-304.

5. Grishin NV, Osterman AL, Brooks HB, Phillips MA & Goldsmith EJ. 1999. X-ray structure of ornithine decarboxylase from Trypanosoma brucei: the native structure and the structure in complex with alpha-difluoromethylornithine. *Biochemistry* 38:15174-84.

6. Vledikov V, Blagova L, Bose N, Momany C. 2002. Diaminopimelate decarboxylase uses a versatile active site for stereospecific decarboxylation (In preparation).

7. Nimura N, Ogura H, and Kinoshita T. 1980. Reverse-phase liquid chromatographic resolution of amino acid enantiomers by derivatization with 2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosyl isothiocyanate. *J. Chromatogr* 202:375-9

8. Kinoshita T, Kasahara Y and Nimura N. 1981. Reverse-phase high-performance liquid chromatographic resolution of nonesterified enantiomeric amino acids by

derivatization with 2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosyl isothiocyanate and 2,3,4-tri-O-acetyl-alpha-D-arabinopyranosyl isothiocyanate. 1981. *J. Chromatogr* 210:77-81.

9. Tian Z, Hrinyo-Pavlina T, Roseke RW and Rao PN. 1991. Resolution of 2,3,4,6-tetra-O-acetyl-beta-D-glucopyranosylisothiocyanate derivatives of alpha-methyl amino acid enantiomers by high-performance liquid chromatography. J Chromatogr. 541:297-302.

10. Tymiak AA, McCormick TJ, Unger SE.1988. Structure determination of lysobactin, a macrocyclic peptide lactone antibiotic. *J Org Chem* 54:1149-57.

11. Scott CG, Petrin MJ, McCorkle T.1976. The liquid chromatographic separation of some acyclic isoprenoid acid enantiomers via diastereomer derivatization. *J Chromatogr* 125:157-161

12. Furukawa H, Mori Y, Takeuchi and Ito K. 1977. Separation of L- and D-amino acids as diastereomeric derivatives by high-performance liquid chromatography. *J Chromatogr* 136:428-431.

13. Goto J, Hasegawa S, Nakamura K, Shimada K and Nambara T. 1977. Sensitive derivatization reagents for the resolution of carboxylic acid enantiomers by high-performance liquid chromatography. *Analytica Chimica Acta* 120:187-192.

14. Takahashi HK, Takeda N, Nimura N and Ogura H.1979. Reaction of amino acids with D-glycosyl and D-gluconyl isothiocyanates.Chem Pharm Bull 27:1137-1141.

15. Nambara T, Ikegawa S, Hasegawa M and Goto J.1978. High pressure liquid chromatographic resolution of amino acid enantiomers by derivatization with new chiral reagents. *Anal Chim Acta* 101:111-116.

16. Work E. 1964. Variation of activity of bacterial diaminopimelate decarboxylase under different conditions of growth. *Biochemical J* 91:600-610.

17. Kuntz ID. 1992. Structure based strategies for drug design and discovery. *Science* 257:1078-82.

18. *www.cmpharm.ucsf.edu/kuntz/dock.html*

19. *http://cbcg.lbl.gov/ssi-csb/Chapter5.html*

20. Shaw JP, Petsko GA.and Ringe D. 1997. Determination of the structure of alanine racemase from Bacillus stearothermophilus at 1.9-A resolution. *Biochemistry* 36:1329-42

21. Greenberg B, Okaya Y. 1969. Crystal and molecular structure of 2-diazonium-4-phenolsulfonate monohydrate *Acta Crystallogr*, Sec B. 25:2101-3

22. Delaere IM et al. 1995. N-γ-L-glutamyl-β-cyano-L-alanine, an antinutritional factor *ex Vicia sativa L.*, as its ammonium salt. *Acta Crystallogr* Sect C. 51:289:91

23. Bleri JH, Humme WP, Visortini M. 1976. The crystal structure of xanthopterine hydrochloride. *Helv Chim Acta* 59:2374-77.