M.Sc. Chanakan Tongsook

Biochemical characterization and kinetic studies on *S*-adenosylmethionine-dependent methyltransferases

Dissertation

Zur Erlangung des akademischen Grades einer

Doktorin der technischen Wissenschaften (Dr. *techn*.)

Erreicht an der

Technischen Universität Graz

Betreuer: Univ.-Prof. Mag. *rer. nat.* Dr. *rer. nat.* Peter Macheroux

Institut für Biochemie

Technische Universität Graz

Graz, 2016

ACKNOWLEDGEMENTS

Abstract

Methyltransferases catalyze methylation processes involved in metabolism, signal transduction, protein/DNA repair, and biosynthesis. *S*-adenosylmethionine (SAM) is the major biological methyl donor in reactions catalyzed by methyltransferases. Several human SAM-dependent methyltransferases in genes encoding polymorphisms are connected to inherited diseases. Moreover, many small-molecule SAM-dependent methyltransferases are applied for industrial bioprocessing *i.e.* biosynthesis of antibiotics. This work focuses on the functional and structural characteristics of two SAM-dependent methyltransferases and their genetic polymorphisms linked to intellectual disability and on a small-molecule SAM-dependent methyltransferase essential for antibiotic biosynthesis.

Histamine *N*-methyltransferase (HNMT) catalyzes the degradation of histamine by transferring a methyl group from SAM to histamine. HNMT is critically important for the maintenance of neurological processes. Recently, two mutations in the encoding human gene were reported to give rise to dysfunctional protein variants (G60D and L208P) leading to intellectual disability. Determination of biochemical and structural properties of the wild-type and variants were conducted and we confirmed that G60D disrupts the SAM binding site leading the variant catalytically inactive. On the other hand, the L208P variant showed reduced protein stability. This finding rationalizes the loss of enzymatic activity observed in the L208 variants.

METTL23 is a putative SAM-dependent methyltransferase. Nonsense mutation and a 5 bp frameshifting deletion lead to truncation of the putative METTL23 protein disrupting the predicted catalytic domain and altering the cellular localization. Expression analysis of METTL23 indicated a strong association with heat shock proteins, which suggests that these may act as putative substrates for methylation by METTL23. Disruption of METTL23 presented here supports the importance of methylation processes for intact neuronal function and brain development.

N,*N*-8-demethyl-8-amino-d-riboflavin dimethyltransferase (RosA) catalyzes the final dimethylation of 8-demethyl-8-amino-d-riboflavin (AF) to the antibiotic roseoflavin (RoF) in *S. davawensis*. In the present study, we have solved the X-ray structure of RosA, and determined its biochemical properties. The structure of RosA is similar to that of previously described SAM-dependent methyltransferases featuring two domains: a mainly

α-helical 'orthogonal bundle' and a Rossmann-like domain. Both substrates, AF and SAM, bind independently to their respective binding pocket. This finding was confirmed by kinetic experiments that demonstrated a random-order 'bi-bi' reaction mechanism. Both products, RoF and *S*-adenosylhomocysteine (SAH), bind more tightly to RosA compared to the substrates, AF and SAM. This suggests that RosA may contribute to roseoflavin resistance in *S. davawensis*. The tighter binding of products is also reflected by the results of inhibition experiments, in which RoF and SAH behave as competitive inhibitors for AF and SAM, respectively. We also showed that formation of a ternary complex of RosA, RoF and SAH (or SAM) leads to drastic spectral changes that are indicative of a hydrophobic environment.

Kurzfassung

Methyltransferasen katalysieren wichtige Methylierungsprozesse im Metabolismus, wie Z.B. Signalübertragungen, Protein/DNA Reparaturmechanismen oder auch in verschiedenen Biosynthesewegen. *S*-Adenosylmethionin (SAM) ist hierbei der vorwiegende biologische Methylgruppendonor für die zuvor genannten Reaktionen. Polymorphismen bei wichtigen menschlichen genen, die für SAM-abhängigen Methyltransferasen kodieren wurde bereits mit verschiedenen Krankheiten in Verbindung gebracht. SAM-abhängige Methyltransferasen werden außerdem für industrielle Zwecke, z.B. zur Herstellung von Antibiotika, verwendet. Der Fokus dieser Arbeit liegt auf der funktionellen und strukturellen Charakterisierung zweier SAM-abhängiger Methyltransferasen, deren Polymorphismus zu starken geistigen Beeinträchtigungen der Betroffenen führt. Außerdem wurde im Rahmen dieser Arbeit auch eine SAM-abhängige Methyltransferase in der Antibiotikabiosynthese genauer untersucht.

Histamin-*N*-Methyltransferase (HNMT) katalysiert den Abbau von Histamin durch Übertragung einer Methylgruppe von SAM auf Histamin. Dadurch besitzt HNMT eine ausgesprochen wichtige Rolle in der Erhaltung neurologischer Prozesse. Vor Kurzem wurde berichtet, dass zwei Mutationen im dem für HNMT kodierenden menschlichen Gen zu veränderten Proteinvarianten (G60D und L208P) und intellektuellen Beeinträchtigungen der betroffenen Personen führen. Im Rahmen dieser Arbeit wurden biochemische und strukturelle Untersuchungen sowohl am Wildtyp als auch den beiden Proteinvarianten durchgeführt. Der Austausch von G60 zu D60 beeinträchtigt die SAM-Bindeeigenschaften und die enzymatische Aktivität der Proteinvariante, während für die L208P Variante eine im Vergleich zum Wildtyp geringere Proteinstabilität und ein Verlust der enzymatischen Aktivität festgestellt wurde.

METTL23 ist eine mutmaßliche SAM-abhängige Methyltransferase bei der verkürzte Proteinvarianten ebenfalls mit starker mentaler Beeinträchtigung in Verbindung gebracht werden. Bei diesen verkürzten Proteinvarianten kommt es zu einer Störung der vorhergesagten katalytischen Domäne und zu einer Änderung der Lokalisierung des Proteins innerhalb der Zelle. Expressionsversuche zeigten eine starke Assoziierung von METTL23 mit "Heat Shock" Proteinen. Dies könnte darauf hindeuten, dass Heat Shock Proteine potentielle Substrate für eine Methylierung durch METTL23 sind.

N,N-8-Demethyl-8-amino-d-riboflavin dimethyltransferase (RosA) katalysiert den letzten Dimethylierungsschritt von 8-Demethyl-8-amino-d-riboflavin (AF) zum Antibiotikum Roseoflavin (RoF) in *S. davawensis*. Im Rahmen dieser Arbeit wurden sowohl die Kristallstruktur als auch biochemische und kinetische Parameter von RosA untersucht. Die Struktur von RosA ähnelt der Struktur zuvor beschriebener SAM-abhängiger Methyltransferasen. Die beiden Substrate AF und SAM binden unabhängig voneinander in ihren jeweiligen Bindetaschen, was durch einen "Random-Order" „Bi-Bi" Reaktionsmechanismus bestätigt werden konnte. Außerdem wurden Dissoziationskonstanten für Substrate und Produkte mit Isothermischer Titratrationskalorimetrie und UV/Vis Absorptions Spektromtrie ermittelt und Inhibierungsexperimente durchgeführt. Hierbei ergab sich, dass die beiden Produkte RoF und *S*-Adenosylhomocystein (SAH) stärker an RosA binden als die Substrate AF und SAM, und RoF und SAH als kompetitive Inhibitoren wirken. Dieses Ergebnis deutet auf eine mögliche Bedeutung von RosA für die Roseoflavin-Resistenz in *S. davawensis* hin.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die die den benutzten Quellen wörtlich und Inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by contend from the used sources.

Graz,…………………………..   …………………………………………..

Table of Contents

CHAPTER 1

1. Introduction to methyltransferases and methyl group donors

1.1 General aspects of methyltransferases and methyl group donors

Methyltransferases (MTase; EC 2.1.1.X) catalyze methylation reactions occurring in all living organisms. Methylation is the process of methyl transfer from methyl donors to a wide variety of methyl acceptors including DNA, RNA, proteins, lipids, polysaccharides, and a range of small molecules, leading to changes in physicochemical properties of the methyl acceptors [1]. Methylation is the key reaction involved in several biological processes *i.e.* biosynthesis, metabolism, detoxification, signal transduction, protein sorting and repair, nucleic acid processing and epigenetic regulation [1,2].

1.2 Methyl group donors

There are several methyl group donors, which are utilized to drive methylation reaction. Their structures are shown in Fig. 1.
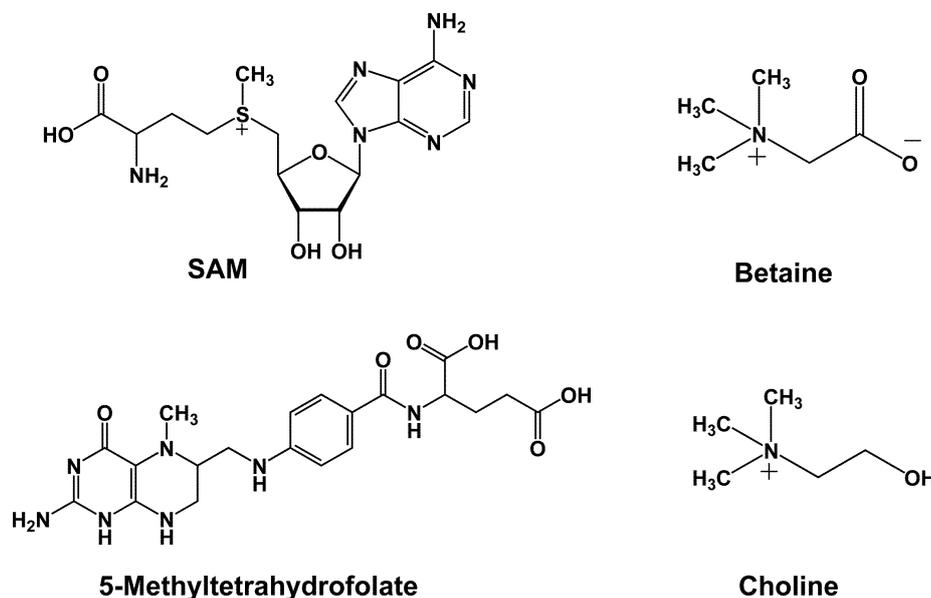
Figure 1: Biological methyl donors

1.2.1 *S*-Adenosylmethionine

S-Adenosylmethionine (Adomet or SAM) is a biological sulfonium compound and the major methyl donor for methylation reactions catalyzed by SAM-dependent methyltranserases (MTases) [3,4]. SAM-dependent methyltransferases catalyze the transfer

of a methyl group from SAM to methyl acceptors such as proteins, nucleic acids, phospholipids, and small molecules, yielding a methylated product and *S*-adenosylhomocysteine (SAH) [5]. SAM consists of a high-energy sulfonium center (Fig. 1), which activates the attached carbon (methyl group of SAM) toward nucleophilic attack [6]. SAM is an ubiquitous cofactor and takes the second place after ATP [3]. Other methyl group donors *i.e.* folate, betaine are also used by a few methyltransferases [7]. The strong preference for SAM over the other methyl donors reflects the highly favorable thermodynamics resulting from the methylsulfonium center which the $\Delta G°$ for SAM hydrolysis in the reaction of SAM + homocysteine (Hcy) $\rightarrow$ SAH + methionine (Met) is -17 kcal/mol; in comparison the $\Delta G°$ for adenosine triphosphate (ATP) hydrolysis to adenosine diphosphate and P_i is -7.3 kcal/mol [8,9]. SAM is generated from methionine and ATP by methionine adenosyltransferase (MAT, EC 2.5.1.6) [10]. SAM itself exists in two stereoisomers, (*S*, *S*)- and (*R*, *S*)-SAM. The former enantiomer is the biological active form, but the latter is a potent inhibitor of methyltransferases [11]. Generally, humans produce about 6-8 g of SAM and most of the generated SAM is utilized in methylation reactions [5,12]. After the methylation, SAH is hydrolyzed to adenosine and homocysteine by SAH hydrolase (EC 3.3.1.1) [5,12]. Both SAM and SAH function as metabolic effectors to SAM-dependent methylation reactions. In addition, both of them appear to have particular importance in the coordination and regulation of the cycle that converts methionine to homocysteine (see section 1.3) [5,13].

1.2.2 Folate

Folate is a water-soluble B vitamin that is naturally found in a variety of foods and also available in nutritional supplements. Folate is naturally occurring in the diet as folate, while folic acid is a fully oxidized monoglutamate form that is used as a dietary supplement [14]. Green leafy plants are rich sources of folate, but the vitamin is also found in fruits, nuts, meats, seafood, eggs, and dairy products. Folate is essential for several biological processes *i.e.* purine and pyrimidine synthesis and amino acid production. Moreover, folate in the form of methylenetetrahydrofolate is used as a methyl donor for methylation reaction of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP) by thymidylate synthase. The other form of folate, 5-methyl-tetrahydrofolate (MTHF) (Fig. 1) also provides methyl groups for remethylation of homocysteine, generating methionine in the methionine cycle [14,15]. There are three

enzymes, serine hydroxymethyltransferase (SHMT), methyltetrahydrofolate reductase (MTHFR), and methionine synthase that are required for production of the methyl group donors via the folate cycle. The latter enzyme is required to produce methionine, the precursor of SAM in the methionine metabolic pathway. Besides its function as a methyl donor, folate also has antioxidant effects that are relevant to asthma and allergy responses [16,17]. Deficiency of folate can cause the impairment of several metabolic pathways in the cell including hyperhomocysteinemia, low level of SAM and reduction of DNA methylation [18]

1.2.3 Betaine and Choline

Betaine is a water-soluble compound, which plays roles in a variety of biological processes, including DNA methylation. Betaine contains three chemically reactive methyl groups connected to the central nitrogen atom (Fig. 1). Moreover, betaine functions as a methyl donor, transferring its methyl group to homocysteine for generation of methionine in the homocysteine remethylation system which the reaction is catalyzed by betaine-homocysteine methyltransferase (BHMT) [15]. It is then converted into dimethylglycine, which is eliminated consequently via the kidney [15]. Betaine is normally found in several kinds of food sources, especially in wheats, grains, beets, spinach, seafood, and broccoli [14]. In addition, betaine is produced by choline dehydrogenase through the irreversible oxidation of choline. Choline (Fig. 1) also supports indirectly homocysteine methylation. From a previous study of postmenopausal women, supplemental choline increases levels of betaine and slightly attenuated levels of total homocysteine after six months [19]. Choline functions in several biological processes [20]. It is a precursor of phosphatidylcholine and sphingomyelin which are phospholipids involved in the maintenance of cell structure in biological membranes. In addition, choline also helps movements of lipids in cells [21,22]. Choline is also a precursor of platelet-activating factor and acetylcholine [14]. Choline is produced in small quantities by liver. However, it is found in many kinds of food, including soybeans, egg, peanuts, cauliflower, lentils, and flax seeds [23]. In animal models, choline supplementations reduced the inflammatory responses in patients with inflammatory arthritis [24], suppressed inflammatory responses to ovalbumin in a mouse model of allergic airway disease, and decreased generation of the reactive oxygen species [25]. In contrast, choline deficiency caused the loss of membrane phosphatidylcholine inducing cell apoptosis [26]. Increase of oxidative stress and changes of antioxidant

properties were presented in a choline-deficient diet fed mouse model [27]. In addition, a choline-deficient diet lowered formation of methionine in animal livers by 20-25%, probably because less choline was available and supplied for conversion into betaine. Consequently, the effect of choline deficiency also leads the reduction of SAM by 60% and the increase of SAH by 50% in rat liver [28].

1.3 Metabolism of SAM

SAM is a key molecule that is found in all organisms and involved in several biological processes. Generally, SAM has been known as the principle biological methyl donor. Besides being a key methyl donor, SAM is the precursor of aminopropyl groups utilized in the biosynthesis of polyamines and in the liver, it is also a precursor of glutathione, the major cellular anti-oxidant (see details in section 1.4) [5,6,12]. SAM participates in a wide variety of biological reactions encompassing three major types of reactions; transmethylation, transsulfuration, and aminopropylation (biosynthesis of bioamines) (Fig. 2) [6,29].

In transmethylation reactions (the methionine cycle), SAM is synthesized from methionine and ATP in the reaction catalyzed by methionine adenosyltransferase (MAT) [10], which is encoded by two different genes, *MAT1A* and *MAT2A*, for two homologous MAT catalytic subunits, α1 and α2 [5]. The expression of each type of MAT depends on the type of tissue. *MAT1A* is expressed solely in the liver and it encodes the α1 subunit found in two native MAT isozymes, which are either a dimer of MAT III or a tetramer of MAT I. While *MAT2A* encodes the α2 subunit found in the native MAT isozyme (MAT II) present in almost all tissues [5,6,30]. SAM functions primarily as the universal methyl donor, transferring methyl group to a wide range of acceptors from small molecules to macromolecules [5]. Although each of the SAM-dependent methyltransferases catalyzed the methyl transfer to their specific substrates, the common product of all methylation reaction is SAH. Both SAM and SAH are potent effectors. Most of SAM-dependent methyltransferases are inhibited by increases in SAH and decreases in SAM concentrations [5,31]. Thus, the removal of SAM is required to maintain the flow of the methionine cycle. *S*-Adenosylhomcysteine (SAH) hydrolase (SAHH; EC 3.3.1.1) catalyzes the hydrolysis of SAH to adenosine and homocysteine, which is a reaction occurring in all cells. This hydrolysis reaction is reversible and the equilibrium of SAH hydrolase favors the

formation of SAH. Thus, to prevent accumulation of SAH, both homocysteine and adenosine need to be metabolized or transported out of the cells [5,12,32].
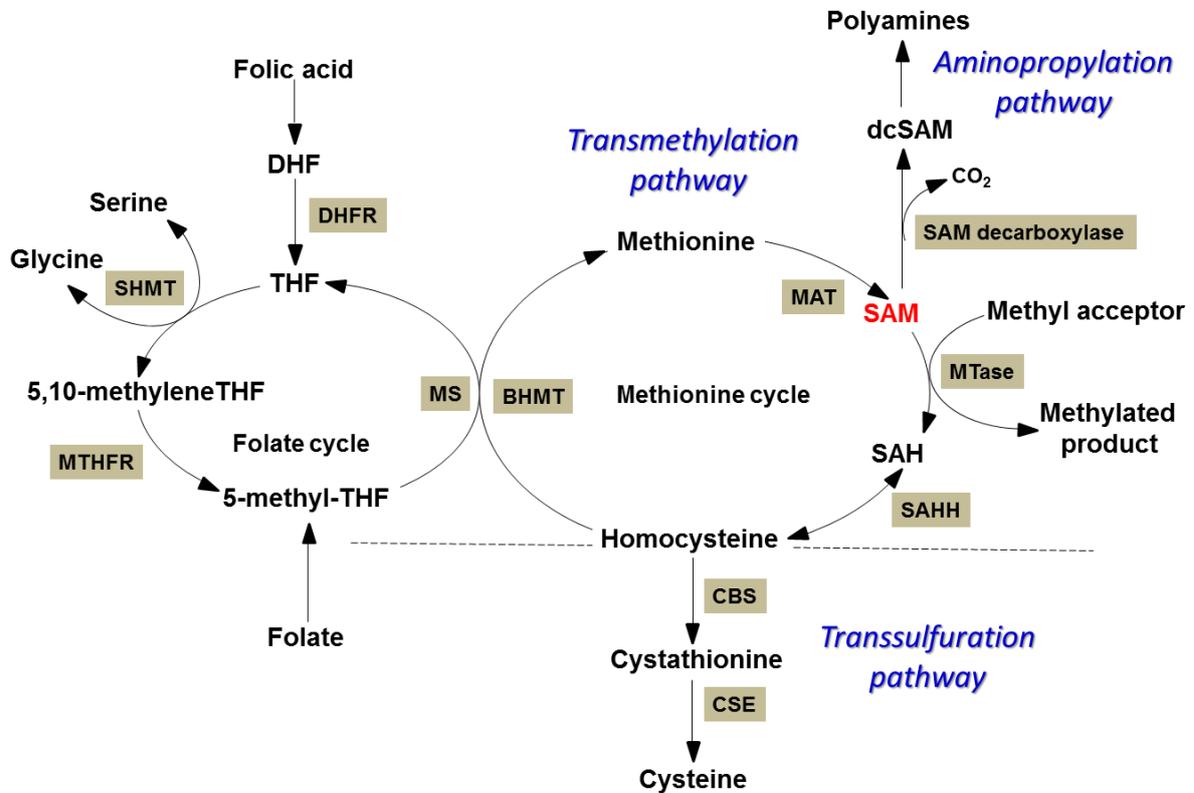
Fig. 2: Synthesis and metabolic pathways (transmethylation, transsulfuration, and aminopropylation) for SAM. MAT: methionine adenosyltransferse; MTase: methyltransferase; SAHH: *S*-adenosylhomocysteine hydrolase; MS: methionine synthase; BHMT: betaine-homocysteine methylransferase; DHFR: dihydrofolate reductase; SHMT: serine hydroxymethyltransferase; MTHFR: methylenetetrahdrofolate reductase; CBS: cystathionine β-synthase; CSE: cystathionine γ-lyase. In aminopropylation, SAM is decarboxylated and then the decarboxylated SAM (dcSAM) is channeled into the polyamine biosynthesis pathway [6,29].

There are two pathways, remethylation and transsulfuration that are utilized for homocysteine metabolism (Fig. 2) [29,31]. The remethylation pathway in the methionine cycle is the conversion of homocysteine to methionine where homocysteine acquires a methyl group from either betaine or 5-methyltetrahydrofolate. Betaine, the metabolite from choline oxidation, transfers the methyl group to homocysteine for regeneration of methionine by the catalysis of betaine-homocysteine methyltransferase (BHMT) (Fig.2), which is mainly expressed in liver and kidney [33]. The other remethylation of homocysteine is catalyzed by methionine synthase (MS), occurring in all mammalian

tissues. MS is an enzyme in the folate cycle. It utilizes a methyl group provided from 5-methyltetrahydrofolate (methylTHF). This methyl donor is produced by methylenetetrahydrofolate reductase (MTHFR). In the remethylation, MS requires methylcobalamin as its cofactor and this reaction is a part of folate and vitamin B_{12} metabolism (Fig 2). The other pathway of homocysteine metabolism is transsulfuration, which is the process to irreversibly degrade homocysteine to cysteine (Fig. 2). Condensation of homocysteine with serine to form cystathionine is catalyzed by the vitamin B_6-dependent cystathionine β-synthase (CBS). Consequently, vitamin B_6-dependent cystathionine γ-lyase (CSE) catalyzes the hydrolysis of cystathionine to cysteine and α-ketobutyrate. Cysteine reacts with glutamate and then glycine through two consecutive reactions to produce glutathione, which is an endogenous cellular antioxidant and an essential compound for detoxification of several xenobiotics [6,34]. CBS in human is expressed in liver, kidneys, muscle, brain, and ovary and also during early embryogenesis in the neural and cardiac systems [35]

In the aminopropylation pathway SAM is utilized as a precursor for polyamine biosynthesis. SAM is decarboxylated by SAM decarboxylase (EC 4.1.1.50) and its aminopropyl group is subsequently transferred, first to putrescine and then to spermidine to form polyamines (Fig. 2). Compared to transmethylation, ca. 5% of the available SAM is required for polyamine biosynthesis [5]. However, increased polyamine synthesis can be markedly induced during liver regeneration [36] and early hepatocarcinogenesis [37].

SAM is a metabolic effector for the regulation of activities of enzymes involved in the transmethylation and transsulfuration pathways. SAM functions as an allosteric inhibitor of MTHFR and as an activator of CBS. In the folate cycle (Fig. 2), increasing level of SAM suppresses the synthesis of MTHF, which is an important substrate for remethylation of homocysteine to methionine [13,31]. In the methionine cycle, SAM inhibits the activity of MAT II and also down-regulates MAT II gene expression in hepatocytes. Conversely, low concentration of SAM induces the expression of this gene in liver cells [38]. In addition to the metabolic effector SAM, increased SAH also affects both transmethylation and transsulfuration pathways. SAH inhibits methionine synthase and also betaine-homocysteine methyltransferase activities [13]. SAH limits its own synthesis by inhibiting virtually all SAM-dependent methyltransferases. Concurrently, SAH increasing shifts SAHH reaction in the direction of SAH hydrolysis. As described above, SAH inhibits the reactions of both methionine synthase and betaine-homocysteine methyltransferase, which

may increase the concentration of homocysteine. To maintain homocysteine at adequate levels, SAH and increased homocysteine activate the CBS activity to convert homocysteine through the transsulfuration pathway [13].

1.4 Other functions of SAM

SAM is a biological sulfonium compound known as the main biological methyl donor for methylation reaction, transferring methyl group(s) to wide ranging of substrates from small molecules to macromolecules. However, SAM is not only a methyl donor, but also a source of a variety of chemical groups, that are generated from several reactions as shown in Fig. 3. Interestingly, these reactions are driven by the electrophilic character of the carbon center of SAM that is adjacent to the positively charged sulfur atom [39].
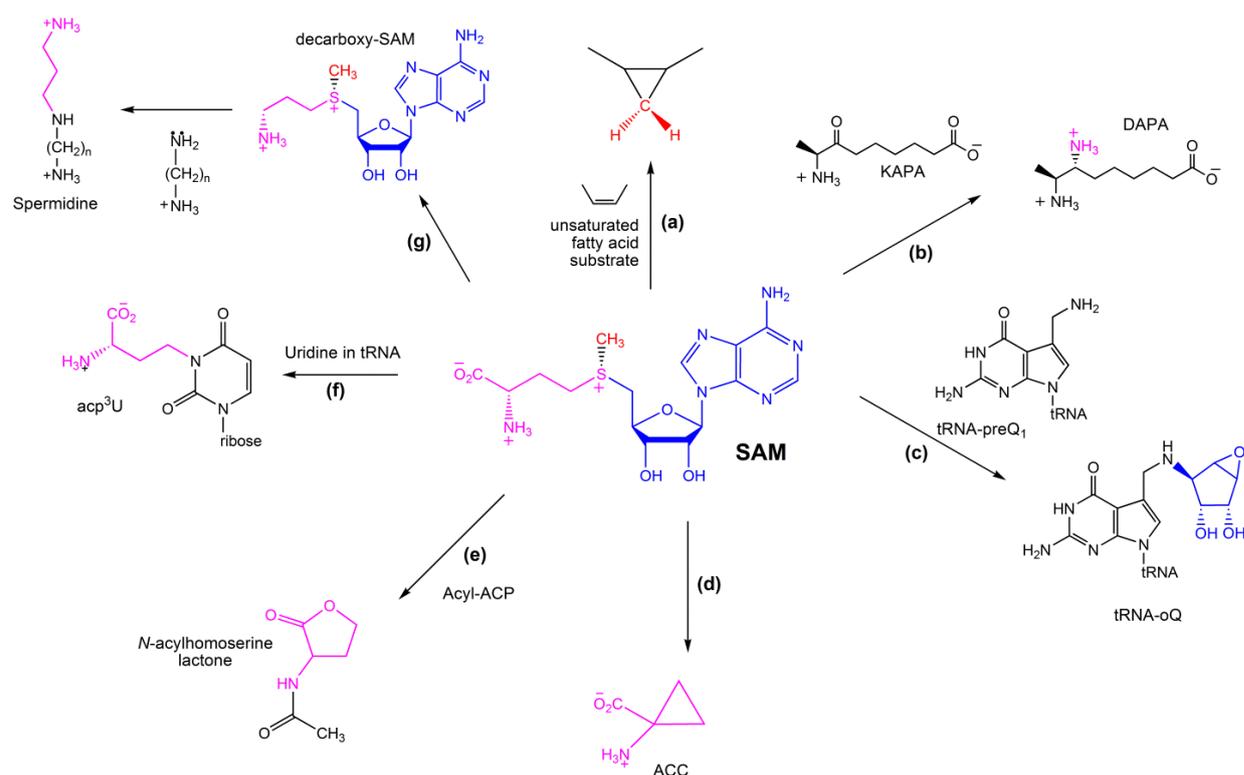
Fig. 3: *S*-Adenosylmethionine, a source of chemical groups. The reactions shown are catalyzed by (a) cyclopropane fatty acid (CFA) synthase, (b) 7,8-diaminopelargonic acid (DAPA) synthase, (c) SAM-tRNA ribosyltransferase-isomerase (QueA), (d) 1-aminocyclopropane-1-carboxylic acid (ACC) synthase, (e) acylhomoserine lactone synthase (ACP: acyl carrier protein), (f) 3-(3-amino-3-carboxypropyl)uridine (Acp^{3}U) synthase, (g) SAM decarboxylase and spermidine synthase [39].

1.4.1 SAM as a source of methylene groups

Cyclopropane fatty acid (CFA) synthase catalyzes the formation of CFA formed by methylenation of the double bond of unsaturated fatty acids using the methyl group of SAM (Fig. 3(a)) [39,40]. CFAs are found in the phospholipids of eubacteria and some eukaryotic organisms [40,41]. The mechanism of CFA formation catalyzed by CFA synthase involves the nucleophilic attack from fatty-acid double bond to the electrophilic methyl group of SAM, resulting in formation of SAH and a carbocationic form of the methylated substrate (Fig. 3(a)). A basic residue in the active site facilitates deprotonation of the methylated substrate and then ring closure follows [40].

1.4.2 SAM as a source of amino groups

7,8-Diaminopelargonic acid (DAPA) is produced by the transfer of the amino group of SAM to 7-keto-8-amino pelargonic acid (KAPA) (Fig. 3(b)), which is the antepenultimate reaction step in biotin biosynthesis catalyzed by DAPA synthase, a pyridoxal 5'-phosphate (PLP)-dependent aminotransferase [42]. The three-dimensional structure of DAPA synthase belongs to the type-I fold of aminotransferases which was first described for aspartate aminotransferase [43].

1.4.3 SAM as a source of ribosyl groups

SAM also donates its ribosyl group in the penultimate step of the biosynthesis of queuosine, a hypermodified tRNA nucleoside (Fig. 3(c)) [44]. Queuosine, a 7-deazaguanosine containing an aminomethyl cyclopentenediol sidechain occurs exclusively at position 34 in the anticodon loop of bacterial and eukaryotic tRNAs, coding for asparagine, aspartic acid, histidine, and tyrosine. SAM-tRNA ribosyltransferase-isomerase (QueA) catalyzes the transfer of the ribose moiety of SAM to 7-aminomethyl-7-deaza-guanosine in tRNA-preQ1, following the rearrangement to the epoxy-carbocycle (tRNA-oQ) [45].

1.4.4 SAM as a source of aminoalkyl groups

1-Aminocyclopropane-1-carboxylic acid (ACC), the precursor of plant hormone ethylene, is synthesized from conversion of SAM to ACC and 5'-methylthioadenosine (5'MTA) by the PLP-dependent ACC synthase (Fig. 3(d)) [46]. The transformation of SAM to ACC and 5'MTA initiates the formation of a PLP-substrate aldimine adduct and deprotonation of the α carbon (C2) of the amino acid moiety of SAM. The 5'MTA moiety of SAM is a good leaving group that facilitates the intramolecular nucleophilic attack of the C4 carbon by the activated C2 carbanion. ACC is released after hydrolysis of the PLP-ACC adduct. Consequently, ACC is oxidized to ethylene by ACC oxidase [47]. In addition, *N*-acylhomoserine lactone, which is produced in luminescent bacteria and participates in inter-bacterial communication adopts the aminobutyryl group derived from SAM and the acyl group provides by acyl carrier protein (Fig. 3(e)) [48]. The other example is the synthesis of 3-(3-amino-3-carboxypropyl)uridine (acp^{3}U) in phenylalanine tRNA by alkylation with the aminocarcoxypropyl group of SAM (Fig. 3(f)) [49].

1.4.5 SAM as a source of 5'-deoxyadenosyl radicals

Radical SAM enzymes catalyze reductive cleavage of SAM, yielding methionine and a highly oxidized 5'-deoxyadenosyl radical intermediate [50]. The latter radical then abstracts a hydrogen atom from a substrate to generate 5'-deoxyadenosine and a substrate radical, which is then converted to the final reaction product [51]. For instance, RlmN, a radical SAM methyltransferase catalyzes the methylation of adenosine 2503 at C2 in 23*S* rRNA [52] (see section 1.5.2).

1.5 Methyltransferases

Methylation, the addition of a methyl group to biologically active molecules, is an important biological process, which is involved in biosynthesis, detoxification, signal transduction, nucleotide and protein modifications [1]. In the 1950s discovered by Cantoni's group, SAM serves as the actual methyl donor [53] and at that time SAM was thought to methylates only small molecules. Later in the 1960s the nature of protection of bacterial DNA by modification enzymes against cognate restriction *i.e.* methylation was discovered [54]. Presently, hundreds of methyltransferases (about 323 MTases) catalyzing the diversity of methylation have been discovered (http://enzyme.expasy.org/EC/2.1.1.-).

Although several classes of methyltransferases are known, the great majority of methylation reaction is catalyzed by *S*-adenosylmethionine methyltransferase (SAM-MTase) [9]. Recently, methylation reactions catalyzed by radical SAM methyltransferases have been discovered. These radical SAM-MTases methylate their substrates at unreactive non-nucleophilic centers through a radical mechanism [55] (see section 1.5.2). Besides the major methyl donor SAM, various forms of folate donate methyl groups to several targets. A well-known folate-dependent methyltransferase is thymidylate synthase [56] (see section 1.5.3).

1.5.1 SAM-dependent methyltransferases

A wide variety of substrates, ranging from small- to macro-molecules, are the targets for methylation by SAM. The addition of a methyl group occurs at C, O, N, S atoms of the targets or even at halide and arsenic atoms.

For *O*-methylation, the most widely studied enzyme is catechol-*O*-methyltransferase (COMT; EC 2.1.1.6), which catalyzes the methylation of small molecules containing a catechol structure, involved in the elimination of catechol-based neurotransmitters, catechol steroids, and xenobiotic catechols [57]. In addition, plants appear to be rich sources of *O*-MTases. For instance, caffeic acid and caffeoyl-CoA *O*-methyltransferases (C-3-OMT and CCoAOMT, respectively) are the enzymes involved in lignin biosynthesis (Fig. 4) [58]. Besides small molecule *O*-methyltransferase, *O*-methylation of carbohydrate derivatives and dicarboxylic amino acids are abundant in nature [9].

N-methyltransferases such as protein arginine and protein lysine *N*-methyltransferases, play an important role for post-translational modifications [55]. In chromosomal DNA, N6-methyladenine (N6mA) and N4-methylcytosine (N4mC) are found to be produced by DNA methyltransferases [9]. Moreover, RNA methylation at N7-guanosine, N1-guanosine and N1-adenine are also found. *N*-methyltransferases also play a prominent role in natural product biosynthesis and are particular prevalent in the biosynthesis of many alkaloids produced by plants [55]. *N*-methyltransferases are also common in the secondary metabolism of bacteria [55]. For example, Asm10 methylates the amide NH group of a macrolactam precursor in the ansamitocin biosynthesis, a potent antitumour agent produced in *Actinosynnema pretiosum* (Fig. 4) [59].

DNA and RNA methylation at C-atoms of nucleotides such as C5-methylcytosine, C5-methyluridine, and C2-methyladenine are catalyzed by C-methyltransferases. For biosynthesis, C-methyltransferases generally catalyze the methylation of nucleophilic atoms of ketones and phenolic substrates [55]. Some bacterial C-methyltransferases also catalyze stereoselective methylation of α-keto acids derived from glutamate, phenylalanine, arginine, and tryptophan. An example of this stereoselective methylation catalyzed by GlmT yields (3R)-methyl-2-oxoglutaric acid, which is then transaminated by the branched chain amino acid transaminase IlvE (SCO5523) to produce (2S,3R)-methylglutamic acid, a precursor in the biosynthesis of calcium-dependent antibiotics (Fig. 4) [60].
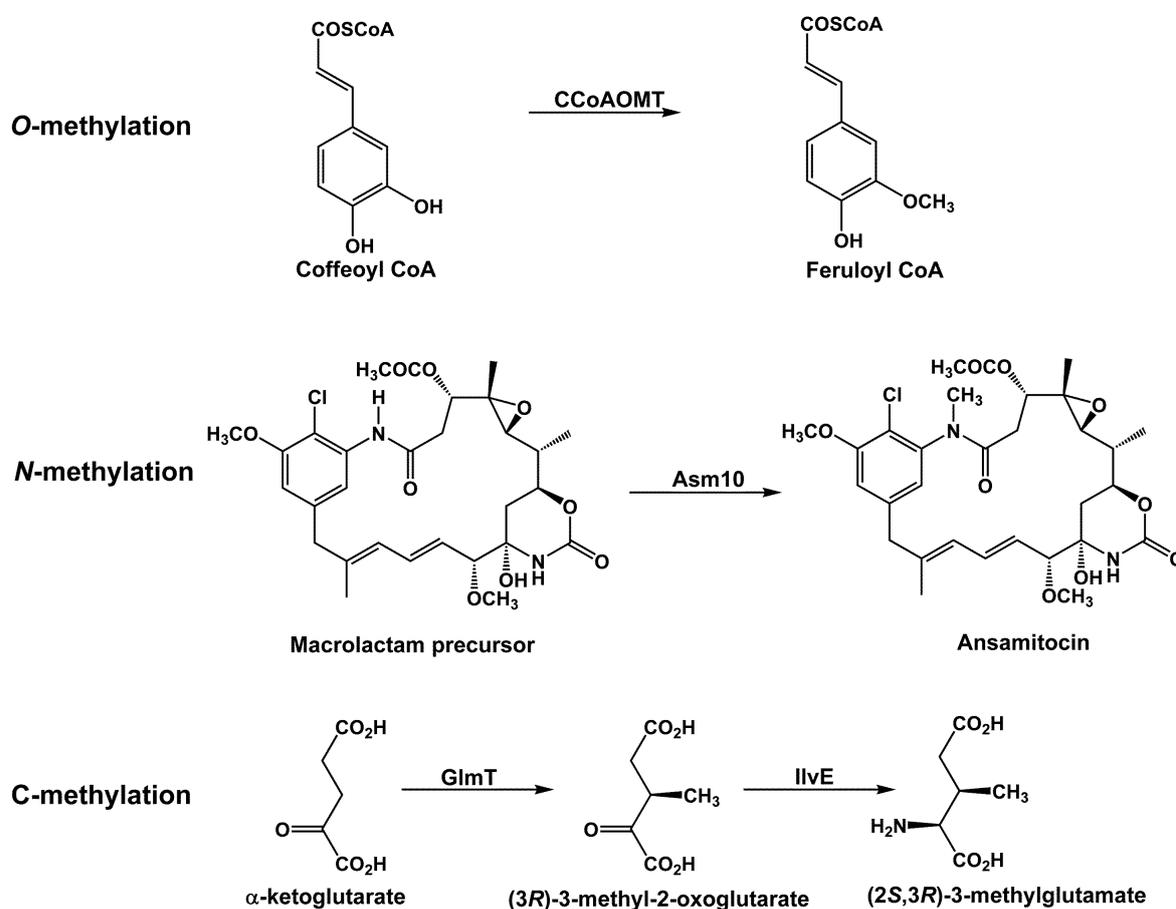
Fig. 4: Methylation reactions by SAM-dependent methyltransferases at O-, N-, C-, atoms of the target substrates [55].

The other lesser-known methyltransferases are halide methyltransferases from plants and microorganisms, which add a methyl group to either chloride, bromide, or iodide, resulting in the formation of methylhalides [61]. Selenium (Se) and arsenic (As) methylation are required for detoxification pathways in plants and bacteria [9,55]. Conversion of the toxic

selenocysteine to Se-methylselenocystein which is then metabolized to dimethyldiselenide ($CH_3SeSeCH_3$, a volatile compound) is catalyzed in plants by selenocysteine methyltransferase to reduce the selenocysteine toxicity and emit it into the atmosphere [62]. In addition, microorganisms and plants also metabolize arsenic by methyltransferases to generate the volatile trimethylarsine [63].

1.5.1.1 Mechanism of SAM-dependent methyltransferases: methyl group transfer

Since SAM contains a high-energy sulfonium ion (Fig. 1) that facilitates the nucleophilic attack on the attached carbon atoms. Thus, SAM-dependent methyltransferases catalyze the transfer of the methyl group from SAM through the S_N2 mechanism, which requires a linear rearrangement of the nucleophilic methyl acceptor, the methyl carbon and the thioester leaving group of SAM in the transition state. The methylation reaction catalyzed by SAM-MTases is the nucleophilic attack (S_N2) at sp^3 of the methyl carbon of SAM [9]. In Fig. 5, an example is presented for the reaction mechanism of the N6-adenine methylation catalyzed by a DNA SAM-MTase, RsrI [64]. The nucleophilic attack of N6-adenine toward the methyl group of SAM is facilitated by H-bond interaction to the peptide backbone between Pro66 and Pro67 and deprotonation of adenine by a proximal basic residue Asp65 [64]. Even though, methyl transfer reactions by SAM-dependent methyltransferases mainly occur via the nucleophilic attack of target atom of substrates, there are still several SAM-dependent methyltransferases defined as radical SAM methyltransferases which employ methylation through a radical-mediated mechanism (see section 1.5.2)
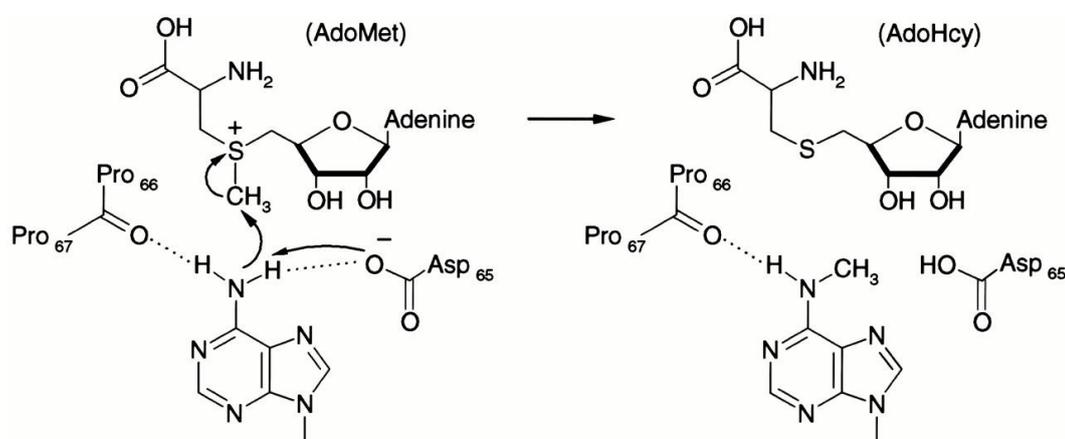
Fig. 5: Proposed catalytic mechanism of N6-adenine methylation by DNA methyltransferase, RsrI [64].

In the case of *O*-, *N*-, and *S*-methylation, oxygen, nitrogen, or sulfur acts as the nucleophile that attacks the methyl group of SAM following a direct methyl transfer in the S_N2 transition state [9]. In contrast to *O*-, *N*-, and *S*-methylation, methyl transfer for C-methylation requires the generation of a carbanion on the target substrate to accept the methyl group from SAM, since the carbon atom is not a readily polarized nucleophile and there are normally no lone-pair electrons available. Generating a carbanion of a methyl acceptor for C-methylation is achieved when that carbon is a part of a resonance structure, *i.e.* ketone and phenolic compounds. Thus, C-methylations occur to the target carbon next to −C=C− double bond. In addition, C-methylation needs additional steps during catalysis to generate a carbanion at the target carbon [9]. For instance, C5 cysteine methylation catalyzed by DNA methyltransferases, the sp^2-hybridized carbon requires nucleophilic assistance by a cysteine residue at the active site [65].

1.5.1.2 Classification of SAM-dependent methyltransferases

Currently, SAM-dependent methyltransferases are categorized in five classes based on their structural features [66]. The tertiary structures of the five classes of SAM-dependent methyltransferases are presented in Fig. 6 [66].
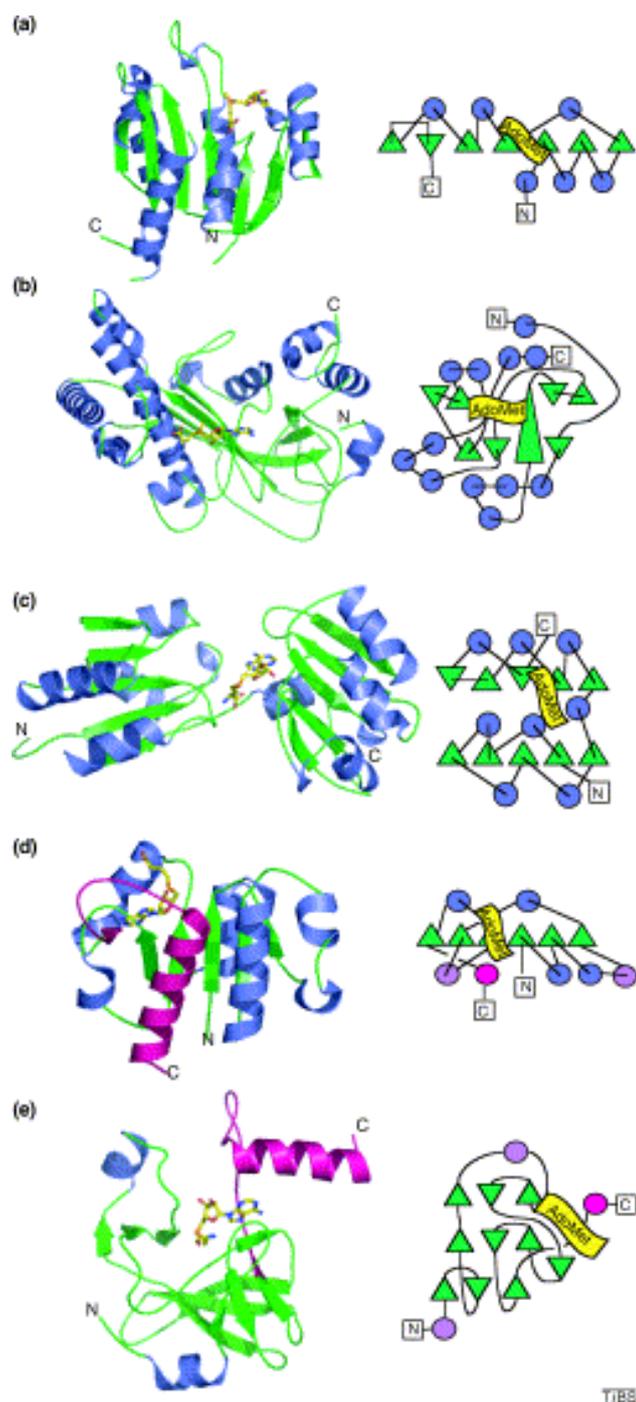
Fig. 6: The tertiary structures of the five classes of SAM-dependent methyltransferases, represented ribbon models (left) and topology diagrams (right). (a) Class I: the major group of SAM-dependent methyltransferases consisting of the MTase core, a seven-stranded β sheet with three helices flanking each side ((a): the structure of M.Hhal). (b) Class II: the methionine synthase domain containing a series of long β strands and the SAM binding pocket in a shallow groove on the surface ((b): the structure of MetH). (c) Class III: the CbiF structure consists of the SAM binding pocket located between the two αβα domains and the active site located in the N-terminal domain. (d) The SPOUT family, the SAH binding pocket located at the C-terminus of a knot structure (highlighted in magenta) ((d): the structure of Yibk). (e) Class V: SET domain of histone-lysine-*N*-methyltransferase [66].

Class I

Class I is the major group of SAM-dependent methyltransferases. The structure of enzymes belonging to this class is remarkably similar, containing the shared MTase fold (incorporation of β strands, β1-β7, and α helices, αZ and αA-αE) comprising a seven-stranded β sheet with a central topological switch-point and a reversed β hairpin at the carboxyl end of the sheet (6↑ 7↓ 5↑ 4↑ 1↑ 2↑ 3↑) (Fig. 6(a)) [66]. The seven-stranded β sheet is flanked with six α helices which are located on either side of the β sheet, forming a doubly wound open αβα sandwich (Fig. 6(a)) [1,9,66]. The MTase fold is remarkably similar to the Rossmann fold that consists of a continuous six-stranded parallel β sheet. The only difference between these two folds is the addition of the antiparallel β7 strand between β5 and β6 in the MTase fold [9]. It should be noted that class I members have a remarkable structural consistency, even though they catalyze the methylation of a wide range of substrates and have only 10% similarity out of the sequence level [55]. Despite this low sequence similarity (~10%), there are still three highly conserved regions; the loop next to β1 strand of the MTase fold ("the G loop", exhibiting the motif GxGxGx or GxG), an acidic residue at the carboxyl end of the β2 strand, and the loop (a hydrophobic Val/Ile/Leu motif) next to the β4 strand (in the P loop) [9]. Generally, in SAM-dependent methyltransferases, the SAM-binding region is localized at the amino-terminal part of the β sheet of the MTase fold and is formed in part by residues from loops following β1-β3 strands [1]. The three highly conserved regions, the G loop, the acidic residue at the carboxyl end of the β2 strand and the P loop, participate in SAM binding with interactions to the methionine carboxylate, the hydroxyl groups of the ribose, and the methionine amino group of SAM, respectively [9]. The substrate binding region is localized at the carboxyl terminal end of the β sheet of the MTase fold [1]. With the diversity of sizes, shapes and chemical properties of substrates, additional domains are inserted in the MTase fold for substrate recognition [1,66].

The three dimensional structure of catechol *O*-methyltransferase (COMT), one of the smallest structures of SAM-dependent methyltransferases, with SAM and the inhibitor 3,5-dinitrocatechol contains only the consensus MTase core and exhibits SAM- and substrate-binding sites at the carboxyl and amino terminal ends, respectively (Fig. 7) [67]. The amino acid residue G66 in the glycine motif (in the G loop), and E90, the conserved acidic residue at the carboxyl end of β2, play a role for SAM binding. In the substrate binding

pocket with the native substrate (catecholamine) modelling, proton abstraction at the hydroxyl catechol is induced by K144, facilitating a direct nucleophilic attack of the phenolate substrate to the methyl group of SAM. The other catechol hydroxyl group is stabilized by E199 through H-bond interaction. W38 and W143 contribute hydrophobic contacts for dihydrocatechol binding (Fig. 7) [67].
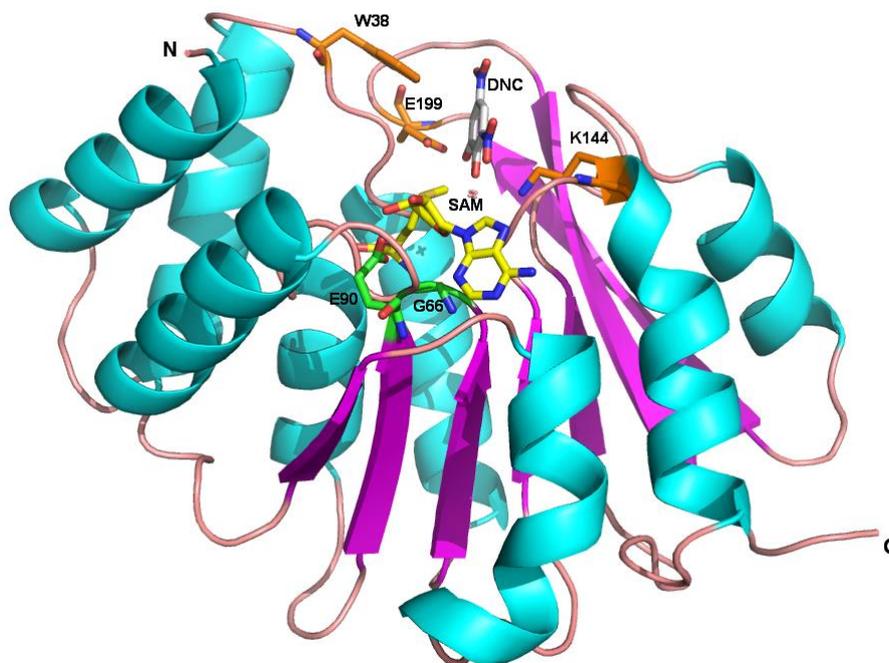
Fig. 7: The three-dimensional structure of COMT with SAM and 3,5-dinitrocatechol (DNC) (pdb: 1VID). Cartoon representation of COMT with binding of SAM, and 3,5-DNC, showing α helices in cyan, β sheets in magenta, loops in tint, SAM in yellow, and 3,5-DNC in grey. Amino acid residues that interact with SAM and 3,5-DNC are shown as green and orange sticks, respectively [67].

Class II

Methionine synthase (MetH), a cobalamin (vitamin B12)-dependent methyltransferase, catalyzes the transfer of a methyl group from methyltetrahydrofolate to homocysteine, generating methionine and tetrahydofolate [68]. The catalytic cycle of methionine synthase is initiated by the transfer of a methyl group of the folate derivative to the cobalamin cofactor and subsequently to homocysteine. Occasionally, the cobalamin cofactor undergoes oxidation and needs a reductive reactivation by SAM and an electron source. In this case, SAM serves as a methyl donor for the cobalamin reactivation and is bound to the

C-terminal domain of MetH. The source of electron is either flavodoxin/NADPH-flavodoxin oxidoreductase in *E. coli* or methionine synthase reductase in humans [69].
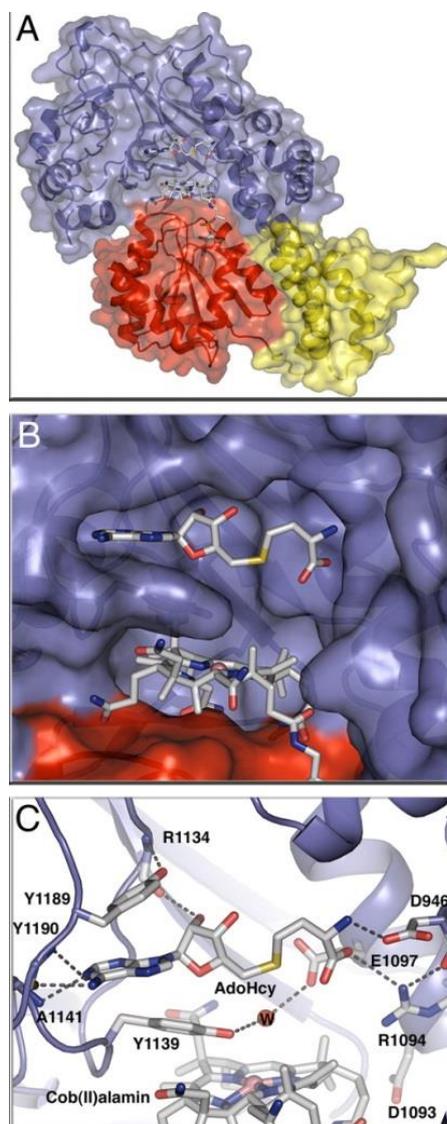
Fig 8: (a) The 3D-structure of complex of MetH, SAH, and cobalamin (Co(II)Cbl); showing the SAM binding domain in blue, the cobalamin binding domain in red, and the Cap domain in yellow. (b) Surface performance of the SAM binding domain in blue, the cobalamin binding domain in red. (c) Interactions of residues at SAM binding domain to SAH and cobalamin. Dashed lines represent H-bond interactions [71].

The MetH reactivation at the C-terminal domain contains a long, central, antiparallel β sheet flanked by groups of helices at either end (Fig. 6(b)) [66,70]. The SAM binding site is localized in a shallow grove along the edges of the β strands, forming H-bonds to a conserved RxxxGY motif [66,70]. As seen in Fig. 8 [71], two invariant residues; Y1139, the residue in the conserved motif of RxxxGY, and E1097 interact with the water

molecule, that coordinates to Co(II) of the cobalamin. Reposition of these residues induced by SAM (SAH) binding generates their H-bond interactions to the water formerly bound to Co of cobalamin which stabilizes the reactivation conformation and thereby facilitates reduction of cobalamin and consequently methyl transfer from SAM.

Class III

Cobalt-precorrin-4 methyltransferase (CbiF), a SAM-dependent methyltransferase, catalyzes the methylation of carbon atoms in large, planar precorrin during cobalamin biosynthesis [72]. CbiF is a homodimer comprising two α/β domains, each containing a five stranded β sheet flanked by four α helices. The two domains are linked by a single coil (Fig. 6(c)) [66,72]. However, the two domains have no topological similarity. Unlike the class I SAM-MTase, a conserved GxGxG (GAGPG, residues 27-31) motif at the carboxyl terminal end of β1 strand does not contact the SAM binding site but instead is localized in a large trough between the two domains (Fig. 6(c)). In the SAM-binding pocket, E103 forms H-bonding interactions to the carboxyl and amino group of homocysteine. T101, M106 and a water molecule also interact with the amino group of homocysteine amine through H-bonding. While there are hydrogen bonds between the homocysteine moiety of SAH and the partially conserved T131 and S132 through the side chain hydroxyl groups (Fig. 9) [72].
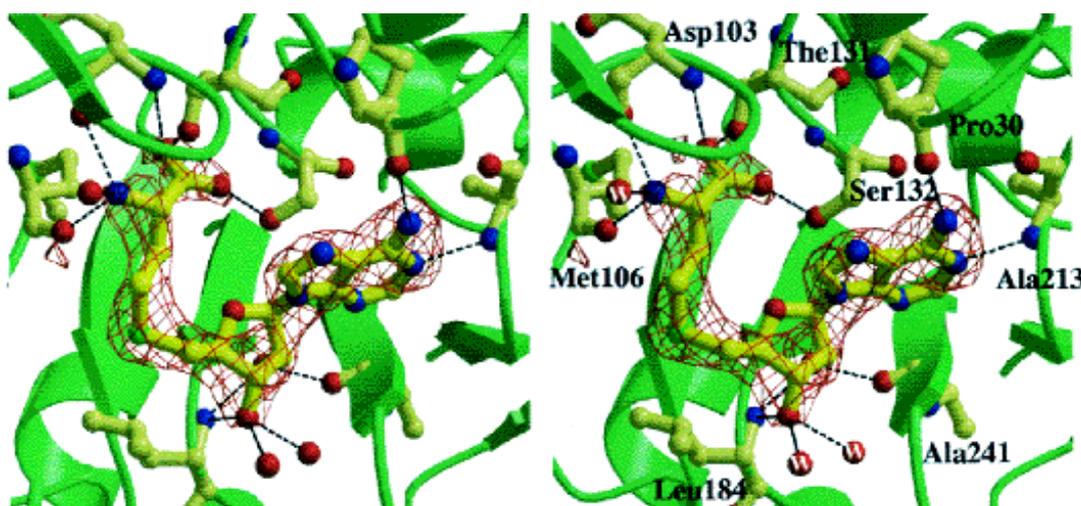
Fig. 9: Schematic representation of the SAH binding pocket with surrounding residues. Dashed lines represent H-bond interaction [72].

SPOUT methyltransferases (**spo**U and **t**rmD RNA methylase superfamilies) are a class of SAM-dependent methyltransferases. They present an unusual fold with a very deep topology knot (classified as the α/β knot superfamily) at the C-terminus, formed by the last ~30 residues and residues in the last α helix (Fig. 6(d)) [66,73]. SAM-dependent methyltransferases, i.e. RNA methyltransferases, in the SPOUT family share a common fold characterized by a 5-stranded parallel β sheet sandwiched between two layer of helices. Moreover, the SPOUT domain can be subdivided into the variable N-terminal subdomain (Rossmanoidal α/β fold) and the conserved C-terminal trefoil knot [74]. Insertion of at least an α/β module in many enzymes occurs at the N-terminal subdomain and the latter subdomain plays an important role for SAM binding. The cofactor binding loop in the knot of one protomer is stabilized by interactions with the other protomer. Moreover, structures of SPOUT members revealed that the active site is formed from participation of residues of both protomers. It should be noted that SPOUT members are generally present as dimers, which is essential for the methylation activity. There are two modes for protein dimerization where the two β sheets from the two monomers are in a 'perpendicular' (TrmL) [75] or 'antiparallel' (TrmD) [76] mode (Fig. 10).

*Ec*TrmL
'perpendicular' mode

TrmD
'antiparallel' mode
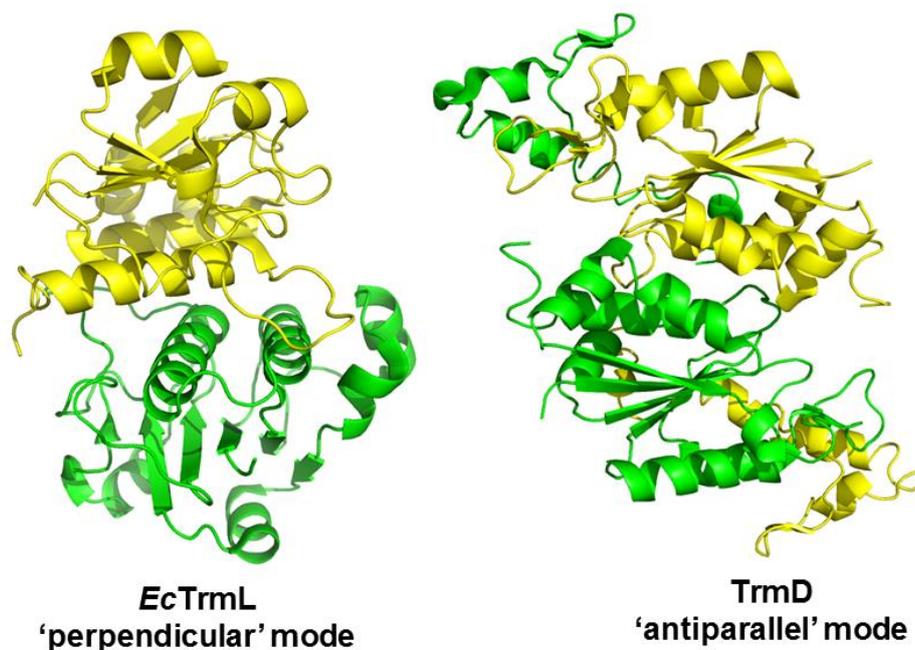
Fig. 10: Two modes of dimerization observed among SPOUT methyltransferases: (left) 'perpendicular' mode dimerization of TrmL (pdb: 4JAL) from *E. coli* [75], (right) 'antiparallel' mode dimerization of TrmD (pdb: 1OY5) from *Aquifex aeolicus* [76].

Class V

Protein methylation of lysine or arginine side chains is characterized as epigenetic targets, since it can regulate activities on histones and transcription [77]. Protein lysine methyltransferases (PKMTs) contains a catalytic SET [originally derived from three *Drosophila* genes involved in epigenetic processes: Su(var)3-9; En(zeste); Trithorax (78)] domain of approximately 130 amino acids. SET-domain methyltransferases catalyze the methylation of lysine in the flexible N-termini of histones, including ribulose-1,5-bisphosphate carboxylase (Rubisco) [79]. Four highly conserved motifs, which are SET motif I (GxG), SET motif II (YxG), SET motif III (RFINHxCxPN) and SET motif IV (ELxFDY) have been found in SET-domain proteins [80]. The four conserved motifs participate in the three steps of methylation, *i.e.* binding of SAM (motif I, and the first half of motif III), methyl transfer (the Tyr of motif II), formation of the hydrophobic target lysine-binding channel (the second half of motif III and motif IV) [80]. Typically, SET-domain methyltransferases contain a series of eight curved β strands forming three small sheets (Fig. 6(e)) and the C-terminus tucked underneath a surface loop forming an unusual knot, not found in the other classes of SAM-dependent methyltransferases [66,72]. Formation of the knot structure leads the two conserved domains, motif III and IV, to form an active site, which is located next to the SAM-binding pocket and the hydrophobic target lysine-binding channel [80]. Besides the SET-domain, I-SET (immunoglobulin-SET) and post-SET domains also participate in the lysine-binding channel and SAM-binding site [81]. Upon the methyl transfer, a target lysine residue binds to the hydrophobic channel within the SET domain and is reoriented by hydrogen-bond interaction of the terminal ε amine to the hydroxyl of the catalytic Tyr (in SET motif II, Tyr245 in SET7/9) (Fig. 11) [80,82]. Moreover, a water channel facilitates deprotonation during the methyl transfer (Fig. 11). Normally, lysine may possibly acquire one, two, or three methyl groups from SAM. Some SET-domain enzymes, for instance, the Tyr305 in SET7/9 (Fig. 11) plays role for substrate specificity since the hydroxyl of Tyr305 may act as a source of steric hindrance, limiting the methylation, Thus, SET7/9 catalyzes only lysine monomethylation because steric hindrance from Tyr305sterically block further methylation [80,82].

Fig. 11: Schematic representation showing interactions of the amino group of lysine 4 in SET7/9 with the surrounding residues and a water channel in the presence of SAM [82].

The other protein SAM-dependent methyltransferase is protein arginine methyltransferase (PRMT), which catalyzes methyl transfer to flexible N termini of arginine of histones. PRMTs are structurally unique and different from SET-domain protein lysine methyltransferases. At their catalytic core, the proteins comprise three domains: MTase domain (the methyltransferase core, the Rossmann-like fold) for SAM binding, the β-barrel domain, and a dimerization domain [81]. The interface between the MTase and the β-barrel domain provides the substrate-binding pocket for peptide substrates while the dimerization domain is a conserved feature in the PRMTs family. The dimerization domain is required to engage residues in SAM-binding site, generating the proper binding site for SAM [80]. PRMTs were assigned to class I (a seven-stranded β sheet flanked with helices), since they also contain the MTase core [81].

1.5.2 Radical SAM-dependent methyltransferases

Typically, as described in section 1.5.1.1, the electron-deficient sulfonium ion of SAM renders its carbon (the methyl group) electrophilic and nucleophilic methyl donor attacks directly that methyl group, yielding a methylated product and SAH. Thus, methylation reactions catalyzed by SAM-dependent methyltransferases proceed through a S_N2 displacement mechanism [9]. To complement the S_N2 methylation, a radical-mediated mechanism for methylation of non-nucleophilic substrates is required and catalyzed by radical SAM-methyltransferases [83]. These enzymes need two equivalents of SAM, one as a methyl donor and the other as a highly oxidizing 5'-deoxyadenosyl 5'-radical (5'-dA·). The latter and homocysteine are produced from the reductive cleavage of SAM. 5'-dA·

initiates radical-based methylations (including other transformation) by abstracting a substrate hydrogen atom (H· radical) [84]. All radical SAM-dependent methyltransferases contain at least one [4Fe-4S] cluster, typically coordinated by three cysteine residues in a conserved CxxxCxxC motif where the [4Fe-4S] cluster acts as an electron supplier for the reductive cleavage of SAM, generating 5'-dA· [85]. After the reductive cleavage of SAM, in *E coli*., the [4Fe-4S] cluster is reduced and regenerated by an electron from the flavodoxin/flavodoxin reductase reducing system or in vitro by reducing chemicals (dithionite or illuminated deazaflavin) [86]. Structures of SAM radical enzymes commonly exhibit a full or partial β-TIM (triose phosphate isomerase) barrel fold consisting of eight α helices alternating with eight β strands (($\beta\alpha$)$_8$-barrel), forming a barrel-like structure with the β strands on the interior and the α helices proximal aligning on the protein surface [84]. The [4Fe-4S] cluster and the substrate-binding sites localize inside the full or partial β-TIM barrel fold [87]. Currently, based on the sequence, cofactor requirement, and reaction mechanism of the radical SAM-methyltransferases, they can be cataloged into four classes [83]

Class A of radical SAM-dependent methyltransferases

RlmN and Cfr from several organisms are the radical SAM-dependent methyltransferases in class A. RlmN methylates the C2 atom of adenosine 2503 (A2503) in 23*S* rRNA, and C2 of adenosine 37 which plays an important role for enhancing translational fidelity and the nascent peptide response [89]. Cfr preferentially catalyzes methylation of C8 of the same nucleotide A2503, conferring bacterial resistance to at least five classes of antibiotics acting upon the peptidyl transferase center [90]. Class A have the conserved motif of three cysteines CxxxCxxC and two additional conserved cysteines (Cys118 and Cys355 in RlmN and Cys105 and Cys338 in Cfr), which are the distinctive characters essential for catalysis [85]. In addition, class A enyzmes catalyze methylation at sp^2-hybridized carbon atoms. Studies of single-turnover reactions of RlmN and Cfr revealed that the reactions proceed via a ping-pong mechanism.
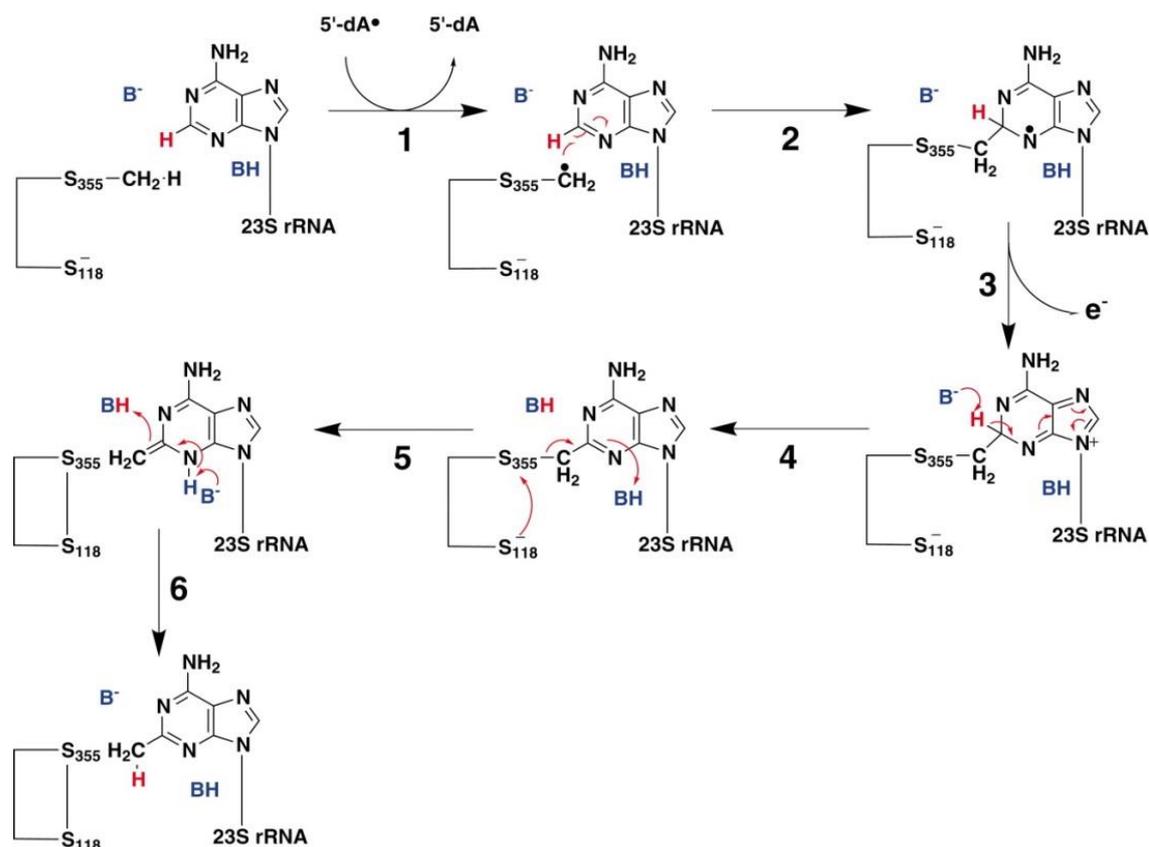
Fig. 12: Proposed mechanism of methylation of adenosine 2503 (A2503) in 23S rRNA catalyzed by RlmN [88].

As discussed in details by Grove's group [85,88], catalysis of enzymes on this class is initiated by the methylation of a conserved cysteine (Cys355 in RlmN) by SAM through a common S_N2 displacement mechanism (Fig. 12). A second SAM binds at the same binding site [91] and is then reductively cleaved to a 5'-dA· radical that subsequently abstracts a hydrogen atom of the methylthio group, yielding a protein-based methylene radical (Fig. 12). Attack of the protein radical to the C2 adenine ring (for RlmN) causes the loss of an electron from the adenine ring to the 5'-dA· and generating 5'-deoxyadenosine and the adenine-protein adduct. The other conserved cysteine (Cys118 in RlmN) initiates reductive cleavage of the thioether covalent adduct, resulting in the formation of a disulfide bond of both cysteines and the release of the methylated product (Fig. 12) [83,85,88]

Class B of radical SAM-dependent methyltransferases

Currently, class B is the largest class of radical SAM-dependent methyltransferases. Unlike class A, enzymes of class B catalyze methylation not only of sp^2-hybridized carbon atoms,

but also sp^3-hybridized forms, including phosphinate phosphorous atoms. The other distinctive character of these enzymes is the N-terminal addition of cobalamin-binding domain (CBD). An enzyme classified in the class B of radical SAM-dependent methyltransferases, PhpK, catalyzes the methylation of the phosphinate group of 2-acetylamino-4-hydroxyphosphinylbutanoate to generate 2-acetylamino-4-hydroxymethylphosphinylbutanoate, an analog of glutamate, which acts as an inhibitor of bacterial and plant glutamine synthetases [88]. Another class B enzyme that methylates its substrate at sp^3-hybridized carbon is GenK. In *Micromonspora echinospora*, GenK is expressed to respond to the methylation of C-6' of gentamicin X_2, an intermediate in the biosynthesis of the aminoglycoside antibiotic gentamicin C1 [88,92]. Besides GenK, Fom3 methylates stereoselectively at the sp^3-hybridized C2 of 2-hydroxyethylphosphonate to *S*-2-hydroxypropylphosphonate, in the penultimate step in the biosynthesis of the antibiotic fosfomycin [88,93]. Unlike the requirement of two conserved cysteines (*i.e.* Cys118 and Cys355 in RlmN) of class A radical SAM-dependent methyltransferases, class B needs methylcobalamin (MeCbl) for methylation catalysis. A proposed mechanism for class B methylation proceeds initially by the reductive cleavage of SAM, yielding a 5'-dA·, which then abstracts a hydrogen atom of the substrate. This generates a substrate radical intermediate and the byproduct, 5'-deoxyadenosine, is released. Subsequently, the former intermediate reacts to MeCbl, transferring the methyl group from MeCbl to the intermediate to obtain Cob(II)alamin and the product G418. The second SAM is utilized for reductive methylation of Cob(II)alamin, regenerating MeCbl for the next turnovers by adopting an electron supplied by the flavodoxin/flavodoxin reductase reducing system [83,93,94].

Class C of radical SAM-dependent methyltransferases

Radical SAM-dependent methyltransferases in class C catalyze the methyl transfer of the sp^2-hybridized carbon atoms, similar to the methyl transfer catalyzed by class A. However, class C members do not exhibit the two conserved cysteines *i.e.* Cys118 and Cys355 in the class A RlmN [88]. So far, all class C radical methyltransferases catalyze the methylation reactions involving in biosynthesis of complex secondary metabolites, presenting antitumor and antibiotic behaviors [88]. For example, two class C members, NosN and NocN, catalyze the methylation of C4 of 3-methyl-2-indolic acids (MIA) moiety in the biosynthesis of nosiheptide and nocathiacin [95-97]. TpdI, TpdL, and TpdU are class C

enzymes involved in methylation of thiazole heterocycles in the biosynthesis of GE2270 and thiomuracin [98]. Class C SAM-dependent methyltransferases have significant sequence similarity with coproporphyrinogen III oxidase (HemN), a radical SAM-dependent enzyme that converts coproporphyrinogen III to protoporpyrinohen IX, involved in the pathway for heme biosynthesis [99]. HemN structure reveals that it contains a similar TIM barrel radical SAM core like the other classes of radical SAM-dependent methyltransferases, but it also possesses two additional domains at the C-terminus and the beginning of the N-terminus (called 'trip-wire' domain) [100]. With these two additional domains, the former is predicted to be possibly participating in SAM/substrate binding [83]. The crystal structure of HemN also contains two simultaneously bound SAM molecules, one of which is coordinated to the [4Fe-4S] cluster producing a 5'-dA· radical and the other lies immediately adjacent to the first one [99]. In comparison to HemN, structures of class C enzymes adopt the TIM barrel fold and the unique C-terminals domains are similar to that of HemN, the N-terminal trip-wire domain is missing. The presence of two simultaneously SAM molecules in the structures of HemN suggests that this binding arrangement may be operative in the radical SAM-dependent methyltransferases of class C. The first SAM bound to the [4Fe-4S] cluster is the precursor to the 5'-dA· radical, whereas the second SAM is the source of the methyl group [88].

Class D of radical SAM-dependent methyltransferases

A novel class of radical-SAM methyltransferases [88] comprises the bacterial class D enzyme from *Methanocaldococcus jannaschii* [101]. This enzyme is responsible for methylation at C7 and C9 of the pterin ring in the methanopterin (MPT) and its analogs which are cofactors utilized in C1 metabolism during methanogenesis and methylotrophy in specialized microorganisms (archaea and bacteria). Class D radical-SAM methyltransferase features the presence of the two conserved radical-SAM motifs (two of CxxxCxxC), instead of one motif as found in the other classes [101]. Moreover, these two motifs possibly participate to distinctively respond to methylation at the different carbon atoms of the pterin. Interestingly, in C7 methylation, the adopted methyl group is not derived from methionine, indicating that enzyme activity is not stimulated by cobalamin. In this case, the methyl donor is derived from methylenetetrahydrofolate instead. In the proposed mechanism for the C7 methylation, the 5'-dA· radical from reductive cleavage of the first SAM abstracts an H· from C7 of pterin ring of MPT, which is followed by a

radical addition of the substrate onto methylenetetrahydrofolate. After injection of an electron and a proton, elimination of tetrahydrofolate results a substrate containing a C7 exocyclic methylene group. Hydride transfer from tetrahydrofolate to the exocyclic methylene leads to in formation of the methylated product [88,101].

1.5.3 Folate-dependent methyltransferases

Folate and its derivatives and betaine, also act as methyl donors [56]. In the case of RNA methylation, methylation reaction at the C5-position of deoxyuridine (2'-deoxyuridine-5'-monophosphate; dUMP) to produce thymidylate (2'-deoxythymidine-5'-monophosphate; dTMP) is catalyzed through the reductive methylation by thymidylate synthase (TSase, EC 2.1.1.45), which utilizes methylenetetrahydrofolate (CH_2THF) as the methyl donor and a reductant [56,102]. Besides the production of dTMP, dihydrofolate (H_2folate) is also a product of TSase reaction. H_2folate is subsequently reduced to tetrahydrofolate (H_4folate) by dihydrofolate reductase. Then H_4folate is regenerated to CH_2THF by serine transhydroxymethylase [56].
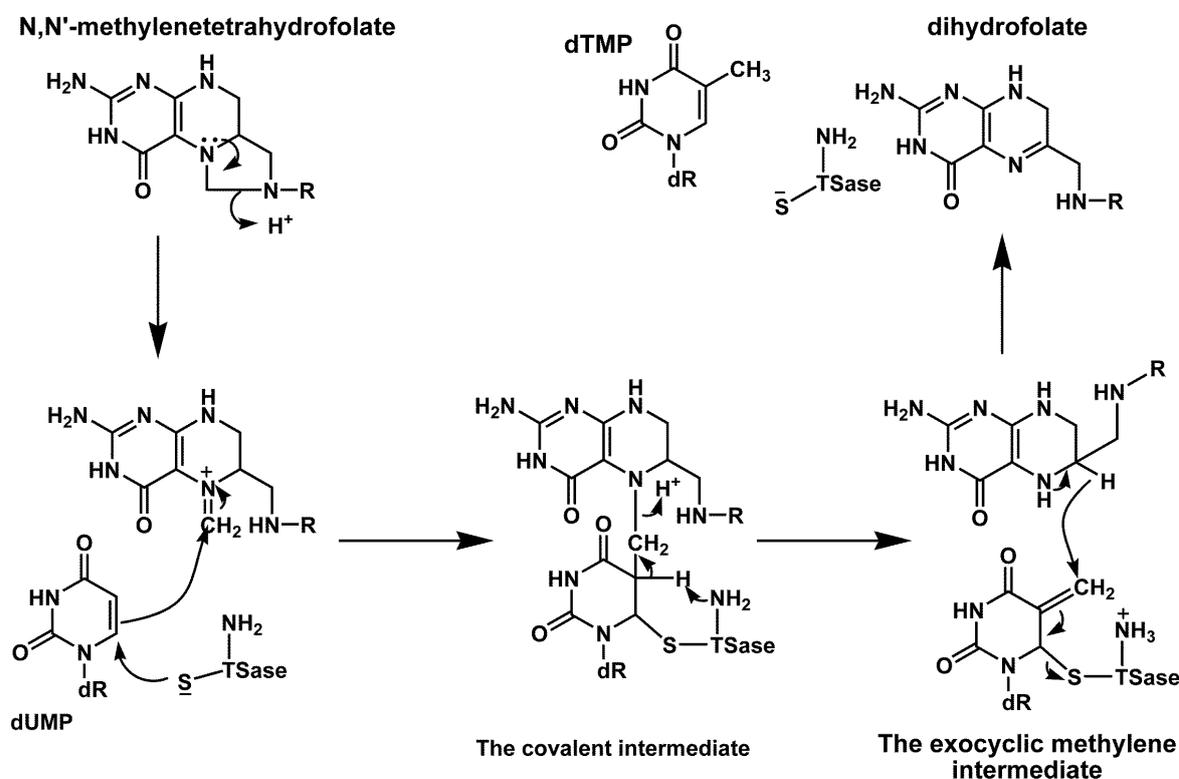

Fig. 13: Mechanism of thymidylate synthase

The catalytic mechanism of TSase is shown in Fig. 13 [56]. After dUMP and CH_2THF are bound to the enzyme, the 5-membered imidazolidine ring of CH_2THF is opened to form the iminium ion at N5 position induced by protonation of N10 with an acidic residue (Glu60). Then nucleophilic attack by a side chain of a basic residue of TSase (Cys198 in *Lactobacillus casei*, [103]) at C6 of dUMP results in the C5 nucleophilic enol(ate) formation and a covalent bond is formed between Cys198 and C6 of dUMP. The covalent intermediate is formed by nucleophilic attack of C5 dUMP to the N iminium ion of CH_2THF in which both of them are connected via a methylene bridge (Fig. 13). The proton of the C5 covalent intermediate is eliminated resulting in the formation of $H_2folate$ and the exocyclic methylene intermediate which is then reduced by hydride transfer to C6 of $H_2folate$, producing the final product, dTMP.

The three-dimensional structure of TSase revealed that native TSase is a symmetric dimer with the subunits binding to one dUMP and the quinazoline antifolate (CB3717) [104]. Each subunit contains a series of eight α helices and ten β strands connected with several coiled segments. The interface between two subunits is formed by a six-stranded twisted β-sheet. Each subunit has a deep active site cavity [104]. The ternary complex of TSase with folate and 5-fluoro-2'-deoxyuridylate (FdUMP) or with the anti-folate CB3717 and dUMP [104 -106] revealed that a covalent bond is formed between Cys198 and C6 of dUMP, phosphate moiety ligated by a quartet arginine and Ser219, and H-bonds formed among 3-hydroxyl group of 2-deoxyribose, Tyr261 and His259. In the ternary complex structure of TSase, dUMP and the anti-folate, a water molecule closed to the pyrimidine ring interacts to the carbonyl oxygen of Al96 and the side-chain hydroxyl of Tyr146 via H-bonding, facilitating proton transfer to and from the pyrimidine ring [104-106]. The ring opening of the 5-membered imidazolidine ring is activated by protonation at N10 of the ring by Glu60, resulting in the formation of the active 5-iminium ion. The pterin ring is packed in several nonpolar residues, including Trp82 and Trp85. There is an interaction (H-binding) through a bridging water molecule between Asp221 and N3 of the pterin ring. The backbone carbonyl oxygen of Ala315 forms H-bonds directly to the 2-NH_2 group and through a mediated water molecule to the 3-NH group. The terminal carboxylate interacts with Trp85 and Arg23 via H-bonding [104-106].

Some microorganisms *i.e. Mycobacterium tuberculosis and Helicobacter pylori*, are still able to survive in thymidine-deprived conditions even though genes encoding the enzymes involved in thymidylate cycle; thymidylate synthase, dihydrofolate reductase, thymidine

kinase, are missing [107]. These organisms still have the alternative enzymes that also catalyze tRNA methylation of dUMP and other RNAs in order to compensate the TSase function. Flavin-dependent thymidylate synthase (ThyX) catalyzes the same reaction as thymidylate synthase [108]. While FAD/folate-dependent RNA methyltransferases, TrmFO [109] and RlmFO [110], catalyze reductive methylation at the C5 position of uridine 54 and of tRNAs and C5 position of uridine 1939 in 23S rRNA, producing m^5U_{54} and m^5U_{1939}, respectively. There is no sequence or structural homologues of these proteins similar to the classic TSase [108]. Unlike TSase, using CH_2THF as a methylene-moiety donor and a reductant, the flavin-dependent thymidylate synthase and FAD/folate-dependent RNA methyltransferases utilize the flavin cofactor (the reduced form, $FADH_2$) for providing hydride transfer and CH_2THF as a source of the methylene moiety [111].

For the proposed mechanism of TrmFO reaction [109], a 5-membered ring opening and iminium ion formation at N5-position of CH_2THF, activated by nucleophilic attack of Cys53 at N10, is formed and then the N5-position iminium ion is attacked by the N5 isoalloxazine ring of reduced FAD to form a bridging methylene between the folate and the reduced FAD. Reorientation of the flavin and folate causes elimination of dihydrofolate and the N5-alkylated reduced flavin. The activated dUMP enolate, forming a covalent bond to Cys266 at C6 of dUMP, is attacked and the methylene bridge is formed between the N5-alkylated reduced flavin and C5-position of dUMP. Deprotonation at C5 of dUMP of the covalent intermediate induces dissociation of the reduced flavin and generation of the exocyclic methylene intermediate. Then a hydride from the reduced flavin is transferred to the exocyclic methylene intermediate, resulting in the formation of dTMP [109].

Recently, it was proposed that an arginine residue (Arg174, forming H-bonds with the C4 and C2 carbonyl of dUMP, ThyX from *Thermotoga maritima*) may act as a mediator for methylene transfer from CH_2THF to dUMP at C5-position [112]. However, the three dimensional structure (Fig. 14) of ThyX with FAD, dUMP, and CH_2THF [113] showed that Arg174 play an important role for binding and orientation of dUMP, but does not act as a methylene transfer mediator. His53 is a highly conserved residue of flavin-dependent thymidylate synthase for stabilization of substrate and the folate binding. Interestingly, data from the crystal structure of ThyX with FAD, dUMP, and CH_2THF and molecular simulation presented that methylene from the folate is transferred directly to dUMP at the C5-position and that the protein requires the extended stacking between the pyrimidine,

isoalloxazine, pterin, and imidazole rings of dUMP, FAD, CH_2THF, and His53, respectively (Fig. 14B), facilitating the CH_2THF iminium ion in direct contact with dUMP [113].
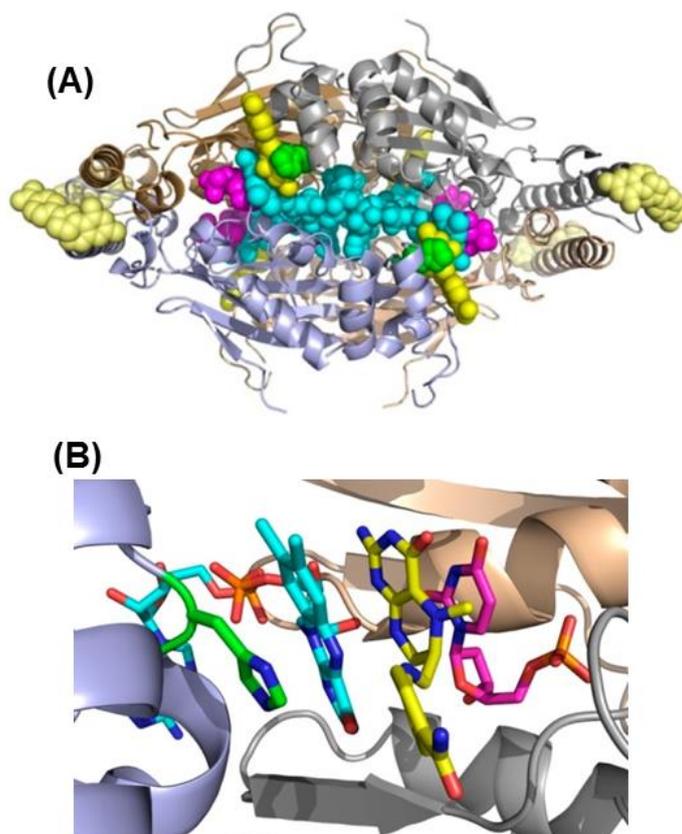
Fig. 14: (A) Schematic representation of the three-dimensional ThyX structure with FAD (cyan), CH_2THF (yellow), and dUMP (magenta) (pdb: 4GT9), including the emphasized His53 residue in green sphere. Pale yellow spheres represent four other CH_2THF, located apart from the active site. (B) A ThyX model at the active site showing the extended stacking of His53 (green), FAD (cyan), the iminium CH_2THF intermediate (yellow), and dUMP (magenta) [113].

1.6 Histamine and Histamine *N*-methyltransferase (HNMT)

1.6.1 Introduction to histamine

Histamine, [2-(1H-imidazol-4-yl)ethanamine], is a biogenic amine that plays important roles in several physiological and pathological processes including the regulation of inflammatory and immune responses, gastric acid secretion and bronchial asthma [114-117]. Histamine also functions as a neurotransmitter in the central nervous system (CNS) where it contributes to the regulation of many processes such as the sleep/wake cycle, the

circadian rhythm, thermoregulation, stress and fluid homeostasis [118-120]. Histamine was isolated initially from a variety of fresh tissues, thus implying that histamine is a common constituent of the body [121]. Hence, the name is derived from 'histos' the Greek word for 'tissue', indicating an amine occurring in several tissues [118]. Histamine is synthesized in mast cells, basophils, platelets, histaminergic neurons, and enterochromaffine cells by decarboxylation of histidine catalyzed by PLP-dependent histidine decarboxylase (HDC) [122]. It is also produced by actions of microorganisms in the course of food processing and spoilage. Thus, several fermented foods and beverages may contain substantial amounts of histamine [123]. Consumption of histamine-rich food *i.e.* matured cheese, pickled and canned food, sausages, and alcohol, may provoke diarrhea, headache, congestion of the nose, asthma, hypotension, arrhythmia, urticarial, pruritus, flushing and other conditions in patients with histamine intolerance [122]. Thus, histamine levels need to be balanced *i.e.* the synthesis, release, and elimination of histamine, to avoid undesirable conditions which may cause pathological effects.

In the mammalian central nervous system, besides dopamine, norepinephrine, epinephrine, and serotonin, histamine is a biogenic-amine neurotransmitter. Generally, the biogenic amines in the CNS are synthesized in a relatively small number of neurons, but these neurons can send long projections to the whole brain and spinal cord to modulate neuronal functions [124]. Histamine is produced in the brain by mast cells and neurons catalyzed by HDC. The tuberomammillary nucleus (TMN) located in the hypothalamus is the only site of synthesis of neuronal histamine in the adult mammalian brain [125,126]. About 64,000 histaminergic neurons are present in the human brain in and around the TMN area between the supraoptic nucleus and ventral tegmentum [125,127]. During mammalian fetal development, histamine production and storage are different from those processed in the adult brain. In rat embryonic development, histamine is present at maximal level between days E14-16, which then reduces steadily until birth (Fig. 15) [128,129]. The concentration of histamine detected at E20 is similar to the levels in the adult organism. The peak level of histamine coincides with the period in which neuronal differentiation takes place in brain regions. Moreover, expression of HDC is initiated at E14 and E15 in neurons located in the choroid plexus and mesencephalon, respectively [130]. Thus, during brain development histamine is a neurogenic and an important neuromodulator [129,131].
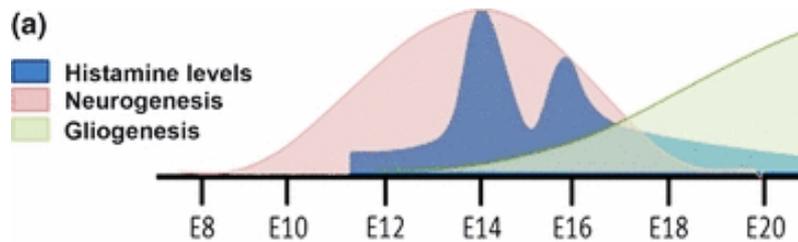
Fig. 15: Histamine levels during neural differentiation. Neurogenesis (pink) reaches the maximum at E14 during the fetal brain development, concomitant with the peak level of histamine (blue) in which gliogenesis (green) is initiated [128,129,131].

Release of histamine from cells is triggered by cross-linking surface-bound immunoglobulin E (IgE) to allergens [132]. The actions of released histamine are mediated through the four histamine receptors (H1R-H4R), which belong to the rhodopsin-like family of G protein-coupled receptors [133]. H1R, H2R and H3R are expressed in abundance in the brain whereas H4R is expressed mainly in peripheral tissues (e. g. the gut and connective tissue) [125,134]. H1R in human contains 487 amino acids encoded by the gene on chromosome 3p25. H1R is not restricted to the brain but it is found throughout the body. The second histamine receptor, H2R, contains 359 amino acids (40 kDa). The human gene encoding H2R is located on chromosome 5q35.5. In the rodent brain distribution of H2R is widespread and in some area colocalization with H1R [129]. H2R couples to $G_s\alpha$ to activate adenylyl cyclase, resulting in the increase of cAMP, which stimulates protein kinase A and the transcription factor cAMP-response element-binding protein (CREB) [125]. Both of them (protein kinase A and CREB are key modulators of neuronal physiology and plasticity [135]. The gene (*hrh3*) encoding human H3R (the third histamine receptor) containing 445 amino acids with 70 kDa is located on chromosome 20q13.333. This gene comprises two or three introns, thus yielding several splice-derived receptor isoforms with different distribution and pharmacology [136]. Distribution of H3R is heterogeneous in brain areas. High density is found in anterior parts of the cerebral cortex, hippocampus, amygdala, nucleus accumbens, striatum, olfactory tubercles, cerebellum, substantia nigra, and brainstem [137]. The forth histamine receptor, H4R, plays a role responsible for inflammation and allergy. It is expressed mainly in peripheral tissues, and also in hematopoietic cells [125,138,139].

1.6.2 Histamine storage and release

In aqueous solution, histamine is soluble, as it is mainly protonated to a single charge cation under physiological pH conditions [140]. Histamine contains two basic centers, which are the aliphatic amino group and one of the two nitrogens of the imidazole [140,141]. Because of the positive charge of histamine, permeability through membranes is limited by diffusion. Thus, histamine carriers are required to transport histamine between cells. To date, there are three different carrier proteins, classified by their properties and transport mechanism [140]: the organic cation transporter (OCT), the plasma membrane monoamine transporter (PMAT) and the vesicular monoamine transporter (VMAT). Moreover, histamine transport is achieved by endo/exocytosis of histamine-containing cellular vesicles or granules. The distribution of these transporters depends on the tissue and organ.

Histamine is synthesized by decarboxylation of L-histidine catalyzed by L-histidine decarboxylase [140]. It is normally present in mast cells, basophils, enterochromaffin-like cells in gastric mucosa, and histaminergic neurons (TMN) in which histamine is stored in secretory granules [118]. Upon appropriate stimulation, cells rapidly release large amounts of histamine leading to simultaneous fusion of several storage granules with the cell surface [142]. Degranulation of histamine storage granules is stimulated by the activation of phospholipase C and protein kinase C mediated by increased mobilization of Ca^{2+} [143]. In the process of exocytotic histamine release, SNARE proteins assist to break up membranes of storage granules and plasma membranes and fuse both membranes together [144]. Amounts of histamine released from storage granules depend on stimuli and is proportional to the number of storage granules and histamine density in the granules [145].

1.6.3 Metabolism of histamine

1.6.3.1 Histamine synthesis

Histidine decarboxylase (HDC) has pyridoxal 5'-phosphate (PLP), a vitamin B6 derivative, as a cofactor. Typically, the cofactor PLP acts as a universal ligand to react at the α-NH_2 group of α-amino acids in several reactions catalyzed by racemases, amino transferases, α- and β-decarboxylases [147]. A single gene encoding HDC is located on chromosome 15q21-22. The encoding region comprises twelve exons [148]. However, a nonsense mutation of W317X, causing a flexible loop region close to the active site of human HDC

is associated with Tourette syndrome (a neuropsychiatric disorder involving in repetitive movements or uncontrollable vocalization in the patients) [149]. Expression of HDC is upregulated by gastrin [150]. Conversely, HDC expression is repressed by Kruppel-like factor 4 by interfering with the downstream gastrin responsive elements [151].
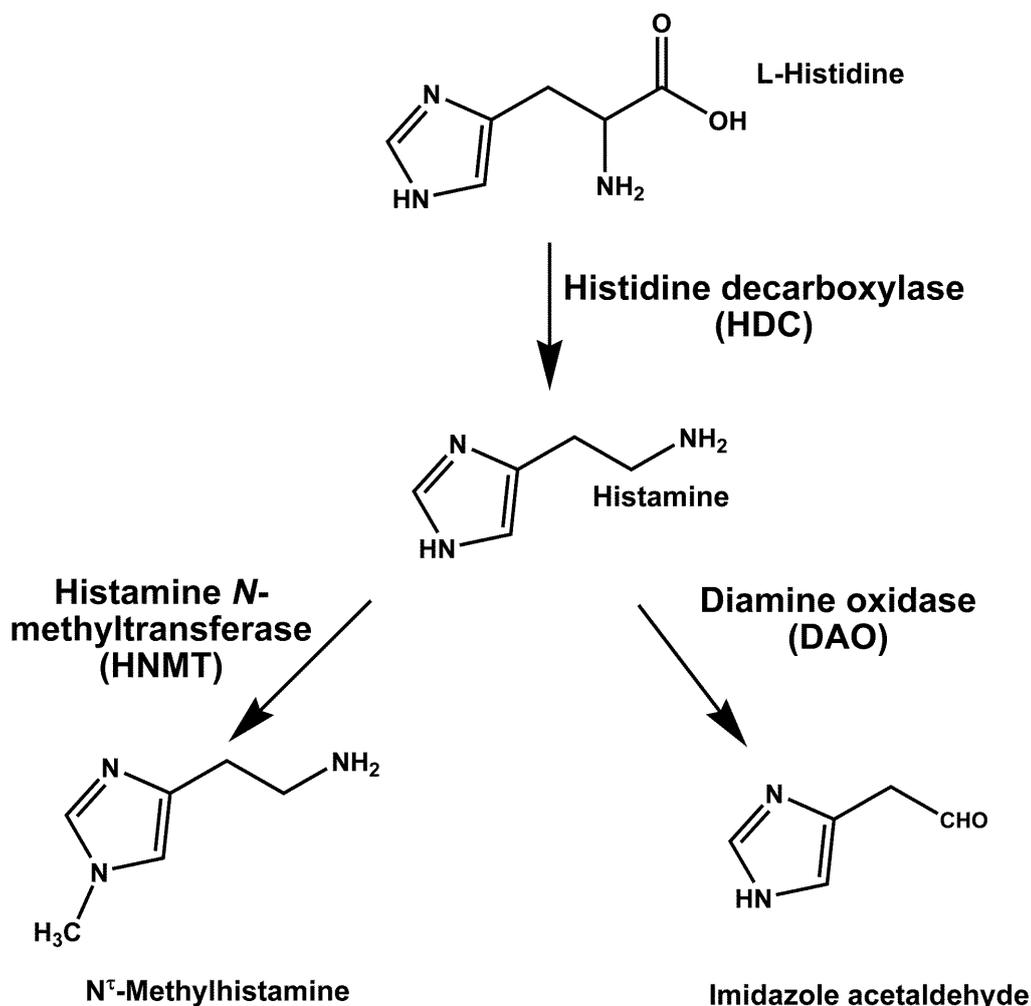
Fig. 16: Schematic representation of histamine metabolism [140]

HDC is synthesized as a 74 kDa precursor, containing 656 and 622 amino acid residues in the rat and human HDC, respectively. HDC requires to be processed by proteolytic removal of the carboxy-terminal fragment comprising 150 amino acids generating a fully active enzyme that forms a homodimer with approximate molecular mass of 110 kDa. Since the enzyme is not stable with a half-life of only a few hours in *vivo* [152], it is synthesized only when cells need to produce histamine. HDC is then degraded immediately when sufficient histamine has been generated.
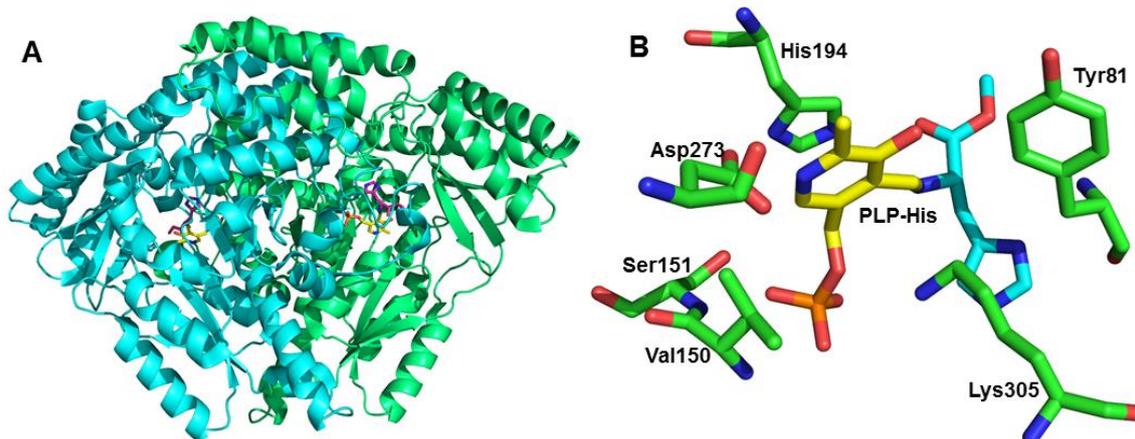
Fig. 17: The three-dimensional structure of human HDC (pdb: 4E1O). (A) Cartoon representation of the human HDC dimer with one protomer shown in cyan and the other in limegreen. PLP and histidine are shown in yellow and magenta sticks, respectively. (B) Interactions of the external Schiff base between PLP and histamine with surrounding residues in the active site. The PLP-histidine adduct is represented in yellow for PLP and cyan for histidine. The surrounding residues are represented in green [153].

The three dimensional structure of human HDC revealed that each subunit contains three domains, an N-terminal domain (residues 2-71), a large domain (residues 71-371), and a small domain (residues 372-477) (Fig. 17) [153]. The N- terminal and large domains participate in the dimer interface, producing a stable dimeric structure. The N-terminal domain engages in hydrophobic interactions with the other subunit, while the large domain forms principally water-mediated electrostatic interactions with the other subunit [153]. In addition, the N-terminal domain contains a PEST sequence, which is a degradation promoter motif of Pro, Glu, Ser, and Thr amino acids, resulting in the protein degradation by proteolytic systems [154]. A flexible loop between residues 330-340, including a short 3_{10}-helix is located proximal to the active site and plays an important role in PLP binding and the formation of the active site entrance. Moreover, this loop is involved in the external aldimine intermediate stabilization and protection of the protein from proteolytic cleavage [155]. In the human HDC active site, Ser354 plays a key role for substrate specificity. The single amino acid exchange at this position from Ser to Gly enlarges the substrate-binding pocket and leads to lower affinity for histidine [153].
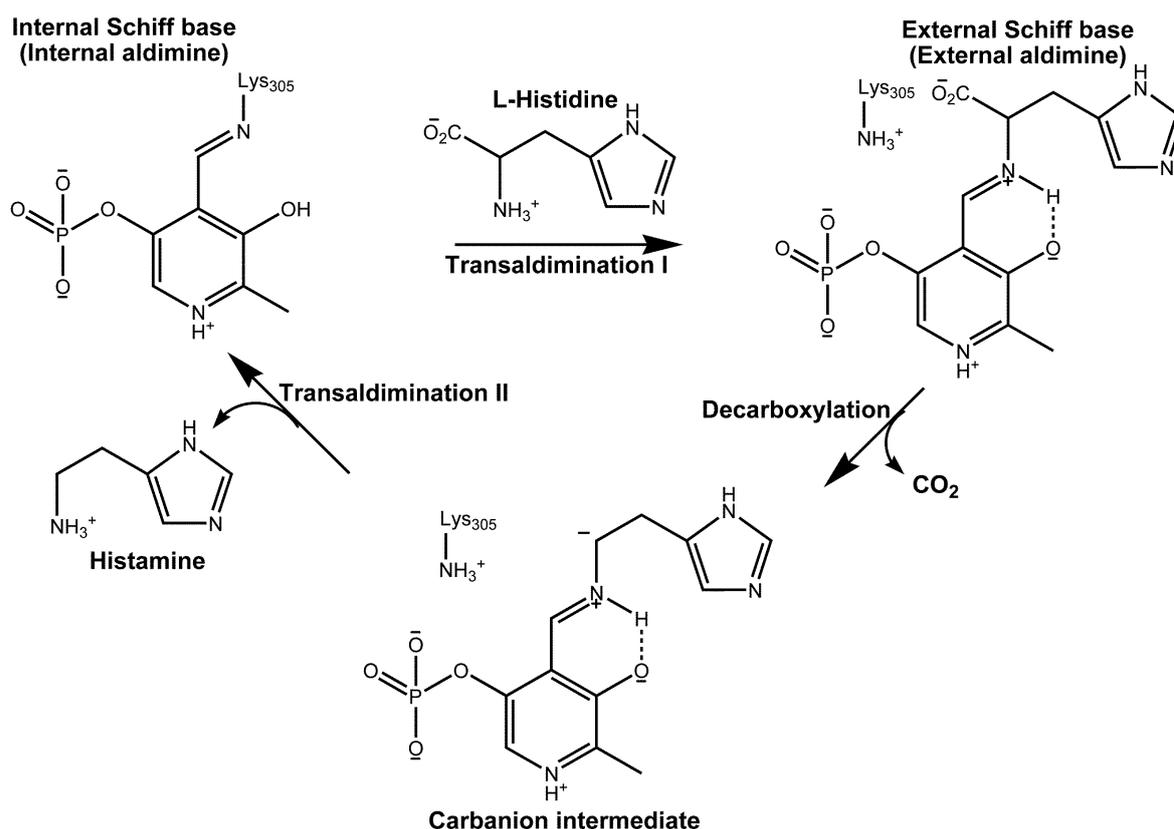
Fig. 18: The reaction mechanism of HDC (human HDC) [140,156].

The reaction mechanism of HDC involves Schiff base formation with the ε-amino group of a specific lysine (Lys308 in rat HDC and Lys305 in human HDC) in the active site (Fig. 17 and 18) [140,156]. In the free form of human HDC, the carbonyl group of PLP forms an internal Schiff base with the ε-amino group of Lys305. When L-histidine binds to the active site, it acts as a nucleophile that attacks the Schiff base resulting in the displacement of Lys305 to form the external Schiff base [140]. Deprotonation at Cα of the intermediate results in the formation of a carbanionic intermediate, which is stabilized by the extended π system of the PLP pyridine ring [147]. Then, this Cα of the carbanionic intermediate is broken by the attack of Lys305, resulting in the release of histamine and re-transaldimination between Lys305 and PLP.

1.6.3.2 Histamine inactivation

In mammals, the action of histamine is terminated by either one of two mechanisms that involve the methylation of the imidazole ring or the oxidative deamination of the primary amino group. The former reaction is catalyzed by histamine *N*-methyltransferase (HNMT) and the latter reaction by diamine oxidase (DAO) [157]. Products generated from these two

enzymatic reactions are virtually inactive at histamine receptors and are further converted for transport and secretion

Diamine oxidase (DAO)

Diamine oxidase (DAO) (EC 1.4.3.22), an enzyme of the class of copper-containing amine oxidases, is the main histamine-degrading enzyme in peripheral tissue including termination of exogenous histamine from ingestion (food and alcohol) [125]. Molecular oxygen (O_2) is utilized by DAO to oxidatively deaminate histamine, yielding imidazole acetaldehyde, ammonia, and hydrogen peroxide (Fig. 16) [158]. DAO is encoded by a single gene, ABP1 or AOC1, which contains five exons and is located on chromosome 7q36.1 [159]. Expression of DAO is restricted to specific tissues and cells. DAO is detectable at high levels in intestine, placenta, small bowel, colon ascendens, kidney, and body fluids (lymph pregnancy plasma, seminal plasma) [157,160,161]. The highest level of histamine is present in intestine. It may imply that the most important function of DAO is to prevent resorption of histamine and other diamines from ingestion [140]. A decreased DAO activity and an increased histamine level in the intestinal mucosa (localized in the cytoplasm of the mature enterocytes of the small and large bowel) reflect that DAO is an important factor involved in inflammatory bowel diseases [162]. DAO is produced and stored in vesicle structures associated with the basolateral plasma membrane in proximal tubular epithelial cells, where it is secreted by specific signal stimulation. One example of the stimulation is secretion of heparin which is released together with histamine by activated mast cells [140]. In all DAO proteins from different species, amino acid residues playing important catalytic and structural functions are conserved. Proteolytic cleavage of a 19 amino acid N-terminal signal peptide from a 751 amino acid DAO precursor, protein dimerization, and attachment of asparagine-linked oligosaccharide chains are required to generate a mature human DAO [163]. Human DAO is a homodimer, consisting of two 85 kDa subunit (Fig. 19). Moreover, its structure was found to be similar to that of previously described copper-containing amine oxidase (AOC) that has in the active site a Cu^{2+} ion bound by three conserved histidine residues (His510, His512, His675) and the cofactor 2,4,5-trihydroxyphenylalanine quinone (TPQ), generated post-translationally from a conserved tyrosine residue [158,163]. Cys736 of both subunits form an intermolecular disulfide to link and stabilize the dimeric protein. Typically, DAO can catalyze the oxidative deamination of several diamine substrates with primary amino groups *i.e.*

aliphatic amines (1,4-diaminobutane (putrescine), 1,5-diaminopentane (cadaverine)) and polyamines (agmatine), but not monoamines [158]. The conserved Asp373 residue is the catalytic base, which interacts with the two amino groups of substrates and is responsible for substrate specificity [163].
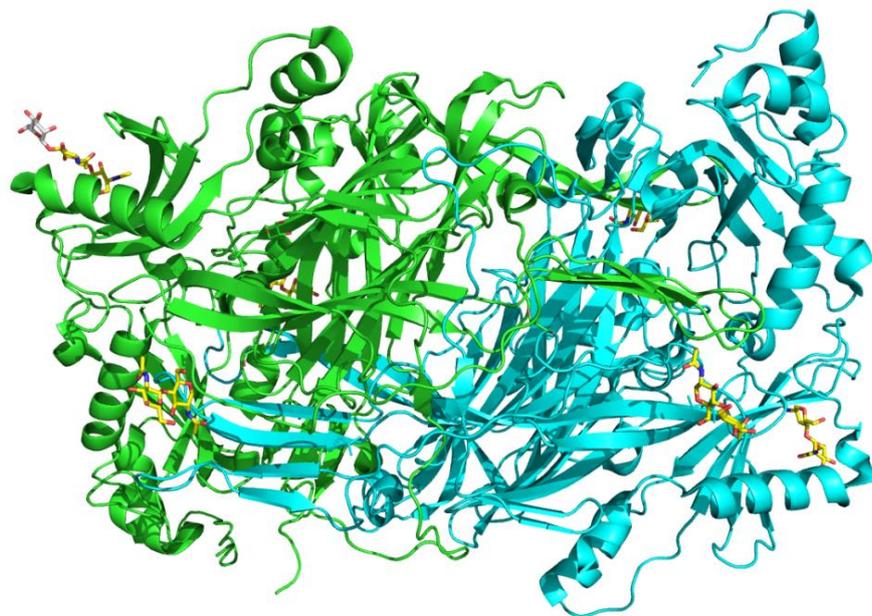
Fig. 19: Schematic representation of homodimeric human DAO (pdb: 3HI7) consisting of $2Cu^{2+}$ ions, $4Ca^{2+}$ ions, and 12 molecules of *N*-acetyl-D-glucosamine (yellow sticks) [163].

After the oxidative deamination of histamine catalyzed by DAO, the product imidazole acetaldehyde is subsequently oxidized to imidazole acetic acid by aldehyde dehydrogenase (EC 1.2.1.3). Conjugation of imidazole acetic acid with phosphoribosyl pyrophosphate to form imidazole acetic acid ribotide is catalyzed by imidazole acetic acid phosphoribosyltransferase (EC 6.3.4.8). The product phosphoribosyl imidazole acetic acid is then eliminated by the kidney [140].

Histamine N-methyltransferase (HNMT)

HNMT (EC 2.1.1.8) inactivates histamine by transferring a methyl group from *S*-adenosylmethionine (SAM), to the $N_{\varepsilon 2}$ atom of the imidazole ring, generating *N*-methylhistamine and *S*-adenosylhomocysteine (SAH) [164,165]. HNMT is the primary metabolizing enzyme for released histamine in bronchial epithelial cells, endothelial cells of human airways and the human stomach [166,167]. In addition, HNMT is also the only

enzyme responsible for the termination of histamine action in the mammalian brain, since the other histamine-metabolizing enzyme, DAO, is not expressed in the CNS [118,168]. Enzyme activity and mRNA expression were found in most human tissues with high expression levels in kidney and liver, with substantial expression in spleen, prostate, ovary, intestine and the spinal cord, along with lower-level expression in heart, brain, placenta, lung, stomach and the thyroid gland [140,169]. The *hnmt* gene is approximately 34 kb in length, located on chromosome 2q22.1, and consists of 6 exons [170-172]. Expression and characterization of human HNMT showed that cytoplasmic HNMT has 292 amino acid residues with a total molecular mass of 33 kDa [140]. Several three-dimensional structures of HNMT with bound SAH and histamine as well as various HNMT inhibitors revealed that HNMT is a monomeric protein and consists of two domains (Fig. 20) [173,174]. The large domain adopts a canonical methyltransferase (MTase) fold with a SAM binding pocket which contains a seven-stranded β sheet (6↑ 7↓ 5↑ 4↑ 1↑ 2↑ 3↑) flanked on each side by three helices (αZ, αA, αB and αC, αD, αE). Whereas the small domain contains three short β strands (↓β8 ↑β7' ↓β6'), two α helices and two single-turn 3_{10} helices, contributing part of the histamine binding pocket and may be involved in protein-protein interactions (Fig. 20) [173]. The bound histamine is buried in a hydrophobic pocket (surrounded by 14 aromatic amino acid residues) at the interface of the two domains. In the SAM-binding pocket, the ribose hydroxyl groups interact with the carboxylic atom of Glu89 and the side-chain $N_{\varepsilon 2}$ atom of Gln94. The former interaction is the most highly conserved SAM interaction among SAM-dependent methyltransferases [9]. The terminal carboxyl oxygen atom forms hydrogen bonds to His29 and water-mediated hydrogen bonds to Glu28 and Glu65. The backbone oxygen atom of Gly60 and Ile 142 interact with the amino group through H-bonding. In addition, the amino group forms two water-mediated interactions to the side-chain Asp67 and His140 [173]. Therefore, in vertebrates, amino acid residues that interact with SAH and histamine are highly conserved [173,174]. Methylation of histamine is catalyzed via nucleophilic attack from the $N_{\varepsilon 2}$ atom of histamine to the methyl group adjacent to SAM sulfonium ion (S_N2 reaction mechanism) of SAM. The H-bond network between Gln143, a water molecule, and Glu28 assists proton abstraction at $N_{\varepsilon 2}$ to subsequently enable attack at the methyl group of SAM, yielding N^{τ}-methylhistamine and SAH (Fig. 20) [173]. Kinetic studies of human HNMT revealed that the methylation reaction is processed via the ordered sequential bi-bi reaction mechanism with SAM being the first substrate to bind to HNMT and N^{τ}-methylhistamine being the first product to dissociate [173,175]. The product *N*-methylhistamine is

oxidatively deaminated by monoamine oxidases (MAO, EC 1.4.3.4) similar to other biogenic amines. In the case of *N*-methylhistamine, MAO catalyzes oxidative deamination to form *N*-methylimidazole acetaldehyde, which then is oxidized by aldehyde dehydrogenase, yielding *N*-methylimidazoleacetic acid. The product from the reaction of aldehyde dehydrogenase is then eliminated by secretion [176].
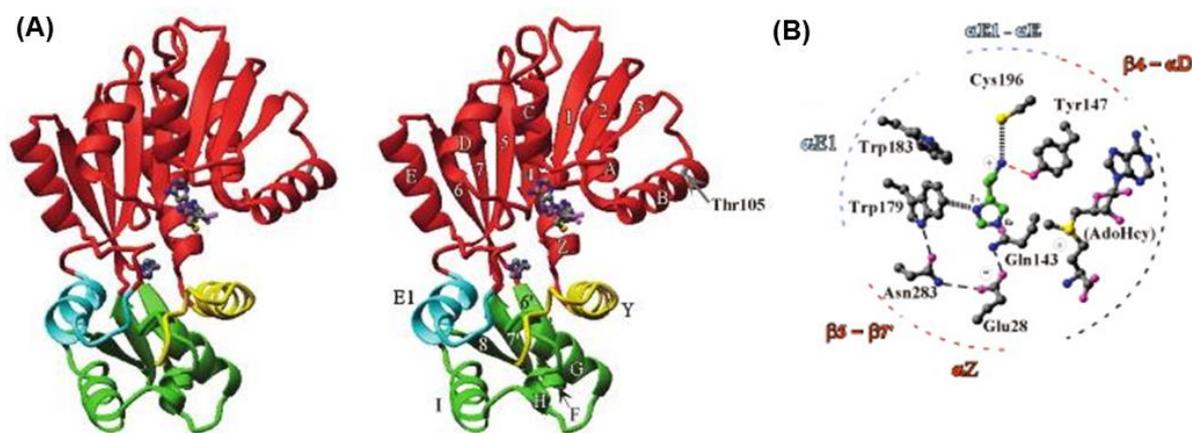
Fig. 20: Schematic representation of monomeric human HNMT (pdb: 1JQD) in the presence of SAH and histamine (the purple sticks) (Adopted from Fig. 2a and 3c of [173]).

Since disturbances of histamine metabolism have been related to several diseases, *i.e.* asthma, bronchial hyper-responsiveness (BHR), neurological disorders, the lack of expression or decrease in HNMT activity have been proposed as a causative factor for these diseases. A common single nucleotide polymorphism (SNP), C314T transition (prevalence 0.1), resulting in the replacement of Thr105 by isoleucine was associated with decreased enzyme activity, immunoreactive protein, and thermal stability [169]. Steady-state kinetics revealed that the apparent Michaelis constants (K_M) increased 1.8 and 1.3 fold for SAM and histamine, respectively, and showed a reduction of the specific activity of about 16% [173]. Moreover, from multiple molecular dynamic (MD) simulations, isoleucine 105 apparently more strongly interacts with neighboring residues leading to a disordering of several key residues responsible for SAM binding and lower hydrophobicity of the substrate-binding site [177]. Thus, the polymorphic T105I variant has been proposed to associate with several diseases including asthma, allergic rhinitis, essential tremor and Parkinson's disease [178-181].

Intellectual disability (ID) is a common neurodevelopmental disorder characterized by an intelligence quotient of 70 or below with deficits in adaptive, daily living skills and intellectual functioning manifesting before the age of 18 [182]. With a prevalence of about 1% of children worldwide ID is one of the major socioeconomic problems [183]. Recently, two other homozygous missense mutations in the *hnmt* gene were identified in patients affected with nonsyndromic autosomal recessive intellectual disability (ND-ARID) from two unrelated consanguineous families of Turkish and Kurdish ancestry [184]. The first mutation (*hnmt* c.179G>A) from the Turkish family resulting in a change of Gly60 to aspartate (G60D) occurs in the conserved MTase region I, which is part of the SAM-binding pocket (the conserved G-loop) (Fig. 21). This amino acid exchange completely disrupts binding of SAM and as a consequence compromises catalytic activity, resulting in reduced histamine inactivation. The other mutation (*hnmt* c.632T>C) found in the Kurdish family results in the replacement of Leu208 with proline (L208P). This invariant residue is located in helix E of the MTase domain and forms several hydrophobic interactions with neighboring residues (Fig. 21). Interestingly, L208 is located in a considerable distance from the active site (ca. 18 Å from the substrate binding pocket) (Fig. 21) suggesting that the adverse effect of the amino acid replacement is caused by perturbations of the protein structure leading to compromised enzyme function. Moreover, a substantial loss of enzymatic activity and binding affinity for histamine from the variations of the L208 position, *i.e.* the L208R, and L208K, revealed that this position is extremely sensitive to amino acid replacements resulting in a deterioration of stability and enzyme function (Unpublished data, see Chapter 2.2) [185]. Molecular dynamics simulations showed that amino acid replacements in position 208 perturb the helical character and disrupt interactions with the adjacent β-strand, which is involved in the binding and correct positioning of histamine.
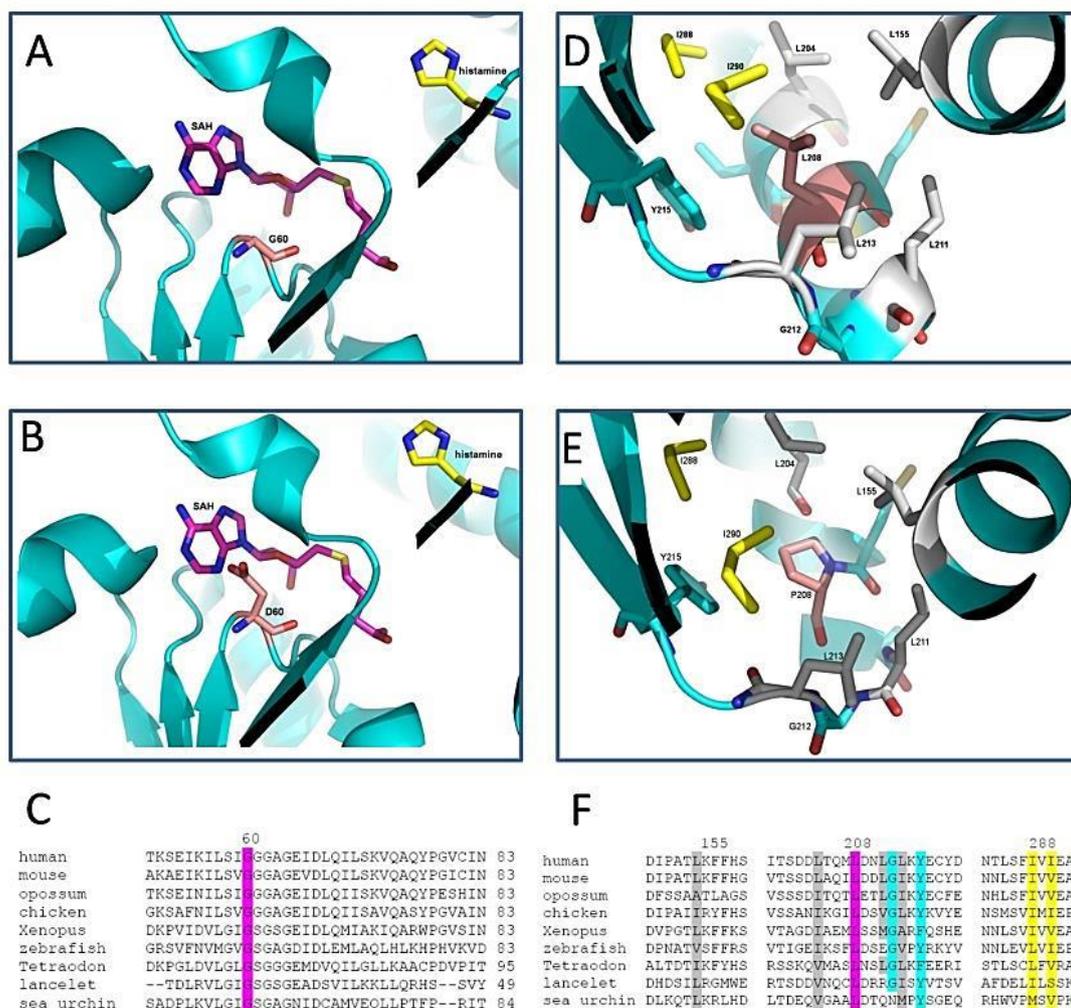
Fig. 21: Three-dimensional HNMT structure (pdb: 1JQD) and ClustalW2 analysis: Structures of HNMT at the catalytic domain for Gly60 (in pink) (A), Gly60Asp (in pink) (B) with histamine (yellow) and SAH (magenta). Structures at the hydrophobic pocket around residue 208 were presented: Leu208 (in pink) (D) and Leu208Pro (in pink). (E) Leu 155, Leu204, Leu211, and Leu213 were labeled in grey. Ile288 and Ile290 were labeled in yellow. Tyr215 and Gly212 were labeled blue. (C) and (F) are the amino acid sequence alignments around the residues Gly60 and Leu208, respectively, using ClustalW2 [184].

1.7 A putative SAM-dependent methyltransferase (METTL23)

METTL23 (methyltransferase-like 23) gene is an important gene involved in human cognition [186,187]. The gene located in chromosome 17q25.1 encodes two METL23 isoforms, consisting of 190 (isoform 1) and 123 (isoform 2) amino acids [186,187]. The former isoform is mainly localized in the endoplasmic reticulum membrane, whereas the latter isoform is localized in comparable concentrations in nucleus and in cytoplasm. The protein localization of METLL23 implies that the isoform 1 has a transmembrane domain [187]. Moreover, expression of transfected METLL23 in HEK239T, HeLa, and N2A cells

occured in both the cytoplasm and nucleus, highly present in the nucleus [186]. In the developing human brain, METTL23 is expressed at low-to-moderate level [186]. Based on METTL23 amino acid sequence, it is predicted to be a putative methyltransferase (Fig. 22a) [187]. The recombinant METTL23 was expressed in *E. coli.* and yielded high amounts of cytoplasmic inclusion bodies. Interestingly, METTL23 was tightly bound to chaperones *e.g.* GroEL, resulting in the soluble METTL23:GroEL complex. These findings are consistent with the presence of transmembrane domain of METTL23 isoform 1. Moreover, chaperones may act as a putative substrate for methylation by METTL23 [187]. Based on sequence similarity, the other METTL family members *i.e.* METLL22, METTL21A-D were strongly associated with chaperones. They catalyze trimethylation of lysine residues of chaperones (KIN/Kin17, HSPA8/Hsc70, Hsp70, and VCP/p97) [188], consistent with the tight-binding of METTL23:GroEL found in Bernkopf and coworkers' results [187].

Nonsense mutation and 4- or 5- bp frameshifting deletion lead to truncation of the putative METTL23. Moreover, these truncating mutations are also associated to autosomal recessive intellectual disability (ARID) [186,187]. Reiff and colleagues revealed that METLL23 acts as a transcription regulator and physically interacts with the transcriptional factor GABPA (GA-binding protein transcription factor, α-unit), which functionally affects the expression of GABP target genes, *THPO* (thrombopoietin) and *ATP5B* (ATP synthase, H^+ transporting, mitochondria F1 complex, β-polypeptide) [189,190]. Both genes are involved in several processes in brain *i.e.* neuroprotection, apoptosis, development, and neural cell differentiation [191,192]. In ARID patients, disruption of *METTL23* gene caused from the 4-bp frameshift deletion has functional effects on the transcriptional factor GABPA. Consequently, there is alteration of expression of the GABPA target genes, *THPO* and *ATP5B*, which affects on several processes in brain [186]. In addition, the nonsense mutation and 5- bp frameshifting deletion of *METTL23* was found in members with autosomal recessive intellectual disability of consanguineous Pakistani and Austrian families. These mutations result two isoforms of the truncating METTL23 (Gln133* and Gln94Hisfs*) (Fig. 22b) [187]. These two variants disrupt the predicted catalytic domain and alter the cellular localization. Overexpression of the two variants resulted in the formation of protein aggregates of isoforms 1 and 2 in the cytoplasm [187]. Therefore, expression of aggregate METTL23 variants leads to the loss of protein function, resulting autosomal recessive intellectual disability [187].
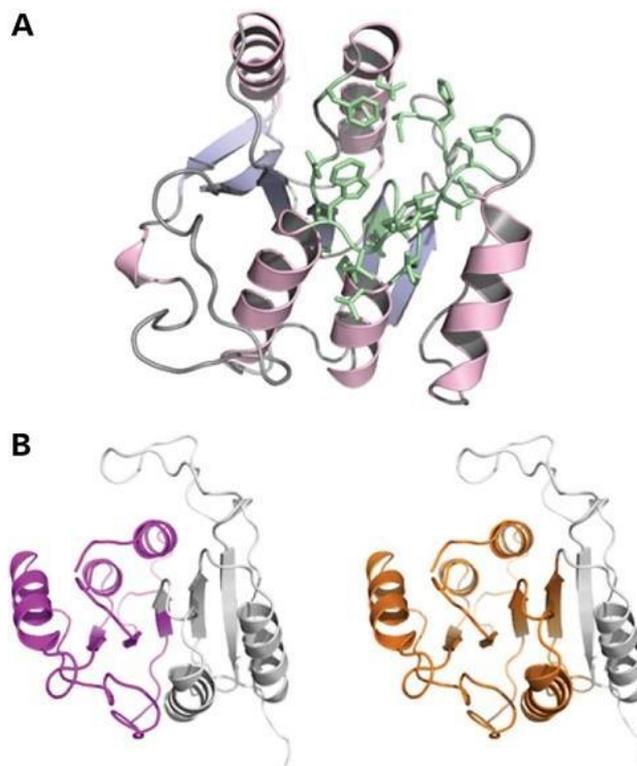
Fig. 22: The modelled three-dimensional structures of METTL23: (A) wild-type METTL23 (isoform 1, 190 amino acids), (B) two truncating METTL23 variants (left; residue 1-132 Gln133* and right; residue 1-98 Gln94Hisfs*) [187].

1.8 Roseoflavin and *N*,*N*-8-amino-8-demethyl-D-riboflavin dimethyltransferase (RosA)

1.8.1 Introduction to roseoflavin (RoF)

Riboflavin, vitamin B$_2$, is the precursor of flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN, riboflavin-5′-monophosphate) in many catalytic reactions. Riboflavin is converted to FAD and FMN by flavokinases (EC 2.7.1.26) and FAD synthetases (EC 2.7.7.2), respectively [193]. Flavoproteins are able to catalyze a wide variety of different biochemical processes [194]. To date, some flavin analogs have the potential to serve as basic structures for the development of novel antimicrobials [195,196]. Currently, the natural vitamin analogs: bacimethrin (a vitamin B$_1$ analog) [197], ginkgotoxin (a vitamin B$_6$ analog) [198], and roseoflavin (a vitamin B$_2$ analog) [199] are examples of antimicrobials [200]. Since about 0.1-3.5% of predicted proteins in organisms are flavin-dependent proteins [201], these proteins are considered to be cellular targets of the antimicrobial flavin analogs [195]. Biosynthesis of antimicrobial flavin analogs in microorganisms is a versatile process that can be utilized for many advantages. First of all,

the strategy for the biosynthesis of an antimicrobial, converting vitamins (*e.g.* vitamin B_1, B_2 and B_6) to antibiotic vitamin analogs, seems to be very economical because the precursors for antibiotic synthesis are readily available in the cytoplasm of a producer cell. Furthermore, several microorganisms contain efficient vitamin transporters (*i.e.* flavin transporter; Rib protein), which process the vitamin uptake, but also vitamin analogs. Consequently, delivery of the antibiotic vitamin analogs to cellular targets is ensured. Moreover, the antimicrobial vitamin analogs have multiple cellular targets, because several vitamins (precursors of enzyme cofactors) are active at more than one site in cells [202].

8-dimethylamino-8-demethyl-D-riboflavin, roseoflavin or RoF, is the only known natural riboflavin analog with antimicrobial activity [199]. RoF is synthesized in only two Gram-positive bacteria, *Streptomyces davawensis* and *Streptomyces cinnabarinus* [199,203,204]. Similar to riboflavin, roseoflavin (RoF) is converted to roseoflavin 5'-monophosphate (RoFMN) and roseoflavin adenine dinucleotide (RoFAD) by flavokinases and FAD synthetases [205]. The isoalloxazine ring of RoF loses its oxidizing ability because of an intramolecular charge transfer from the 8-dimethylamino group to the pteridine moiety [206]. Although the molecular mechanism of action of RoF and the resistance mechanism of the RoF-producing strains are not clear yet, FAD/FMN-dependent proteins were found to be less active or completely inactive upon the replacement of the native flavin cofactor with RoFMN or RoFAD [205,207]. FMN-dependent azobenzol reductase (AzoR, EC 1.7.1.6) from *E. coli* was less active when its apo-form was reconstituted with RoFMN [208]. Interestingly, RoFMN bound to AzoR with higher affinity than the native FMN and redox potential of AzoR:RoFMN was reduced (-145 mV for AzoR:FMN vs -223 mV for AzoR:RoFMN) [208]. The difference in the redox potential was also observed in the less active FMN/RoFMN reconstituted L-lactate oxidase from *Aerococcus viridans* [209]. It should be noted that alteration of reactivity of the flavoproteins which bind to flavin analogs, including RoF, RoFAD, or RoFMN is related to the different physicochemical properties of the unnatural ligands [202]. Furthermore, FMN riboswitches, genetic elements that regulate riboflavin biosynthesis and transport, have been reported to be turned off in the presence of RoFMN, thereby causing riboflavin deficiency presenting additional targets for RoF [200].

The biosynthesis of RoF was postulated to begin from riboflavin via 8-amino- (AF) and 8-methylamino-8-demethyl-D-riboflavin (MAF) [210,211]. The occurrence of the intermediates AF and MAF was confirmed by the discovery of the first enzyme, SAM-

dependent dimethyltransferase RosA (*N,N*-8-amino-8-demethyl-D-riboflavin dimethyltransferase) [212]. This enzyme synthesizes RoF from AF via MAF in two consecutive methylation reactions [212]. The *rosA* gene encoding RosA is localized on a cluster consisting of nine genes (on gene cluster 7). Heterologous expression of this gene cluster 7 was performed in *S. albus* and *S. lividans*. However, both strains could not produce RoF. Thus, the remaining genes on this gene cluster 7 are not involved in RoF biosynthesis [212]. Detection of [uniformly-^{15}N] labeled RoF (all four N atoms labeled) from the culture supernatant of *S. davawensis* after supplementation of [uniformly-^{15}N] labeled riboflavin confirmed that riboflavin is a starting-point precursor and is utilized directly to produce RoF [210]. At that time, the pathway from riboflavin to AF and the corresponding enzymes was unknown.
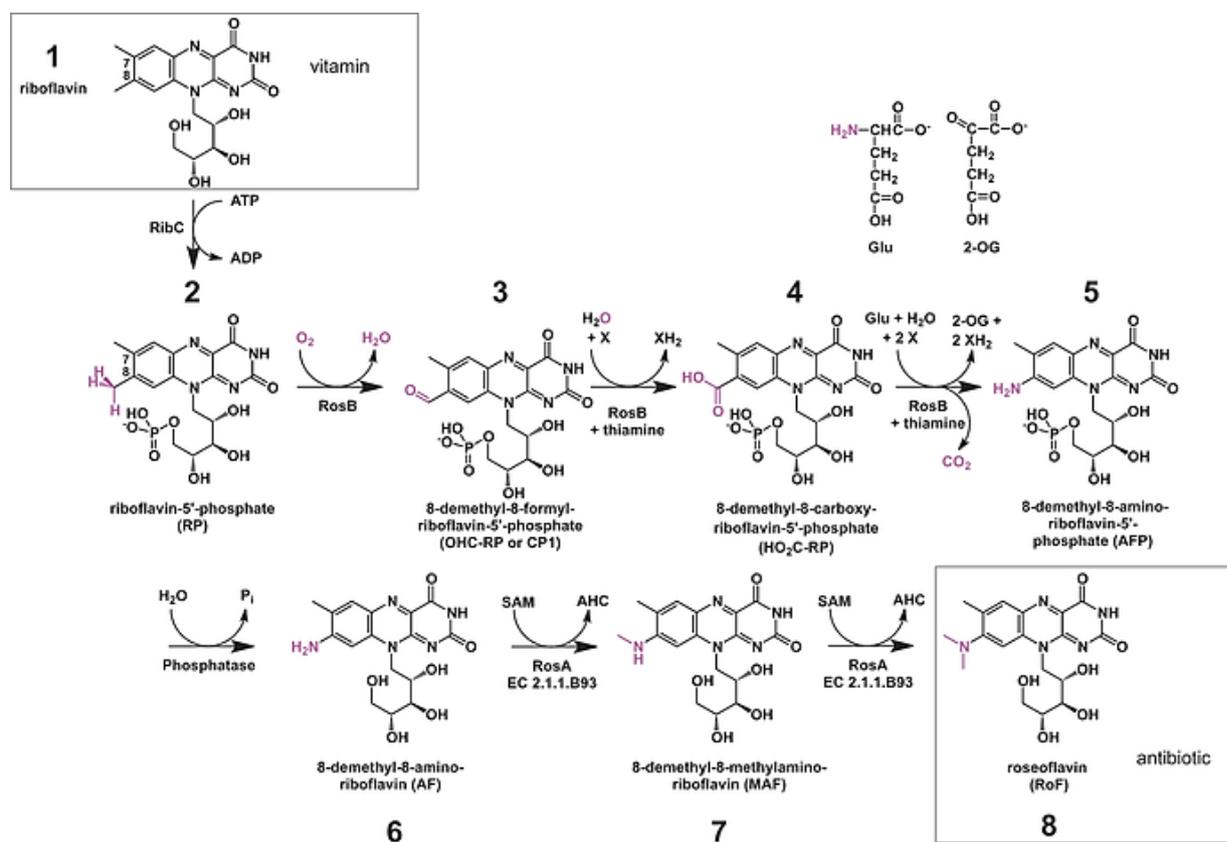
Fig. 23: Roseoflavin biosynthesis pathway in *Streptomyces davawensis*. The pathway is initiated by flavokinase (RibC), resulting in the formation of FMN (or RP). RosB, the second discovered enzyme of RoF biosynthesis, catalyzes three consecutive reactions as the following steps of oxidation, decarboxylation and amino transfer to yield 8-demethyl-8-amino-riboflavin-5'-phosphate (AFP). Removal of the AFP phosphate group is catalyzed by a phosphatase to produce AF The last step is the utilization of AF as the substrate of RosA, which catalyzes two consecutive methylation reactions, generating the antibiotic RoF [200].

Recently, Schwarz and coworkers discovered the second key enzyme in RoF biosynthesis from *S. davawensis* and *S. cinnabarunus*, termed RosB [200]. Moreover, they also showed that four enzymes are required to synthesize RoF as shown in Fig. 23. Starting with riboflavin as the first substrate of RoF biosynthesis, it is converted to FMN, by the flavokinase, RibC [213]. Then, RosB, which is the second enzyme of RoF synthesis, converts FMN to 8-demethyl-8-amino-riboflavin-5'-phosphate (AFP) via 8-demethyl-8-formyl-riboflavin-5'-phosphate (OHC-RP or CP1) and 8-demethyl-8-carboxy-riboflavin-5'-phosphate (HO_2C-RP) [200]. The full name of RosB is 8-demethyl-8-aminoriboflavin-5'-phosphate synthase. The gene *BN159_7989* encoding RosB is located in the gene cluster 2 of *S. davawensis*. Based on the amino acid sequence and compared to the other proteins, RosB exhibits a putative FMN binding site. It is suggested that FMN (not riboflavin) is the starting substrate for RoF synthesis [200]. To convert FMN to AFP by RosB (Fig. 23 the methyl group at C8 of isoalloxazine ring of FMN is oxidized to form OHC-RP (CP1). Addition of thiamine could drive the oxidation of OHC-RP to HO_2C-RP. The role of thiamine is still unclear in the oxidation step of OHC-RP to HO_2C-RP [200]. Synthesis of the final product of RosB, AFP, needs transfer of an amino group from a donor to HO_2C-RP. In this case, glutamate acts as a putative amino donor to transfer an amino group to HO_2C-RP with elimination of CO_2 to form the final product AFP. The third enzyme, an unknown phosphatase hydrolyses AFP to AF which subsequently is dimethylated by RosA to form the antibiotic RoF [200].

1.8.2 *N,N*-8-amino-8-demethyl-D-riboflavin dimethyltransferase (RosA): structure and reaction kinetics

N,N-8-demethyl-8-amino-D-riboflavin dimethyltransferase (RosA) catalyzes the final dimethylation of 8-demethyl-8-amino-D-riboflavin (AF) to the antibiotic roseoflavin (RoF) in *S. davawensis* and *S. cinnabarinus* [199,204]. RosA is the first enzyme discoverd and involved in RoF biosynthesis [212]. RosA consists of 347 amino acid residues with a subunit mass of 38 kDa [212]. Although RosA catalyzes a *N,N*-dimethylation (two consecutive methylations), its primary structure is similar to several SAM-dependent *N*-methyl, and O-methyltransferases. RosA shares low similarities to several characterized *N,N*-dimethyltransferase *e.g.* ErmSF, TylM1 and DesVI [212,214,215].

The three dimensional structure of RosA exhibits that the RosA protomer comprises two domains, a mainly α-helical "orthogonal bundle domain" and a Rossmann-like domain

(α/β twisted open-sheet) (Fig. 24). The orthogonal bundle has five α -helices and one antiparallel β-sheet (residues 1-98). An interdomain part contains five helices (residues 99-178), that connect the Rossmann motif with the orthogonal bundle. The Rossmann motif (residues 179 to 353) comprises a seven-stranded β-sheet core consisting of five parallel and two antiparallel β-strands, which are connected by two pairs of α-helices [216]. The latter is similar to MTase fold of SAM-dependent methyltransferases, which is the domain providing the SAM binding pocket and catalytic sites [9,66]. RosA is homodimeric protein. Dimerization of the protein is processed via a head-to-head arrangement (Figure 1B Fig. 24). Attempts to obtain structural information for complexes with AF, SAM, and SAH were unsuccessful. Therefore, using structural bioinformatics methods to locate putative binding sites for substrates revealed that AF was located at the interface between the Rossmann domain and the N-terminal domain, whereas SAM was located in the C-terminal Rossmann domain (Fig. 24).
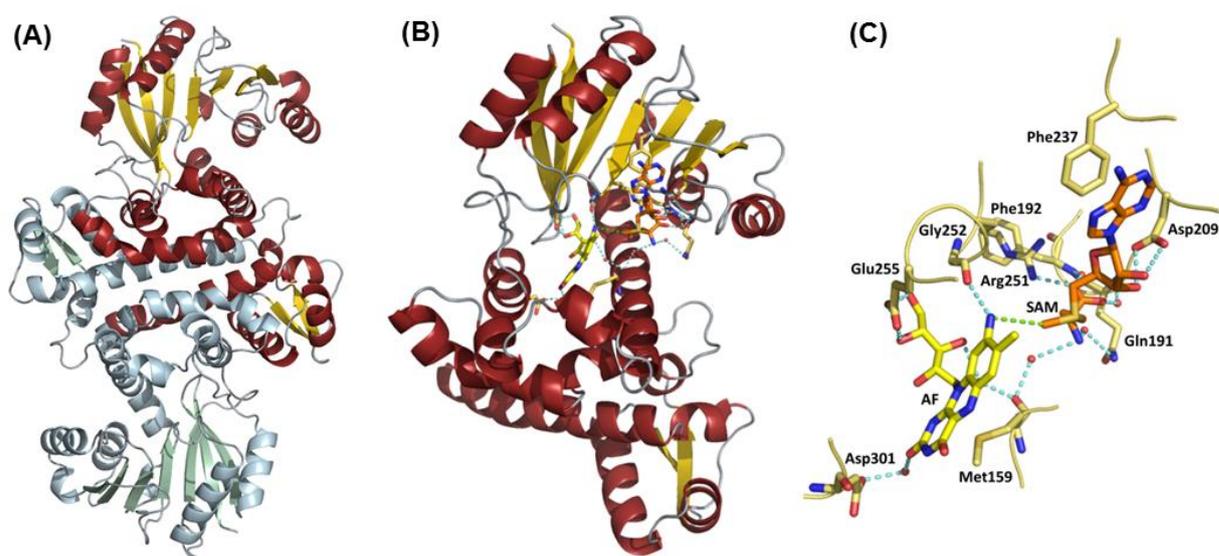
Fig. 24: Structure of RosA (A) Schematic representation of dimeric RosA structure (B) Schematic representation of RosA protomer with the predicted substrate (SAM in orange and AF in yellow) binding sites. (C) Close-up view of substrate binding sites of RosA [216].

Studies to investigate the binding properties of the substrates (SAM, AF) and the products (SAH and RoF) to RosA exhibited that both products bind more tightly to RosA (about 20 times) than SAM and AF do. Furthermore, RosA tends to form the ternary complex RosA:RoF:SAH in the presence of RoF and SAH. This ternary complex leads to drastic

spectral changes that are indicative of a hydrophobic environment. In addition, the tight binding of products is reflected by inhibition of the reaction since both RoF and SAH act as competitive inhibitors for AF and SAM, respectively. Thus, it is plausible that the tight binding of RoF (and SAH) to RosA constitutes a mechanism to confer resistance toward RoF of RoF-producing strains [216]. Predicting the potential SAM and AF binding sites in RosA suggested that both substrates, AF and SAM, can bind independently to their respective binding pockets. This finding was confirmed by kinetic studies that demonstrated a random-order bi-bi reaction mechanism [216]. However, RosA is a dimethylase performing two consecutive methylation reactions in the same active site. Consequently, after the first methylation, SAH needs to dissociate from the active site to provide room for another SAM. But, as mentioned above, SAH binds tightly to RosA and its affinity is even enhanced in the presence of RoF. It can be assumed that the binding affinity of SAH in the presence of MAF is in the same range between that of SAH in the presence/absence of RoF. Thus, the second methylation reaction will depend on the rate of dissociation of SAH, the limiting step for the conversion of MAF to RoF (Scheme 3 in Chapter 4) [216].

1.9 References

1. Martin, J.L. and McMillan, F.M. (2002) SAM (dependent) I AM: the S-adenosylmethioninedependent methyltransferase fold. *Curr Opin Struct Biol*, 12, 783-793

2. Fontacave, M., Atta, M., and Mulliez, E. (2004) S-adenosylmethionine: nothing goes to waste. *Trends Biochem Sci*, 29, 243-249

3. Cantoni, G.L. (1975) Biological methylation: selected aspects. *Annu Rev Biochem*, 44, 435-451

4. Chiang, P.K. Gordon, R.K., Tal, J., Zeng, G.C., Doctor, B.P., Pardhasaradhi, K., and McCann, P.P. (1996) S-adenosylmethionine and methylation. *FASEB J*, 10, 471-480

5. Mato, J.M., Alvarez, L., Ortiz, P., and Pajares, M.A. (1997) S-adenosylmethionine synthesis: molecular mechanisms and clinical implications. *Pharmacol Ther*, 73, 265-280

6. Lu, S.C. (2000) S-adenosylmethionine. *Int J Biochem Cell Biol*, 32, 391-395

7. Awad, W.M. Jr., Whitney, P.L., Skiba, W.E., Mangum, J.H., and Wells, M.S. (1983) Evidence for direct methyl transfer in betaine: homocysteine S-methyl-transferase. *J Biol Chem*, 258, 12790-12792

8. Boyer, P.D., Stokes, B.P., Wolcott, R.G., and Degani, C. (1975) Coupling of "high-energy" phosphate bonds to energy transductions. *Fed Proc*, 34, 1711-1717

9. Fauman, E.B., Blumenthal, R.M., and Cheng, X., (1999) Structure and evolution of AdoMet-dependent MTases. In Cheng, X. Blumenthal, R.M. (eds) S-Adenosylmethionine-dependent methyltransferases: structures and functions, World Scientific Inc, Singapore, 1-38

10. Mudd, S.H. and Cantoni, G.L. (1958) Activation of methionine for transmethylation. III. The methionine-activating enzyme of Bakers' yeast. *J Biol Chem*, 231, 481-492

11. Borchardt, R.T. and Wu, Y.S. (1976) Potential inhibitors of S-adenosyl-methionine-dependent methyltransferases. 5. Role of the asymmetric sulfonium pole in the enzymatic binding of S-adenosyl-L-methionine. *J Med Chem*, 19, 1099-1103

12. Finkelstein, J.D. (1990) Methionine metabolism in mammals. *J Nutr Chem*, 1, 228-237

13. Finkelstein, J.D. (2007) Metabolic regulatory properties of S-adenosylmethionine and S-adenosylcysteine. *Clin Chem Lab Med*, 2007, 45, 1694-1699

14. Sharma, S. and Litonjua, A. (2014) Asthma, allergy, and responses to methyl donor supplements and nutrients. *J Allergy Clin Immunol*, 133, 1246-1254

15. Obeid, R. (2013) The metabolic burden of methyl donor deficiency with focus on the betaine homocysteine methyltransferase pathway. *Nutrients*, 5, 3481-3495

16. Lin, L.H., Matsui, W., Aloe, C., Peng, R.D., Diette, G.B., Breysse, P.N., et al. (2013) Relationships between folate and inflammatory features of asthma. *J Allergy Clin Immunol*, 131, 918-920

17. Moens, A.L., Champion, H.C., Claeys, M.J., Tavazzi, B., Kaminski, P.M., Wolin, M.S. et al. (2008) High-dose folic acid pretreatment blunts cardiac dysfunction during ischemia coupled to maintenance of high-energy phosphates and reduces postreperfusion injury. *Circulation*, 117, 1810-1819

18. Duthie, S.J., Narayanan, S., Blum, S., Pirie, L., and Brand, G.M. (2000) Folate deficiency in vitro induces uracil misincorporation and DNA hypomethylation and inhibits DNA excision repair in immortalized normal human colon epithelial cells. *Nutr Cancer*, 37, 245-251

19. Wallace, J.M., McCormack, J.M., McNulty, H., Walsh, P.M., Robson, P.J., Bnham, M.P., Duffy, M.E., Ward, M., Molloy, A.M., Scott, J.M. et al. (2012) Choline supplementation and measures of choline and betaine status: a randomised, controlled trial in postmenopausal women. *Br J Nutr*, 108, 1264-1271

20. Zeisel, S.H. and da Costa K.A. (2009) Choline: an essential nutrient for public health. *Nutr Rev*, 67, 615-623

21. Li, Z. and Vance, D.E. (2008) Phosphatidylcholine and choline homeostasis. *J Lipid Res*, 49, 1187-1194

22. Blusztajn, J.K. (1998) Choline, a vital amine. *Science*, 281, 794-795

23. Konstantinova, S.V., Tell, G.S., Vollset, S.E., Ulvik, A., Drevon, C.A., and Ueland, P.M. (2008) Dietary patterns, food groups and nutrients as predictors of plasma choline and betaine in middle-aged and elderly men and women. *Am J Clin Nutr*, 88, 1663-1669)

24. Ganley, O., Graessle, O., and Robinson, H. (1958) Anti-inflammatory activity of components obtained from egg-yolk, peanut, oil, and soya-bean lecithin. *J Lab Clin Med*, 51, 709-714

25. Mehta, A.K., Arora, N., Gaur, S.N., and Singh, B.P. (2009) Choline supplementation reduces oxidative stress in mouse model of allergic airway disease. *Eur J clin Invest*, 39, 934-941

26. Yen, C.L., Mar, M.H., and Zeisel, S.H. (1999). Choline deficiency-induced apoptosis in PC12 cells is associated with diminished membrane phosphatidylcholine and sphingomyelin, accumulation of ceramide and diacylglycerol, and activation of a caspase. *FASEB J*, 13, 135-142

27. Yoshida, Y., Itoh, N., Hayakawa, M., Habuchi, Y., Inoue, R., Chen, Z.H. et al. (2006) Lipid peroxidation in mice fed a choline-deficient diet as evaluated by total hydroxyoctadecadienoic acid. *Nutrition*, 22, 303-311

28. Zeisel, S.H., Zola, T., da Costa, K.A., and Pomfret, E.A. (1989) Effect of choline deficiency on S-adenosylmethionine and methionine concentrations in rat liver. *Biochem J*, 259, 725-729

29. Mandavia, P.R., Stolk, L., and Heil, S.G. (2014) Homocysteine and DNA methylation: a review of animal and human literature. *Mol Genet Metab*, 113, 243-252

30. Kotb, M., Mudd, S.H., Mato, J.M., Geller, A.M., Kredich, N.M., Chou, J.Y., and Cantoni, G.L. (1997) Consensus nomenclature for the mammalian methionine adenosyltransferase genes and gene products. *Trends Genet*, 13, 51-52

31. Lieber, C.S. and Packer, L (2002) S-adenosylmethionine: molecular, biological, and clinical aspects-an introduction. *Am J Clin Nutr*, 76 (suppl), 1148S-1150S

32. Blom, H.J. and Smulders, Y. (2011) Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. *J Inherit Metab Dis*, 34, 75-81

33. Sunden, S.L., Renduchintala, M.S., Park, E.I., Miklasz, S.D, and Garrow, T.A. (1997) Betaine-homocysteine methyltransferase expression in porcine and human tissues and chromosomal localization of the human gene. *Arch Biochem Biophys*, 345, 171-174

34. Selhub, J. (1999) Homocysteine metabolism. *Annu Rev Nutr*, 19, 217-246

35. Quere, I., Paul, V., Rouillac, C., Janbon, C., London, J., Demaille, J., Kamoun, P., Dufier, J.L., Abitbol, M., and Chasse, J.F. (1999) Spatial and temporal expression of the cystathionine beta-synthase gene during early human development. *Biochem Biophys Res Commun*, 254, 127-137

36. Mudd, S.H. and Poole, J.R. (1975) Lablie methyl balances for normal humans on various dietary regimens. *Metabolism*, 24, 721-735

37. Garcea, R., Pascale, R., Diano, L., Frassetto, S., Cozzolina, P., Ruggiu, M.E., Vannini, M.G., Gaspa, L., and Feo, F. (1987) Variations of ornithine decarboxylase activity and S-adenosyl-Lmethionine and 5'-methylthioadenosine contents during the development of diethylnitrosamine-induced liver hyperplastic nodules and hepatocellular carcinoma. *Carcinogenesis*, 8, 653-658

38. Garcia-Trevijano, E.R., Latasa, M.U., Carretero, M.V., Berasain, C., Mato, J.M., and Avila, M.A. (2000) S-adenosylmethionine regulates MAT1A and MAT2A gene expression in cultured rat hepatocytes: a new role for S-adenosylmethionine in the maintenance of the differentiated status of the liver. *FASEB J*, 14, 2511-1518

39. Fontecave, M., Atta, M., and Mulliez, E. (2004) S-adenosylmethionine: nothing goes to waste. *Trends Biochem Sci*, 29, 243-249

40. Grogan, D.W., and Cronan, J.E. Jr (1997) Cyclopropane ring formation in membrane lipids of bacteria. *Miocrobiol Mol Biol Rev*, 61, 429-441

41. Taylor, F.R., and Cronan, J.E. Jr (1979) Cyclopropane fatty acid synthase of *Escherichia coli*. Stabilization, purification, and interaction with phospholipid vesicles. *Biochemistry*, 18, 3292-3300

42. Stoner, G.L. and Eisenberg, M.A. (1975) Purification and properties of 7,8-diaminopelargonic acid aminotransferase. *J Biol Chem*, 250, 4029-4036

43. Käck, H., Sandmark, J., Gibson, K., Schneider, G., and Lindqvist, Y. (1999) Crystal structure of diaminopelargonic acid synthase: evolutionary relationship between pyridoxal-5'-phosphate- dependent enzymes. *J Mol Biol*, 291, 857-876

44. Iwata-Reuyl, D. (2003) Biosynthesis of the 7-deazaguanosine hypermodified nucleosides of transfer RNA, *Bioorg Chem*, 31, 24-43

45. Van Lanen, S.G., Kinzie, S.D., Matthieu, S., Link, T., Culp, J., and Iwata-Reuyl, D. (2003) tRNA modification by S-adenosylmethionine: tRNA ribosyltransferase-isomerase. *J Biol Chem*, 278, 10491-10499

46. White, M.F., Vasquez, J., Yang, S.F., Kirsch, J.F. (1994) Expression of apple 1-aminocyclopropane-1-carboxylate synthase in *Escherichia coli*: kinetic characterization of wild-type and active-site mutant form. *Proc Natl Acad Sci USA*, 91, 12428–12432

47. Ayub, R., Guis, M., Ben Amor, M., Gillot, L., Roustan, J.P., Latché, A., Bouzayen, M., and Pech, J.C. (1996) Expression of ACC oxidase antisense gene inhibits ripening of cantaloupe melon fruits. *Nat Biotechnol*, 14, 862-866

48. Val, D.L. and Cronan, J.E. Jr (1998) In vivo evidence that S-adenosylmethionine and fatty acid synthesis intermediates are the substrate for the LuxI family of autoinducer synthetases. *J Bacteriol*, 180, 2644-2651

49. Nishimura, S., Taya, Y., Kuchino, Y., and Oashi, Z. (1974) Enzymatic synthesis of 3-(3-amino-3-carboxypropyl)uridine in *Escherichia coli* phenylalanine transfer RNA: transfer of the 3-amino-carboxypropyl group from S-adenosylmethionine. *Biochem Biophys Res Commun*, 57, 702-708

50. Jarrett, J.T. (2003) The generation of 5'-deoxyadenosyl radicals by adenosylmethionine-dependent radical enzymes. *Curr Opin Chem Biol*, 7, 174-182

51. Roje, S. (2006) S-adenosyl-L-methionine: beyond the universal methyl group donor. *Phytochemistry*, 67, 1686-1698

52. Toh, S.M., Xiong, L., Bae, T., and Mankin, A.S. (2008) The methyltransferase YfgB/RlmN is responsible for modification of adenosine 2503 in 23S rRNA. *RNA*, 14, 98-106

53. Cantoni, G.L. (1952) The nature of the active methyl donor formed enzymatically from l-methionine and adenosinetriphosphate. *J Amer Chem Soc*, 74, 2942-2943

54. Loenen, W.A.M. (2006) S-adenosylmethionine: jack of all trades and master of everything? *Biochem Soc Trans*, 34, 330-333

55. Struck, A.W., Thompson, M.L., Wong, L.S., and Micklefield, J. (2012) S-adenosyl-methionine-dependent methyltransferases: highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. *ChemBioChem*, 13, 2642-2655

56. Carreras, C.W. and Santi, D.V. (1995) The catalytic mechanism and structure of thymidylate synthesis. *Annu Rev Biochem*, 64, 721-762

57. Lautala, P., Ulmanen, I., and Taskinen, J. (2001) Molecular mechanisms controlling the rate and specificity of catechol O-methylation by human soluble catechol O-methyltransferase. *Mol Pharmacol*, 59, 393–402

58. Joshi, C., and Chiang, V.L. (1998) Conserved sequence motifs in plant S-adenosyl-L-methionine-dependent methyltransferases. *Plant Mol Biol*, 37, 663–674

59. Wu, Y.Y., Kang, Q.J., Shang, G.D., Spiteller, P., Carroll, B., Yu, T.W., Su, W.J., Bai, L.Q., and Floss, H.G. (2011) N-Methylation of the Amide Bond by Methyltransferase Asm10 in Ansamitocin Biosynthesis. *ChemBioChem*, 12, 1759-1766

60. Mahlert, C., Kopp, F., Thirlway, J., Micklefield, J., and Marahiel, M.A. (2007) Stereospecific enzymatic transformation of α-ketoglutarate to (2S,3R)-3-methyl glutamate during acidic lipopeptide biosynthesis. *J Am Chem Soc*, 129, 12011–12018

61. Schmidberger, J.W., James, A.B., Edwards, R., Naismith, J.H., O'Hagan, D. (2010) Halomethane biosynthesis: structure of a SAM-dependent halide methyltransferase from Arabidopsis thaliana. *Angew Chem Int Ed Engl*, 49, 3646-3648

62. LeDuc, D.L., Tarun, A.S., Montes-Bayon, M., Meija, J., Malit, M.F., Wu, C.P., AbdelSamie, M., Chiang, C.Y., Tagmount, A., deSouza, M., Neuhierl, B., Böck, A., Caruso, J., and Terry, N. (2004) Overexpression of selenocysteine methyltransferase in Arabidopsis and Indian mustard increases selenium tolerance and accumulation. *Plant Physiol*, 135, 377-383

63. Qin, J., Lehr, C.R., Yuan, C.G., Le, X.C., McDermott, T.R., and Rosen, B.P. (2009) Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga. *Proc Natl Acad Sci USA*, 106, 5213-5217

64. Scavetta, R.D., Thomas, C.B., Walsh, M.A., Szegedi, S., Joachimiak, A., Gumport, R.I., and Churchill, M.E.A. (2000) Structure of RsrI methyltransferase, a member of the N6-adenine β class of DNA methyltransferases. *Nucleic Acid Res*, 28, 3950-3961

65. Chen, l., MacMillan, A.M., Chang, W., Ezaz-Nikpay, K., Lane, W.S., and Verdine, G.L. (1991) Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. Biochemistry, 30, 11018–11025

66. Schubert, H.L., Blumenthal, R.M., and Cheng, X. (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci*, 28, 329-335

67. Vidgren, J., Svensson, L.A., and Liljas, A. (1994) Crystal structure of catechol O-methyltransferase. *Nature*, 368, 354-358

68. Goulding, C.W., Postigo, and D., Matthews, R.G. (1997) Cobalamin-dependent methionine synthase is a modular protein with distinct regions for binding homocysteine, methyltetrahydrofolate, cobalamin, and adenosylmethionine. *Biochemistry*, 36, 8082–8091.

69. Matthews, R.G., Koutmos, M., and Datta, S. (2008) Cobalamin-dependent and cobamide-dependent methyltransferases. *Curr Opin Struct Biol*, 18, 658–666.

70. Dixon, M.M., Huang, S., Matthews, R.G., and Ludwig, M. (1996) The structure of the C-terminal domain of methionine synthase: presenting S-adenosylmethionine for reductive methylation of B12. *Structure*, 4, 1263–1275

71. Koutmos, M., Datta, S., Pattridge, K.A., Smith, J.L., Matthews, R.G. (2009) Insights into the reactivation of cobalamin-dependent methionine synthase. *Proc Natl Acad Sci USA*, 106, 18527-18532

72. Schubert, H.L., Wilson, K.S., Raux, E., Woodcock, S.C., and Warren, M.J. (1998) The X-ray structure of a cobalamin biosynthetic enzyme, cobalt-precorrin-4 methyltransferase. *Nat Struct Biol*, 5, 585-592

73. Michel, G., Sauvé, V., Larocque, R., Li, Y., Matte, A., and Cygler, M. (2002) The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot. *Structure*, 10, 1303-1315

74. Tkaczuk, K.L., Dunin-Horkawicz, S., Purta, E., and Bujnicki, J.M. (2002) Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. *BMC Bioinformatics*, 8:73

75. Liu, R.J., Zhou, M., Fang, Z.P., Wang, M., Zhou, X.L., and Wang, E.D. (2013) The tRNA recognition mechanism of the minimalist SPOUT methyltransferase, TrmL. *Nucleic Acids Res*. 2013, 41, 7828-7842

76. Liu, J., Wang, W., Shin, D.H., Yokota, H., Kim, R., and Kim, S.H. (2003) Crystal structure of tRNA (m1G37) methyltransferase from Aquifex aeolicus at 2.6 Å resolution: a novel methyltransferase fold. *Proteins*, 53, 326-328

77. Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, 128, 693-705

78. Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998) SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci*, 54, 80-93

79. Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng X. (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol*, 6: 227

80. Cheng, X., Collins, R.E., Zhang, X. (2005) Structural and sequence motifs of protein (histone) methylation enzymes. *Annu Rev Biophys Biomol Struct*, 34, 267-294

81. Boriack-Sjodin, P.A. and Swinger K.K. (2016) Protein Methyltransferases: A Distinct, Diverse, and Dynamic Family of Enzymes. *Biochemistry*, 55, 1557-1569

82. Zhang, X. and Bruice, T.C. (2008) Enzymatic mechanism and product specificity of SET-domain protein lysine methyltransferases. *Proc Natl Acad Sci USA*, 105, 5728-5732

83. Zhang, Q., van der Donk, W.A., and Liu, W. (2012) Radical-mediated enzymatic methylation: a tale of two SAMS. *Acc Chem Res*, 45, 555-564

84. Broderick, J.B., Duffus, B.R., Duschene, K.S., and Shepard, E.M. (2014) Radical S-adenosylmethionine enzymes. *Chem Rev*, 114, 4229–4317

85. Grove, T.L., Benner, J.S., Radle, M.I., Ahlum, J.H., Landgraf, B.J., Krebs, C., and Booker, S.J. (2011) A radically different mechanism for S-adenosylmethionine-dependent methyltransferases. *Science*, 332, 604-607

86. Frey, P.A. and Magnusson, O.Th. (2003) S-Adenosylmethionine: A Wolf in Sheep's Clothing, or a Rich Man's Adenosylcobalamin? *Chem Rev*, 103, 2129–2148

87. Wang, S.C. and Frey, P.A. (2007) S-adenosylmethionine as an oxidant: the radical SAM superfamily. *Trends Biochem Sci*, 32, 101-110

88. Bauerle, M.R., Schwalm, E.L., and Booker, S.J. (2015) Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *J Biol Chem*, 290, 3995-4002

89. Vázquez-Laslop, N., Ramu, H., Klepacki, D., Kannan, K., and Mankin, A. S. (2010) The key role of a conserved and modified rRNA residue in the ribosomal response to the nascent peptide. *EMBO J*, 29, 3108–3117

90. Long, K.S., Poehlsgaard, J., Kehrenberg, C., Schwarz, S., and Vester, B. (2006) The Cfr rRNA methyltransferase confers resistance to Phenicols, Lincosamides, Oxazolidinones, Pleuromutilins, and Streptogramin A antibiotics. *Antimicrob Agents Chemother*, 50, 2500-2505

91. Grove, T.L., Radle, M.I., Krebs, C., and Booker, S.J. (2011) Cfr and RlmN contain a single [4Fe-4S] cluster, which directs two distinct reactivities for S-

adenosylmethionine: methyl transfer by S_N2 displacement and radical generation. *J Am Chem Soc*, 133, 19586–19589

92. Kim, H.J., McCarty, R.M., Ogasawara, Y., Liu, Y.N., Mansoorabadi, S.O., LeVieux, J., and Liu, H.W. (2013) GenK-catalyzed C-6' methylation in the biosynthesis of gentamicin: isolation and characterization of a cobalamin-dependent radical SAM enzyme. *J Am Chem Soc*, 135, 8093-8096

93. Woodyer, R.D., Li, G., Zhao, H., van der Donk, W.A. (2007) New insight into the mechanism of methyl transfer during the biosynthesis of fosfomycin. *Chem Commun (Camb)*, 28, 359-361

94. van der Donk, W.A. (2006) Rings, radicals, and regeneration: the early years of a bioorganic laboratory. *J Org Chem*, 71, 7561-7571

95. Yu, Y., Duan, L., Zhang, Q., Liao, R., Ding, Y., Pan, H., Wendt-Pienkowski, E., Tang, G., Shen, B., and Liu, W. (2009) Nosiheptide biosynthesis featuring a unique indole side ring formation on the characteristic thiopeptide framework. *ACS Chem Biol*, 4, 855-864

96. Ding, Y., Yu, Y., Pan, H., Guo, H., Li, Y., and Liu, W. (2010) Moving posttranslational modifications forward to biosynthesize the glycosylated thiopeptide nocathiacin I in Nocardia sp. ATCC202099. *Mol Biosyst*, 6, 1180-1185

97. Zhang, Q., Chen, D., Lin, J., Liao, R., Tong, W., Xu, Z., and Liu, W. (2011) Characterization of NocL involved in thiopeptide nocathiacin I biosynthesis: a [4Fe-4S] cluster and the catalysis of a radical S-adenosylmethionine enzyme. *J Biol Chem*, 286, 21287-21294

98. Morris, R.P., Leeds, J.A., Naegeli, H.U., Oberer, L., Memmert, K., Weber, E., LaMarche, M. J., Parker, C.N., Burrer, N., Esterow, S., Hein, A.E., Schmitt, E.K., and Krastel, P. (2009) Ribosomally synthesized thiopeptide antibiotics targeting elongation factor Tu. *J Am Chem Soc*, 131, 5946-5955

99. Layer, G., Grage, K., Teschner, T., Schünemann, V., Breckau, D., Masoumi, A., Jahn, M., Heathcote, P., Trautwein. A.X., and Jahn, D. (2005) Radical S-adenosylmethionine enzyme coproporphyrinogen III oxidase HemN: functional features of the [4Fe-4S] cluster and the two bound S-adenosyl-L-methionines. *J Biol Chem*, 280, 29038-29046

100. Layer, G., Moser, J., Heinz, D.W., Jahn, D., and Schubert, W.D. (2003) Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. EMBO J, 22, 6214-6224

101. Allen, K.D., Xu, H., White, R.H. (2014) Identification of a unique radical S-adenosylmethionine methylase likely involved in methanopterin biosynthesis in *Methanocaldococcus jannaschii*. *J Bacteriol*, 196, 3315-3323

102. Finer-Moore, J.S., Santi, D.V., and Stroud, R.M. (2003) Lessons and conclusions from dissecting the mechanism of bisubstrate enzyme: thymidylate synthase mutagenesis, function, and structure. *Biochemistry*, 42, 248-256

103. Finer-Moore, J., Fauman, E.B., Foster, P.G., Perry, K.M., Santi, D.V., and Stroud, R.M. (1993) Refined structures of substrate-bound and phosphate-bound thymidylate synthase from *Lactobacillus casei*. *J Mol Biol*, 232, 1101-1116

104. Montfort, W.R., Perry, K.M., Fauman, E.B., Finer-Moore, J.S., Maley, G.F., Hardy, L., Maley, F., and Stroud, R.M. (1990) Structure, multiple site binding, and segmental accommodation in thymidylate synthase on binding dUMP and an antifolate. *Biochemistry*, 29, 6964-6977

105. Matthews, D.A., Villafranca, J.E., Janson, C.A., Smith, W.W., Welsh, K., and Freer, S. (1990) Stereochemical mechanism of action for thymidylate synthase based on the X-ray structure of the covalent inhibitory ternary complex with 5-fluoro-2'-deoxyuridylate and 5,10-methylenetetrahydrofolate. *J Mol Biol*, 214, 937-948

106. Matthews, D.A., Appelt, K., Oatley, S.J., and Xuong, Ng.H. (1990) Crystal structure of Escherichia coli thymidylate synthase containing bound 5-fluoro-2′-deoxyuridylate and 10-propargyl-5,8-dideazafolate. *J Mol Biol*, 214, 923-936

107. Myllykallio, H., Leduc, D., Filee, J., and Liebl, U. (2003) Life without dihydrofolate reductase FolA. *Trends Microbiol*, 11, 220-223

108. Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., and Liebl, U. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. *Science*, 297, 105-107

109. Hamdane, D., Argentini, M., Cornu, D., Golinelli-Pimpaneau, B., and Fontecave, M. (2012) FAD/folate-dependent tRNA methyltransferase: flavin as a new methyl-transfer agent. *J Am Chem Soc*, 134, 19739-19745

110. Lartigue, C., Lebaudy, A., Blanchard, A., El Yacoubi, B., Rose, S., Grosjean, H., and Douthwaite, S. (2014) The flavoprotein Mcap0476 (RlmFO) catalyzes m5U1939 modification in Mycoplasma capricolum 23S rRNA. *Nucleic Acids Res*, 42, 8073-8082

111. Agrawal, N., Lesley, S.A., Kuhn, P., and Kohen, A. (2004) Mechanistic studies of a flavin-dependent thymidylate synthase. *Biochemistry*, 43, 10295-10301

112. Griffin, J., Roshick, C., Iliffe-Lee, E., and McClarty, G. (2005) Catalytic mechanism of Chlamydia trachomatis flavin-dependent thymidylate synthase. *J Biol Chem*, 280, 5456-5467

113. Koehn, E.M., Perissinotti, L.L., Moghram, S., Prabhakar, A., Lesley, S.A., Mathews, I.I., and Kohen, A. (2012) Folate binding site of flavin-dependent thymidylate synthase. *Proc Natl Acad Sci USA*, 109, 15722–15727

114. Falus, A., Grosman, N., and Darvas, Z. (2004) Histamine: biology and medical aspects. Budapest: SpringMed Publishing

115. Rosenwasser, L. (2007) New insights into pathophysiology of allergic rhinitis. *Allergy Asthma Proc*, 28, 10-15

116. Loisella, J. and Wollin, W. (1993) Mucosal histamine elimination and its effect on acid secretion in rabbit gastric mucosa. *Gastroenterology*, 104, 1013-1020

117. Zhang, M. Thurmond, R.L., and Dunford, P.J. (2007) The histamine H(4) receptor: A novel modulator of inflammatory and immune disorders. *Pharmacol Ther*, 113, 594-606

118. Schwartz, J.C., Arrang, M.M., Garbarg, M., Pollaed, H., and Ruat, M. (1991) Histaminergic transmission in the mammalian brain. *Physiol Rev*, 71, 1-51

119. Hough, L.B. (1988) Cellular localization and possible functions for brain histamine: recent progress. *Prog Neurobiol*, 30, 469-505

120. Wada, H., Inagaki, N., Yamatodani, A., and Watanabe, T. (1991) Is the histaminergic neuron system a regulatory center for whole-brain activity? *Trends Neurosci*, 14, 415-418

121. Best, C.H., Dale, H.H., Dudley, H.H., and Thorpe, W.V. (1927) The nature of the vaso-dilator constituents of certain tissue extracts. *J Physiol*, 62, 397–417

122. Maintz, L. and Novak, N. (2007) Histamine and histamine intolerance. *Am J Clin Nutr*, 85, 1185-1196

123. Sarkadi, L. (2004) Histamine in food. In Falus, A. (Ed.), Histamine: biology and medical aspects (pp. 176-185), Budapest: SpringMed Publishing

124. Brown, R.E., Stevens, D.R., and Haas, H.L. (2001) The physiology of brain histamine. *Prog Neurobiol*, 63, 637-672

125. Haas, H.L., Sergeeva, O.A., and Selbach, O. (2008) Histamine in the nervous system. *Physiol Rev*, 88, 1183-1241

126. Panula, P., Yang, H.Y., and Costa, E. (1984) Histamine-containing neurons in the rat hypothalamus. *Proc Natl Acad Sci USA*, 81, 2572-2576

127. Panula, P. and Nuutinen, S. (2013) The histaminergic network in the brain: basic organization and role in disease. *Nat Rev Neurosci*, 14, 472-487

128. Vanhala, A., Yamatodani, A., and Panula P. (1994) Distribution of histamine-, 5-hydroxytryptamine-, and tyrosine hydroxylase-immunoreactive neurons and nerve fibers in developing rat brain. *J Comp Neurol*, 347, 101-114

129. Molina-Hernández, A., Díaz, N.F., Arias-Montaño, J.A. (2012) Histamine in brain development. *J Neurochem*, 122, 872-882

130. Sauvageot, C. M. and Stiles, C. D. (2002) Molecular mechanisms controlling cortical gliogenesis. *Curr Opin Neurobiol*, 12, 244–249

131. Molina-Hernandez, A. and Velasco, I. (2008) Histamine induces neural stem cell proliferation and neuronal differentiation by activation of distinct histamine receptors. *J Neurochem*, 106, 706–717

132. Barnes, P.J. (2000) Anti-IgE therapy in asthma: rationale and therapeutic potential, *Int. Arch. Allergy Immunol*. 123, 196-204

133. Hill, S.J., Ganellin, C.R., Timmerman, H., Schwartz, J.C., Shankley, N.P., Young, J.M., Schunack, W., Levi, R., and Haas, H.L. (1997) International union of pharmacology. XIII. Classification of histamine receptors, *Pharmacol. Rev*, 49, 253-278

134. de Esch, I.J., Thurmond, R.L., Jongejan, A., and Leurs, R. (2005) The histamine H4 receptors as a new therapeutic target for inflammation. *Trends Pharmacol Sci*, 26, 462-469

135. Widnell, K.L., Chen, J.S., Iredale, P.A., Walker, W.H., Duman, R.S., Habener, J.F. and Nestler, E.J. (1996) Transcriptional regulation of CREB (cyclic AMP response element-binding protein) expression in CATH a cells. *J. Neurochem*, 66, 1770–1773

136. Drutel, G., Peitsaro, N., Karlstedt, K., Wieland, K., Smit, M.J., Timmerman, H., Panula, P. and Leurs, R. (2001) Identification of rat H3 receptor isoforms with different brain expression and signaling properties. *Mol Pharmacol*, 59, 1–8

137. Pollard, H., Moreau, J., Arrang, J.M. and Schwartz, J.C. (1993) A detailed autoradiographic mapping of histamine H3 receptors in rat brain areas. *Neuroscience*, 52, 169–189

138. O'Reilly, M., Alpert, R., Jenkinson, S., Gladue, R.P., Foo, S., Trim, S., Peter, B., Trevethick, M. and Fidock, M. (2002) Identification of a histamine H4 receptor on human eosinophils-role in eosinophil chemotaxis. *J Recept Signal Transduct Res*, 22, 431–448

139. Dijkstra, D., Stark, H., Chazot, P.L., Shenton, F.C., Leurs, R., Werfel, T. and Gutzmer, R. (2008) Human inflammatory dendritic epidermal cells express a functional histamine H4 receptor. *J Invest Dermatol*, 128, 1696–1703

140. Schwelberger, H.G., Ahrens, F., Fogel, W.A., and Sánchez-Jiménez, F. (2013) Chapter 3: Histamine metabolism, in Stark, H. (Ed.), Histamine H4 receptor: a novel drug target for immunoregulation and inflammation (pp 63-102), Versita, London

141. Levy, M. (1935) A note on the titration constants of imidazole derivatives. *J Biol Chem*, 109, 361-363

142. Dvorak, A.M. (1998) Histamine content and secretion in basophils and mast cells. *Prog Histochem Cytochem*, 33, 169-320

143. Gilfillan, A.M. and Tkaczyk, C. (2006) Integrated signalling pathways for mast-cell activation. *Nat Rev Immunol*, 6, 218-30

144. Lorentz, A., Baumann, A., Vitte, J., and Blank U. (2012) The SNARE Machinery in Mast Cell Secretion. *Front Immunol*, 3:143

145. Travis, E.R., Wang, Y.M., Michael, D.J., Caron, M.G., and Wightman, R.M. (2000) Differential quantal release of histamine and 5-hydroxytryptamine from mast cells of vesicular monoamine transporter 2 knockout mice. *Proc Natl Acad Sci USA*, 97, 162-167

146. Darvas, Z. and Falus, A. (2004) Histidine decarboxylase (HDC) enzyme and gene. In Falus, A. (Ed.), Histamine: biology and medical aspects (pp. 31-42). Budapest: SpringMed Publishing

147. Hayashi, H. (1995) Pyridoxal enzymes: mechanistic diversity and uniforminity. *J Biochem*, 118, 463-473

148. Yatsunami, K., Ohtsu, H., Tsuchikawa, M., Higuchi, T., Ishibashi, K., Shida, A., Shima, Y., Nakagawa, S., Yamauchi, K., Yamamoto, M., Hayashi, N., Watanabe, T., and Ichikawa A. (1994). Structure of the L-histidine decarboxylase gene. *J Biol Chem*, 269, 1554-1559

149. Ercan-Sencicek, A.G., Stillman, A.A., Ghosh, A.K., Bilguvar, K., O'Roak, B.J., Mason, C.E., Abbott, T., Gupta, A., King, R.A., Pauls, D.L., Tischfield, J.A., Heiman, G.A., Singer, H.S., Gilbert, D.L., Hoekstra, P. J., Morgan, T.M., Loring, E., Yasuno, K., Fernandez, T., Sanders, S., Louvi, A., Cho, J.H., Mane, S., Colangelo, C.M., Biederer, T., Lifton, R.P., Gunel, M., and State, M.W. (2010) L-histidine decarboxylase and Tourette syndrome. *N Engl J Med*, 362, 1901–1908

150. Hocker, M. (2004) Molecular mechanisms of gastrin-dependent gene regulation. *Ann N Y Acad Sci*, 1014, 97-109

151. Ai, W., Liu, Y., Langlois, M., and Wang, T.C. (2004) Kruppel-like factor 4 (KLF4) represses histidine decarboxylase gene expression through an upstream Sp1 site and downstream gastrin responsive elements. *J Biol Chem*, 279, 8684-8693

152. Dartsch, C., Chen, D., Håkanson, R, Persson, L. (1999) Histidine decarboxylase in rat stomach ECL cells: relationship between enzyme activity and different molecular forms. *Regul Pept*, 81, 41-48

153. Komori, H., Nitta, Y., Ueno, H., and Higuchi, Y. (2012) Structural study reveals that Ser-354 determines substrate specificity on human histidine decarboxylase. *J Biol Chem*, 287, 29175-29183

154. Moya-García, A.A., Medina, M.A., and Sánchez-Jiménez, F. (2005) Mammalian histidine decarboxylase: from structure to function. *Bioessays*, 27, 57-63

155. Pino-Angeles, A., Morreale, A., Negri, A., Sánchez-Jiménez, F., and Moya-García, A.A. (2010) Substrate uptake and protein stability relationship in mammalian histidine decarboxylase. *Proteins*, 78, 154-161

156. Olmo, M.T., Sánchez-Jiménez, F., Medina, M.A., and Hayashi, H. (2002) Spectroscopic analysis of recombinant rat histidine decarboxylase. *J Biol Chem*, 132, 433-439

157. Maslinski, C. and Fogel, W.A. (1991) Catabolism of histamine, in: Uvnäs, B. (Ed.), Handbook of experimental pharmacology, Vol 97, Histamine and histamine antagonists (pp. 165-189). Springer, Berlin- Heidelberg

158. Schwelberger, H.G. (2004) Diamine oxidase (DAO) enzyme and gene. In Falus, A. (Ed.), Histamine: biology and medical aspects (pp. 44-52), Budapest: SpringMed Publishing

159. Chassande, O., Renard, S., Barbry, P., and Lazdunski, M. (1994) The human gene for diamine oxidase, an amiloride binding protein. Molecular cloning, sequencing, and characterization of the promoter. *J Biol Chem*, 269, 14484-14489

160. Schwelberger, H.G., Hittmair, A., and Kohlwein, S.D. (1998) Analysis of tissue and subcellular localization of mammalian diamine oxidase by confocal laser scanning fluorescence microscopy. *Inflamm Res*, 47, Suppl 1:S60-61

161. Klocker, J., Mätzler, S.A., Huetz, G.N., Drasche, A., Kolbitsch, C., and Schwelberger, H.G. (2005) Expression of histamine degrading enzymes in porcine tissues. *Inflamm Res*, 54, Suppl 1:S54-57

162. Schmidt, W.U., Sattler, J., Hesterberg, R., Röher, H.D., Zoedler, T., Sitter, H., and Lorenz, W. (1990) Human intestinal diamine oxidase (DAO) activity in Crohn's disease: a new marker for disease assessment? *Agents Actions*, 30, 267-270

163. McGrath, A.P., Hilmer, K.M., Collyer, C.A., Shepard, E.M., Elmore, B.O., Brown, D.E., Dooley, D.M., and Guss, J.M. (2009) Structure and inhibition of human diamine oxidase. *Biochemistry*, 48, 9810-9822

164. Brown, D.D., Tomchick, R., and Axelrod, J. (1959) The distribution and properties of a histamine-methylating enzyme. *J Biol Chem*, 234, 2948-2950

165. Lindahl, K.M. (1960) The histamine methylating enzyme system in liver. *Acta Physiol Scand*, 49, 119-138

166. Okinaga, S., Ohrui, T., Nakazawa, H., Yamauchi, K., Sakurai, E., Watanabe, T., Sekizawa, K., and Sasaki, H. (1995) The role of HMT (histamine N-methyltransferase) in airways: a review. *Methods Find Exp Clin Pharmacol*, 17, Suppl C: 16-20

167. Chen, G.L., Zhu, B., Nie, W.P., Xu, Z.H., Tan, Z.R., Zhou, G., Liu, J., Wang, W., and Zhou, H.H. (2004) Single nucleotide polymorphisms and haplotypes of histamine N-methyltransferase in patients with gastric ulcer. *Inflamm Res*, 53, 484-488

168. Burkard, W.P., Gey, K.F., and Pletscher, A. (1963) Diamine oxidase in the brain of vertebrates. *J Neurochem*, 10, 183-186

169. Preuss, C.V., Wood, T.C., Szumlanski, C.L., Raftogianis, R.B., Otterness, D.M., Girard, B., Scott, M.C., and Weinshilboum, R.M. (1998) Human histamine N-methyltransferase pharmacogenetics: common genetic polymorphisms that alter activity. *Mol Pharmacol*, 53, 708-717

170. Girard, B., Otterness, D.M., Wood, T.C., Honchel, R., Wieben, E.D., and Weinshilboum, R.M. (1994) Human histamine N-methyltransferase pharmacogenetics: cloning and expression of kidney cDNA, *Mol Pharmacol*, 45, 461-468

171. Yamauchi, K., Sekizawa, K., Suzuki, H., Nakazawa, H., Ohkawara, Y., Katayose, D., Ohtsu, H., Tamura, G., Shibahara, S., Takemura, M., et al. (1994) Structure and function of human histamine N-methyltransferase: critical enzyme in histamine metabolism in airway. *Am J Physiol*, 267, L342-349

172. Aksoy, S., Raftogianis, R., and Weinshilboum, R.M. (1996) Human histamine N-methyltransferase gene: structural characterization and chromosomal localization, *Biochem Biophys Res Commun*, 219, 548-554

173. Horton, J.R., Sawada, K., Nishibori, M., Zhang, X., and Cheng, X. (2001) Two polymorphic forms of human methyltransferase: structure, thermal, and kinetic comparisons. *Structure*, 9, 837-849

174. Horton, J.R., Sawada, K., Nishibori, M., Cheng, X. (2005) Structure basis for inhibition of histamine N-methyltransferase by diverse drugs. *J Mol Biol*, 353, 334-344

175. Gitomer, W.L. and Tipton, K.F. (1986) Purification and kinetic properties of ox brain histamine N-methyltransferase. *Biochem J*, 233, 669-676

176. Gitomer, W.L. and Tipton, K.F. (1983) The role of cytoplasmic aldehyde dehydrogenase in the metabolism of N-tele-methylhistamine. *Pharmacol Biochem Behav*, 18, Suppl 1:113-116

177. Rutherford, K., Parson, W.W., and Daggett,V. (2008) The histamine N-methyltransferase T105I polymorphism affects active site structure and dynamics. *Biochemistry*, 47, 893-901

178. Yan, L., Galinsky, R.E., Bernstein, J.A., Liggett, S.B., and Weinshilboum, R.M. (2000) Histamine N-methyltransferase pharmacogenetics: association of a common functional polymorphism with asthma. *Pharmacogenetics*, 10, 261-266

179. García-Martín, E., García-Menaya, J., Sánchez, B., Martínez, C., Rosendo, R., and Agúndez, J.A.G. (2007) Polymorphisms of histamine-metabolizing enzymes and clinical manifestations of asthma and allergic rhinitis. *Clin Exp Allergy*, 37, 1175-1182

180. Ledesma, M.C., García-Martín, E., Alonso-Navarro, H., Martínez, C., Jiménez-Jiménez, F.J., Benito-León, J., Puertas, I., Rubio, L., López-Alburquerque, T., and Agúndez, J.A.G. (2008) The nonsynonymous Thr105Ile polymorphism of the histamine N-methyltransferase is associated to the risk of developing essential tremor. *Neuromol Med*, 10, 356-361

181. Palada, V., Terzić, J., Mazzulli, J., Bwala, G., Hagenah, J., Peterlin, B., Hung, A.Y., Klein, C., and Krainc, D. (2012) Histamine N-methyltransferase Thr105Ile polymorphism is associated with Parkinson's disease. *Neurobiol Aging*, 33, 836.e1-836.e3

182. Kaufman, L., Ayub, M., and Vincent, J.B. (2010) The genetic basic of non-syndromic intellectual disability: a review. *J Neurodevelop Disord*, 2, 182-209

183. Maulik, P.K., Mascarenhas, M.N., Mathers, C.D., Dua, T., and Saxena, S. (2011) Prevalence of intellectual disability: a meta-analysis of population based studies. *Res Dev Disabil*, 32, 419-436

184. Heidari, A., Tongsook, C., Najafipour, R., Musante, L., Vasil, N., Garshasbi, M., Hu, H., Mittal, K., McNaughton, A.J., Sritharan, K., Hudson, M., Stehr, H., Talebi, S., Moradi, M., Darvish, H., Arshad Rafiq, M., Mozhdehipanah, H., Rashidinejad, A., Samiei, S., Ghadami, M., Windpassinger, C., Gillessen-Kaesbach, G., Tzschach, A., Ahmed, I., Mikhailov, A., Stavropoulos, D.J., Carter, M.T., Keshavarz, S., Ayub, M., Najmabadi, H., Liu, X., Ropers, H.H., Macheroux, P., and Vincent, J.B. (2015) Mutations in the histamine N-methyltransferase gene, HNMT, are associated with nonsyndromic autosomal recessive intellectual disability. *Hum Mol Genet*, 24, 5697-5710

185. Tongsook, C., Niederhauser, J., Kronegger, E., Straganz, G., and Macheroux, P. Leucine 208 in human histamine N-methyltransferase emerges as a hotspot for protein stability rationalizing the role of the L208P variant in intellectual disability. *Unpublished data*

186. Reiff, R.E., Ali, B.R., Baron, B., Yu, T.W., Ben-Salem, S., Coulter, M.E., Schubert, C.R., Hill, R.S., Akawi, N.A., Al-Younes, B., Kaya, N., Evrony, G.D., Al-Saffar, M., Felie, J.M., Partlow, J.N., Sunu, C.M., Schembri-Wismayer, P., Alkuraya, F.S., Meyer, B.F., Walsh, C.A., Al-Gazali, L., and Mochida, G.H. (2014) METTL23, a transcriptional partner of GABPA, is essential for human cognition. *Hum Mol Genet*, 23, 3456-3466

187. Bernkopf, M., Webersinke, G., Tongsook, C., Koyani, C.N., Rafiq, M.A., Ayaz, M., Müller, D., Enzinger, C., Aslam, M., Naeem, F., Schmidt, K., Gruber, K. Speicher, M.R., Malle, E., Macheroux, P., Ayub, M., Vincent, J.B., Windpassinger, C., Duba, H.C. (2014) Disruption of the methyltransferase-like 23 gene METTL23 causes mild autosomal recessive intellectual disability. *Hum Mol Genet*, 23, 4015-4023

188. Cloutier, P., Lavallée-Adam, M., Faubert, D., Blanchette, M., and Coulombe, B. (2013) A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity. *PLoS Genet*, 9:e1003210

189. Kruiswijk, T., Kunst, A., Planta, R.J., and Mager, W.H. (1978) Modification of yeast ribosomal proteins. Methylation. *Biochem J*, 175, 221–225

190. Lee, S.W., Berger, S.J., Martinovic, S., Pasa-Tolic, L., Anderson, G.A., et al. (2002) Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc Natl Acad Sci USA*, 99, 5942–5947

191. Diederich, K., Schabitz, W.R., and Minnerup, J. (2012) Seeing old friends from a different angle: novel properties of hematopoietic growth factors in the healthy and diseased brain. *Hippocampus*, 22, 1051-1057.

192. Anitha, A., Nakamura, K., Thanseem, I., Matsuzaki, H., Miyachi, T., Tsujii, M., Iwata, Y., Suuki, K., Sugiyama, T., and Mori, N. (2013) Downregulation of the expression of mitochondrial electron transport complex genes in autism brains. *Brain Pathol*, 23, 294-302

193. Bacher, A. (1991) Riboflavin kinase and FAD synthetase, In Muller, F. (Ed.), Chemistry and biochemistry of flavoenzymes (pp. 349-370), vol. 1. CRC Press, Boca Raton, Florida

194. Ghisla, S. and Massey, V. (1986) New flavins for old: artificial flavins as active site probes of flavoproteins. *Biochem J*, 239, 1-12

195. Mack, M. and Grill, S. (2006) Riboflavin analogs and inhibitors of riboflavin biosynthesis. *Appl Microbiol Biotechnol*, 71, 265-275

196. Pedrolli, D.B., Jankowitsch, F., Schwarz, J., Langer, S., Nakanishi, S., Frei, E., and Mack M. (2013) Riboflavin analogs as antiinfectives: occurrence, mode of action, metabolism and resistance. *Curr Pharm Des*, 19, 2552-2560

197. Cooper, L.E., O'Leary, S.E., and Begley, T.P. (2014) Biosynthesis of a thiamin antivitamin in Clostridium botulinum. *Biochemistry*, 53, 2215-2217

198. Fiehe, K., Arenz, A., Drewke, C., Hemscheidt, T., Williamson, R.T., and Leistner, E. (2000) Biosynthesis of 4'-O-methylpyridoxine (Ginkgotoxin) from primary precursors. *J Nat Prod*, 63, 185-189

199. Otani, S., Takatsu, M., Nakano, M., Kasai, S., and Miura, R. (1974) Roseoflavin, a new antimicrobial pigment from Streptomyces. *J Antibiot (Tokyo)*, 27, 86-87

200. Schwarz, J., Konjik, V., Jankowitsch, F., Sandhoff, R., and Mack, M. (2016) Identification of the Key Enzyme of Roseoflavin Biosynthesis. *Angew Chem Int Ed Engl*, 55, 6103-6106

201. Macheroux, P., Kappes, B., and Ealick, S.E. (2011) Flavogenomics--a genomic and structural view of flavin-dependent proteins. *FEBS J*, 278, 2625-2634

202. Pedrolli, D.B., Jankowitsch, F., Schwarz, J., Langer, S., Nakanishi, S. and Mack, M. (2014) Natural riboflavin analogs. *Meth Mol Biol*, 1146, 41-63

203. Otani, S., Kasai, S., and Matsui, K. (1980) Isolation, chemical synthesis, and properties of roseoflavin. *Methods Enzymol*, 66, 235-241

204. Jankowitsch, F., Schwarz, J., Rückert, C., Gust, B., Szczepanowski, R., Blom, J., Pelzer, S., Kalinowski, J., and Mack, M. (2012) Genome sequence of the bacterium Streptomyces davawensis JCM 4913 and heterologous production of the unique antibiotic roseoflavin. *J Bacteriol*, 194, 6818-6827

205. Walsh, C., Fisher, J., Spencer, R., Graham, D.W., Ashton, W,T., Brown, J.E., Brown, R.D., and Rogers, E.F. (1978) Chemical and enzymatic properties of riboflavin analogues. *Biochemistry*, 17, 1942-1951

206. Shinkai, S., Kameoka, K., Honda, N., Ueda, K., Manabe, O., and Lindsey, J. (1986) Spectral and reactivity studies of roseoflavin analogs: correlation between reactivity and spectral parameters. *Bioorg Chem*, 14, 119-133

207. Grill, S., Busenbender, S., Pfeiffer, M., Kohler, U., and Mack, M. (2008) The bifunctional flavokinase/flavin adenine dinucleotide synthetase from Streptomyces davawensis produces inactive flavin cofactors and is not involved in resistance to the antibiotic roseoflavin. *J Bacteriol*, 190, 1546–1553

208. Langer, S., Nakanishi, S., Mathes, T., Knaus, T., Binter, A., Macheroux, P., Mase, T., Miyakawa, T., Tanokura, M., and Mack, M. (2013) The flavoenzyme azobenzene reductase AzoR from Escherichia coli binds roseoflavin mononucleotide (RoFMN) with high affinity and is less active in its RoFMN form. *Biochemistry*, 52, 4288-4295

209. Yorita, K., Misaki, H., Palfey, B.A., and Massey, V. (2000) On the interpretation of quantitative structure-function activity relationship data for lactate oxidase. *Proc Natl Acad Sci USA*, 97, 2480-2485

210. Matsui, K., Juri, N., Kubo, Y., and Kasai, S. (1979) Formation of roseoflavin from guanine through riboflavin. *J Biochem*, 86, 167-175

211. Juri, N., Kubo, Y., Kasai, S., Otani, S., Kusunose, M., and Matsui, K. (1987) Formation of roseoflavin from 8-amino- and 8-methylamino-8-demethyl-D-riboflavin. *J Biochem*, 101, 705-711

212. Jankowitsch, F., Kühm, C., Kellner, R., Kalinowski, J., Pelzer, S., Macheroux, P., Mack, M. (2011) A novel *N,N*-8-amino-8-demethyl-D-riboflavin Dimethyltransferase (RosA) catalyzing the two terminal steps of roseoflavin biosynthesis in Streptomyces davawensis. *J Biol Chem*, 286, 38275-38285

213. Pedrolli, D.B., Nakanishi, S., Barile, M., Mansurova, M., Carmona, E.C., Lux, A., Gärtner, W., and Mack, M. (2011) The antibiotics roseoflavin and 8-demethyl-8-amino-riboflavin from Streptomyces davawensis are metabolized by human flavokinase and human FAD synthetase. *Biochem Pharmacol*, 82, 1853-1859

214. Zalacain, M. and Cundliffe, E. (1989) Methylation of 23S rRNA caused by tlrA (ermSF), a tylosin resistance determinant from Streptomyces fradiae. *J Bacteriol*, 171, 4254-4260

215. Chen, H., Yamase, H., Murakami, K., Chang, C.W., Zhao, L., Zhao, Z., and Liu, H.W. (2002) Expression, purification, and characterization of two N,N-dimethyltransferases, tylM1 and desVI, involved in the biosynthesis of mycaminose and desosamine. *Biochemistry*, 41, 9165-9183

216. Tongsook, C., Uhl, M.K., Jankowitsch, F., Mack, M., Gruber, K., and Macheroux, P. (2016) Structural and kinetic studies on RosA, the enzyme catalysing the methylation of 8-demethyl-8-amino-d-riboflavin to the antibiotic roseoflavin. *FEBS J*, 283, 1531-1549

CHAPTER 2

2. Biochemical characterization and kinetic studies on histamine *N*-methyltransferase

2.1 Mutations in the histamine *N*-methyltransferase gene, *HNMT*, are associated with nonsyndromic autosomal recessive intellectual disability

AUTHOR CONTRIBUTIONS

OXFORD

ORIGINAL ARTICLE

Mutations in the histamine N-methyltransferase gene, HNMT, are associated with nonsyndromic autosomal recessive intellectual disability

Abolfazl Heidari[1,2], Chanakan Tongsook[4,†], Reza Najafipour[2,†], Luciana Musante[5], Nasim Vasli[1], Masoud Garshasbi[5,6], Hao Hu[5], Kirti Mittal[1], Amy J. M. McNaughton[7], Kumudesh Sritharan[1], Melissa Hudson[7], Henning Stehr[9], Saeid Talebi[10], Mohammad Moradi[2], Hossein Darvish[11], Muhammad Arshad Rafiq[1], Hossein Mozhdehipanah[3], Ali Rashidinejad[12], Shahram Samiei[13], Mohsen Ghadami[14], Christian Windpassinger[15], Gabriele Gillessen-Kaesbach[16], Andreas Tzschach[5,‡], Iltaf Ahmed[1,17], Anna Mikhailov[1], D. James Stavropoulos[18], Melissa T. Carter[19], Soraya Keshavarz[2], Muhammad Ayub[8], Hossein Najmabadi[20,21], Xudong Liu[7], Hans Hilger Ropers[5], Peter Macheroux[4] and John B. Vincent[1,22,23,*]

[1]Molecular Neuropsychiatry and Development (MiND) Lab, The Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada M5T 1R8, [2]Cellular and Molecular Research Center, [3]Department of Neurology, Bou Ali Sina Hospital, Qazvin University of Medical Sciences, Qazvin 34197/59811, Iran, [4]Institute of Biochemistry, Graz University of Technology, Graz 8010, Austria, [5]Max Planck Institute of Molecular Genetics, Berlin D-14195, Germany, [6]Department of Medical Genetics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran 14117-13116, Iran, [7]Ongwanada Genomics Lab, [8]Division of Developmental Disabilities, Department of Psychiatry, Queen's University, Kingston, ON, Canada K7L7X3, [9]Department of Medicine, Stanford University, Stanford, CA 94305-5101, USA, [10]Department of Medical Genetics, Medical University of Tehran, Tehran 14167-53955, Iran, [11]Department of Medical Genetics, Shahid Beheshti University of Medical Sciences, Tehran 4739, Iran, [12]Maternal, Fetal and Neonatal Research Center, Tehran University of Medical Sciences, Tehran 1419733141, Iran, [13]Blood Transfusion Research Center, Tehran 1449613111, Iran, [14]Department of Medical Genetics, Tehran University of Medical Sciences, Tehran 1417613151, Iran, [15]Institute of Human Genetics, Medical University of Graz, Graz 8010, Austria, [16]Institut für Humangenetik, Universität zu Lübeck, Lübeck 23562, Germany, [17]Atta-ur-Rehman School of Applied Biosciences, National University of Sciences and Technology, H-12, Islamabad, Pakistan, [18]Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, ON, Canada, [19]Division of Clinical and Metabolic Genetics, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada,

[20]Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran 19857, Iran, [21]Kariminejad-Najmabadi Pathology and Genetics Center, Tehran 14667, Iran, [22]Department of Psychiatry, University of Toronto, Toronto, ON, Canada M5T 1R8 and [23]Institute of Medical Science, University of Toronto, Toronto, ON, Canada M5S 1A8

*To whom correspondence should be addressed at: The Campbell Family Mental Health Research Institute, R-32, Centre for Addiction and Mental Health, 250 College Street, Toronto, ON, Canada M5T 1R8. Tel: +1 416 535 8501; Fax: +1 416 979 4666; Email: john.vincent@camh.ca

Abstract

Histamine (HA) acts as a neurotransmitter in the brain, which participates in the regulation of many biological processes including inflammation, gastric acid secretion and neuromodulation. The enzyme histamine N-methyltransferase (HNMT) inactivates HA by transferring a methyl group from S-adenosyl-L-methionine to HA, and is the only well-known pathway for termination of neurotransmission actions of HA in mammalian central nervous system. We performed autozygosity mapping followed by targeted exome sequencing and identified two homozygous *HNMT* alterations, p.Gly60Asp and p.Leu208Pro, in patients affected with nonsyndromic autosomal recessive intellectual disability from two unrelated consanguineous families of Turkish and Kurdish ancestry, respectively. We verified the complete absence of a functional HNMT in patients using *in vitro* toxicology assay. Using mutant and wild-type DNA constructs as well as *in silico* protein modeling, we confirmed that p.Gly60Asp disrupts the enzymatic activity of the protein, and that p.Leu208Pro results in reduced protein stability, resulting in decreased HA inactivation. Our results highlight the importance of inclusion of *HNMT* for genetic testing of individuals presenting with intellectual disability.

Introduction

Intellectual disability (ID) is a neurodevelopmental disorder, characterized by considerable limitation of intellectual functioning, adaptive behavior, or daily living skills, and with an onset before 18 years of age. It is one of the most important challenges in healthcare, with significant life-long socio-economic burden. ID is genetically heterogeneous and may result from chromosomal aberrations, or from either autosomal recessive (AR), autosomal dominant, X-linked or mitochondrial mutations. With the prevalence of ∼1% of children worldwide (1), ID can be divided into two main groups: nonsyndromic (NS) ID, where it might present as the sole clinical feature, whereas in syndromic ID additional clinical or dysmorphological features may also be present. Over the past few years, next-generation sequencing technologies have led to the identification of a number of ID-associated genes, emphasizing the considerable genetic heterogeneity of ID (2). Studies into the molecular basis of autosomal recessive forms of ID (ARID) are lagging some way behind studies of X-linked ID, in part because the larger families needed for gene mapping are rare in North American and European populations. However, a recent review suggests that ARID is not rare, and in outbred populations as many as 13–24% of ID may be due to AR genes (2).

Histamine (HA), a biogenic amine, plays a key role in the regulation of gastric acid secretion (3), and is a neurotransmitter in the central nervous system (CNS) (4). HA is produced and stored in airway mast cells, basophils and in the synaptic vesicles of HAergic neurons. In response to immune allergens, HA releases from storage granules and rapidly diffuses into surrounding tissues. Released HA is rapidly inactivated and disappears from the bloodstream within minutes [reviewed in Schwartz et al. (4)].

Histamine N-methyltransferase (HNMT; MIM 605238) is a cytoplasmic protein that belongs to the methyltransferase superfamily and is one of two enzymes involved in the metabolism of HA. N-methylation catalyzed by HNMT and oxidative deamination catalyzed by diamine oxidase (DAO; encoded by amino oxidase, copper containing 1 (AOC1)) are the two major pathways for HA biotransformation in mammals (5,6). HNMT catalyzes the methylation of HA in the presence of S-adenosylmethionine (SAM), forming N-*tele*-methylhistamine (7,8). HNMT is widely expressed in human tissues; with the greatest expression in kidney and liver, followed by spleen, colon, prostate, ovary, spinal cord cells, bronchi and trachea. HNMT is the key enzyme for HA degradation in the bronchial epithelium (9). Since DAO is not expressed in the CNS, N-methylation through HNMT is the main pathway responsible for termination of the neurotransmitter actions of HA in the brain (4).

Here, we report the identification of *HNMT* as a novel gene responsible for ID and discuss the consequences of the identified missense mutations on the protein function.

Results

Family A (Iranian family)

Ascertainment and clinical evaluation

We ascertained a consanguineous family with Turkish background from the Avaj area within Qazvin province in Iran, in which the first-cousin parents had nine children, four of them were affected with NS ID; two males and two females (Fig. 1). The study was approved by the Research and Ethics Board of Qazvin Medical University and appropriate written informed consent was obtained from the parents. The affected family members were assessed by an experienced neurologist and standard clinical assessment forms were used to document the findings. The clinical descriptions of the patients are summarized in Table 1. The affected females showed profound to severe ID and their speech was limited to just a few words, whereas in affected males the condition was milder. A mild degree of regression after about 5 years of age was reported for affected members. The patients did not have any neurological problems, autistic features, congenital malformations or facial dysmorphisms. Body height, weight and head circumference were normal in all patients. Wechsler Intelligence Scales for Children (WISC) were used to assess the IQ in patients. For patient IV:I, we performed a magnetic resonance imaging scan which revealed no morphological brain abnormalities.
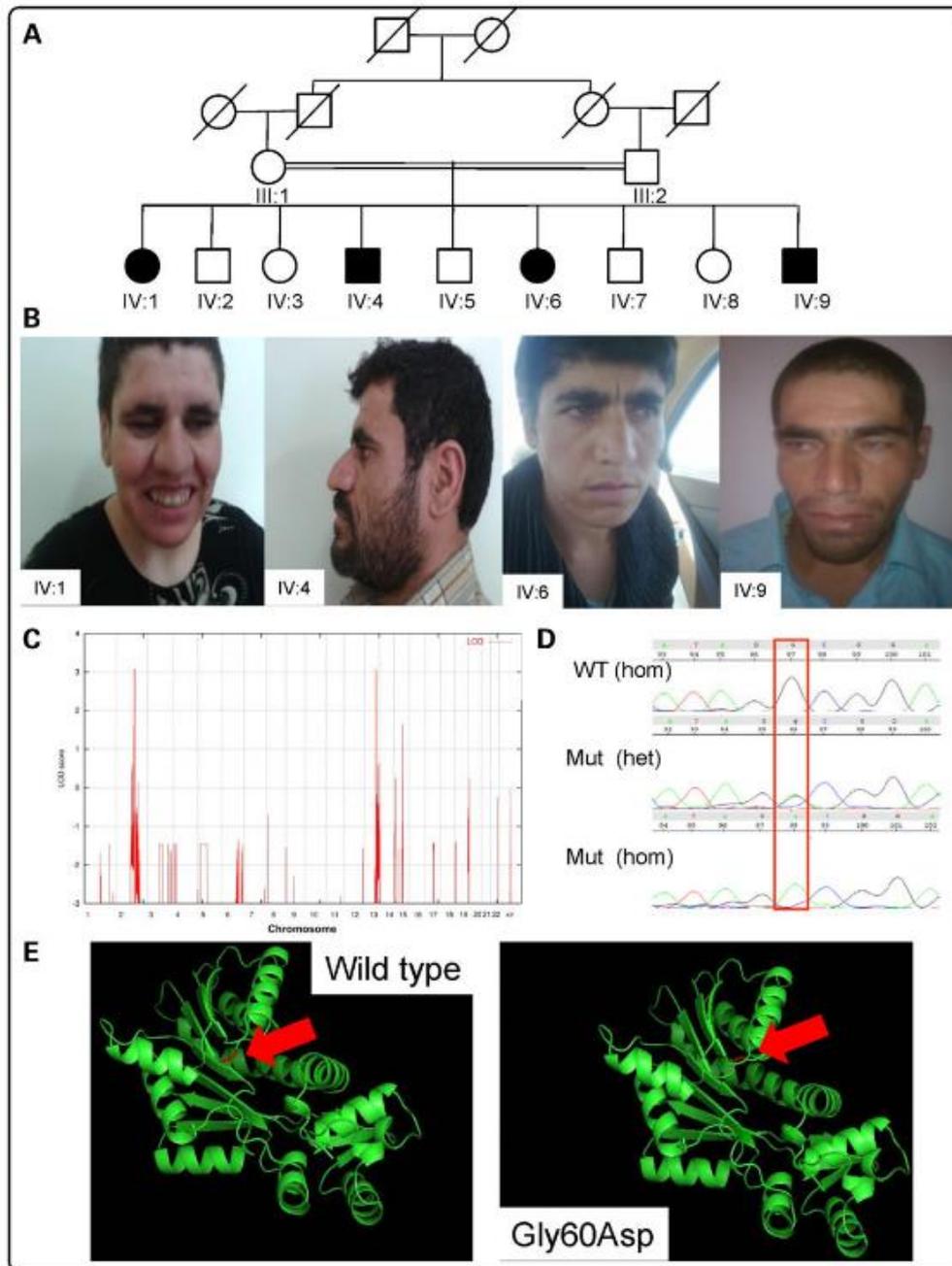
Figure 1. Analysis of Family A (Iranian family): (**A**) Pedigree. Black-shaded symbols indicate affected individuals (IV:1, IV:4, IV:6 and IV:9). (**B**) Photos of affected individuals: from left to right: IV:1, IV:4, IV:6 and IV:9. (**C**) Homozygosity mapping data analysis indicates peaks (LOD = 3.0) on chromosomes 2 and 13. (**D**) Electropherograms from Sanger confirmation in family members showing NM_006895.2 (HNMT): c.179G>A; p.Gly60Asp WT, heterozygous and homozygous sequence. (**E**) *In silico* modeling of p.Leu208Pro within HNMT predicted protein structure for p.Gly60Asp WT and mutant using PDB file 2AOT and Pymol software. The red arrow indicates the location of residue 60 within the protein.

Homozygosity-by-descent (HBD) mapping and mutation identification
HBD mapping led to the identification of a 14-Mb autozygous locus on 2q21.3 (single-nucleotide polymorphisms, SNPs: rs1869829–rs7573156), and a 3-Mb autozygous within region 13q33.1 (SNPs:

rs1336666–rs1475276) with a significant LOD (logarithm (base 10) of odds) score of 3.13 (Fig. 1C). Additionally, existence of copy number variations (CNVs) exclusive to the affected individuals was also ruled out. Exome target enrichment was performed within the

Table 1. Clinical and biometric features for the Iranian family (Family A) and the Kurdish family (Family B)

	Family A: IV:1	Family A: IV:4	Family A: IV:6	Family A: IV:8	Family B: III:1	Family B: III:3	Family B: III:4
HNMT mutation (NM_006895.2)	c.179G>A (p.Gly60Asp)	c.179G>A (p.Gly60Asp)	c.179G>A (p.Gly60Asp)	c.179G>A (p.Gly60Asp)	c.632T>C (p.Leu208Pro)	c.632T>C (p.Leu208Pro)	c.632T>C (p.Leu208Pro)
Gender	F	M	F	M	M	F	M
Parental consanguinity	First cousins	First cousins	First cousins	First cousins	First cousins	First cousins	First cousins
Ethnic origin	Turkish	Turkish	Turkish	Turkish	Kurdish	Kurdish	Kurdish
Age at examination (years)	35	33	31	27	18	15	13
Height (cm) (SD)	151 (−2.5 SD)	172 (−0.8 SD)	153 (−2.2 SD)	170 (−1.2 SD)	165 (−1.8 SD)	151 (−2 SD)	138 (−2.1 SD)
Weight (kg) (SD)	65 (+1.5 SD)	71 (+0.7 SD)	63 (+1.2 SD)	69 (+0.6 SD)	56 (+0.2 SD)	44 (+0.2 SD)	40 (+0.8 SD)
Head circumference (cm) (SD)	55 (Mean)	54.5 (−1.46 SD)	55 (Mean)	54 (−1.46 SD)	54 (−1.5 SD)	52 (−1.7 SD)	52 (−1.7 SD)
ID	IQ: 28 (severe)	IQ: 49 (moderate)	IQ: 25 (severe)	IQ: 54 (mild)	IQ: 20–34 (severe)	IQ: 20–34 (severe)	IQ: 20–34 (severe)

Height and head circumference are given in cm; weight in kg. Standard deviation (SD) from mean values is given in parentheses.

linkage intervals by using a custom Agilent SureSelect array, followed by sequencing 8.3 Gb of 101 bp paired-end reads using the Illumina Genome Analyzer II platform with 100% coverage and average depth of 202 reads. After filtering the variants with dbSNP130, and 1000 Genome data, we annotated the remaining mutations with the RefSeq gene model. Analysis of prospective changes from the critical regions indicated DNA variants in three genes: METTL21C [chr13:103346806C>G, NM_001010977.1: c.43G>C; p.Gly60Arg], ZRANB3 [chr2:136107663T>C, NM_032143.2:c.482A>G; p.Tyr161Cys] and HNMT [chr2:138727776G>A; NM_006895.2:c.179G>A; p.Gly60Asp] (coordinates used in hg19). The evolutionary conservation of the relevant nucleotides, as defined by the PhyloP44 score and the pathogenicity of these variants, as predicted by PolyPhen2 and SIFT were calculated (Table 2). Analysis with Condel—an integrated analysis that uses prediction using five different algorithms, including PolyPhen2, SIFT and Mutation Assessor (10), was also performed (Table 2). The co-segregation pattern of the three variants was checked in the family, and only the HNMT variant segregated correctly. The parents of the patients were both heterozygous for the HNMT missense variant, which was not found in either in a homozygous or heterozygous form among 100 unrelated healthy Iranian and 200 Pakistani individuals.

We identified a potentially pathogenic missense mutation (HNMT c.179G>A [p.Gly60Asp]) in HNMT (RefSeqNC_000002.11), which encodes a two-domain protein; MTase, which is the larger domain, is composed of a seven-stranded β-sheet and is mainly responsible for interaction with cofactor (SAM) and substrate (HA), and the S domain of HNMT, which may interact with other proteins for its *in vivo* function. The homozygous c.179G>A (p.Gly60Asp) variant occurs in the conserved MTase region I (Ile56-Gly64), which is part of the SAM-binding pocket (Fig. 3), highlighting the role of Gly60 in the interaction of HNMT with its cofactor. The predicted 3D structure of HNMT appears to be altered by the substitution (Fig. 1E). This variant was not present in either dbSNP138, nor the Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA, USA (http://evs.gs.washington.edu/EVS/; accessed January 2014), nor in 1000 Genomes (http://browser.1000genomes.org/index.html), but was present in 2 of 125 604 alleles in the Exome Aggregation Consortium database (ExAC; Cambridge, MA, USA; http://exac.broadinstitute.org; October 2014). Importantly, it was shown to be absent from a panel of 200 ethnically matched control chromosomes, as well as an in-house database of 521 exomes/genomes from unrelated individuals of Middle Eastern origin. The c.179G>A (p.Gly60Asp) mutation showed complete segregation with the disease in all affected family members; parents were heterozygous carriers and unaffected children available for genetic screening were either carriers or did not harbor the mutation.

A known C-to-T SNP (rs11558538) in HNMT changes the amino acid at position 105 from threonine to isoleucine, with the frequencies of the Thr105 and Ile105 alleles being ~90 and 10%, respectively. The Ile105 allele is correlated with diminished HNMT enzymatic activity, which could result in reduced HA inactivation and increased sensitivity to this amine (11,12). We checked this polymorphism in Family A and verified that Ile105 allele is not present in the patients.

Family B (Kurdish family)

Ascertainment and clinical evaluation

We ascertained a three-generation consanguineous Kurdish family (G016), originally from Iraq, recruited in Germany. The first-cousin parents had seven children, three of them presented

Table 2. Analysis of variants identified by targeted exome sequencing in the Iranian and Kurdish families (A and B, respectively)

Family	Gene (accession #)	Variant	SNP?	Polyphen2	SIFT	PROVEAN	PhyloP44 (Mean)	Condel	Segregates with phenotype, Y/N
Family A (Iranian)	HNMT (NM_006895.2)	Gly60Asp		1 (probably damaging)	0.003 (damaging)	−6.303 (damaging)	3.092	1 (deleterious)	Y
	METTL21C (NM_001010977.1)	Gly15Arg	rs2390760	0.002 (benign)	0.429 (tolerated)	0.523 (neutral)	−0.444	0 (neutral)	N
	ZRANB3 (NM_032143.2)	Tyr161Cys	rs181335970	1 (probably damaging)	0.001 (damaging)	−6.847 (damaging)	2.268	0.990 (deleterious)	N
Family B (Kurdish)	HNMT (NM_006895.2)	Leu208Pro		1 (probably damaging)	0.001 (damaging)	−6.534 (damaging)	2.475	1 (deleterious)	Y

Scores and predicted effect are indicated from Polyphen2, SIFT, PROVEAN, PhyloP44 and Condel (which is itself an integration of five separate prediction algorithms).

with NS ID; the pedigree is shown in Figure 2. The probands were examined by experienced clinical geneticists who assessed their physical and mental status. The study was approved by the local institutional ethics committee, and appropriate informed consent was obtained from the parents. To exclude chromosomal abnormalities, karyotype analysis by G-banding was performed in all affected individuals; karyotypes were found to be normal. Clinical descriptions of the patients are summarized in Table 1. On assessment, the 18-year-old boy (III:1) showed severe ID. He started to walk at 1 year of age. His speech was severely delayed and he attended special school for intellectually disabled children. His sister, III:3, a 15-year-old, was also severely intellectually disabled. She had a normal motor development, but active speech started at the age of about 2 years and she attended, as with the older brother, a special school for intellectually disabled children. The younger affected male (III:4), age 13 years, was more delayed than his older siblings. The mother reported that at birth the child was thin and hypotonic. He began walking at 2½ years. He started to speak at 4–5 years of age. He presented with hyperactive behavior. At the age of 12 years, he was diagnosed with myelodysplastic syndrome. The patients did not present with congenital malformations or facial dysmorphisms. Physical measurements are also reported in Table 1.

HBD mapping and mutation identification

Analysis of genotypes for the available individuals (III:1, III:2, III:3 and II:1) of family G016 revealed a large interval of autozygosity on 2q21.2–q24.3. This 30 Mb interval was flanked by the heterozygous SNPs, rs10928469 and rs6705268. Four additional autozygous loci were identified: a 3.4-Mb interval on 10q26 (rs11244548–rs4751029), a 1.7-Mb HBD on 11p12–p13 (rs672597–rs10836780), a 12.6-Mb region on 11q22.3–q23.3 (rs10895742–rs521609), and finally an 8.5-Mb locus in 18q12.1–q12.3 (rs1021598–rs12971263; Fig. 2C).

We performed targeted next-generation sequencing by enrichment of the exonic regions within the linkage intervals. After filtering the variants, a single homozygous variant in HNMT [chr2: 138771444T>C; NM_006895.2:c.623T>C; p.Leu208Pro] (coordinates used in hg19) was ranked as potentially pathogenic. The mutation has not been reported in dbSNP 138, 1000 Genomes or the NHLBI Exome Variant Server, and was absent in our in-house databases of 521 exomes/genomes from unrelated individuals of Middle Eastern origin, but was present in 1 of 126 358 alleles in the ExAC database. Direct Sanger sequencing analysis was performed for all available family members, and demonstrated co-segregation with the disease in the family according to a recessive mode of inheritance. The predicted 3D structure of HNMT does not appear to be significantly altered by the substitution (Fig. 2E). The mutation affects a highly conserved amino acid (Fig. 3) and, in line with this, the PhyloP score (13) was found to be 4425, Grantham score 98 (14). *In silico* analyses with SIFT, PolyPhen-2, ConDel and others all predicted the amino acid substitution to be damaging (Table 2).

ID and ASD cohorts from outbred population

Cohorts of N = 991 ID and N = 1000 autism spectrum disorder (ASD) subjects were screened using a pooled targeted sequencing approach. No potentially damaging rare homozygous or compound heterozygous variants were identified. A rare heterozygous variant, NM_006895.2:c.430-1G>A; Chr2:138762701G>A, which would potentially alter splicing of exon 5, was identified in one ASD and one ID individual. This variant was not present in the ExAC database (~65 000 exomes), nor in a database of 521 Middle Eastern exomes. Although potentially damaging heterozygous
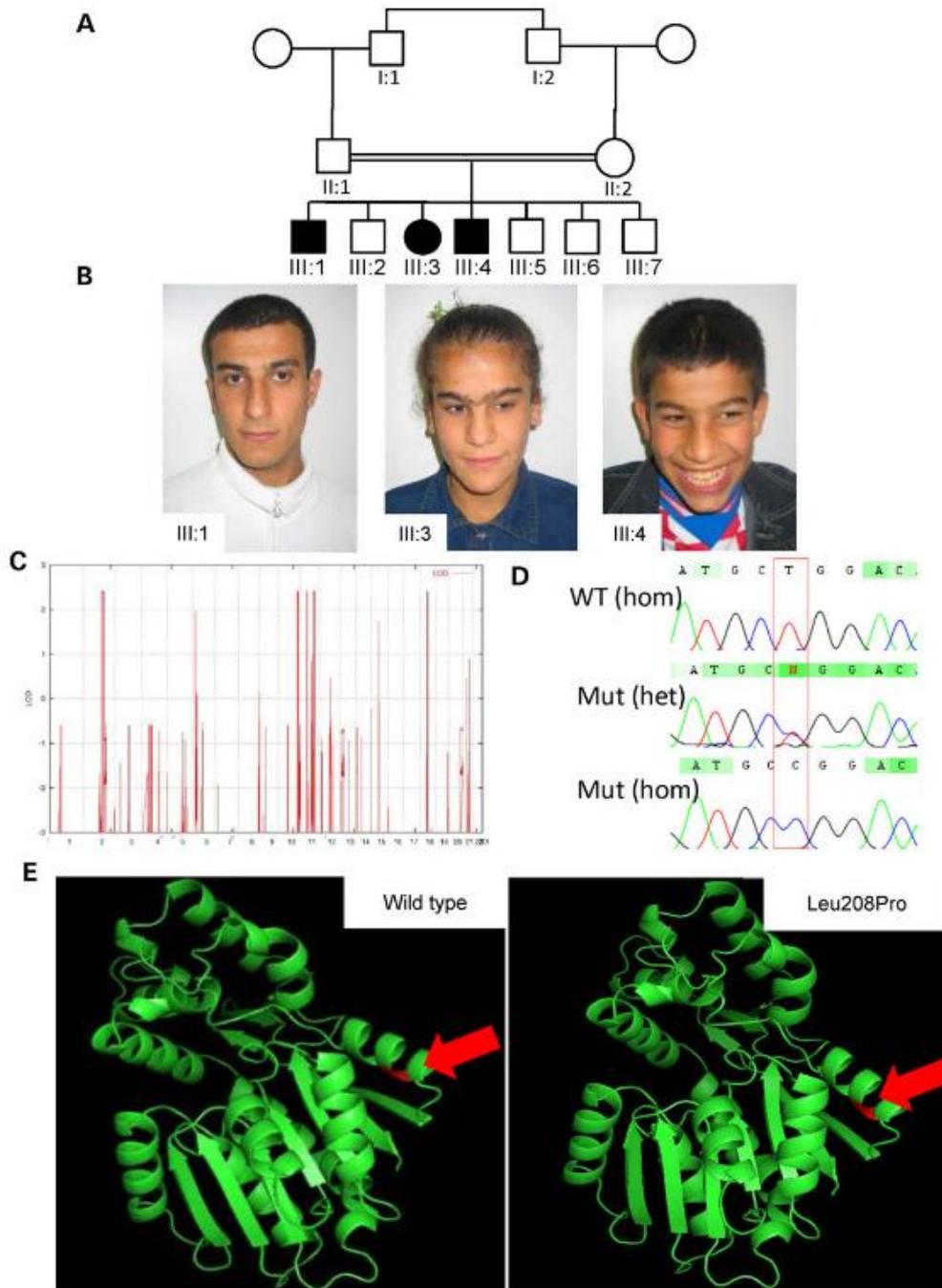
Figure 2. Analysis of Family B (Kurdish family): (**A**) Pedigree. Black-shaded symbols indicate affected individuals (III:1, III:3 and III:4). (**B**) Photos of affected individuals: from left to right: III:1, III:3 and III:4. (**C**) Homozygosity mapping data analysis indicates peaks (LOD = 2.4) including a large interval (30 Mb) on chromosome 2. (**D**) Electropherograms from Sanger confirmation in family members showing NM_006895.2 (HNMT):c.632T>C; p.Leu208Pro WT, heterozygous and homozygous sequence. (**E**) *In silico* modeling of p.Leu208Pro within HNMT. Modeling was performed with PDB file 2AOT and PyMol software. The red arrow indicates the location of residue 208 within the protein.

single-nucleotide variants (SNVs) and CNVs are reported elsewhere for schizophrenia, bipolar disorder, ID and autism spectrum disorder (Supplementary Material, Fig. S1), loss of function (LoF) variants are reported for HNMT in more than 40 individuals in the ExAC control data, and thus, it seems unlikely that heterozygous LoF variants are associated with these disorders.

Figure 3. Three-dimensional Protein structure and ClustalW2 analysis: structures of HNMT (pdb 1jqd) at the catalytic domain for (**A**) Gly6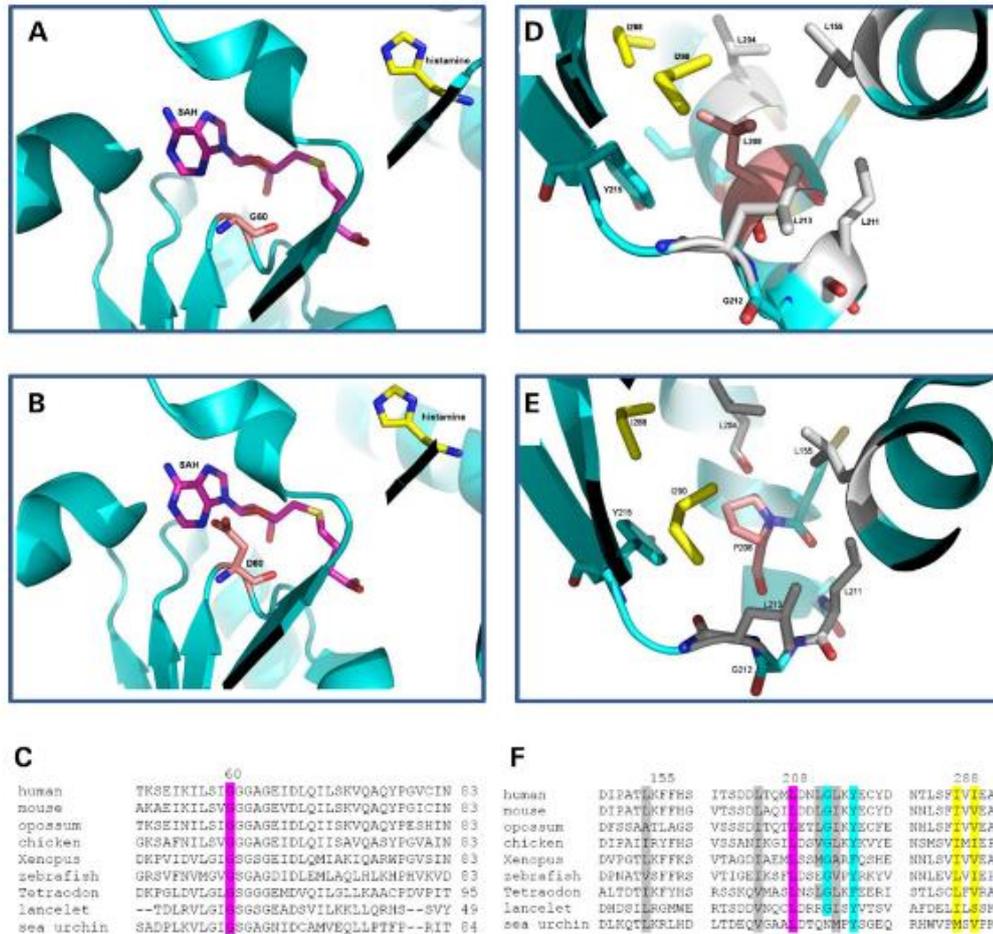0, (**B**) Asp60, with HA in yellow and S-adenosyl homocysteine (SAH) in pink, and Gly60 and Asp60 in light pink, and (**C**) ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2) alignment/comparison of HNMT across vertebrate species showing conservation at Gly60Asp (highlighted in pink). Structures at the hydrophobic pocket around residue 208 (labeled pink) for (**D**) Leu208 and (**E**) Pro208. Leu155, Leu204, Leu211 and Leu213 were labeled in gray. Ile288 and Ile290 were labeled in yellow. Tyr215 and Gly212 were labeled in blue. (**F**) ClustalW2 alignment of HNMT for the hydrophobic pocket surrounding Leu208, with Leu208 highlighted in pink, Leu155, Leu204, Leu211 and Leu213 labeled in gray, Ile288 and Ile290 labeled in yellow. Tyr215 and Gly212 labeled in blue. Sequences used for the HNMT alignment included human (NP_008826.1), mouse (NP_536710.1), opossum (from N-SCAN and Genscan gene predictions, UCSC browser), chicken (NP_001264802.1), *Xenopus laevis* (NP_001080614.1), zebrafish (NP_001003636.1), *Tetraodon nigroviridis* (Q4SBY6.1), lancelet (*Branchiostoma floridae*: predicted from mRNA XM_002613293.1) and sea urchin (*Strongylocentrotus purpuratus*: from mRNAs CX698504, CD312314 and CX689147).

In silico modeling of HNMT Leu208Pro predicts functional impairment

To gain insights into the structural consequence of the p.Leu208-Pro mutation in HMNT, we performed structure-based *in silico* three-dimensional modeling of the mutant protein (Fig. 2E). The non-conservative leucine to proline change at position 208 occurs in one of the six regions with an alpha helix conformation (α-E) which flanks the seven-stranded β-sheet within the MTase fold (15). The analysis indicated that although the mutation is not proximal to the catalytic site, the rigidity introduced by the proline residue would likely alter the helical conformation, destabilizing the protein and therefore affecting the enzymatic function. Proline is well known as an alpha helix breaker due to its side chain and steric constraints. In fact, the hydrogen bonding network and conformation of the helix would be disrupted by the side chain forced into the space occupied by the helix

backbone, and by the methyl group at the position normally occupied by an amide proton (16). We have quantified this effect by performing *in silico* energy calculations with PoPMuSiC 2.1 and IMutant 2.0 (17), which calculate stability changes upon mutation. In both cases, the simulation revealed a decrease in stability, 3.73 and 2.09 kcal/mol, respectively. An additional analysis using FoldX (18) revealed a more stark decrease in stability of 8.98 kcal/mol. A leucine to proline substitution is causing the most destabilizing effect when all permutations are considered. Thus, we hypothesize that the p.Leu208Pro mutant protein is either unstable or the function is severely impaired.

Gly60Asp and Leu208Pro alterations do not affect HNMT protein localization

To check if Gly60Asp and Leu208Pro changes disturb HNMT protein localization, COS-7 cells were transfected with mutant and

wild-type (WT) HNMT constructs with GFP Tag. Visualization of the fusion protein was performed using a confocal laser scanning system (data not shown). HNMT-Gly60Asp HNMT-GFP, like WT, is localized to the cytoplasm of the cells, suggesting that there is no difference in cellular localization between mutant and WT form of the HNMT protein. For Leu208Pro, the transformed cells appear to show a punctuate distribution of HNMT-GFP, possibly indicative of the formation of protein aggregates; however, consistent results were not achieved (data not shown). We conclude that Gly60Asp does not impair the proper cellular localization of HNMT.

Patients' lymphoblasts are considerably more vulnerable to HA than the controls

We examined the vulnerability of patients' lymphoblasts (available for Family A, but not Family B) to HA when compared with lymphoblasts from unrelated healthy controls. Cells were first treated with different concentrations of HA and then using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) assay, the viability of cells was determined. Cells in both groups were found to rapidly undergo cell death upon culture *in vitro* after treating with high concentrations of HA (at 500 μM), which is in line with findings suggesting that HA induces neutrophil apoptosis at the sites of allergic inflammation (19). While control cells were able to survive at lower concentration of HA (125 μM), patients' lymphoblast cells were still not tolerant to the cytotoxic effects of HA, supporting our hypothesis that HNMT, as the defense mechanism against HA, is defective in the patients (P-value, 1×10^{-5}) (Fig. 4).

No significant difference in HA and N-*tele*-methylhistamine levels between patients carrying Gly60Asp variant and healthy unrelated controls

The primary goal of HA inactivation is its conversion to metabolites that will not activate HA receptors, and this is achieved either by methylation or by oxidation. HNMT catalyzes the transfer of a methyl group from SAM to the secondary amino group of the imidazole ring of HA, forming N-*tele*-methylhistamine. Therefore, if HNMT is defective, we would expect to detect high levels of HA and low levels of methylhistamine. We measured the HA and methylhistamine levels in patients' and controls' lymphoblast cell lines using an ELISA assay. However, we could not detect any significant difference between them (data not shown).

Gly60Asp alteration does not affect HNMT expression at the RNA and protein level

To examine whether the Gly60Asp alteration disrupts the expression of HNMT, quantitative RT-PCR and western blotting analysis were performed using patients' and controls' lymphoblast cell lines. We did not detect any significant difference in HNMT expression between patients and controls at both the RNA and protein level (data not shown).

Gly60Asp variant affects thermal stability of HNMT protein

Melting temperatures are the best descriptor of thermal stability. A difference in the melting temperature corresponds to a difference in energy between the pair of proteins. As indicated in Figure 5, glycine to aspartic acid replacement decreases the melting temperature of the HNMT protein from 62.3 ± 1.7 to $58.3 \pm 1.1°C$ for WT and the Gly60Asp variant, respectively, suggesting that amino acid change at position 60 will lead to reduced thermal stability of HNMT.

Gly60Asp variant disturbs the affinity of HNMT protein for binding to SAM

The dissociation constant (K_d) is commonly used to describe the affinity and binding property between a ligand and a protein, i.e. how tightly a ligand binds to a particular protein. Figure 6 presents the calorimetric titration of HNMT (WT) and the Gly60Asp variant with SAM by isothermal titration calorimetry. Figure 6A indicates the isothermal binding curve for the titration of 30 μM HNMT (WT) with 0.42 mM SAM. The isothermal binding curve (Fig. 6B) yields a K_d of 49.2 ± 2.0 μM. Figure 6C demonstrates a significant decrease in affinity of the Gly60Asp variant for SAM,
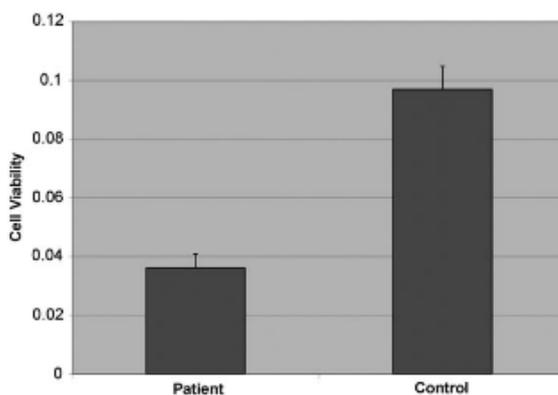
Figure 4. Cell viability using the MTT assay. Lymphoblast cells from 4 patients from Family A and 12 healthy unrelated controls were treated with 125 μM concentration of HA for 2 h. Reconstituted MTT in an amount equal to 10% of the culture medium volume was added to the cells. The plate was read at 570 nm. Values are mean ± SD for three independent experiments, and averaged across all the individuals per group.

Figure 5. Thermal stability of HNMT (WT) and HNMT (Gly60Asp) using Thermofluor®. The experiment was performed by mixing 5 μl of 1 mg/mL of HNMT (WT) or HNMT (Gly60Asp) with 5 μl 200-fold diluted SYPRO Orange and a buffer containing 50 mM Tris–HCl and 100 mM NaCl, pH 8.0, in a 98-well RT-PCR plate. Reaction mixtures were heated at 0.5°C/min from 20 to 95°C. A plot between d(fluorescence)/dT versus temperature of HNMT (WT) (red line) and HNMT (Gly60Asp) (blue line) is shown. A melting temperature of 62.3 ± 1.7 and 58.3 ± 1.1°C for WT and the Gly60Asp variant, respectively, was obtained.
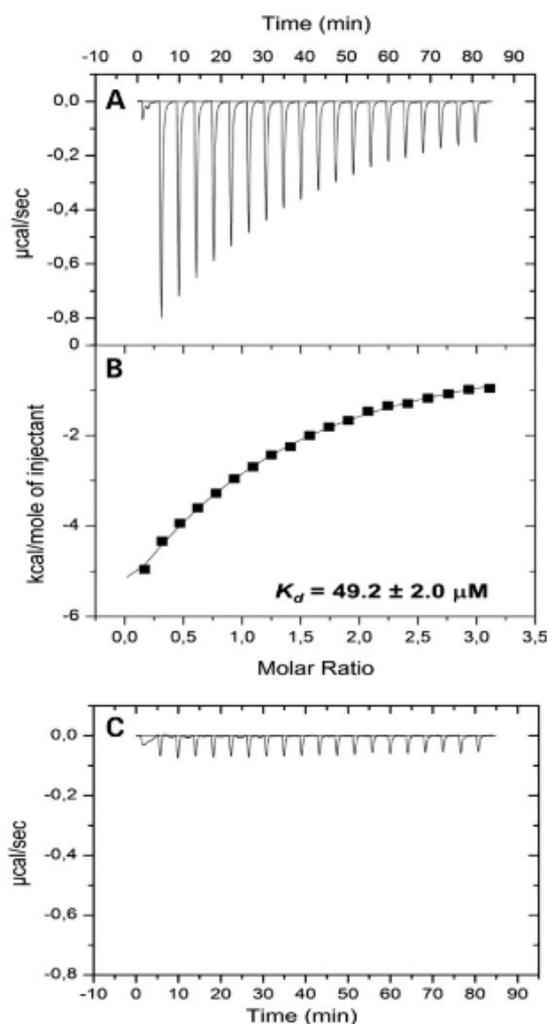
Figure 6. Isothermal calorimetric titration of HNMT (WT) and Gly60Asp variant with SAM and HA. The experiments for HNMT (WT) consisted of 20 consecutive injections of 15 µl of 0.42 mM SAM into 30 µM HMNT (WT) at 25°C. (**A**) The released heats of injection after baseline correction by subtracting the heat of the reference measurement. (**B**) Integrated data and data analysis using non-linear least square fitting in Origin 7. The dissociation constant of HNMT (WT) for SAM was 49.2 ± 2.0 µM. (**C**) Released heats of 20 consecutive injections of 15 µl of 2 mM SAM into 30 µM HNMT (Gly60Asp) at 25°C.

providing another line of evidence for the effects of the amino acid substitution at Gly60.

Gly60Asp variant significantly disrupts catalytic activity of HNMT protein

The isothermal titration calorimetric (ITC) method was used to assess the interaction between WT and Gly60Asp variant HNMT with SAM. As indicated in Table 3, the K_m (substrate concentration at which the reaction rate is half its maximum value) of WT HNMT lies in the normal range of most enzymes. However, this value was not measurable for Gly60Asp variant (Fig. 7). This can be explained by the failure of mutant HNMT

Table 3. Apparent K_m for HA from HNMT reaction

Enzyme	K_m (µM)	V_{max} (µM s^{-1})	k_{cat} (s^{-1})
HNMT (WT)	5.47 ± 1.41	8.21 ± 0.16 × 10^{-3}	8.21 ± 0.16 × 10^{-3}

WT and Gly60Asp HNMT constructs were used; however, only results for WT are given, as the catalytic activity of HNMT-Gly60Asp was too low to record using the VP-ITC system (MicroCal, GE Healthcare).

to bind with SAM due to the dramatic change of its substrate-binding pocket.

Leu208Pro destabilizes HNMT

Attempts were made to generate the Pro208 variant of HNMT; however, soluble protein was not detectable from the constructs, suggesting that misfolding occurs, leading to rapid degradation of the protein. Addition of N-terminal GST tag, or expression in the presence of chaperones, was not successful, and the variant protein was only detectable in the inclusion body fraction (data not shown). Interestingly, constructs for additional variants at Leu208 (Leu208Val, Leu208Phe, Leu208Arg, Leu208Thr, Leu208Asp and Leu208Asn) have also been generated in *Escherichia coli*, and only substitutions with the first two (apolar) amino acids, valine and phenylalanine, and to a much lower extent the polar, positively charged arginine (i.e. 5 mg/20 g cells instead of 20 mg/20 g cells), lead to soluble and active HNMT protein (P. Macheroux and C. Tongsook, personal communication). For this reason, we were unable to perform parallel experimentation to the Gly60Asp mutation for thermal stability, SAM-binding affinity and ITC measurements, and would also be a plausible explanation for the punctate cellular distribution of Leu208Pro-HNMT-GFP.

Discussion

HA in adult brain acts as a neurotransmitter in several biological processes. Moreover, it has been reported that, during the development of the rat brain in the fetus, HA concentration reaches its maximum level throughout the period where neuronal differentiation takes place in several brain regions, suggesting that HA, acting as a neurogenic factor, is an important modulator in the developing brain [20,21]. Even at very low concentrations, HA has strong pharmacological activity [22]. HA-induced apoptosis is mediated by caspase activation and PKC-γ signaling [23]. Therefore, synthesis, release and degradation of HA have to be carefully regulated to avoid undesirable reactions. HAergic neurons synthesize substantial quantities of HA and store it in special storage granula inside the cell [24]. Basal plasma HA concentrations of 0.3–1.0 ng/ml are considered normal and exceeding these HA levels give rise to concentration-dependent HA-mediated symptoms [25]. The level of HA in brain is slightly lower than that of other biogenic amines; however, its turnover is considerably faster and once released, it must be inactivated within a few minutes [19]. HNMT is the sole enzyme in the CNS, inactivating the neurotransmitter actions of HA [26].

We identified two homozygous missense mutations in the HNMT gene segregating with NS ID in an Iranian family of Turkish origin (Family A) and in a Kurdish family (Family B). For Family A, the condition had a relatively similar pattern during the childhood of the patients, starting at around 5 years of age and gradually worsening. Males were relatively mildly affected, but the condition in females was profound. This could be at least in part due to gender differences in neurotransmitter activity of
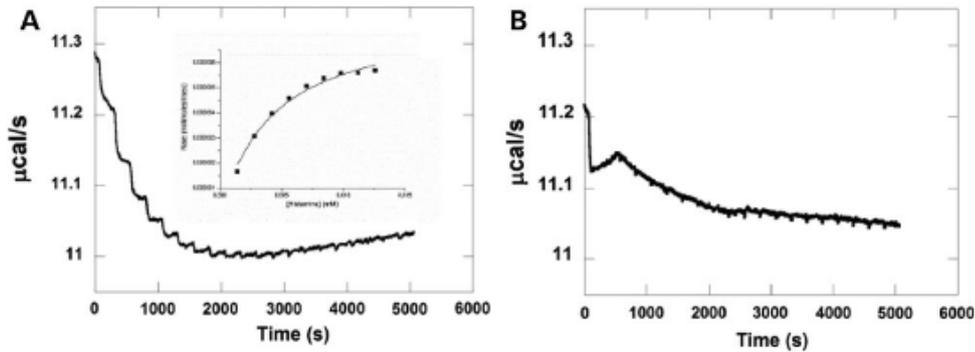
Figure 7. Catalytic activity of HNMT (WT) and HNMT (Gly60Asp) as determined by isothermal titration calorimetry. (**A**) Raw calorimetric data for the methylation of HA catalyzed by HNMT (WT) using SAM as a methyl group donor in 50 mM Tris–HCl buffer, pH 8.0, containing 100 mM NaCl at 25°C. The experiment was performed by injection of 2 mM HA (2 µl/injection; 20 injections). The cell contained 1 µM HNMT (WT) and 100 µM SAM. A Michaelis–Menten plot for methylation by HNMT (WT) is shown in the inset of A. (**B**) Raw calorimetric data for the reaction with the Gly60Asp variant (same conditions as for WT).

HA. It has been reported that in certain areas of the brain, neurotransmitter synthesis, content and metabolism are sexually differentiated and under the influence of sex steroids in development and adulthood (27,28). As mentioned above, once HA is released in the brain, it would typically be cleared from the intracellular space within a few minutes by HNMT. Therefore, malfunctioning HNMT would be expected to result in a high residual level of HA in the region. Using ELISA, we tested to see if there is any difference in the level of HA between patients' and controls' lymphoblast cell lines. However, no significant difference was detected. Since the alternative pathway of HA inactivation, i.e. DAO is active in the lymphoid lineage (but not in the CNS), therefore we measured HA level after blocking the DAO pathway and again no difference was detected (data not shown). As discussed above, during rat embryonic development, HA concentration reaches its maximum level at embryo days 14–16 when neuronal differentiation takes place and then steadily decreases until birth, yielding the concentration present in the adult organism. This may explain why differences in the level of HA among the patients' and controls' lymphoblast cell lines were not detected.

HNMT has a two-domain structure—the MTase domain and the S domain. The MTase domain primarily carries out cofactor binding and probably catalysis, while both domains contribute to HA binding (11,26). The Gly60Asp variant occurs within the conserved MTase region I (Ile56-Gly64), which is part of the SAM-binding pocket, highlighting the role of Gly60 in the interaction of HNMT with its cofactor. From our protein modeling analysis, we predict that the Gly60Asp variant does not bind SAM. Inspection of the active site also shows that the aspartate side chain would interfere with the ribose ring of SAM, rationalizing the lack of binding. Therefore, Gly60Asp does not affect overall protein stability, but compromises SAM in the active site of the enzyme. This interpretation is fully supported by the results obtained with the recombinant HNMT-Gly60Asp variant.

Leu208, as part of a very compact hydrophobic pocket, has several contacts with the residues in the same α-helix, the adjacent α-helix and the two β-strands. As a result, the leucine-to-proline substitution has a disruptive effect on the hydrophobic pocket, and it is likely that the substitution breaks the α-helix, and this in turn disrupts the hydrophobic packing in this area leading to exposure of hydrophobic residues, and thus favoring misfolding and aggregation of the protein.

Expression studies of the Gly60Asp variant revealed that the mutation does not disrupt the expression of HNMT at either the mRNA or protein level, suggesting that HNMT malfunction is not a result of reduced enzyme concentration. Moreover, observations from protein localization studies of HNMT proved that Gly60Asp and Leu208Pro variants do not disrupt the normal distribution of protein throughout the cytoplasm of cell. Results from thermal stability, binding affinity and catalytic activity investigations of HNMT WT and the Gly60Asp variant, and failed attempts to stabilize and solubilize Leu208Pro, support the deleterious effects of these substitutions on HNMT function in these two ID families.

A short isoform of HNMT has also been reported that includes the N-terminal 126 amino acids (Genbank ID: NM_001024075.1; NP_001019246.1 (Q8IU56)). The function of this isoform is unknown. The mutation identified in Family A (Gly60Asp) would potentially affect this as well as the long isoform, whereas the mutation identified in Family B (Leu208Pro) would only impact the longer isoform. Also, it is predicted that the Leu208Pro substitution is extremely destabilizing for the protein. The effects of different mutations on protein stability and function, as well as on different isoforms, could account for different clinical aspects and severities for the two families.

Decreased HA levels have been found in the brain of patients with Alzheimer disease (AD) (29), including reduction in the neuronal pool of hippocampus, hypothalamus and temporal cortex of AD patients (30), but, in contradicting reports, increased HA levels in cerebrospinal fluid and brain tissue were demonstrated (31,32). HA measurements *per se* may be, however, unreliable due to many confounding factors as, e.g. HA levels increase with *postmortem* interval (30), but may be valuable when accompanied by determination of activity or levels of HA-related enzymes.

While we can postulate about the effect of these mutations on levels of HA in the CNS, and possible detrimental effects on neurodevelopment, it cannot be ignored that a decrease in levels of the catabolite and product of HNMT activity, N-*tele*-methylhistamine may also play an important role in the disease etiology in these families. It has been suggested that there are HA/glutamate functional interactions in the brain (33), and it has already been shown that HA can potentiate NR2B-type N-methyl-D-aspartate (NMDA) receptors in hippocampal neurons, and that N-*tele*-methylhistamine also produces an equipotent enhancement of NMDA currents (34). In addition, N-*tele*-methylhistamine has been reported to be an agonist for β_3 γ-butyric acid (GABA) receptors (35). HA has recently been reported to increase neural differentiation to FOXP2 neurons in cultured cells (36). Mutation of the FOXP2 gene has been reported as the cause of a rare speech/

language disorder (MIM 602081) (37). Given the potential for N-*tele*-methylhistamine to act as an analog of HA, it would be interesting to see whether it would have a similar effect, in which case one could postulate that some of the deficits identified in the patients reported here could be the result of such a mechanism.

Conclusion

In the current study, we have shown that mutations resulting in HNMT LoF are associated with a NS form of ARID. We also confirmed that the p.Gly60Asp substitution results in complete loss of HNMT enzymatic activity, resulting in reduced HA inactivation and increased sensitivity to this amine. For p.Leu208Pro, although the mutation does not lie in the catalytic region of the protein, the rigidity introduced by the proline residue would most likely alter the helix conformation, destabilizing the protein and therefore affecting protein stability and the substrate-binding site.

Collectively, these findings indicate that HNMT plays an important role in human neurodevelopment. Our results indicate that HNMT should be included in genetic testing of individuals presenting with ID in consanguineous populations and, given estimates of a role for AR genes in 13–24% of ID in non-consanguineous populations (2), in outbred populations.

Materials and Methods

This study was approved by the Research Ethics Committees of the Qazvin University of Medical Sciences, Qazvin, Iran, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran, Max Planck Institute of Molecular Genetics, Berlin, Germany, the Centre for Addiction and Mental Health in Toronto, Canada, and Queen's University, Kingston, Canada. Informed written consent was obtained for all participating subjects.

Gene mapping

Family A (Iranian family)

Genomic DNA was extracted from peripheral blood leukocytes by standard methods. We used the Affymetrix GeneChip Mapping SNP 6.0 array (950K SNPs and 950 K CNV markers) to analyze DNA samples of all affected individuals, parents and one healthy sibling. Approximately 200 000 markers with good quality genotypes (based on their array hybridization confidence score) were selected for linkage analysis. Appropriate input files for the linkage analysis programs Merlin (38) and Allegro (39) were generated by ALOHOMORA software (40) with subsets of 300–500 markers in a sliding window mode based on mapping information from DeCode and Caucasian allele frequencies. Quality control checks such as gender check and verifying the relationships between individuals within the family were performed (38). Mendelian inconsistencies and unlikely genotypes were detected by the PedCheck (41) and Merlin (38) programs, respectively, and they were excluded from genotyping data prior to linkage analysis. Parametric linkage analysis was carried out based on an AR mode of inheritance, and assuming complete penetrance.

Family B (Kurdish family)

Genomic DNA was extracted from blood samples using standard protocols. Genotyping (SNP analysis) was performed using the Affymetrix 250k Genome-Wide Human SNP Array (Affymetrix, Santa Clara, CA, USA) for available individuals. We used ALOHOMORA software (40) for SNP array quality controls, as described previously (42). The program Merlin was applied for parametric multipoint

linkage analysis, consistent with an AR mode of inheritance, a disease allele frequency of 10^{-3} and complete penetrance.

Exon enrichment and high-throughput sequencing

Custom-made Agilent SureSelect DNA Capture Arrays (Agilent Technologies, Inc., Santa Clara, CA, USA) were used for the enrichment of exons from homozygous intervals, including, on average, 60 bp of flanking sequences on either side of the exon. Enriched exons were sequenced on an Illumina Genome Analyzer II, generating 76-bp single reads with 98.9% coverage.

Sequence alignment, variant calling, annotation and verification

Raw sequence reads were prescreened to remove low-quality reads, and then aligned to the human reference genome (*hg19*, GRCh37) with SOAP (version 2.20). Aligned and unaligned reads were used to call the SNVs and Indels, respectively. Variant lists were filtered against dbSNP137, whole genomes from 185 healthy individuals (1000 Genomes Project), and 200 exomes from Danish individuals, and 6500 exomes present in the Exome Variant Server (NHLBI GO Exome Sequencing Project, Seattle, WA, USA; http://evs.gs.washington.edu/EVS/; 30 August 2013; v.0.0.21). In addition, variants were compared with an in-house database containing more than 521 exomes from individuals of Middle Eastern origin. Variants were ranked as potential candidates as previously described (43), using an improved version of Medical Re-sequencing Analysis Pipeline (MERAP) (44). The OMIM catalog (http://www.ncbi.nlm.nih.gov/omim) and the Human Gene Mutation Database (HGMD, http://www.hgmd.org/) were used as a filter to identify all previously described pathogenic changes. Sanger sequencing was used to confirm the co-segregation of the final candidate variants in the family.

Targeted exome sequencing for HNMT

Ninety-nine genes were included in the custom design: 64 for known or suspected NS ARID genes, including HNMT; 4 for known or suspected NS AR autism genes; 7 for known or suspected NS X-linked ID or ASD genes; 7 genes for known syndromic ARID, also reported in ASD cohorts and 17 known or suspected NS autosomal dominant ID or ASD genes. N = 1000 ASD and N = 991 ID-unrelated individuals from an outbred population (Canada) were included. Samples were measured using the RNaseP assay by quantitative PCR, using ViiA™7 Real-Time PCR (Life Technologies, Carlsbad, CA, USA), and pooled at equimolar concentrations in pools of 20. Seven pools were barcoded and run using the custom Ampliseq (Life Technologies) primer pools on a single Proton P1v2 chip. Samples were analyzed using the Ion Torrent software (Life Technologies), and Bam files generated and runs visualized using the Integrated Genome Viewer (Broad Institute: http://www.broadinstitute.org/igv/), and for the purpose of this study, focusing just on HNMT coding regions.

Expression studies of HNMT

Quantitative reverse transcriptase-PCR

Primers were designed to amplify the coding sequence of the HNMT gene. Coding sequence primers for β-Actin (ACTB) and HPRT were used as an internal reference for all the runs. Reverse transcriptase (RT)-PCR was performed in quadruplicates using a 384-well optical plate, with a final reaction volume of 16 μl. Universal SYBR Green PCR conditions were used, consisting of 95°C

for 2 min and 30 s and 40 cycles at 95°C for 4 s and 60°C for 20 s. Each reaction contained 2 µl of cDNA in a 16-µl volume run in 384-well optical plates on a ViiA™7 (Life Technologies). For each gene analyzed, all samples were run on one plate, to avoid interplate variability, and in quadruplicate. Furthermore, each plate contained H_2O, RT-minus and RNA-minus as negative controls. The C_t for all reactions was calculated automatically by the ViiA™7 (Life Technologies) software. Gene expression analysis was calculated using the comparative C_t method.

Western blot analysis

Protein samples from patients and controls were separated by SDS page (4–20% Mini-PROTEAN TGX polyacrylamide gel, Bio-Rad Laboratories, Hercules, CA, USA) and electrically transferred to a nitrocellulose membrane (BioTrace NT nitrocellulose membrane, PALL Life Sciences, Ann Arbor, MI, USA). The membrane was blocked for 1 h using 5% skimmed milk in TBS-Tween, incubated overnight with the primary rabbit anti-HNMT (1 : 1000, Bethyl Laboratories, Inc., Montgomery, TX, USA) and then incubated with the secondary antibody horseradish peroxidase-conjugated donkey anti-rabbit IgG (1 : 1000, GE Healthcare UK Limited, Little Chalfont, UK). The immunoblots were developed by an enhanced chemiluminescence western blot detection system (GE Healthcare).

Heterologous expression of WT HNMT, p.Gly60Asp and p.Leu208Pro variants in E. coli

The open reading frames encoding the target proteins (WT HNMT; p.Gly60Asp and p.Leu208Pro variants) were amplified by PCR from the pcDNATM3.3-TOPO TA vector using primers 5′-TCTCATATGATGGCATCTTCCATGAGGAGC-3′ (sense strand) and 5′-TCAGCGGCCGCTTAATGATGATGATGATGATGTGCCTCA ATCACTATGAAACTCAGA G-3′ (anti-sense strand) using Phusion High-Fidelity DNA polymerase (Thermo Scientific, Waltham, MA, USA). The amplified genes were then cloned into the pET-21a vector using the *Nde*I and *Not*I restriction sites for expression as C-terminally hexa-histidine tagged proteins. Similarly, glutathione S-transferase (GST) fusions were generated by cloning the open reading frames into pGEX-6p2 using the *Bam*HI and *Not*I restriction sites. Induction of expression of HNMT WT, and Gly60Asp, and Leu208Pro variants with C-terminal His6 or N-terminal GST tag was carried out at an OD_{600} ~1, by addition of 0.1 mM IPTG (final concentration). The culture was maintained at 18°C for 14 h prior to harvest, resulting in a yield of 5 g of cell paste per liter of cell culture.

Purification of HNMT WT and the Gly60Asp variant

Frozen cell paste (~20 g) was thawed and resuspended in lysis buffer (50 mM Tris–HCl, pH 8.0 containing 100 mM NaCl, 10 mM imidazole, 1 mM DTT and 100 µM PMSF). Cells were disrupted by ultrasonication with 50% amplitude for 15 min. After ultrasonication, the suspension was centrifuged at 18 000 rpm at 4°C for 1 h and the pellet was discarded. The clear crude extract was filtered (0.22 µm) prior to loading onto a 5-ml Ni-Sepharose™ High-Performance HisTrap™ HP column equilibrated with lysis buffer. The column was washed with 100 ml of 50 mM Tris–HCl buffer, pH 8.0, containing 100 mM NaCl and 20 mM imidazole. Elution of HNMT was initiated using 50 mM Tris–HCl buffer, pH 8.0, containing 100 mM NaCl and 300 mM imidazole. Fractions containing HNMT were pooled, concentrated using a Centricon YM10 and loaded onto a Superdex 200 prep grade column previously equilibrated with 50 mM

Tris–HCl buffer, pH 8.0, containing 100 mM NaCl. Fractions containing target protein were pooled and concentrated as before. The purified protein was stored at −80°C.

Determination of melting temperature of HNMT (WT) and the Gly60Asp variant

To investigate the thermal stability of HNMT (WT) and the Gly60Asp variant, the melting temperature was determined using a fluorescence-based thermal shift assay. Experiments were performed in a real-time PCR detection system (Bio-Rad) with a 96-well plate in a FRET scan mode. In the experiments, 5 µl of 1 mg/ml HNMT (WT) or the Gly60Asp variant was mixed with 5 µl of 200-fold diluted SYPRO Orange solution and 15 µl of a buffer containing 50 mM Tris–HCl, pH 8.0 and 100 mM NaCl to a final volume of 25 µl in a 96-well plate. Thermal unfolding of the protein was monitored using a temperature gradient from 20 to 95°C, measuring fluorescence emission at 0.5°C increments with a 60-s hold for signal stabilization. The melting temperature of the protein was derived from the peak of the derivatives of the experimental data.

Determination of dissociation constant (K_d) of HNMT (WT) and Gly60Asp variant

Dissociation constants for binding of SAM and HA to HNMT (WT) and Gly60Asp variant were determined using the VP-ITC system (MicroCal, GE Healthcare). Experiments were performed at 25°C in 50 mM Tris–HCl buffer, pH 8.0, containing 100 mM NaCl. All solutions were degassed before measurements. Titration experiments for SAM and HA were performed by using 20 injections of 15 µl of 0.42 mM SAM or 4 mM HA (duration time 29.9 s and spacing time 250 s) into the cell containing 30 µM HNMT (WT). In the case of the Gly60Asp variant, 30 µM was titrated with SAM (2 mM, 20 injection, 15 µl/injection, duration time 29.9 s and spacing time 250 s) or HA (200 µM, 20 injection, 15 µl/injection, duration time 29.9 s and spacing time 250 s). To determine K_d values of HNMT (WT) and Gly60Asp variant, one set of sites fitting with Origin version 7.0 (MicroCal) for ITC data analysis was used.

Determination of catalytic activity of HNMT WT and Gly60Asp variant

Kinetic parameters for HNMT (WT) and the Gly60Asp variant were determined using the VP-ITC system (MicroCal) at 25°C in 50 mM Tris–HCl buffer, pH 8.0, containing 100 mM NaCl in multi-injection mode. All solutions were degassed before the measurements. The HNMT analysis consisted of 20 injections of 2 µl (duration time 4 s and spacing time 250 s) of 2 mM HA into the cell containing 1 µM of HNMT (WT) or Gly60Asp variant and 100 µM SAM. The experimental data were fitted with Origin version 7.0 (MicroCal) for ITC data analysis to obtain kinetic parameters.

In vitro toxicology assay

The cytotoxic effects of HA on patients' lymphoblast cells carrying the p.Gly60Asp variation and lymphoblasts from unrelated healthy controls were determined using the MTT assay. This method measures the metabolic reduction of MTT to a colored water-insoluble formazan salt by mitochondrial dehydrogenases. Lymphoblast cells were available for all four affected individuals from Family A. For comparison, lymphoblast cell lines from 12 unrelated unaffected individuals were selected,

and matched for age, gender and number of passages for the cells. Cells for each line were seeded at 10 000 cells per well in 96-well plates (in triplicates for each line) and cultured in serum-free DMEM for 2 h. Cells were subsequently incubated for 2 h after which reconstituted MTT in an amount equal to 10% of the culture medium volume was added. After incubation of the plates at 37°C for 1 h, the cells were then washed with phosphate-buffered saline and the formazan salts dissolved in 200 ml of dimethyl sulfoxide with gentle shaking for 10 min at room temperature. The plates were read at 570 nm using a Tecan Spectra Fluor plate reader, and data were averaged across replicates and across both the affected and unaffected group.

HNMT protein localization

We generated two different sets of genetic constructs for HNMT protein localization. RNA was extracted from one of the patients' (Family A) fibroblast cells, and then HNMT full-length cDNA (CCDS2181.1) was PCR amplified and cloned into the pcDNA3.1/CT-GFP-TOPO expression vector (Invitrogen, Carlsbad, CA, USA). Using the Q5® Site-Directed Mutagenesis Kit (New England Bio-Labs, Ipswich, MA, USA), an A>G change was introduced to the HNMT mutant plasmid (c.179G>A) in order to generate the HNMT WT construct. Subsequently, a T>C change was introduced to the WT construct to generate a mutant HNMT construct for the c.623T>C, (p.Leu208Pro) substitution. Thus, we generated three constructs that are identical except for the mutation changes, in order that we would be able to exclude any potential effects of intervening SNPs, as it has been well documented that genetic variation among individuals can result in as much as 5-fold differences in HNMT activity [20]. Orientation of the inserts and correct sequences were finally confirmed by Sanger sequencing. To examine the consequences of mutations on cellular localization of the HNMT protein, we transiently transfected COS-7 cells with 2 µg of purified constructs (p.Gly60Asp, p.Leu208Pro and WT) with Polyfect (Qiagen, Germantown, MD, USA). Twenty-four h after transfection, we visualized the HNMT-GFP fusion protein in transfected cells using a Zeiss Axioplan 2 imaging microscope (Carl Zeiss AG, Oberkochen, Germany), equipped with the LSM510 array confocal laser scanning system, and the Zeiss LSM510 version 3.2 SP2 software package.

Measurement of HA/methylhistamine levels in patients' and controls' lymphoblast cells' by ELISA

Cell-free supernatants were used to measure the HA and *N-tele*-methylhistamine levels in patients' lymphoblast cell lines and in a group of unrelated healthy controls using an ELISA assay, performed according to the manufacturer's instructions (GenWay Biotech, Inc., San Diego, CA, USA) through a Fluoroskan Ascent microplate reader (Thermo Scientific) at 450 nm. Standard reagents were employed to draw a calibration curve. The samples OD values and the calibration curve were used to calculate the HA and *N-tele*-methylhistamine concentrations in cases and controls.

Bioinformatic analysis of the HNMT protein

Protein structure modeling was conducted with the Phyre2 server with the 3D structure image realized by JMOL. In addition, mutant protein structures were modeled with the FoldX package Version 3.0 Beta3 [18]. Models were built based on the WT crystal structure [PDB 2AOT [26]; available in the Protein Data Bank PDB] as the template using the BuildModel function after energy

minimization of the WT structure using the RepairPDB function. Figures were generated with PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.1, Schrodinger, LLC).

Supplementary Material

Supplementary Material is available at *HMG* online.

References

1. Maulik, P.K., Mascarenhas, M.N., Mathers, C.D., Dua, T. and Saxena, S. (2011) Prevalence of intellectual disability: a meta-analysis of population based studies. *Res. Dev. Disabil.*, **32**, 419–436.
2. Musante, L. and Ropers, H.H. (2014) Genetics of recessive cognitive disorders. *Trends Genet.*, **30**, 32–39.
3. Loiselle, J. and Wollin, W. (1993) Mucosal histamine elimination and its effect on acid secretion in rabbit gastric mucosa. *Gastroenterology*, **104**, 1013–1020.
4. Schwartz, J.C., Arrang, M.M., Garbarg, M., Pollaed, H. and Ruat, M. (1991) Histaminergic transmission in the mammalian brain. *Physiol. Rev.*, **71**, 1–51.
5. Maslinski, C. (1975) Histamine and its metabolism in mammals. Part II. Catabolism of histamine and histamine liberation. *Agents Actions*, **5**, 183–225.
6. Hough, L.B. and Green, J.P. (1984) Histamine and its receptors in the nervous system. In Lajtha, A. (ed.), *Handbook of Neurochemistry, Vol. 6: Receptors in the Nervous System*, Springer, USA, pp. 145–211.
7. Lindahl, K.M. (1960) The histamine methylating enzyme system in liver. *Acta Physiol. Scand.*, **49**, 114–138.
8. Brown, D.D., Tomchick, R. and Axelrod, J. (1959) The distribution and properties of a histamine-methylating enzyme. *J. Biol. Chem.*, **234**, 2948–2950.
9. Galinsky, R.E., Bernstein, J.A., Liggett, S.B. and Weinshilboum, R.M. (2000) Histamine N-methyl transferase pharmacogenetics: association of a common functional polymorphism with asthma. *Pharmacogenetics*, **10**, 261–266.
10. González-Pérez, A. and López-Bigas, N. (2011) Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am. J. Hum. Genet.*, **88**, 440–449.
11. Pang, Y.P., Zheng, X.E. and Weinshilboum, R.M. (2001) Theoretical 3D model of histamine N-methyltransferase: insights into the effects of a genetic polymorphism on enzymatic

Supplementary information

for

Mutations in the histamine *N*-methyltransferase gene, *HNMT*, are associated with nonsyndromic autosomal recessive intellectual disability

Abolfazl Heidari[1,2], Chanakan Tongsook[4,†], Reza Najafipour[2,†], Luciana Musante[5], Nasim Vasli[1], Masoud Garshasbi[5,6], Hao Hu[5], Kirti Mittal[1], Amy J. M. McNaughton[7], Kumudesh Sritharan[1], Melissa Hudson[7], Henning Stehr[9], Saeid Talebi[10], Mohammad Moradi[2], Hossein Darvish[11], Muhammad Arshad Rafiq[1], Hossein Mozhdehipanah[3], Ali Rashidinejad[12], Shahram Samiei[13], Mohsen Ghadami[14], Christian Windpassinger[15], Gabriele Gillessen-Kaesbach[16], Andreas Tzschach[5,‡], Iltaf Ahmed[1,17], Anna Mikhailov[1], D. James Stavropoulos[18], Melissa T. Carter[19], Soraya Keshavarz[2], Muhammad Ayub[8], Hossein Najmabadi[20,21], Xudong Liu[7], Hans Hilger Ropers[5], Peter Macheroux[4] and John B. Vincent[1,22,23,]*

†These authors contributed equally

*To whom correspondence should be addressed at: The Campbell Family Mental Health Research Institute, R-32, Centre for Addiction and Mental Health, 250 College Street, Toronto, ON, Canada M5T 1R8. Tel: +1 416 535 8501; Fax: +1 416 979 4666; Email: john.vincent@camh.ca

[1]Molecular Neuropsychiatry and Development (MiND) Lab, The Campbell Family Mental Health Research, Institute, Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada M5T 1R8, [2]Cellular and Molecular Research Center, [3]Department of Neurology, Bou Ali Sina Hospital, Qazvin University of Medical Sciences, Qazvin 34197/59811, Iran, [4]Institute of Biochemistry, Graz University of Technology, Graz 8010, Austria, [5]Max Planck Institute of Molecular Genetics, Berlin D-14195, Germany, [6]Department of Medical Genetics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran 14117-13116, Iran, [7]Ongwanada Genomics Lab, [8]Division of Developmental Disabilities, Department of Psychiatry, Queen's University, Kingston, ON, Canada K7L7X3, [9]Department of Medicine, Stanford University, Stanford, CA 94305-5101, USA, [10]Department of Medical Genetics, Medical University of Tehran, Tehran 14167-53955, Iran, [11]Department of Medical Genetics, Shahid Beheshti University of Medical Sciences, Tehran 4739, Iran, [12]Maternal, Fetal and

Neonatal Research Center, Tehran University of Medical Sciences, Tehran 1419733141, Iran, [13]Blood Transfusion Research Center, Tehran 1449613111, Iran, [14]Department of Medical Genetics, Tehran University of Medical Sciences, Tehran 1417613151, Iran, [15]Institute of Human Genetics, Medical University of Graz, Graz 8010, Austria, [16]Institut für Humangenetik, Universität zu Lübeck, Lübeck 23562, Germany, [17]Atta-ur Rehman School of Applied Biosciences, National University of Sciences and Technology, H-12, Islamabad, Pakistan, [18]Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, ON, Canada, [19]Division of Clinical and Metabolic Genetics, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada, [20]Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran 19857, Iran, [21]Kariminejad-Najmabadi Pathology and Genetics Center, Tehran 14667, Iran, [22]Department of Psychiatry, University of Toronto, Toronto, ON, Canada M5T 1R8 and [23]Institute of Medical Science, University of Toronto, Toronto, ON, Canada M5S 1A8

Supplemental information

Suppl Fig. S1: Point mutations and copy number variants (CNV) at the HNMT locus:
Here we indicate in an ideogrammatic representation of the HNMT locus (from the UCSC Genome Browser, hg19) the two homozygous missense mutations segregating with NS-ARID identified in this study (blue arrows), as well as heterozygous canonical splice acceptor mutations in two Canadian individuals diagnosed with ASD and ID and a Swedish schizophrenia individual with a heterozygous canonical splice donor mutation (Schizophrenia Exome Sequencing Genebook: //atgu.mgh.harvard.edu/~spurcell/genebook/genebook.cgi) (red arrows). This data was retrieved by searching the DECIPHER and ISCA databases, and data from published CNV studies for neuropsychiatric and development disorders [45-47]. Control comparison CNV data was checked through the Database for Genomic Variants (http://dgv.tcag.ca/dgv/app/home), however only a single CNV was identified within a control cohort, that span exons and potentially disrupts *HNMT*. SNV controls can be accessed through the ExAC site (http://exac.broadinstitute.org).
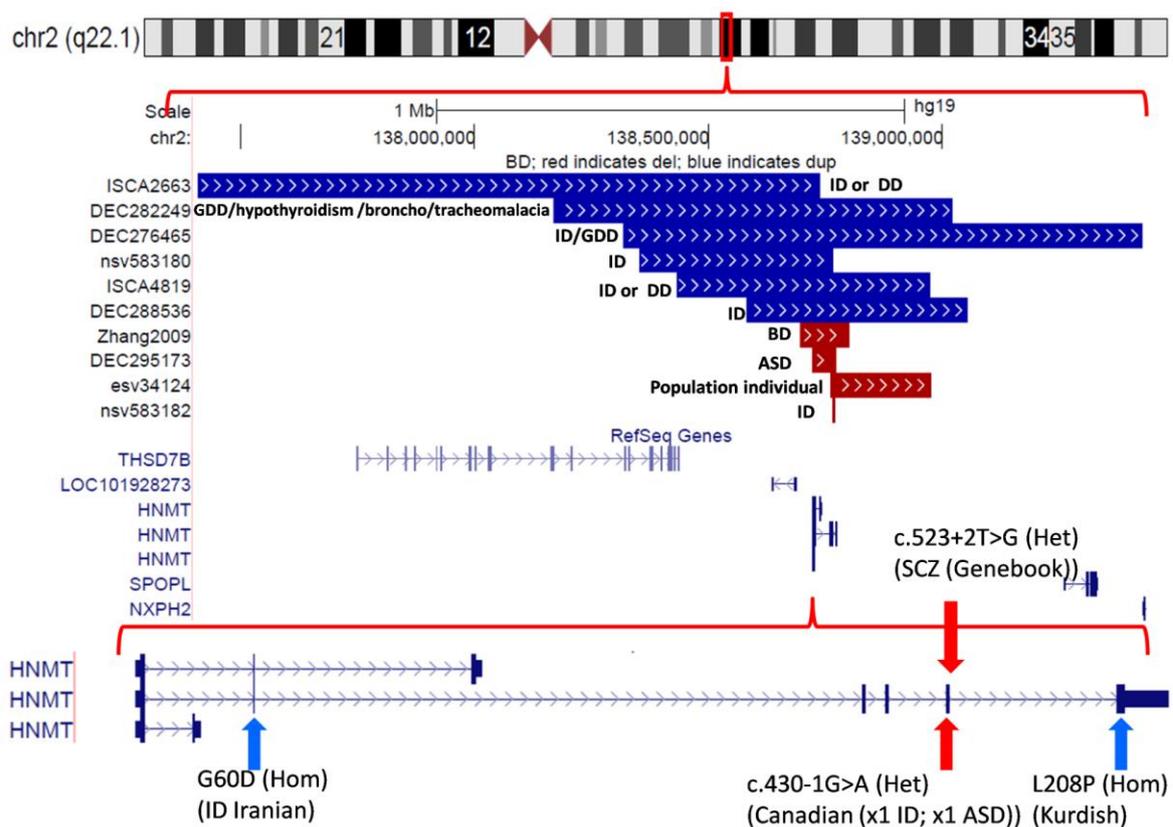
Additional References

45. Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader P, McCracken E, Niyazov D, Leppig K, Thiese H, Hummel M, Alexander N, Gorski J, Kussmann J, Shashi V, Johnson K, Rehder C, Ballif BC, Shaffer LG, Eichler EE. (2011) A copy number variation morbidity map of developmental delay. *Nat. Genet.* **43**, 838-846.

46. Zhang, D., Cheng, L., Qian, Y., Alliey-Rodriguez, N., Kelsoe, J.R., Greenwood, T., Nievergelt, C., Barrett, T.B., McKinney, R., Schork, N., Smith, E.N., Bloss, C., Nurnberger, J., Edenberg, H.J., Foroud, T., Sheftner, W., Lawson, W.B., Nwulia, E.A., Hipolito, M., Coryell, W., Rice, J., Byerley, W., McMahon, F., Schulze, T.G., Berrettini, W., Potash, J.B., Belmonte, P.L., Zandi, P.P., McInnis, M.G., Zöllner, S., Craig, D., Szelinger, S., Koller, D., Christian, S.L., Liu, C., Gershon, E.S. (2009) Singleton deletions throughout the genome increase risk of bipolar disorder. *Mol. Psychiatry* **14,** 376-380.

47. Gusev, A., Lowe, J.K., Stoffel, M., Daly, M.J., Altshuler, D., Breslow, J.L., Friedman, J.M., Pe'er, I. (2009) Whole population, genome-wide mapping of hidden relatedness. *Genome Res.* **19**, 318-326.

2.2 Leucine 208 in human histamine *N*-methyltransferase emerges as a hotspot for protein stability rationalizing the role of the L208P variant in intellectual disability

Leucine 208 in human histamine *N*-methyltransferase emerges as a hotspot for protein stability rationalizing the role of the L208P variant in intellectual disability

Chanakan Tongsook[\$], Johannes Niederhauser[\$], Elena Kronegger[\$], Grit Straganz[\$], and Peter Macheroux[\$]

From the [\$]Institute of Biochemistry, Graz University of Technology, Graz, Austria

Running title: Histamine *N*-methyltransferase

To whom correspondence should be addressed: Prof. Dr. Peter Macheroux, Graz University of Technology, Institute of Biochemistry, Petersgasse 12/II, A-8010 Graz, Telephone: +43-(0)316-873 6450; FAX: +43 (0)316-873 6952; E-mail: peter.macheroux@tugraz.at

Abbreviations

CNS, central nervous system; DAO, diamine oxidase; HNMT, histamine *N*-methyltransferase; MTase, methyltransferase; SAM, *S*-adenosylmethionine; SAH, *S*-adenosylhomocysteine;

Abstract

The degradation of histamine catalyzed by the SAM-dependent histamine *N*-methyltransferase (HNMT) is critically important for the maintenance of neurological processes. Recently, two mutations in the encoding human gene were reported to give rise to dysfunctional protein variants (G60D and L208P) leading to intellectual disability. In the present study, we have expressed eight L208 variants with either apolar (L208F and L208V), polar (L208N and L208T) or charged (L208D, L208H, L208K and L208R) amino acids to define the impact of side chain variations on protein structure and function. We found that the variants L208N, L208T, L208D and L208H were severely compromised in their stability. The other four variants were obtained in lower amounts in the order wild-type HNMT>L208F=L208V>L208K=L208R. Biochemical characterization of the two variants L208F and L208V exhibited similar Michaelis-Menten parameters for SAM and histamine while the enzymatic activity was reduced to 21% and 48%, respectively. A substantial loss of enzymatic activity and binding affinity for histamine was seen for the L208K and L208R variants. Similarly the thermal stability for the latter variants was reduced by 8 and 13 °C, respectively. These findings demonstrate that position 208 is extremely sensitive to side chain variations and even conservative replacements affect enzymatic function. Molecular dynamics simulations showed that amino acid replacements in position 208 perturb the helical character and disrupt interactions with the adjacent β-strand, which is involved in the binding and correct positioning of histamine. This finding rationalizes the gradual loss of enzymatic activity observed in the L208 variants.

Introduction

Histamine, [2-(1H-imidazol-4-yl)ethanamine], is a biogenic amine that plays important roles in several physiological and pathological processes including the regulation of inflammatory and immune responses, gastric acid secretion and bronchial asthma [1-4]. Histamine also functions as a neurotransmitter in the central nervous system (CNS) where it contributes to the regulation of many processes such as the sleep/wake cycle, the circadian rhythm, thermoregulation, stress and fluid homeostasis [5-7]. Furthermore, histamine acts as a key neuromodulator in the developing nervous system. In the embryonic state of the developing rat brain the maximum level of histamine concentration

coincides with the period where neuronal differentiation takes place in several brain regions [8-10]. Histidine decarboxylase, a pyridoxal 5'-phosphate (PLP)-dependent enzyme, catalyzes the decarboxylation of the amino acid L-histidine to generate histamine, which is stored in granules in airway mast cells, basophils and enterochromaffine cells [5,11]. Release of histamine from cells is triggered upon the cross-linking surface-bound immunoglobulin E (IgE) by allergens [12]. The actions of released histamine are mediated through the four histamine receptors (H1R-H4R), which belong to the rhodopsin-like family of G protein-coupled receptors [13-15]. H1R, H2R and H3R are expressed in abundance in the brain whereas H4R is expressed mainly in peripheral tissues (*e. g.* the gut and connective tissue) [14,16]. Released histamine is then rapidly inactivated and disappears from the bloodstream within minutes [5].

In mammals, the action of histamine is terminated by either one of two mechanisms that involve the methylation of the imidazole ring or the oxidative deamination of the primary amino group. The former reaction is catalyzed by histamine *N*-methyltransferase (HNMT) and the latter reaction by diamine oxidase (DAO) [17]. Products generated from these two enzymatic reactions are virtually inactive at histamine receptors and are further converted for transport and secretion [18,19]. HNMT is the primary metabolizing enzyme for released histamine in bronchial epithelial cells, endothelial cells of human airways and the human stomach [20,21]. In addition, HNMT is also the only enzyme responsible for the termination of histamine action in the mammalian brain, since the other histamine-metabolizing enzyme, DAO, is not expressed in the CNS [5,22]. DAO, an enzyme of the class of copper-containing amine oxidases, is the main histamine-degrading enzyme in peripheral tissue including termination of exogenous histamine from ingestion (food and alcohol) [14].

HNMT (EC 2.1.1.8) inactivates histamine by transferring a methyl group from *S*-adenosylmethionine (SAM), to the $N_{\varepsilon 2}$ atom of the imidazole ring, generating *N*-methylhistamine and *S*-adenosylhomocysteine (SAH) [23,24]. Enzyme activity and mRNA expression were found in most human tissues with high expression levels in kidney and liver, with substantial expression in spleen, prostate, ovary, intestine and the spinal cord, along with lower-level expression in heart, brain, placenta, lung, stomach and the thyroid gland [18,25]. The *hnmt* gene is approximately 34 kb in length, located on chromosome 2q22.1, and consists of 6 exons [26-28]. Expression and characterization of human HNMT showed that cytoplasmic HNMT has 292 amino acid residues with a total molecular mass

of 33 kDa [18]. Several three-dimensional structures of HNMT with bound SAH and histamine as well as various HNMT inhibitors revealed that HNMT is a monomeric protein and consists of two domains. The large domain adopts a canonical methyltransferase (MTase) fold with a SAM binding pocket, whereas the small domain contributes part of the histamine binding pocket and may be involved in protein-protein interactions. The bound histamine is buried in a hydrophobic pocket (surrounded by 14 aromatic amino acid residues) at the interface of the two domains. Amino acid residues that interact with SAH and histamine are highly conserved in vertebrates [29,30].

Since disturbances of histamine metabolism have been related to several diseases, *i.e.* asthma, bronchial hyper-responsiveness (BHR), neurological disorders, the lack of expression or decrease in HNMT activity have been proposed as a causative factor for these diseases. A common single nucleotide polymorphism (SNP), C314T transition (prevalence 0.1), resulting in the replacement of Thr105 by isoleucine was associated with decreased enzyme activity, immunoreactive protein, and thermal stability [25]. Steady-state kinetics revealed that the apparent Michaelis constants (K_M) increased 1.8 and 1.3 fold for SAM and histamine, respectively, and showed a reduction of the specific activity of about 16% [29]. Moreover, from multiple molecular dynamic (MD) simulations, isoleucine 105 apparently more strongly interacts with neighboring residues leading to a disordering of several key residues responsible for SAM binding and lower hydrophobicity of the substrate-binding site [31]. Thus, the polymorphic T105I variant has been proposed to associate with several diseases including asthma, allergic rhinitis, essential tremor and Parkinson's disease [32-35].

Intellectual disability (ID) is a common neurodevelopmental disorder characterized by an intelligence quotient of 70 or below with deficits in adaptive, daily living skills and intellectual functioning manifesting before the age of 18 [36]. With a prevalence of about 1% of children worldwide ID is one of the major socioeconomic problems [37]. Recently, two other homozygous missense mutations in the *hnmt* gene were identified in patients affected with nonsyndromic autosomal recessive intellectual disability (ND-ARID) from two unrelated consanguineous families of Turkish and Kurdish ancestry [38]. The first mutation (*hnmt* c.179G>A) from the Turkish family resulting in a change of Gly60 to aspartate (G60D) occurs in the conserved MTase region I, which is part of the SAM-binding pocket. This amino acid exchange completely disrupts binding of SAM and as a consequence compromises catalytic activity, resulting in reduced histamine inactivation.

The other mutation (*hnmt* c.632T>C) found in the Kurdish family results in the replacement of Leu208 with proline (L208P). This invariant residue is located in helix E of the MTase domain and forms several hydrophobic interactions with neighboring residues. Interestingly, L208 is located in a considerable distance from the active site (ca. 18 Å from the substrate binding pocket) (Fig. 1, panels A-C) suggesting that the adverse effect of the amino acid replacement is caused by perturbations of the protein structure leading to compromised enzyme function. In fact, previous attempts to express the L208P variant were not successful [38]. In order to better understand the role and importance of L208 in maintaining protein structure, stability and enzyme function we have engaged in a more detailed analysis by generating a series of variants with either apolar (L208V and L208F), polar (L208T and L208N) or charged (L208R, L208K, L208H, and L208D) amino acid replacements. In the present study we report the biochemical properties of the variants that were successfully expressed. In addition, we have investigated the role of L208 and the impact of variations in this position by a set of molecular dynamic simulations. Taken together these methods have provided detailed insights into the functional and structural role of L208 and provide a rationale for the disease causing effect of mutations affecting this amino acid position.

Materials and Methods

Reagents

All chemicals and reagents were of the highest purity commercially available from Sigma-Aldrich, Merck, and Carl Roth GmbH & Co. KG. [Methyl-^{3}H]-SAM and Ultima Gold liquid scintillation cocktail were from PerkinElmer. Ni-Sepharose (prepacked HisTrap^{TM} HP) and Superdex 200 prep grade (HiLoad 16/600 Superdex 200 pg) columns were from GE Healthcare (Little Chalfont, UK). HNMT (wild-type) and L208 variants were overexpressed in *Escherichia coli* and purified as previously described [38]. The concentrations of the following compounds were determined spectrophotometrically using these extinction coefficients: SAM and SAH, $\varepsilon_{260} = 15.4$ mM^{-1} cm^{-1} [39]; and HNMT (wild-type and L208 variants, based on amino acid sequence), $\varepsilon_{280} = 42.8$ mM^{-1} cm^{-1}.

Site-directed mutagenesis

Site-directed mutagenesis at position L208 was performed using QuickChange$^{®}$ Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA). The mutagenic forward

primers and the reverse primers (Supplemental Table S1) were from VBC-Biotech Service GmbH (Vienna, Austria). The plasmid pET-21a harboring *hnmt* (wild-type) gene was used as the template for mutagenic PCR. PCR reactions (18 cycles of the standardized protocol) were performed using a PCR instrument with GeneAmp PCR system, Applied Biosystem® (Thermo Fisher Scientific, Waltham, MA, USA). The DNA sequences of all variant plasmids were analyzed using the forward primers 5'-TAATACGACTCACTATAGGG-3' (T7-Promoter) and the reverse primer 5'-GCTAGTTATTGCTCAGCGG-3' (T7-terminator) by LGC Genomics (Teddington, UK).

Protein expression and purification of HNMT (wild-type) and L208 variants in Escherichia coli

Protein expression and purification of HNMT (wild-type) and L208 variants (L208V, L208F, L208T, L208N, L208R, L208K, L208H, and L208D) with C-terminal His6 tag in *Escherichia coli* was performed. *E. coli* BL21 (DE3) was transformed with the plasmid pET-21a harboring *hnmt* (wild-type or mutated L208) genes. A colony of the transformants was inoculated in lysogen broth (LB) containing 100 μg/mL of ampicillin to prepare a preculture which was aerobically incubated at 37 °C and 150 rpm for 16 hours. The pre cultured cells (1% of cell inoculation) were transferred to 10 L of LB containing 100 μg/mL of ampicillin. Cells were grown at 37 °C and 150 rpm until the absorption at 600 nm of the cell culture reached ~1.0. Synthesis of the recombinant protein was induced by adding 0.1 mM of isopropyl β-D-thiogalactopyranoside (IPTG). The cell culture was further grown at 18 °C and 150 rpm for 16 hours. Cells were harvested using centrifugation at 5000 *g* for 10 min. Cell paste was kept at -70 °C until used.

To purify HNMT (wild-type) and L208 variants, frozen cell paste of the wild-type or L208 variants was thawed and resuspended in lysis buffer using 2.5 mL of buffer/g of cell paste. This buffer contained 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM imidazole, 1 mM DTT, 100 μM PMSF and supplemented with protease inhibitors (Roche complete EDTA-free protease inhibitor cocktail, Roche). Cells were disrupted using ultrasonication with 50% amplitude for 15 min. Cell debris and unbroken cells were removed from the cell lysate to gain the clear crude extract using centrifugation at 40000 *g* and 4 °C for 1 hour. The clear crude extract was filtrated (0.45 μm) and then applied onto a 5-mL Ni-Sepharose HisTrap™ HP column (GE-Healthcare, Little Chalfont, UK) previously equilibrated with 50 mM Tris-HCl pH 8.0, 300 mM NaCl and 10 mM imidazole. The column was washed

with 50 mM Tris-HCl buffer pH 8.0 containing 300 mM NaCl, 20 mM imidazole. HNMT was eluted using 50 mM Tris-HCl buffer, pH 8.0 containing 300 mM NaCl and 300 mM imidazole. Eluted fractions containing HNMT were pooled, and then concentrated using a Centricon YM10 with 10 kDa cut-off. The concentrated protein fraction was loaded onto a Superdex 200 prep grade column (HiLoad 16/600 Superdex 200 pg, GE-Healthcare, Little Chalfont, UK) previously equilibrated with 50 mM Tris-HCl pH 8.0 containing 100 mM NaCl. HNMT fractions were pooled and then concentrated as described before. The purified HNMT was rapid-frozen and stored at -70 °C until used. To check protein purity and subunit molecular mass of HNMT, aliquots of the protein fractions were analyzed using 12.5% SDS-PAGE and staining with Coomassie Brilliant Blue R250. Catalytic activity for histamine *N*-methyltransferase of the protein fractions was determined using radioactive methylation assay as described below.

Western blotting

Protein samples for western blot analysis were separated using 12.5% SDS-PAGE and then electrically transferred to a nitrocellulose membrane by applying electricity (150 mA) from a power supply for 1 hour. The membrane was blocked for 1 hour using 13% milk powder in TBS buffer (50 mM Tris-HCl buffer pH 8.0 and 150 mM NaCl) containing 0.1% tween-20. The membrane was incubated with the primary rabbit His-Tag antibody (1:1000 dilution, Cell Signaling Technology®, Danvers, MA, USA) for 16 hours, and then incubated with the secondary antibody anti-rabbit IgG, horseradish peroxidase-linked antibody (1:5000 dilution, Cell Signaling Technology®, Danvers, MA, USA) for 1 hour. To detect chemiluminescent signal of antibody-bound His6 tagged HNMT, the immunoblots were developed using enhanced chemiluminescent western blotting substrate solution (Pierce-Thermo Fisher Scientific, Waltham, MA, USA).

Determination of melting temperature of HNMT (wild-type) and L208 variants

To investigate the thermal stability of HNMT (wild-type) and L208 variants in the absence and presence of substrates, SAM and histamine, the melting temperature was determined by a fluorescence-based thermal shift assay, using SYPRO Orange (Sigma-Aldrich, St. Louis, MO, USA). Experiments were performed in a CFX Connect™ real-time PCR detection system (Bio-Rad, Hercules, CA, USA) with a 96-well plate in FRET scan mode. To prepare protein samples for the measurement, 5 μL of 1 mg/mL HNMT (wild-type or L208 variants) were mixed with 5 μL of SYPRO Orange solution (200-fold dilution) and

15 µL of a buffer containing 50 mM Tris-HCl pH 8.0 and 100 mM NaCl to a final volume of 25 µL in a 96-well plate. The protein samples in the presence of substrate (SAM or histamine) were prepared similarly as described above but either SAM or histamine was added to a final concentration of 2 mM. A temperature gradient from 20 °C to 95 °C was set to monitor thermal unfolding of the protein, measuring fluorescence emission at 0.5 °C increments with a 60 s hold for signal stabilization. The melting temperature of the protein in the absence and presence of substrates was obtained from the peak of the derivatives of the experimental data.

Determination of dissociation constants (K_d) of HNMT (wild-type) and L208 variants

All of the experiments to investigate dissociation constants (K_d) for SAM, SAH, and histamine to HNMT (wild-type) and L208 variants and for histamine to the binary complex of HNMT:SAH were performed using a VP-ITC system (MicroCal, Northampton, MA, USA) in 50 mM Tris-HCl buffer pH 8.0 containing 100 mM NaCl at 25 °C. All solutions were degassed before measurements. The HNMT (wild-type) solution (30 µM) was injected with SAM (0.42 mM), SAH (0.38 mM), or histamine (4 mM) (20 injections, 15 µL/injection, duration time 29.9 s, spacing time 250 s). Titration experiments for SAM, SAH and histamine binding to the L208V and L208F variants were set to 20 injections of 15 µL of 0.7 mM SAM, 0.55 mM SAH, or 4 mM histamine (duration time 29.9 s, spacing time 250 s) into the cell containing 30 µM HNMT (L208V and L208F). To determine the K_d of L208R and L208K variants for SAM, SAH, and histamine, 20 µM of the HNMT proteins were titrated with SAM (1 mM), SAH (0.8 mM), or 4 mM histamine (20 injections, 15 µL/injection, duration time 29.9 s, spacing time 250 s). For the determination of the dissociation constants for histamine to the binary HNMT (wild-type, L208V, and L208K):SAH complex (20 µM HNMT:120-250 µM SAH), histamine (0.58 mM for HNMT (wild-type) and 0.50 mM for the L208 variants) was titrated into the complex solution (20 injections, 15 µL/injection, duration time 29.9 s, spacing time 250 s). The K_d values of HNMT (wild-type) and L208 variants for SAM, SAH, histamine, and histamine in the presence of SAH were obtained fitting the data to a single-binding-site model with Origin version 7.0 (MicroCal, Northampton, MA, USA) for ITC data analysis.

Determination of catalytic activity of HNMT (wild-type) and L208 variants

To measure catalytic activity of HNMT (wild-type) and L208 variants; L208V, L208F, L208R, and L208K, radioactive methylation assay was performed at 25 °C. [Methyl-^{3}H]-SAM was utilized and the methylation reaction was monitored by formation of the product, [methyl-^{3}H]-histamine [29,40,41]. The assay reaction in a 1.5 mL Eppendorf tube contained 20 nM HNMT (wild-type or L208 variants), 40 μM unlabeled SAM, 40 μM histamine, 0.158 μCi [methyl-^{3}H]-SAM, 1 mM EDTA, and 0.025% bovine serum albumin in 125 mM bicine buffer pH 8.0 to make up a final volume of 100 μL. The reaction mixture was preincubated at 25 °C for 2 min prior to the addition of histamine to initiate the reaction. Termination of the reaction was conducted by adding 75 μL of 1 M sodium borate pH 11 after the reaction was incubated at 25 °C for 30 min. The product, [methyl-^{3}H]-histamine, was isolated from unincorporated [methyl-^{3}H]-SAM in the reaction by organic extraction by adding 1:1 toluene:isoamyl alcohol (1.25 mL). After vortexing and centrifugation at 21100 ×g for 1 min, 1 mL of the organic phase was transferred to a 1.5 mL Eppendorf tube and then mixed with 250 μL of 0.5 M HCl. The tube was again vortexed and centrifuged to facilitate phase separation. 200 μL of the aqueous phase were transferred to a scintillation vial containing 10 mL of Ultima Gold liquid scintillation cocktail (PerkinElmer, Waltham, MA, USA). After mixing, the vial was quantified by scintillation spectrometry. To determine the variant and wild type enzymes' kinetic parameters, K_M and k_{cat}, for SAM, histamine was fixed at saturating concentration of 40 μM and the concentration of SAM was varied in the range of 5-80 μM. In order to obtain the respective kinetic parameters for histamine, SAM concentration was kept constant at 40 μM and histamine concentration was varied in the range of 1-20 μM. Data analysis to obtain the kinetic parameters was performed by plotting the rate of [methyl-^{3}H]-histamine formation normalized by the HNMT concentration (v/e) versus the substrate concentration (SAM or histamine) and fitting the data points using Marquardt-Lavenberg algorithms in the KaleidaGraph program (Synergy Software, Reading, PA, USA).

Measurements of circular dichroism (CD) spectrometry

To determine the secondary structure of HNMT (wild-type) and L208 variants, circular dichoism (CD) spectrometry was performed using a Jasco J715 spectropolarimeter (JASCO Inc., Tokyo, Japan) equipped with thermostat cell compartments. The CD spectra of HNMT in the far UV region (260-185 nm) were recorded in quartz cells with path

length 0.02 cm (Hellma, Forest Hills, NY, USA) with a scan speed of 100 nm/min at 25 °C, averaged over three scans and corrected by the blank signal. HNMT was prepared in 20 mM Tris-acetate buffer pH 8.0 to give a final concentration in the range of 0.6-1 mg/mL. The secondary structure content of HNMT (wild-type) and L208 variants was analyzed using the Spectra Manager program (JASCO Inc., Tokyo, Japan) and the web tool DichroWeb [42,43].

MALDI-TOF mass spectrometry

Protein extraction from polyacrylamide gels and MALDI-TOF mass spectrometric measurements for HNMT (wild-type) and L208 variants were conducted as previously described [44,45]. The individual protein bands on a polyacrylamide gel were excised and transferred to a 1.5 mL Eppendorf tube. Each excised gel piece was washed with deionized water prior to be vortexed in 10% acetic acid solution for 10 min. The gel piece was consequently washed with deionized water followed by acetonitrile and a final washing step with deionized water. The stained color of Coomassie Brilliant Blue R250 on each gel piece was removed completely by incubating the gel in a destaining solution containing formic acid:deionized water:isopropanol (1:3:2 v/v/v) for 30 min. The gel piece was rinsed with deionized water and dehydrated with acetronitrile. After the destaining procedure, cysteines in the gel piece were alkylated and reduced by iodoacetamide and dithiothreitol (DTT), respectively. Trypsin was then added to digest the protein in the gel at 37 °C overnight. After trypsin digestion, peptide mixtures were extracted and then desalted using ZipTip (Millipore, Darmstadt, Germany). The desalted peptides were spotted onto a MALDI plate with a matrix containing α-cyano-4-hydroxycinnamic acid. The mass spectra of peptides were recorded using a Micromass Tof Spec 2E MALDI-TOF mass spectrometer (Micromass, Cary, NC, USA) in reflectron mode at an accelerating voltage of +20 kV. The instrument was calibrated with a poly(ethylene glycol) mixture (Sigma-Aldrich, St. Louis, MO, USA). ProteoMasS ACTH Fragment 18–39 (Sigma-Aldrich, St Louis, MO, USA) was used as the peptide calibration standard for the instrument. To analyze mass spectra from the measurement, MassLynx 4.1 (Waters, Milford, MA, USA) was utilized to assign peptide mass profiles.

In silico prediction of protein stability of the L208 variants

SNPeffect (version 4.0) (http://snpeffect.switchlab.org) [46,47] is a web tool program which is used to analyze the structural effect of protein variations using various algorithms (TANGO, WALTZ, LIMBO, and FoldX). Protein stability analysis and predicted protein structure for L208 variants were determined and modelled by FoldX. The models for L208 variants were created based on the HNMT wild-type crystal structure (PDB: 1JQD, [29]). The calculation of free-energy change ($\Delta\Delta G$) upon variations in the structural phenotype of the protein using FoldX provides information of protein stability of the L208 variants. In addition, the web-tool CUPSAT was utilized to analyze and predict protein stability upon substitution of Leu208 with other amino acids [48].

Molecular dynamics (MD) simulations of HMNT

The crystal structure of the Thr105 polymorphic variant of HNMT in complex with S-adenosylhomocysteine (SAH) and histamine (PDB:1JQD, chain B) [29] was used as a starting structure for molecular dynamic (MD) simulations. For calculations of HMNT in complex with the methylating cofactor SAM and histamine, which represents the catalytically competent enzymatic complex, a model was created based on the crystal structure. Therefore, a methyl group was tethered to the sulfur atom to give a chiral sulfur center in (S) configuration, which constitutes the biologically active form of SAM [49]. Variants of HMNT were produced by in-silico point mutations of the created HMNT model. MD simulations and analyses were performed using the YASARA Structure suite, version 13.9.8 (YASARA Biosciences) [50]. A periodic simulation cell that enveloped the whole enzyme with an additional 5 Å margin in each dimension, was used with explicit solvent. The AMBER03 force field [51] was applied and long-range electrostatic potentials were calculated using the Particle Mesh Ewald (PME) method and a cutoff of 7.864 Å [52,53]. The AutoSMILES utility was used to attribute force field parameters to SAH and histamine [54]. For SAM Mulliken charges were generated by single point DFT calculations via GAMESS [55], using the B3LYP/6-311G** basis set.

The hydrogen bonding network optimizations were carried out using the method of Hooft and co-workers [56], and pK_a values at pH 7.5 were assigned [57]. The simulation cell was filled with water containing 0.9 % of NaCl and a density of 0.993 g/mL. Solvent was relaxed and the system was subsequently energy minimized using steepest descent minimization to remove conformational stress, followed by a simulated annealing

minimization step to convergence (< 0.05 kJ/mol per 200 steps). Integration time steps were 1.33 and 4 fs for intra- and intermolecular forces, respectively. Subsequently, MD simulations at 310 K were performed, whereby integration time steps for intramolecular and intermolecular forces of 1.25 fs and 2.5 fs were applied, respectively. All MD simulations were carried out in quadruplicate with varying seed numbers and over a time span of 25 ns.

Results

In silico prediction of protein stability

The three-dimensional structure of HNMT shows that the L208 residue is located in one of six helices, termed helix E, flanking a seven-stranded β-sheet of the MTase domain [29]. Although this position is not in or near the active site of HNMT a multiple alignment with vertebrate HNMTs (44-99 % identity) revealed that L208 is invariant indicating a conserved structural role of the residue (Fig. 1A and 1C). In fact L208 has several contacts in the same and the neighboring α-helix as well as the adjacent β-sheet (Fig. 1, panels A and B). Recently, we have reported the drastic destabilizing effect (ca. 8 kcal/mol) of the disease-related L208P variant on protein stability [38]. In order to assess further the importance of residue 208 for protein stability we have extended our previous calculations using FoldX in SNPeffect 4.0 [46,47]. As shown in Fig. 1D, all amino acid replacements result in lower protein stability with the strongest effects found for charged amino acids (Arg, Asp, Lys) whereas hydrophobic residues were predicted to exert only a moderate effect on protein stability (1-3 kcal/mol for L208F and L208V). The results from the calculations with FoldX in SNPeffect 4.0 were consistent with protein stability changes predicted by the web-tool CUPSAT (data not shown). From this analysis, we concluded that side chain packing in this hydrophobic pocket is critical for protein stability. This insight prompted us to further investigate protein stability and enzyme activity for the series of HNMT variants displayed in Fig. 1D.
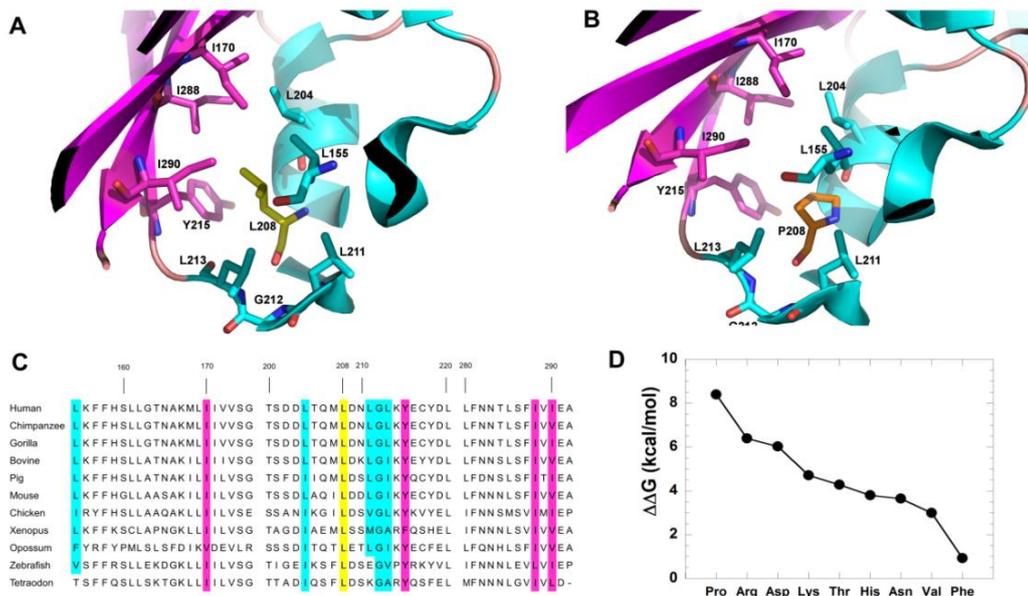
Fig. 1: Structure of HNMT with a focus on position 208, sequence alignment and protein stability analysis. Three-dimensional HNMT structures at the hydrophobic pocket around residue 208 for L208 in yellow (A) and L208P in orange (B) with surrounding apolar residues: I170, Y215, I288, and I290 in purple and L155, L204, L211, G212, L213 in light blue. (C) Protein sequence alignment of HNMT homologs using Clustal Omega 1.2.1 (www.clustal.org/omega) at the hydrophobic pocket surrounding residue L208. The L208 is highlighted in yellow. Residues I170, Y215, I288, and I290 were highlighted in purple and residues L155, L204, L211, G212, L213 in light blue. All HNMT sequences used for the alignment were obtained from NCBI Reference Sequence Database: human (NP_008826.1), chimpanzee (XP_001156140.1), gorilla (XP_004032652.1), bovine (NP_001030511.1), pig (NP_001166431.1), mouse (NP_536710.1), chicken (NP_001264802.1), xenopus (NP_001080614.1), opossum (UniProtKB browser_F6X9A6), zebrafish (NP_001003636.1), and tetraodon (Q4SB6.1). (D) Effect of amino acid substitutions at position 208 on protein stability as calculated by by SNPeffect 4.0 (FoldX) expressed as the difference to wild-type protein ($\Delta\Delta G$ in kcal/mol).

Protein expression and purification of L208 variants

The eight variants of human HNMT, *i.e.* L208V, L208F, L208T, L208N, L208R, L208K, L208H, and L208D were expressed in *E.coli* BL21 (DE3) cells as described in Materials & Methods. The two variants with conservative amino acid exchanges, *i.e.* L208V and L208F were expressed as soluble proteins. In contrast, the variants L208T, L208N, L208H, and L208D were found exclusively in inclusion bodies. Attempts to solubilize these variants by coexpression with chaperones were not successful. Interestingly, the two variants bearing a positively charged amino acid, *i.e.* L208R and L208K, were also detectable as soluble proteins although the stability prediction by FoldX has indicated a profound destabilizing effect (Fig. 1D). The protein variants expressed in the soluble fraction were purified using

a two-step protocol employing Ni-Sepharose affinity and size-exclusion chromatography (see Materials & Methods) as presented in Fig. 2. The yield of purified wild-type protein was 6 mg/g of cell paste. A slightly lower yield of 2 mg/g wet cell paste was achieved with the variants L208V and L208F whereas the yield for the variants L208R and L208K were significantly lower amounting to only 0.2-0.4 mg/ g cell paste, respectively (Table 1).
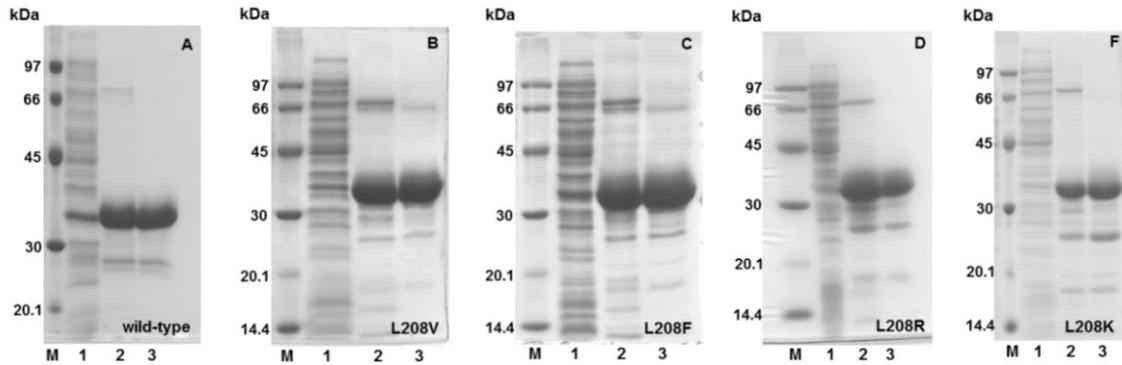

Fig. 2: Protein expression and purification for HNMT (wild-type) and L208 variants. The recombinant HNMT (wild-type) and L208 variants were expressed in *E.coli* and purified using Ni-Sepharose affinity and Superdex 200 size-exclusion chromatography. SDS-PAGE (12.5%) analysis of HNMT (wild-type) (A) and L208 variants; L208V (B), L208F (C), L208R (C), and L208K (D) after each step of the protein purification. Lane M, low molecular mass protein marker; lane 1, crude extract; lane 2, the protein fraction after Ni-Sepharose chromatography; lane 3, the protein fraction after Superdex 200 size-exclusion chromatography.

Table 1: Protein purification table for HNMT (wild-type) and soluble L208 variants

Step of protein purification	Wild-type		L208V		L208F		L208R		L208K	
	Total protein (mg)	% yield	Total protein (mg)	% yield	Total protein (mg)	% yield	Total protein (mg)	% yield	Total protein (mg)	% yield
Crude extract	1727	100	1305	100	1307	100	1243	100	3852	100
Ni-Sepharose	197	11	64	5	68	5	42	3	43	1
Superdex 200	111	6	37	3	41	3	7	0.5	16	0.4
	$(6)^{\delta}$		$(2)^{\delta}$		$(2)^{\delta}$		$(0.2)^{\delta}$		$(0.4)^{\delta}$	

Starting cell paste for protein purification for HNMT (wild-type), L208V, L208F, L208R, and L208K variants was 20, 19, 19, 30, and 45 g respectively.

$^{\delta}$Amount of protein in mg per g of cell paste (mg/g) is shown in parentheses.

Determination of dissociation constants using isothermal titration calorimetry (ITC)

We have employed isothermal titration calorimetry (ITC) to determine the affinity of SAM to wild-type HNMT and the soluble variants L208V, L208F, L208R, and L208K. As shown in Fig. 3 (panels A-E), titration of the proteins with SAM at 25 °C produced exothermic signals. The ITC titration data were fitted to a single-binding-site model yielding dissociation constants of 49 ± 2, 38 ± 2, 42 ± 4, 62 ± 13, and 64 ± 10 µM for the wild-type, L208V, L208F, L208R, and L208K, respectively. The results indicate that affinity of SAM to HNMT (wild-type) and the L208 variants is not significantly affected by the amino acid substitution at position 208. Interestingly, the product SAH, binds twenty times more tightly to the wild-type and six times more tightly to the variants L208V and L208K than SAM, exhibiting dissociation constants of 2 ± 0.2, 6 ± 2, and 11 ± 3 µM, respectively (Fig. 4 and Table 2).
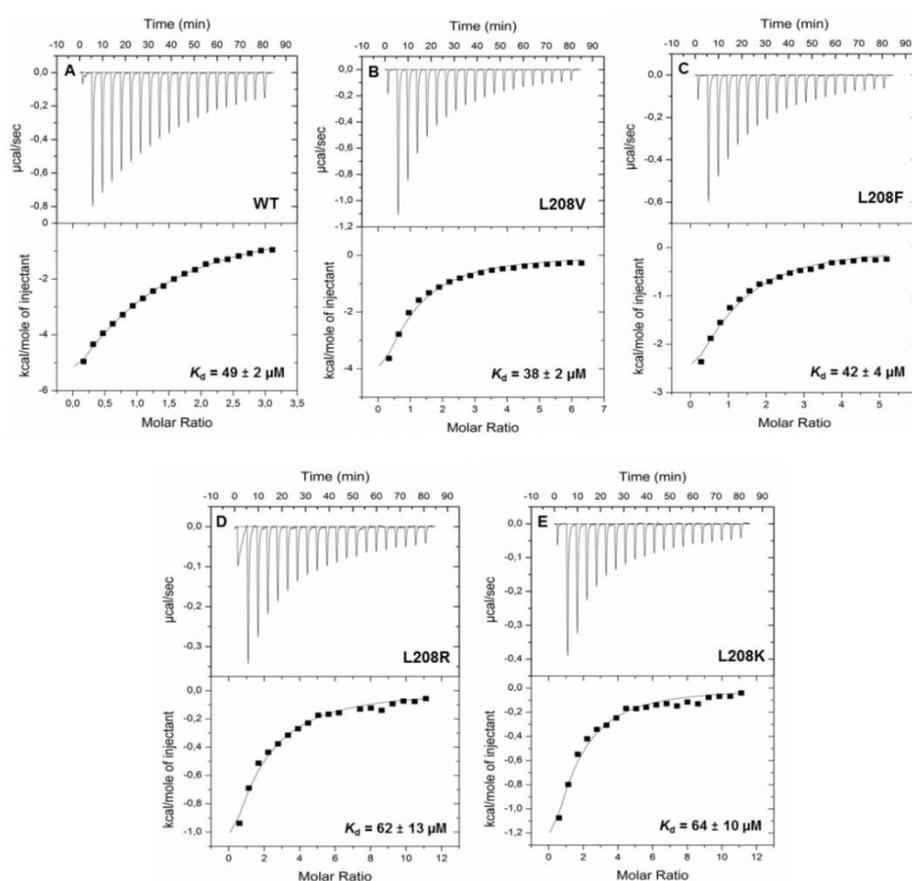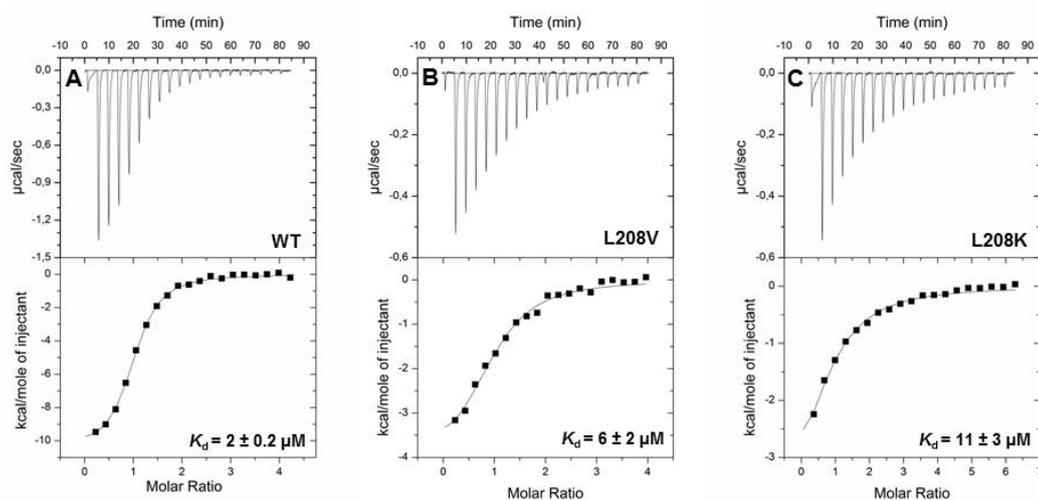

Fig. 3: Binding of SAM to HNMT (wild-type) and L208 variants. ITC experiments were performed in 50 mM Tris-HCl buffer, pH 8.0, containing 100 mM NaCl at 25 °C using a VP-ITC system (MicroCal). Isothermal calorimetric titrations of HNMT (wild-type) (A) and L208 variants for L208V (B), L208F (C), L208R (D) and L208K (E) with SAM are shown. The dissociation constants for SAM binding to HNMT (wild-type), L208V, L208F, L208R, and L208K were 49 ± 2, 38 ± 2, 42 ± 4, 62 ± 13, and 64 ± 10 µM, respectively.

Fig. 4: Binding of SAH to HNMT (wild-type) and L208 variants. ITC experiments were performed in 50 mM Tris-HCl buffer, pH 8.0, containing 100 mM NaCl at 25 °C using a VP-ITC system (MicroCal). Isothermal calorimetric titrations of HNMT (wild-type) (A) and L208 variants for L208V (B), and L208K (C) with SAH were presented. The dissociation constants for SAM binding to HNMT (wild-type), L208V, and L208K were 2 ± 0.2, 6 ± 2, and 11 ± 3 µM, respectively.

Table 2: Dissociation constants (K_d) of HNMT (wild-type) and L208 variants for SAM, SAH and histamine in the presence of SAH by ITC

HNMT	Dissociation constants (K_d, µM)		
	SAM	SAH	Histamine (in the presence of SAH)
Wild-type	49 ± 2	2 ± 0.2	19 ± 3
L208V	38 ± 2	6 ± 2	ND
L208F	42 ± 4	-	-
L208R	62 ± 13	-	-
L208K	64 ± 10	11 ± 3	ND

ND = not detectable and dash (-) = not measured

In contrast to SAM and SAH, the binding of histamine to HNMT is not detectable by ITC. However, in the presence of SAH, exothermal signals were obtained upon titration with histamine enabling the determination of the dissociation constant by fitting the data to a single-binding-site model. This yielded a dissociation constant of 19 ± 3 µM (Fig. 5, panel A). In the case of the L208V and L208K variants no signals were obtained indicating that histamine binding is substantially affected even in the conservative L208V variant (Fig. 5, panels B and C). A summary of the data obtained by ITC is given in Table 2.
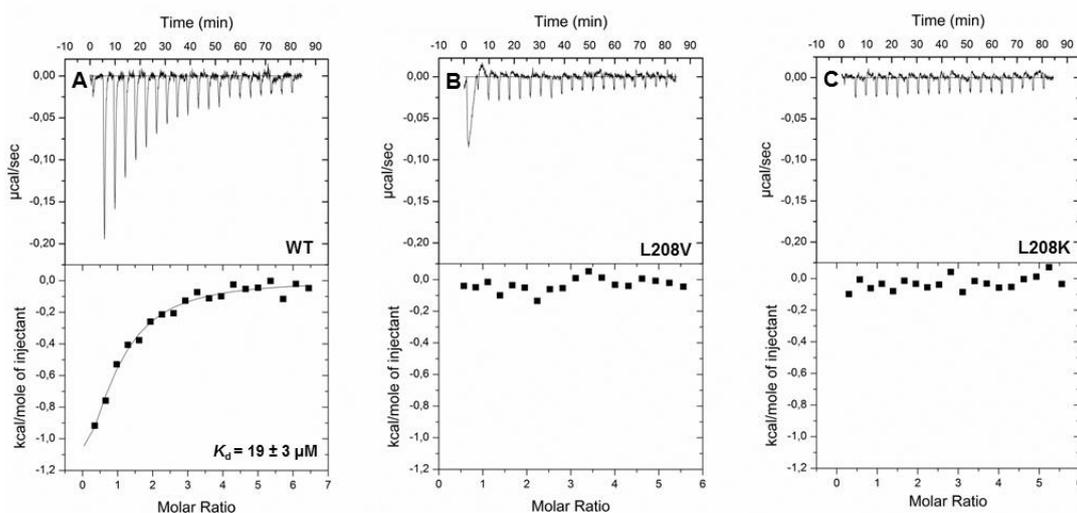

Fig. 5: Binding of histamine to HNMT (wild-type) and L208 variants in the presence of SAH. Isothermal calorimetric titrations of HNMT (wild-type) (A), L208V (B), and L208K (C) variants with histamine in the presence of SAH were performed in 50 mM Tris-HCl buffer, pH 8.0, containing 100 mM NaCl at 25 °C using a VP-ITC system (MicroCal). The ITC experiment for HNMT (wild-type) consisted of 20 consecutive injections of histamine (0.58 mM, 15 µL/injection) into the cell containing 20 µM HNMT (wild-type) and 0.12 mM SAH. For L208V and L208K, a solution of the L208 variants and 0.25 mM SAH in the cell was titrated with histamine (0.50 mM, 15 µL/injection, 20 injections). Data analysis by non-linear least-squares fitting using Origin version 7.0 (Microcal) was conducted to obtain the dissociation constant for histamine binding to HNMT (wild-type) in the presence of SAH of 19 ± 3 µM. There was no released heat observed after 20 consecutive injections of histamine into the L208 variants in the presence of SAH.

Catalytic activity of wild-type HNMT and the soluble L208 variants

To determine the effect of amino acid substitutions at position 208 on the catalytic activity, we employed a radioactive methylation assay as described in Materials & Methods [29,40,41]. As shown in Fig. 6 (panel A), product formation (*N*-methylhistamine) was reduced to 48 ± 5 %, 21 ± 2 %, 6 ± 0.5 % and 1 ± 0.05 % for the L208V, L208F, L208K and L208R variant, respectively, in comparison to wild-type enzyme. To gain further

101

insight, a set of steady-state experiments was performed at various concentrations of SAM and histamine. Kinetic parameters of the methylation reaction were obtained by determining the rate of [methyl-^{3}H]-histamine generation as a function of SAM or histamine concentrations (Fig. 6 B and C, respectively) and fitting of the data to the Michaelis-Menten equation. A summary of the kinetic parameters is shown in Table 3. In the case of the L208V variant kinetic parameters are only marginally affected with both K_M values in a similar range as observed for the wild-type and a 3-times lower k_{cat}. The L208F variant is clearly more affected exhibiting significantly higher K_M values and a 5.5-times lower k_{cat}. This tendency is even more pronounced with the L208R and L208K variants. In the case of the former variant no data was obtained due to the low overall activity (see also Fig. 6, panel A).

Table 3: Apparent kinetic parameters of HNMT reaction for HNMT (wild-type) and L208 variants

HNMT	K_m^{SAM} (µM)	$K_m^{histamine}$ (µM)	k_{cat} (s^{-1})
Wild-type	4.1 ± 1.2	1.9 ± 0.7	0.61 ± 0.08
L208V	6.8 ± 1.6	1.1 ± 0.2	0.23 ± 0.01
L208F	12.4 ± 4.8	3.4 ± 0.5	0.11 ± 0.01
L208R	ND	ND	ND
L208K	15.0 ± 5.8	ND	0.03 ± 0.006

ND = not detectable
Enzymatic activity of HNMT was measured using radioactive methylation assay to determine the product formation of methyl-histamine [methyl-, ^{3}H].
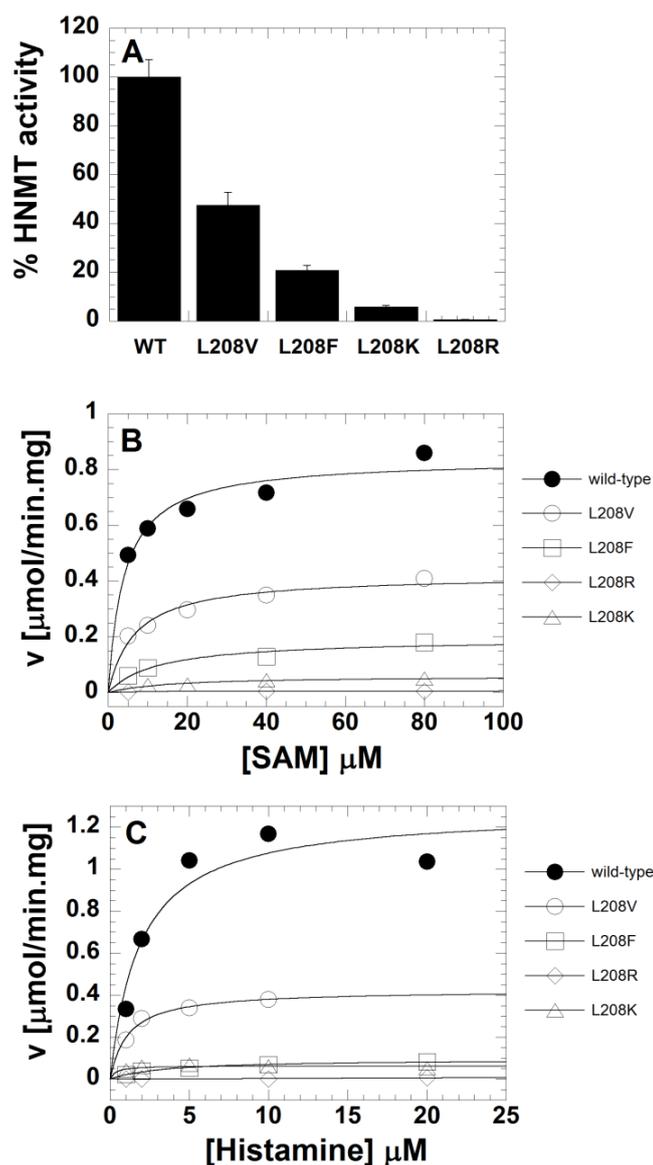
Fig. 6: Catalytic activity and steady-state kinetics of HNMT (wild-type) and L208 variants. (A) Percentage of enzymatic activity of HNMT (wild-type) and L208 variants. (B) Michaelis-Menten plot for HNMT (wild-type) and L208 variant reactions when SAM concentration was varied. Determination of HNMT activity using a radioactive methylation assay was conducted at 25 °C, monitoring the product, methyl-histamine [methyl-, ^{3}H] formation. The reaction contained 20 nM HNMT (wild-type; filled circle, L208V; empty circle, L208F; empty square, L208R; empty diamond, and L208K; empty triangle) in 125 mM bicine buffer pH 8.0 with 0.15 µCi SAM [methyl-, ^{3}H], 0.025% bovine serum albumin, 1 mM EDTA, a fixed concentration of histamine (40 µM), and various concentrations of SAM (5, 10, 20, 40, 80 µM). (C) Michaelis –Menten plot for HNMT (wild-type) and L208 variant reactions when the histamine concentration was varied. The HNMT activity was monitored at 25 °C from the reactions consisting of 20 nM HNMT (wild-type; filled circle, L208V; empty circle, L208F; empty square, L208R; empty diamond, and L208K; empty triangle) in 125 mM bicine buffer pH 8.0, 0.15 µCi SAM [methyl-, ^{3}H], 0.025% bovine serum albumin, 1 mM EDTA, a fixed concentration of SAM (40 µM), and various concentrations of histamine (1, 2, 5, 10, 20 µM).

103

Thermal stability of L208 variants of HNMT in the absence and presence of SAM and histamine

The determination of thermal protein stability was performed by a fluorescence-based thermal shift assay (Thermofluor® technique) to obtain the melting temperatures (T_m) of wild-type and the soluble variants (L208V, L208F, L208R, and L208K). This analysis was also carried out in the presence of either SAM or histamine. A compilation of the data is given in Table 4. In the case of the conservative replacements, *i.e.* the L208V and L208F variants, only a slight decrease of the melting temperature was observed. In contrast, the L208R and L208K variants showed significantly lower melting temperatures amounting to 13 and 8 °C, respectively. In the presence of either SAM or histamine no significant effects on the melting temperature were observed except for a small thermal stabilization of the L208R variant in the presence of SAM (Table 4). In summary, these results confirm that position 208 plays an important role in maintaining the structural integrity of the protein.

Table 4: Melting temperature (T_m, °C) of HNMT (wild-type) and L208 variants in the absence and presence of SAM and histamine

Substrate	Melting temperature (°C)				
	Wild-type	**L208V**	**L208F**	**L208R**	**L208K**
No substrate	61.3 ± 2.0	57.8 ± 2.1	60.4 ± 2.9	48.4 ± 0.6	53.7 ± 1.8
+ SAM	61.5 ± 0.2	58.9 ± 0.5	59.4 ± 1.9	51.7 ± 0.4	53.4 ± 0.9
+ Histamine	60.4 ± 2.8	56.9 ± 1.3	59.8 ± 2.7	48.3 ± 0.3	51.5± 2.5

Secondary structure analysis of HNMT (wild-type) and L208 variants

The effect of amino acid substitutions on the secondary structure and folding properties, circular dichroism spectra of HNMT (wild-type) and the L208 variants (Supplemental Fig. S1) were recorded. Analysis of CD spectrum of the wild-type using the web-tool DichroWeb [42,43] gave a secondary structure composition, as presented in Table 5, which reasonably corresponded to data obtained from the *in-silico* analysis of the crystal structure and showed even better congruence when compared to the average secondary structure composition predicted from 25 ns MD simulations of HNMT (*vide infra*) using the

YASARA suite [58]. In addition, when L208 variants were subjected to the analogous *in-silico* analysis, no significant difference in the content of secondary structure compared to wild type was observed (Table 5).

Table 5: Secondary structure content of HNMT (wild-type) and L208 variants

Secondary - structure content	Wild-type (CD)*	Wild-type (PDB)[#]	Wild-type[$]	L208V[$]	L208F[$]	L208R[$]	L208K[$]	L208P[$]
Helix	38	46	40 ± 1	42 ± 1	42 ± 3	43 ± 3	41 ± 2	42 ± 1
Sheet	17	26	22 ± 1	24 ± 0.5	23 ± 1	24 ± 1	23 ± 1	23 ± 1
Turn	19	10	15 ± 1	13 ± 2	12 ± 3	13 ± 1	15 ± 2	13 ± 1
Coil	27	17	22 ± 2	19 ± 1	21 ± 2	19 ± 2	20 ± 2	22 ± 1
3-10 helix		1	1 ± 1	2 ± 1	2 ± 2	1 ± 1	1 ± 1	1 ± 1

* The secondary structure content of HNMT (wild-type) from the CD measurement were analyzed using the Spectra Manager program and the web tool DichroWeb [42,43]

[#] The secondary structure content from the HNMT crystal structure (pdb: 1JQD:chain B)

[$] The average secondary structure content calculated from quadruplicate 25 ns MD simulations using the YASARA suite.

Molecular dynamics (MD) simulations of HNMT

Using the crystal structure of HNMT in complex with SAH and histamine as a starting point, 25 ns MD calculations were performed. The structures showed over all stability over the observed time frame. Generally, a displacement of the histamine ligand was observed. In three of four simulations the histamine was pushed closer to the SAH cofactor, with the deprotonated $N_{\varepsilon 2}$ pointing towards the sulfur atom of SAH (Supplemental Fig. S2) and thus in a suitable orientation and distance for transfer of a methyl group. In one simulation the histamine dissociated from the binding pocket. Analysis of this set of simulations revealed that during all four simulations direct H-bonding of the histamine ligand to amino acid residues, which was observed in the crystal structure, was lost in lieu of H-bonds to surrounding water molecules. Specifically, the intercalation of water molecules between Asn283, Glu28 and the aliphatic nitrogen of histamine brought the $N_{\varepsilon 2}$ of histamine closer

towards the SAH sulfur (Supplemental Fig. S3). A possible explanation for the repositioning in the MD simulation, as compared to the crystal structure lies in the distinct pH values: While the crystals were obtained at pH 5.6, at which the protein is barely active, MD simulations were performed at pH 7.5, where HNMT shows maximum activity [29].

When simulations were repeated for the catalytically competent complex of HNMT, SAM and histamine, the analogous repositioning and 'anchoring' of histamine by water molecules bridging Asn283 and histamine was observed in all four simulations (Fig. 7, panel A). The positioning of the transfer-methyl group of SAM and the histamine's acceptor atom $N_{\varepsilon 2}$, as seen in trajectory analyses (Fig. 7, panel B), with distances down to ~3.2 Å approximating the van der Waals contact distance, suggests that the MD simulation depicts the catalytically competent enzyme complex. Furthermore, our results suggest that the positioning of histamine in the SAM coordinated HNMT is similar as in the SAH bound complex.
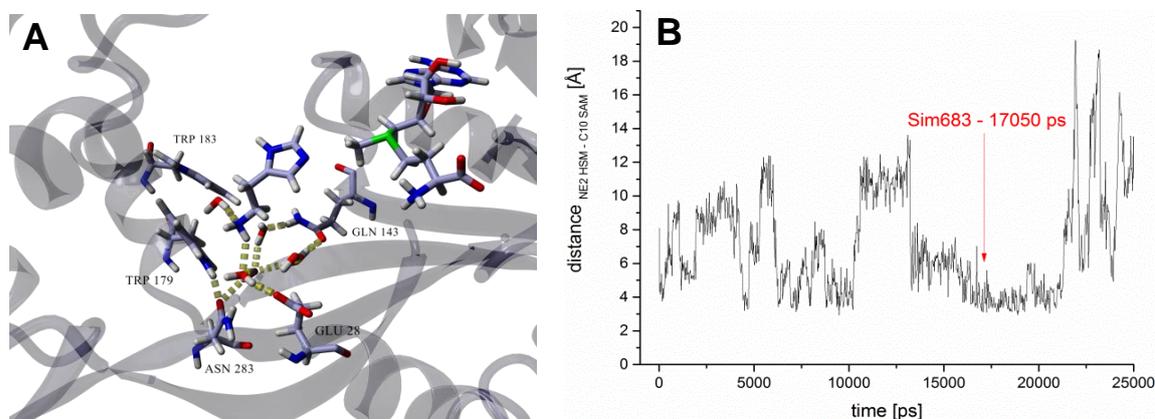

Figure 7: (A) Active site of HNMT in complex with SAM and histamine showing the hydrogen bonding network (yellow) connecting the alphatic N of histamine with the protein structure. The 17050 ps snapshot (snapshot #683) from one MD simulation of the HNMT:SAH:histamine complex at pH 7.5 and 310 K is shown. Nitrogens are in blue oxygen is red, hydrogens are shown in white and sulfur is depicted in green. (B) Trajectory showing the distance of the sulfur bound methyl group of SAM and the recipient atom $N_{\varepsilon 2}$ of histamine. The point of time at which the snapshot from (A) was taken is indicated by a red arrow.

In order to gain insights into the impact of amino acid variations at position 208, 25 ns MD simulations of the respective variants of this study in complex with SAH and histamine were conducted. Substitutions of L208 by V, F, R and K as well as the disease related L208P variant [38] were subjected to MD analysis. During the 25 ns time period of simulation all investigated structures were stable with similar overall secondary structure content. Inspection of the average structures revealed the presence of two principle conformations, namely an open and closed conformation: The crystal structure, which was also the principle starting structure of all simulations, showed a rather narrow cleft between the helix spanning residues 10-25 and the helix constituted by residues 91-105. In several simulations the distance between these two helices widened considerably (Supplemental Fig. S4). This is consistent with previous findings that dynamic conformational changes from closed to open forms occur in apo-HNMT [31]. Here, as a measure of the conformational state (open versus closed), average distances of the Cα of H12 and of Cα of S91 were compiled for all MD simulations (Supplemental Table S2).

While the overall secondary structure content of the structures was not significantly impacted by the investigated amino acid substitutions (Table 5), a significant decrease of the average helical character of helix E was found (Supplemental Table S3), suggesting destabilization of this secondary structure. This effect was also reflected in a local increase in the α-helix's RMSD values (Supplemental Table S2). Notably, the observed effects, a decrease in α-helical secondary structure and increased RMSD values, correlated with the enzyme's activity *in vitro* (Table 3 and Fig. 6), with the strongest impact on the L208R and L208P variants. The destabilization was, however not significantly propagated to other secondary structure elements.

The average histamine's positioning was dramatically disturbed in the variant L208P with the histamine methylation target site $N_{\varepsilon2}$ ($N_{\varepsilon2his}$) more remote from the SAH sulfur than the nitrogen of the substrate's aliphatic amino group N_{his} (Table S3). In contrast, in the wild type the nitrogen of the aliphatic amino group remained at a similar distance and positioning compared to the SAH sulfur in all four simulations, whereas the reorientation of the aromatic ring away from the SAH, resulting in a greater distance of $N_{\varepsilon2}$ to the sulfur only occurred in one instance. The variant L208F showed a similar behavior as the variant L208P with the histamine consistently reoriented and at greater distances than in the wild type. In the variants L208R and L208K deviations from average position of the histamine in the wild type are still observable but less pronounced. Surprisingly, the histamine

appears to be drawn closer to SAH in the L208K variant than in other structures. The data are summarized in Supplemental Table S2 and average data for all four simulations are given in Table 6. Average distances and standard deviations for a set of four MD simulations show that S_{SAH} - N_{his} distances are strongly conserved in the wild type, while distances and their variations fluctuate in the variants, indicating that the 'anchoring' of histamine via its aliphatic nitrogen atom is disturbed. This was further confirmed by the visual inspection of average protein structures.

Table 6: Average distances of $N_{\varepsilon 2}$ and aliphatic N of histamine to the sulfur of SAH (S_{SAH}) and their variance, determined from a set of four MD simulations.

	S_{SAH} - $N_{\varepsilon 2\,his}$		S_{SAH} - N_{his}	
	Average Distance (N=4)	STDEV (N=4)	Average Distance (N=4)	STDEV (N=4)
Wild-Type	8.2	2.1	9.1	0.2
L208V	8.3	1.9	10	0.2
L208F	11.4	2.2	10	1.4
L208R	8.9	2.4	10.7	1.4
L208K	7.9	3	8.7	0.8
L208P	11.5	2.6	10	1

In order to assess how amino acid replacements at position 208 alter the binding site of histamine structures obtained in our MD simulations were analyzed in more detail. In this context, it is noteworthy that no correlation between the open and closed conformation of HNMT and the positioning of histamine was observed. Visual inspection of the protein crystal structure as well as average structures from MD simulations showed that Ser201 in α-helix E forms a hydrogen bond to Ser286 in the adjacent β-strand, which is separated by only two amino acids from Asn283, the residue apparently responsible for the correct anchoring of the nitrogen atom of the aliphatic amino group (Fig. 8). A relative destabilization and greater mobility of the α-helix in the variants, as observed in our simulations (Supplemental Table S2 and S3), may propagate to Asn283 and thus compromise histamine binding. This in turn could reasonably rationalize the observed impaired catalytic activity of the variants that correlates with the increase in α-helix destabilization. Equally, introduction of large amino acids in position 208, *e.g.*

phenylalanine, leads to clashes with Tyr215 (Fig. 8), which is located in the same adjacent β-strand as Ser 286, and may thus disturb the positioning of histamine by Asn283 in a similar fashion.
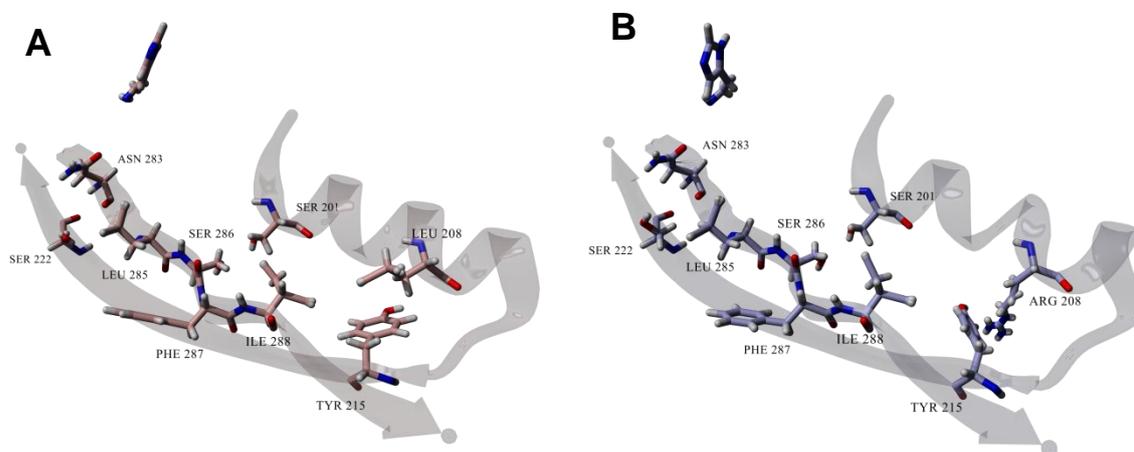

Fig. 8: Cartoon showing the structural features that link the residue 208 bearing α-helix, which spans residues 200-212, with residues lining the histamine binding pocket in HNMT. Average structures from 25 ns simulations of (A) wild type HNMT and (B) the L208R variant are shown. The histamine ligand and selected amino acids, which link the respective secondary structures through side chain interactions are shown. Hydrogen, nitrogen and oxygen atoms are shown in white, blue and red, respectively. Carbon atoms are shown in pink (A) and grey (B).

Discussion

The L208P variant of human HNMT was recently discovered as the cause of intellectual disability although inspection of the crystal structure did not provide a clear rationale for the putative dysfunction of the enzyme in histamine degradation. Analysis of the impact of the replacement of leucine by other amino acids with either charged (K, L, D, H), polar (T, N) or apolar (V, F) side chains suggested a clear relationship disfavoring charged and polar residues in comparison to apolar residues (Fig. 1D). This concept was largely supported by our expression experiments as variants bearing charged or polar side chains were not sufficiently soluble or stable for successful protein purification. Interestingly, the L208K and L208R variant could be isolated and partially characterized although they were predicted to be less stable than the L208T, L208H or L208N variants. This unforeseen stability of the L208K and L208R variant could be due to the length of the side chain that places the positively charged amino or guanidine group closer to the surface and thus alleviates the unfavorable contacts to the hydrophobic amino acid side chains in the vicinity of position 208.

The four variants L208V, L208F, L208K and L208R were isolated in varying amounts. Whereas the yields for the L208V and L208F variants were only reduced three-fold the latter two variants showed a thirty-fold decrease (Table 1). This apparent difference in stability was also reflected by the melting temperatures, which were slightly lower for the L208F and L208V but much more severely reduced for the L208K and L208R variants (Table 4). A similar tendency was observed in steady-state measurements that showed slightly reduced enzymatic activity for the L208V and L208F variants while the L208K and L208R variants exhibited less than 5% residual activity (Fig. 6 and Table 3). This initial characterization of the generated variants clearly suggested that position 208 is very sensitive to amino acid replacements resulting in a deterioration of stability and enzyme function, which strongly depends on the properties of the amino acid side chain.

In order to obtain further insights into the molecular mechanism by which the amino acid in position 208 exerts its influence on enzymatic activity we investigated the binding of SAM, SAH and histamine to the variants. Interestingly, binding of SAM and SAH was apparently not much affected in any of the variants with only a moderately higher dissociation constant for SAM and SAH for the L208K and L208R variants (Fig. 3 and 4; Table 2). In contrast, histamine binding as demonstrated by ITC measurements (in the presence of SAH) was only detectable for wild-type protein and even the L208V variant appeared to be compromised in histamine binding (Fig. 5). This result suggests that the effect on histamine binding is the major cause for the deterioration of enzymatic function in the variants. However, it was still not clear how amino acid exchanges in position 208 affect the binding of histamine. To address this question, we have carried out molecular dynamics (MD) calculations on the basis of the previously reported crystal structure of the HNMT:SAH:histamine complex ([29], PDB: 1JQD).

In our MD simulations we have noticed that histamine repositions in the active site to adopt a suitable position and orientation for the transfer of a methyl group from SAM to the $N_{\varepsilon2}$, $i.e.$ the crystal structure does not reflect a catalytically competent complex. This difference is obviously the consequence of the different pH values used for crystallization and our simulations, respectively. While crystallization was conducted at a low pH where the enzyme is inactive, we have used a physiological pH where the enzyme is active. Thus the observed repositioning of histamine appears to be triggered by the difference in pH used in the MD simulations.

Binding and positioning of histamine appears to be mainly achieved by interactions of the aliphatic amino group with water molecules, which are hydrogen bonding to the side chains of Asn283 and Glu28. Although position 208 is remote from the histamine binding pocket our simulations have indicated that amino acid exchanges destabilize the resident α-helix and this in turn disturbs side chain interactions with the adjacent β-strand that contains Asn283. Thus amino acid replacements in position 208 exert a long-range effect that propagates from α-helix E to the adjacent β-strand and eventually affects amino acids directly involved in histamine binding. This scenario is supported by our observation that the degree of α-helix destabilization and the loss of catalytic activity correlate. The resulting disturbance of histamine binding can potentially impact both histamine affinity (K_d) and the correct positioning of histamine for methylation and, consequently, the rate at which methylation occurs.

This effect from single amino acid exchange was also observed in another polymorphic SAM-dependent methyl transferase, thiopurine S-methyltransferase (TPMT). The amino acid replacement of a tyrosine by a cysteine at position 240 (TPMT*3C, the Y240C variant), which is located in β-strand 9, reduced enzymatic activity [59] and lowered the melting temperature from 50 to 39 °C [60,61]. Although no overall structural differences were observed in the Y240C variant, MD simulations indicated a loss of side chain interactions between β-strand 9 and the adjacent helix α8, which in turn appears to affect the substrate binding site [62]. In addition, catechol O-methyltransferase (COMT) also occurs in the V108M variant. Again, this position is ~16 Å apart from the SAM-binding site, but the amino acid exchange causes the reorientation of a proximal helix, resulting in the perturbation of the SAM and substrate binding sites [63,64].

References

[1] A. Falus, N. Grosman, Z. Darvas, Histamine: Biology and Medical Aspects, SpringMed Publishing, Budapest, 2004.

[2] L. Rosenwasser, New insights into the pathophysiology of allergic rhinitis, Allergy Asthma Proc. 28 (2007) 10-15.

[3] J. Loiselle, A. Wollin, Mucosal histamine elimination and its effect on acid secretion in rabbit gastric mucosa, Gastroenterology 104 (1993) 1013-1020.

[4] M. Zhang, R.L. Thurmond, P.J. Dunford, The histamine H(4) receptor: a novel modulator of inflammatory and immune disorders, Pharmacol. Ther. 113 (2007) 594-606.

[5] J.C. Schwartz, J.M. Arrang, M. Garbarg, H. Pollard, M. Ruat, Histaminergic transmission in the mammalian brain, Physiol. Rev. 71 (1991) 1-51.

[6] L.B. Hough, Cellular localization and possible functions for brain histamine: recent progress, Prog. Neurobiol. 30 (1988) 469-505.

[7] H. Wada, N. Inagaki, A. Yamatodani, T. Watanabe, Is the histaminergic neuron system a regulatory center for whole-brain activity? Trends Neurosci. 14 (1991) 415-418.

[8] A. Molina-Hernández, N.F. Díaz, J.A. Arias-Montaño, Histamine in brain development, J. Neurochem. 122 (2012) 872-882.

[9] P. Panula, M. Lintunen, K. Karlstedt, Histamine in brain development and tumors, Semin. Cancer Biol. 10 (2000) 11-14.

[10] A. Molina-Hernández, I. Velasco, Histamine induces neural stem cell proliferation and neuronal differentiation by activation of distinct histamine receptors, J. Neurochem. 106 (2008) 706-717.

[11] L. Maintz, N. Novak, Histamine and histamine intolerance, Am. J. Clin. Nutr. 85 (2007) 1185-1196.

[12] P.J. Barnes, Anti-IgE therapy in asthma: rationale and therapeutic potential, Int. Arch. Allergy Immunol. 123 (2000) 196-204.

[13] H. Haas, P. Panula, The role of histamine and the tuberomamillary nucleus in the nervous system, Nat. Rev. Neurosci. 4 (2003) 121-130.

[14] H.L. Haas, O.A. Sergeeva, O. Selbach, Histamine in the nervous system, Physiol. Rev. 88 (2008) 1183-1241.

[15] S.J. Hill, C.R. Ganellin, H. Timmerman, J.C. Schwartz, N.P. Shankley, J.M. Young, W. Schunack, R. Levi, H.L. Haas, International union of pharmacology. XIII. Classification of histamine receptors, Pharmacol. Rev. 49 (1997) 253-278.

[16] I.J. de Esch, R.L. Thurmond, A. Jongejan, R. Leurs, The histamine H4 receptors as a new therapeutic target for inflammation, Trends Pharmacol. Sci. 26 (2005) 462-469.

[17] C. Maslinski, W.A. Fogel, Catabolism of histamine, in: B. Uvnäs (ed.), Handbook of experimental pharmacology, Vol 97, Histamine and histamine antagonists, Springer, Berlin- Heidelberg, 1991, pp. 165-189.

[18] H.G. Schwelberger, F. Ahrens, W.A. Fogel, F. Sánchez-Jiménez, Chapter 3: Histamine metabolism, in H. Stark (ed.), Histamine H_4 receptor: a novel drug target for immunoregulation and inflammation, Versita, London, 2013, pp 63-102.

[19] M.A. Beaven, Histamine: its role in physiological and pathological processes, Monogr. Allergy 13 (1978) 1-113.

[20] S. Okinaga, T. Ohrui, H. Nakazawa, K. Yamauchi, E. Sakurai, T. Watanabe, K. Sekizawa, H. Sasaki, The role of HMT (histamine *N*-methyltransferase) in airways: a review, Methods Find. Exp. Clin. Pharmacol. 17 (Suppl C) (1995) 16-20.

[21] G.L. Chen, B. Zhu, W.P. Nie, Z.H. Xu, Z.R. Tan, G. Zhou, J. Liu, W. Wang, H.H. Zhou, Single nucleotide polymorphisms and haplotypes of histamine *N*-methyltransferase in patients with gastric ulcer, Inflamm. Res. 53 (2004) 484-488.

[22] W.P. Burkard, K.F. Gey, A. Pletscher, Diamine oxidase in the brain of vertebrates, J. Neurochem. 10 (1963) 183-186.

[23] D.D. Brown, R. Tomchick, J. Axelrod, The distribution and properties of a histamine-methylating enzyme, J. Biol. Chem. 234 (1959) 2948-2950.

[24] K.M. Lindahl, The histamine methylating enzyme system in liver, Acta. Physiol. Scand. 49 (1960) 119-138.

[25] C.V. Preuss, T.C. Wood, C.L. Szumlanski, R.B. Raftogianis, D.M. Otterness, B. Girard, M.C. Scott, R.M. Weinshilboum, Human histamine *N*-methyltransferase pharmacogenetics: common genetic polymorphisms that alter activity, Mol. Pharmacol. 53 (1998) 708-717.

[26] B. Girard, D.M. Otterness, T.C. Wood, R. Honchel, E.D. Wieben, R.M. Weinshilboum, Human histamine *N*-methyltransferase pharmacogenetics: cloning and expression of kidney cDNA, Mol. Pharmacol. 45 (1994) 461-468.

[27] K. Yamauchi, K. Sekizawa, H. Suzuki, H. Nakazawa, Y. Ohkawara, D. Katayose, H. Ohtsu, G. Tamura, S. Shibahara, M. Takemura, et al. Structure and function of human histamine *N*-methyltransferase: critical enzyme in histamine metabolism in airway, Am. J. Physiol. 267 (1994) L342-349.

[28] S. Aksoy, R. Raftogianis, R.M. Weinshilboum, Human histamine *N*-methyltransferase gene: structural characterization and chromosomal localization, Biochem. Biophys. Res. Commun. 219 (1996) 548-554.

[29] J.R. Horton, K. Sawada, M. Nishibori, X. Zhang, X. Cheng, Two polymorphic forms of human methyltransferase: structure, thermal, and kinetic comparisons, Structure 9 (2001) 837-849.

[30] J.R. Horton, K. Sawada, M. Nishibori, X. Cheng, Structure basis for inhibition of histamine *N*-methyltransferase by diverse drugs, J. Mol. Biol. 353 (2005) 334-344.

[31] K. Rutherford, W.W. Parson, V. Daggett, The histamine *N*-methyltransferase T105I polymorphism affects active site structure and dynamics, Biochemistry 47 (2008) 893-901.

[32] L. Yan, R.E. Galinsky, J.A. Bernstein, S.B. Liggett, R.M. Weinshilboum, Histamine *N*-methyltransferase pharmacogenetics: association of a common functional polymorphism with asthma, Pharmacogenetics 10 (2000) 261-266.

[33] E. García-Martín, J. García-Menaya, B., Sánchez, C. Martínez, R. Rosendo, J.A.G. Agúndez, Polymorphisms of histamine-metabolizing enzymes and clinical manifestations of asthma and allergic rhinitis, Clin. Exp. Allergy 37 (2007) 1175-1182.

[34] M.C. Ledesma, E. García-Martín, H. Alonso-Navarro, C. Martínez, F.J. Jiménez-Jiménez, J. Benito-León, I. Puertas, L. Rubio, T. López-Alburquerque, J.A.G. Agúndez, The nonsynonymous Thr105Ile polymorphism of the histamine *N*-methyltransferase is associated to the risk of developing essential tremor, Neuromol. Med. 10 (2008) 356-361.

[35] V. Palada, J. Terzić, J. Mazzulli, G. Bwala, J. Hagenah, B. Peterlin, A.Y. Hung, C. Klein, D. Krainc, Histamine *N*-methyltransferase Thr105Ile polymorphism is associated with Parkinson's disease, Neurobiol. Aging 33 (2012) 836.e1-836.e3.

[36] L. Kaufman, M. Ayub, J.B. Vincent, The genetic basic of non-syndromic intellectual disability: a review, J. Neurodevelop. Disord. 2 (2000) 182-209.

[37] P.K. Maulik, M.N. Mascarenhas, C.D. Mathers, T. Dua, S. Saxena, Prevalence of intellectual disability: a meta-analysis of population based studies, Res. Dev. Disabil. 32 (2011) 419-436.

[38] A. Heidari, C. Tongsook, R. Najafipour, L. Musante, N. Vasli, M. Garshasbi, H. Hu, K. Mittal, A.J. McNaughton, K. Sritharan, M. Hudson, H. Stehr, S. Talebi, M. Moradi, H. Darvish, M. Arshad Rafiq, H. Mozhdehipanah, A. Rashidinejad, S. Samiei, M. Ghadami, C. Windpassinger, G. Gillessen-Kaesbach, A. Tzschach, I. Ahmed, A. Mikhailov, D.J. Stavropoulos, M.T. Carter, S. Keshavarz, M. Ayub, H. Najmabadi, X. Liu, H.H. Ropers, P. Macheroux, J.B. Vincent, Mutations in the histamine *N*-methyltransferase gene, HNMT, are associated with nonsyndromic autosomal recessive intellectual disability, Hum. Mol. Genet. 24 (2015) 5697-5710.

[39] S.K. Shapiro, D.J. Ehninger, Methods for the analysis and preparation of adenosylmethionine and adenosylhomocysteine, Anal. Biochem. 15 (1966) 323-333.

[40] K.M. Verburg, R.R. Bowsher, D.P. Henry, A new radioenzymatic assay for histamine using purified histamine *N*-methyltransferase, Life Sci. 32 (1983) 2855–2867.

[41] R.R. Bowsher, K.M. Verburg, D.P. Henry, Rat histamine *N*-methyltransferase. Quantification, tissue distribution, purification, and immunologic properties, J. Biol. Chem. 258 (1983) 12215–12220.

[42] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, Biopolymers 89 (2008) 392-400.

[43] L. Whitmore, B.A. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, Nucleic Acids Res. 32 (2004) W668-W673.

[44] W.D. Lienhart, V. Gudipati, M.K. Uhl, A. Binter, S.A. Pulido, R. Saf, K. Zangger, K. Gruber, P. Macheroux, Collapse of the native structure caused by a single amino acid exchange in human NAD(P)H:quinone oxidoreductase, FEBS J. 281 (2014) 4691-4704.

[45] R.L. Gundry, M.Y. White, C.I. Murray, L.A. Kane, Q. Fu, B.A. Stanley, J.E. Van Eyk, Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow, Curr. Protoc. Mol. Biol. Chapter 10 (2009) Unit10.25.

[46] G. De Baerts, J. Van Durme, J. Reumers, S. Maurer-Stroh, P. Vanhee, J. Dopazo, J. Schymkowitz, F. Rousseau, SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants, Nucleic Acids Res. 40 (2012) D935-D939.

[47] J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, L. Serrano, The FoldX web server: an online force field, Nucleic Acids Res. 33 (2005) W382–W388.

[48] V. Parthiban, M.M. Gromiha, D. Schomburg, CUPSAT: prediction of protein stability upon point mutations, Nucleic Acids Res. 34 (suppl 2) (2006) W239-W242.

[49] G. de la Haba, G.A. Jamieson, S.H. Mudd, H.H. Richards, S-adenosylmethionine: the relation of configuration at the sulfonium center to enzymatic reactivity, J. Am. Chem. Soc. 81 (1959) 3975-3980.

[50] YASARA (13.9.8); YASARA Biosciences GmbH, Vienna, Austria, 2012

[51] J. Wang, P. Cieplak, P.A. Kollman, How well does a Restrained Electrostatic Potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. 21 (2000) 1049–1074.

[52] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Jr. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins, nucleic Acids, and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.

[53] U. Essman, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.

[54] A. Jakalian, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, Comput. Chem. 23 (2002) 1623–1641.

[55] M.F. Guest, I.J. Bush, H.J.J. van Dam, P. Sherwood, J.M.H. Thomas, J.H. van Lenthe, R.W.A. Havenith, J. Kendrick, The GAMESS-UK electronic structure package: algorithms, developments and applications, Mol. Phys. 103 (2005) 719-747.

[56] R.W. Hooft, G. Vriend, C. Sander, E.E. Abola, Errors in protein structures, Nature 1996 381, 272.

[57] E. Krieger, J.E. Nielsen, C.A. Spronk, G. J. Vriend, Fast empirical pKa prediction by Ewald summation, Mol. Graph. Model 25 (2006) 481–486.

[58] R.D. King, M.J. Sternberg, Identification and application of the concepts important for accurate and reliable protein secondary structure and prediction, Protein Sci. 5 (1996) 2298-2310.

[59] H.L. Tai, M.Y. Fessing, E.J. Bonten, Y. Yanishevsky, A. d'Azzo, E.Y. Krynetski, W.E. Evans, Enhanced proteasomal degradation of mutant human thiopurine S-methyltransferase (TPMT) in mammalian cells: mechanism for TPMT protein deficiency inherited by TPMT*2, TPMT*3A, TPMT*3B or TPMT*3C, Pharmacogenetics 9 (1999) 641–650.

[60] L. Wang, T.V. Nguyen, R.W. McLaughlin, L.A. Sikkink, M. Ramirez-Alvarado, R.M. Weinshilboum, Human thiopurine S-methyltransferase pharmacogenetics: variant allozyme misfolding and aggresome formation, Proc. Natl Acad. Sci. USA 102 (2005) 9394–9399.

[61] H. Wu, J.R. Horton, K. Battaile, A. Allali-Hassani, F. Martin, H. Zeng, P. Loppnau, M. Vedadi, A. Bochkarev, A.N. Plotnikov, X. Cheng, Structural basis of allele variation of human thiopurine-S-methyltransferase, Proteins 67 (2007) 198–208.

[62] K. Rutherford, V. Daggett, Four human thiopurine S-methyltransferase alleles severely affect protein structure and dynamics, J. Mol. Biol. 379 (2008) 803-814.

[63] K. Rutherford, B.J. Bennion, W.W. Parson, V. Daggett, The 108M polymorph of human catechol O-methyltransferase is prone to deformation at physiological temperatures, Biochemistry, 45 (2006) 2178–2188.

[64] K. Rutherford, V. Daggett, Polymorphisms and disease: hotspots of inactivation in methyltransferases, Trends Biochem. Sci. 35 (2010) 531-538.

Supplementary information

for

Leucine 208 in human histamine *N*-methyltransferase emerges as a hotspot for protein stability rationalizing the role of the L208P variant in intellectual disability

Chanakan Tongsook[$], Johannes Niederhauser[$], Elena Kronegger[$], Grit Straganz[$], and Peter Macheroux[$]

From the [$]Institute of Biochemistry, Graz University of Technology, Graz, Austria

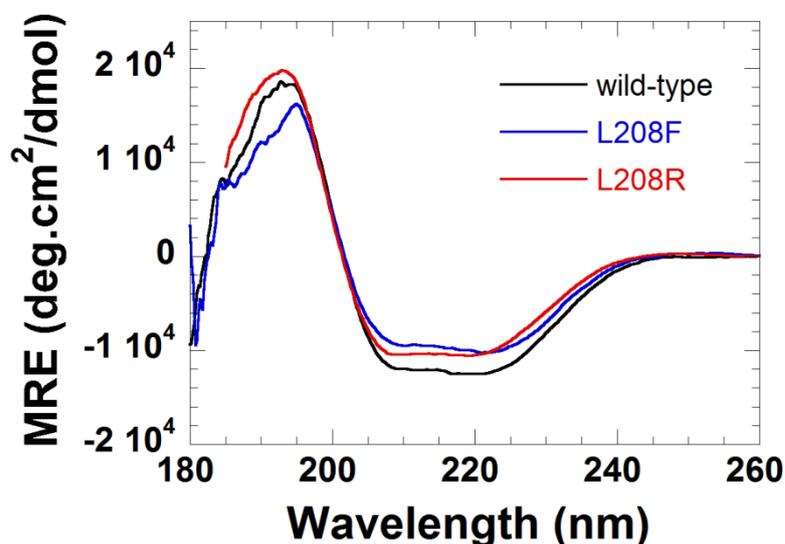Running title: Histamine *N*-methyltransferase

Supplemental Figures

Fig. S1

Fig. S1: Secondary structures of HNMT (wild-type) and L208 variants. Determination of secondary structures of HNMT (wild-type) and L208 variants was performed in 20 mM Tris-acetate buffer, pH 8.0 using a Jasco J715 spectropolarimeter (JASCO Inc., Tokyo, Japan) at 25 °C. CD spectra of HNMT (wild-type; black), L208F (blue), and L208R (red) were presented. Protein concentration was in the range of 0.6-1 mg/mL in a cell of path length 0.02 cm. Buffer spectra have been subtracted.

Fig. S2
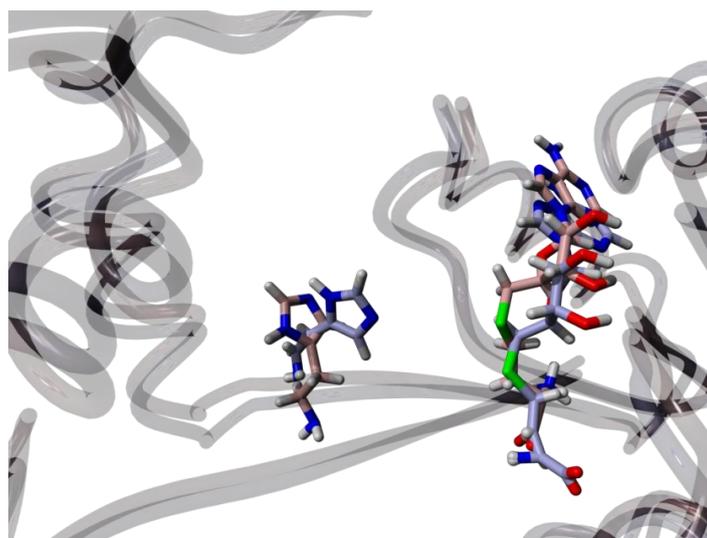
Fig. S2: Overlay of the active sites of the SAH and histamine binding HNMT crystal structure (light pink) and the respective average structure (grey) from one selected 25 ns of simulation at pH7.5 and 310 K. Nitrogens are in blue oxygen is red, hydrogens are shown in white and sulfur is depicted in green.

Fig. S3
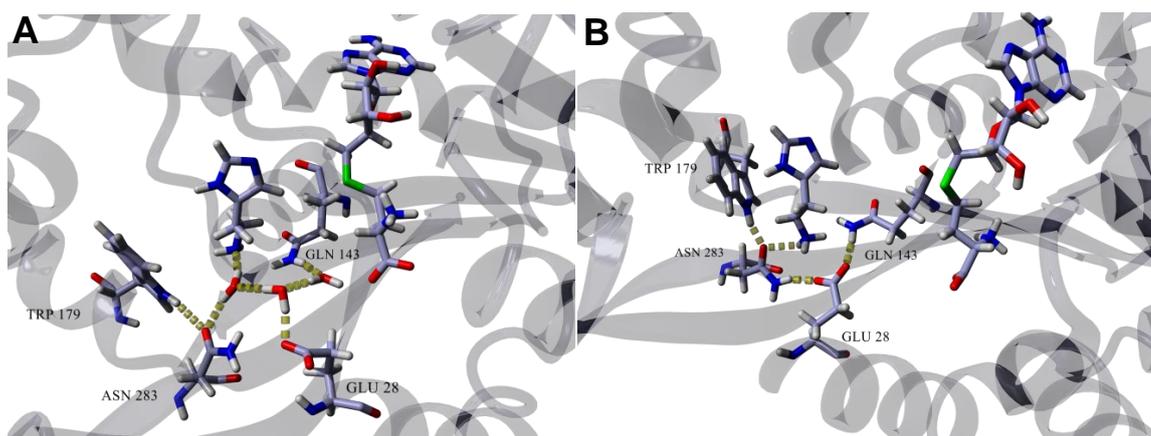
Fig. S3: Active site of HNMT in complex with SAH and histamine showing the hydrogen bonding network (yellow) connecting the alphatic N of histamine with the protein structure. (A) Prototypical, 915 ps snapshot from one MD simulation of the HNMT:SAH:histamine complex at pH 7.5 and 310 K. (B) H-bonding network in the HNMT crystal structure at pH 5.6. Nitrogens are in blue, oxygen is red, hydrogens are shown in white and sulfur is depicted in green.
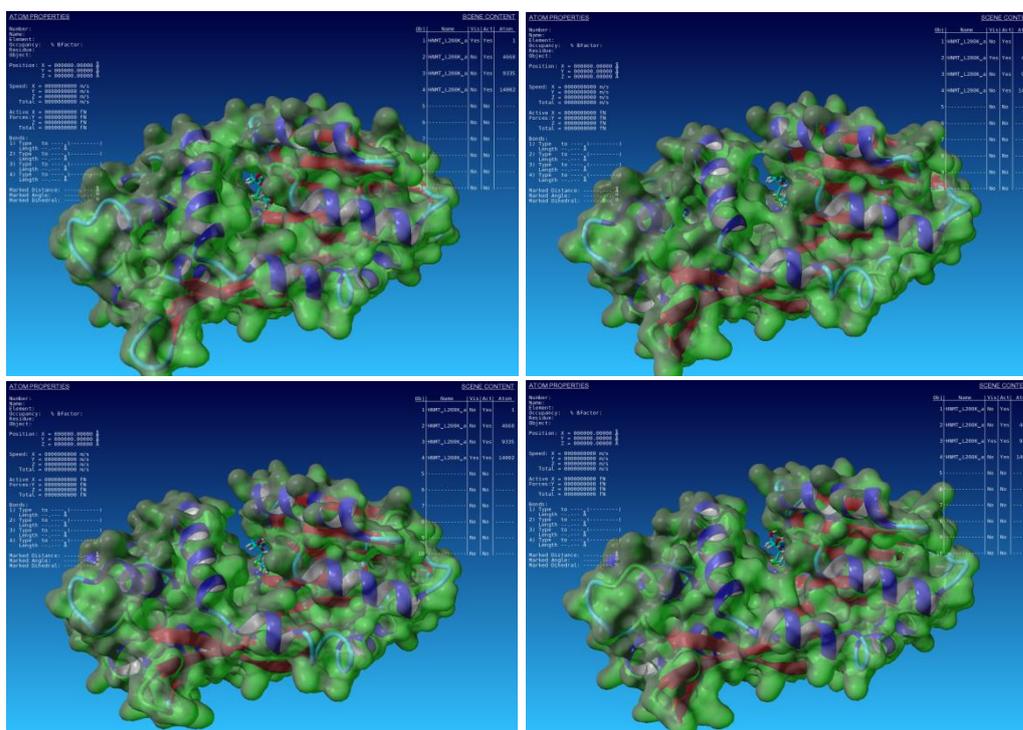
Fig. S4

Fig. S4: Average structures of four 25 ns simulations of the L208K variant of HNMT. Closed (upper left panel), intermediate (upper right and lower left panels) and open (lower right panel) average structures are shown. The molecular surface is given in green, secondary structures are given in blue (α-helix), red (β-sheet) and cyan (coil, turn). SAH carbons are shown in cyan, nitrogens are in blue oxygen is red, hydrogens are shown in white and sulfur is depicted in green.

Supplemental Tables

Table S1: Oligonucleotide primers used in site-directed mutagenesis to generate L208 variant plasmids

Primer	Nucleotide Sequence (5'-3')
HNMT-L208V_fw	GACCTCACTCAGATGGTGGACAACCTAGGGCTTAAG
HNMT-L208V_rev	CTTAAGCCCTAGGTTGTCCACCATCTGAGTGAGGTC
HNMT-L208F_fw	GATGACCTCACTCAGATGTTCGACAACCTAGGGCTTAAG
HNMT-L208F_rev	CTTAAGCCCTAGGTTGTCGAACATCTGAGTGAGGTCATC
HNMT-L208T_fw	GACCTCACTCAGATGACCGACAACCTAGGGCTTAAG
HNMT-L208T_rev	CTTAAGCCCTAGGTTGTCGGTCATCTGAGTGAGGTC
HNMT-L208N_fw	GATGACCTCACTCAGATGAACGACAACCTAGGGCTTAAG
HNMT-L208N_rev	CTTAAGCCCTAGGTTGTCGTTCATCTGAGTGAGGTCATC
HNMT-L208R_fw	GACCTCACTCAGATGCGTGACAACCTAGGGCTTAAG
HNMT-L208R_rev	CTTAAGCCCTAGGTTGTCACGCATCTGAGTGAGGTC
HNMT-L208D_fw	GATGACCTCACTCAGATGGATGACAACCTAGGGCTTAAG
HNMT-L208D_rev	CTTAAGCCCTAGGTTGTCATCCATCTGAGTGAGGTCATC
HNMT-L208K_fw	CAGATGACCTCACTCAGATGAAAGACAACCTAGGGCTTAAG
HNMT-L208K_rev	CTTAAGCCCTAGGTTGTCTTTCATCTGAGTGAGGTCATCTG
HNMT-L208H_fw	GACCTCACTCAGATGCACGACAACCTAGGGCTTAAG
HNMT-L208H_rev	CTTAAGCCCTAGGTTGTCGTGCATCTGAGTGAGGTC

Table S2: **Average fraction of secondary structure elements of residues 200-212**, which constitute the L208 bearing α-helix in the HNMT crystal structure, over the 25 ns time span of the MD simulation. Average helix (H), coil (C), turn (T) content and standard deviation from four MD simulations are given.

	Wild-type	L208V	L208F	L208R	L208K	L208P
H	0.926 ± 0.005	0.894 ± 0.008	0.82 ± 0.028	0.774 ± 0.044	0.711 ± 0.085	0.719 ± 0.019
C	0.055 ± 0.001	0.079 ± 0.001	0.091 ± 0.007	0.125 ± 0.021	0.146 ± 0.027	0.129 ± 0.008
T	0.017 ± 0.005	0.025 ± 0.007	0.08 ± 0.023	0.086 ± 0.017	0.12 ± 0.043	0.128 ± 0.012

Table S3: **Selected average distances and RMSD values extracted from quadruplicate MD simulations in this study are shown**. Average distances of SAH sulfur to the methylation target atom $N_{\varepsilon 2his}$ and to the aliphatic nitrogen N of histamine N_{his} are shown and RMSD values of selected structural features defined by the residue numbers they comprise, are given.

* RMSD values of the protein backbone.

Variant	Distance (Å) S_{SAH} - $N_{\varepsilon 2his}$	Distance (Å) S_{SAH} - N_{his}	Distance (Å) $C\alpha_{His12}$-$C\alpha_{Ser91}$	RMSD* helix 200-212	RMSD* loop 190-199	RMSD* loop 144-148
WT1	10.99 ± 2.1	9.05 ± 1.26	10.83 ± 1.16	0.43 ± 0.08	1.02 ± 0.18	0.4 ± 0.07
WT2	6.16 ± 0.77	9.34 ± 0.71	9.64 ± 1.02	0.46 ± 0.08	0.99 ± 0.12	0.39 ± 0.06
WT3	8.46 ± 2.93	9.3 ± 1.46	12.67 ± 2.35	0.44 ± 0.07	0.94 ± 0.11	0.38 ± 0.08
WT4	7.07 ± 2.27	8.87 ± 0.94	8.86 ± 1.36	0.45 ± 0.07	0.98 ± 0.11	0.46 ± 0.08
L208V 1	8.67 ± 1.36	9.91 ± 1.91	12.89 ± 1.76	0.42 ± 0.07	1 ± 0.16	0.4 ± 0.08
L208V 2	7.07 ± 1.32	10.32 ± 1.18	13.96 ± 3.07	0.45 ± 0.07	0.95 ± 0.14	0.42 ± 0.08
L208V 3	6.52 ± 1.78	9.86 ± 1.54	9.27 ± 1.29	0.46 ± 0.08	1.25 ± 0.29	0.45 ± 0.08
L208V 4	10.84 ± 1.26	9.89 ± 1.22	10.33 ± 1.79	0.45 ± 0.06	0.96 ± 0.16	0.39 ± 0.07
L208F 1	10.79 ± 1.31	9.58 ± 0.98	11.09 ± 0.95	0.63 ± 0.18	1.2 ± 0.14	0.42 ± 0.08
L208F 2	9.3 ± 1.92	9.46 ± 1.61	7.17 ± 1.04	0.66 ± 0.15	1.18 ± 0.21	0.46 ± 0.07
L208F 3	14.43 ± 0.99	9.21 ± 0.75	6.03 ± 0.79	0.68 ± 0.17	1.13 ± 0.16	0.43 ± 0.06
L208F 4	11.14 ± 0.99	12.13 ± 2.32	11.84 ± 1.46	0.63 ± 0.21	0.95 ± 0.13	0.53 ± 0.21
L208R 1	8.01 ± 1.53	9.73 ± 1.5	10.04 ± 1.62	0.77 ± 0.22	0.98 ± 0.1	0.38 ± 0.07
L208R 2	5.93 ± 1.05	9.32 ± 0.97	8.3 ± 0.97	1.04 ± 0.31	0.91 ± 0.1	0.38 ± 0.07
L208R 3	10.26 ± 1.8	11.51 ± 1.51	11.97 ± 1.57	0.75 ± 0.21	1 ± 0.13	0.43 ± 0.09
L208R 4	11.27 ± 3.11	12.24 ± 2.99	14.75 ± 2.03	1.59 ± 0.22	1.1 ± 0.19	0.4 ± 0.09
L208K 1	12.37 ± 2.34	9.48 ± 1.26	9.49 ± 1.72	1.23 ± 0.27	0.95 ± 0.12	0.38 ± 0.07
L208K 2	6.91 ± 1.49	7.58 ± 1.46	9.24 ± 1.11	1.13 ± 0.28	1.03 ± 0.16	0.44 ± 0.11
L208K 3	6.35 ± 1.32	8.86 ± 1.02	10.28 ± 2.42	0.95 ± 0.28	0.94 ± 0.11	0.4 ± 0.09
L208K 4	5.93 ± 1.62	8.74 ± 0.81	9.76 ± 1.06	1.14 ± 0.29	0.93 ± 0.12	0.36 ± 0.07
L208P 1	11.21 ± 2.78	10.67 ± 1.8	12.79 ± 2.13	0.74 ± 0.11	1.1 ± 0.27	0.39 ± 0.08
L208P 2	7.97 ± 1.42	11.12 ± 1.62	12.79 ± 1.96	0.75 ± 0.13	0.96 ± 0.16	0.42 ± 0.07
L208P 3	13.79 ± 1.49	8.86 ± 3.37	10.19 ± 1.88	0.8 ± 0.14	1.02 ± 0.13	0.46 ± 0.13
L208P 4	12.87 ± 3.07	9.9 ± 1.29	9.61 ± 1.74	0.75 ± 0.1	0.95 ± 0.1	0.4 ± 0.07

CHAPTER 3

3. Biochemical characterization and kinetic studies on a putative SAM-dependent methyltransferase (METTL23)

3.1 Disruption of the methyltransferase-like 23 gene *METTL23* causes mild autosomal recessive intellectual disability.

AUTHOR CONTRIBUTIONS

Disruption of the methyltransferase-like 23 gene *METTL23* causes mild autosomal recessive intellectual disability

Marie Bernkopf[1,†], Gerald Webersinke[1,†], Chanakan Tongsook[2], Chintan N. Koyani[3], Muhammad A. Rafiq[4], Muhammad Ayaz[5], Doris Müller[6], Christian Enzinger[7], Muhammad Aslam[5], Farooq Naeem[5,9], Kurt Schmidt[10], Karl Gruber[11], Michael R. Speicher[8], Ernst Malle[3], Peter Macheroux[2], Muhammad Ayub[5,9], John B. Vincent[4,12,13], Christian Windpassinger[8,*] and Hans-Christoph Duba[6]

[1]Laboratory of Molecular Biology and Tumorcytogenetics, Department of Internal Medicine, Krankenhaus der Barmherzigen Schwestern, Linz, Austria, [2]Institute of Biochemistry, Graz University of Technology, Graz, Austria, [3]Institute of Molecular Biology and Biochemistry, Medical University of Graz, Graz, Austria, [4]Molecular Neuropsychiatry and Development (MiND) Lab, The Campbell Family Brain Research Institute, The Centre for Addiction & Mental Health (CAMH), Toronto, ON, Canada, [5]Lahore Institute of Research and Development, Lahore, Punjab Province, Pakistan, [6]Department of Human Genetics, Landes-Frauen und Kinderklinik, Linz, Austria, [7]Department of Neurology and [8]Institute of Human Genetics, Medical University of Graz, Graz, Austria, [9]Division of Developmental Disabilities, Department of Psychiatry, Queen's University, Kingston, ON, Canada, [10]Department of Pharmacology and Toxicology, Karl-Franzens University Graz, Graz, Austria, [11]Institute of Molecular Biosciences, University of Graz, Graz, Austria and [12]Department of Psychiatry and [13]Institute of Medical Science, University of Toronto, Toronto, ON, Canada

We describe the characterization of a gene for mild nonsyndromic autosomal recessive intellectual disability (ID) in two unrelated families, one from Austria, the other from Pakistan. Genome-wide single nucleotide polymorphism microarray analysis enabled us to define a region of homozygosity by descent on chromosome 17q25. Whole-exome sequencing and analysis of this region in an affected individual from the Austrian family identified a 5 bp frameshifting deletion in the *METTL23* gene. By means of Sanger sequencing of *METTL23*, a nonsense mutation was detected in a consanguineous ID family from Pakistan for which homozygosity-by-descent mapping had identified a region on 17q25. Both changes lead to truncation of the putative METTL23 protein, which disrupts the predicted catalytic domain and alters the cellular localization. 3D-modelling of the protein indicates that METTL23 is strongly predicted to function as an *S*-adenosyl-methionine (SAM)-dependent methyltransferase. Expression analysis of METTL23 indicated a strong association with heat shock proteins, which suggests that these may act as a putative substrate for methylation by METTL23. A number of methyltransferases have been described recently in association with ID. Disruption of METTL23 presented here supports the importance of methylation processes for intact neuronal function and brain development.

*To whom correspondence should be addressed at:Institute of Human Genetics, Medical University of Graz, Harrachgasse 21/III, 8010 Graz, Austria. Tel: +43 (0) 3163804114; Email: christian.windpassinger@medunigraz.at
†These authors contributed equally.

123

INTRODUCTION

About 1% of children worldwide are affected by intellectual disability (ID) (1), which can have a devastating impact on many aspects of the lives of the affected individuals, their families and communities, and is a major challenge at the clinical level. The clinical presentation and etiology of ID is complex and has a high degree of heterogeneity, which leads to a poor rate of molecular diagnosis resulting in below satisfactory clinical management. Although recent advances in sequencing technology have accelerated the rate of gene discovery for ID, even where family sizes are small, the majority of ID genes remain undetected. ID can be divided into two groups: nonsyndromic ID is characterized as the only clinical feature in patients, whereas syndromic ID occurs in combination with one or more additional clinical features (2). A recent review suggested that ~40 genes for nonsyndromic autosomal recessive ID (NS-ARID) have been identified to date, but estimates there may be in excess of 2500 autosomal ID genes in total, with the majority being recessive (3). Although the physiological roles of ID genes are heterogeneous, increasing numbers of methyltransferases are being implicated in ID, stressing the essential role that methylation plays in the development and function of the central nervous system. For instance, ID has been linked to mutations in the methyltransferases *NSUN2* (4,5) *EHMT1* (6), *FTSJ1*, (7) and *TRMT1* (8), which have diverse molecular substrates. Here, we report the identification of methyltransferase-like 23 gene (*METTL23*) truncating mutations in two nonsyndromic ID families, one Austrian and the other Pakistani family, and discuss potential substrates for this protein.

RESULTS

The Austrian ID family (LFKK1) has five children. In four affected siblings—two male and two female—a delay in developmental milestones was noted after the first year of life, and ID was diagnosed, whereas one sister was unaffected. All four affected siblings have a relatively high level of functioning, and a detailed clinical description is given in Supplementary Material.

The phenotypes and the degree of ID in the two affected girls of the Pakistani family (PK31) had a striking similarity to the Austrian family (see Supplementary Material for detailed clinical descriptions). Photographs for the two affected Pakistani girls are shown in Figure 1. Magnetic resonance imaging (MRI) was available for individual PK31-II:1 and II:22. For II:2, MRI showed increased volume of the subcallosal gray matter, and decreased delineation of the basal ganglia-regions strongly implicated in affect regulation, which may correlate with the aggression documented for this individual (Fig. 1).

For Austrian family LFKK1, a single large homozygosity-by-descent (HBD) region of about 4.8 Mb on chromosome 17q25.1–q25.3 was identified (Fig. 2); this area contained 107 genes. No other HBD stretch present in all affected family members and not in unaffected members could be identified. One possible candidate gene within the HBD region on 17q was *TSEN54*, which encodes one subunit from the tRNA splicing endonuclease complex.[21] Mutation in this gene causes pontocerebellar hypoplasia type 2A (OMIM #277470) and type 4 (OMIM #225753)-neurodegenerative autosomal recessive disorders with chorea, dystonia or spasticity (9,10). However, Sanger sequencing of all exons of *TSEN54* identified no mutation. After whole-exome sequencing for LFKK1 individual II:5, selection for homozygous exonic mutations in the 17q HBD region and exclusion of entries of the dbSNP131 database, only two genetic changes remained: A synonymous single nucleotide variant (SNV) (chr17: 74392666G>A; NM_022066.3:c.2352C>T; p.Ala784Ala) in the *UBE2O* (ubiquitin-conjugating enzyme E2O) gene and a 5 bp frameshift deletion (chr17:74729256_74729260delAAGAT; NM_001206983.1: c.281_285delAAGAT; p. (Gln94Hisfs*6) resulting in a predicted 98 amino acids METTL23 protein) in exon 3 of the C17orf95 (*METTL23*, methyltransferase-like protein 23) gene (Fig. 2). As the synonymous SNV is a known SNP (rs145605062; dbSNP build 138; MAF = 0.58%), it is not predicted to affect splicing (analyzed through BDGP: www.fruitfly.org), and therefore seems unlikely to be disease related. In contrast, small indels have frequently been associated with severe, sporadic nonsyndromic ID (11), and therefore we focused on investigating *METTL23* further. In addition to this, all known genes associated with NS-ARID were screened for compound heterozygosity in our exome data (12). The homozygous *METTL23* variant was confirmed by Sanger sequencing in all four affected family members, whereas the unaffected sister and the parents were all heterozygous. For controls, we analyzed 200 unrelated, unaffected Austrian individuals for variants in *METTL23* coding exons (exons 2–5) by denaturing high-performance liquid chromatography (DHPLC). In three samples, we identified the known SNV rs138247613 (c.G496A, p.D166N) in exon 5 with a comparable frequency of 2% for the heterozygous G/A genotype in the European population. However, we detected no other variants in this cohort.

In Pakistani family PK31, a 5.1 Mb HBD region was observed on 17q25.1–q25.3, flanked by SNPs rs513643 and rs4313838 (chr17:72 599 418–77 698 582; hg19) (Fig. 2). Because of the overlap of HBD with the Austrian family and the similar phenotypes, we Sanger sequenced the *METTL23* coding exons in family PK31. We identified a nonsense mutation in exon 4, which segregated with the disease (chr17:74729449C>T; NM_001080510.3: c.397C>T; p.(Gln133*) (Fig. 2). We screened 300 healthy Pakistani individuals by restriction fragment length polymorphism-PCR with the enzyme *Mse*I, but did not find this allele. Neither this variant, nor the 5 bp deletion were present in the Exome Variant Server, NHLBI GO Exome Sequencing Project, Seattle, WA, USA (http://eversusgs.wa shington.edu/EVS/) (accessed Dec 2013) or in 1000 Genomes (http://browser.1000genomes.org/index.html).

Marker segregation for the two families was confirmed using microsatellites markers as well as the *METTL23* variants (Fig. 2), and linkage analysis using SimWalk2 (version 2.91) (13) under a recessive model of inheritance gave a maximum LOD score of 2.66.

There are seven different *METTL23* mRNA-transcripts reported (UCSC Genome Browser), six coding and one non-coding. Transcript variants 1, 2 and 3 encode the protein isoform 1 (190 amino acids) and the transcripts 4, 5 and 6 for the shorter protein isoform 2 (123 amino acid). The mRNAs have different length because of differing sizes of the non-coding first exon or because of alternative splicing of exon 2. In order to get a global overview of the expression of these mRNAs, we analyzed the 20 tissue samples from the Human Total RNA
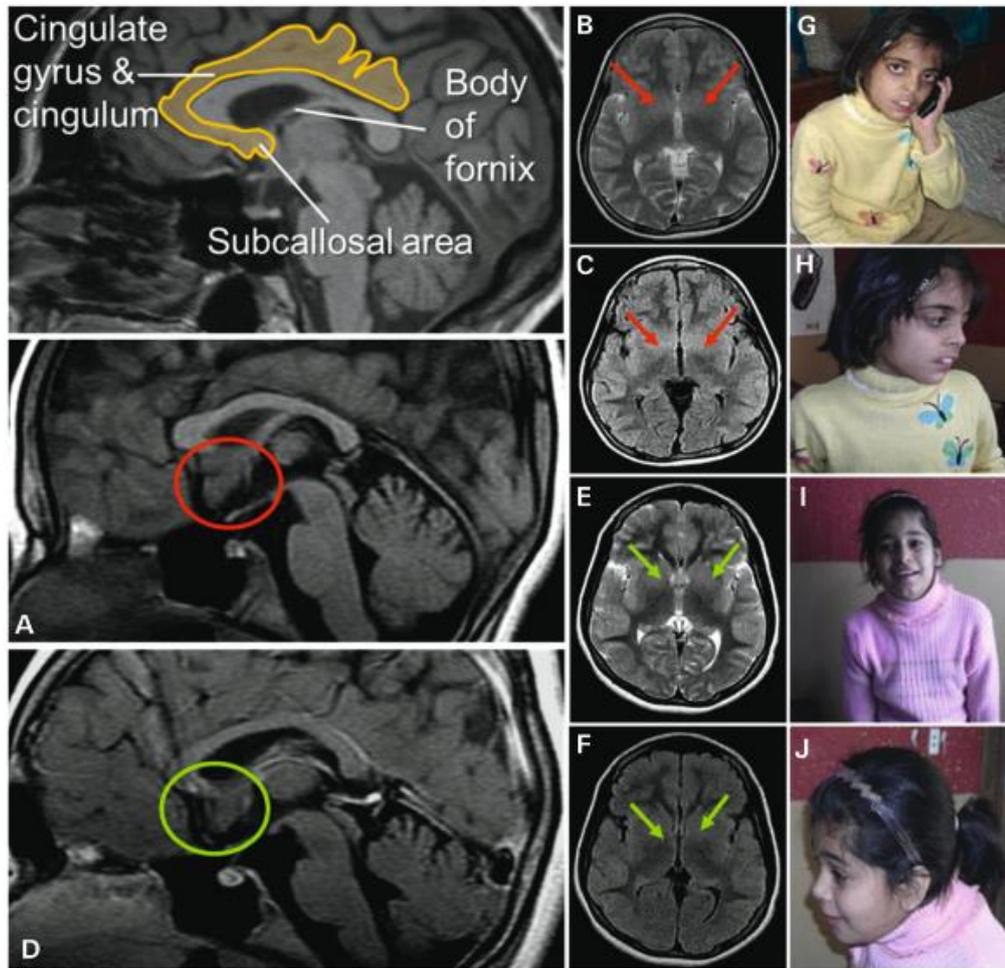
Figure 1. Magnetic resonance images: Two independent experts in neuroimaging (Christian Enzinger, Franz Fazekas) inspected the MRI scans blinded to clinical information and detected increased volume of the subcallosal area (SA), which was particularly evident in patient PK31 II:2 (**A**: red circles; in comparison with patient II:1 in **D**: green circles). On axial T2-weighted and FLAIR sequences, this led to decreased identifiability of the basal ganglia (**B**, **C** red arrows (II:2); compared with **E** and **F** (II:1), green arrows). The SA is commonly attributed to the limbic system, belongs to a phylogenetic old-brain area, is located in the vicinity of strategically important areas such as the hypothalamus and mamillar bodies and has been implicated in affect regulation. Interestingly, clinical information revealed that patient II:2 had demonstrated aggressive behavior and disturbed impulse control. Photos of PK31 II:2 (**G** and **H**) and II:1 (**I** and **J**) are shown.

Panel (Clontech Takara, Madison, WI) by quantitative real-time PCR. We found that *METTL23* is expressed at low levels but ubiquitously in all tissues analyzed, in comparison with the beta-2-microglobulin (*B2M*) and *HPRT1*-mRNA-amplification-levels. Both long- and short-isoform-encoding transcripts of *METTL23* are expressed at very similar levels in all tissues (see Supplementary Material).

Quantitative RT-PCR using mRNA from lymphoblasts from PK31 family members showed similar levels of mRNA for the two main isoforms of *METTL23* NM_001206983.1 (isoform 1) and NM_001206986.1 (isoform 2) in the affected siblings (II:1 and II:2) and a gender-matched unaffected individual, as well as in the tissues from the RNA panel (lymphocyte, brain

and testis). Thus, nonsense-mediated mRNA decay does not appear to be involved in the mechanism of pathogenicity (see Supplementary Material).

To confirm that the nonsense mutation in *METTL23* from patient PK31 II:1 definitively results in a truncated METTL23 protein (position 1−132 amino acids, lacking the catalytic domain) a polyclonal antiserum (raised in our lab, recognizing full-length METTL23 protein expressed in *Escherichia coli*, Fig. 3A) was used. Immunoprecipitation of METTL23 from lymphoblastoid cells and subsequent detection with a sequence-specific (C17orf95) antibody [cross-reacting with the more C-terminal portion (position 137−166 amino acids) of intact METTL23] revealed a strong immunoreactive signal of 21.5 kDa in controls
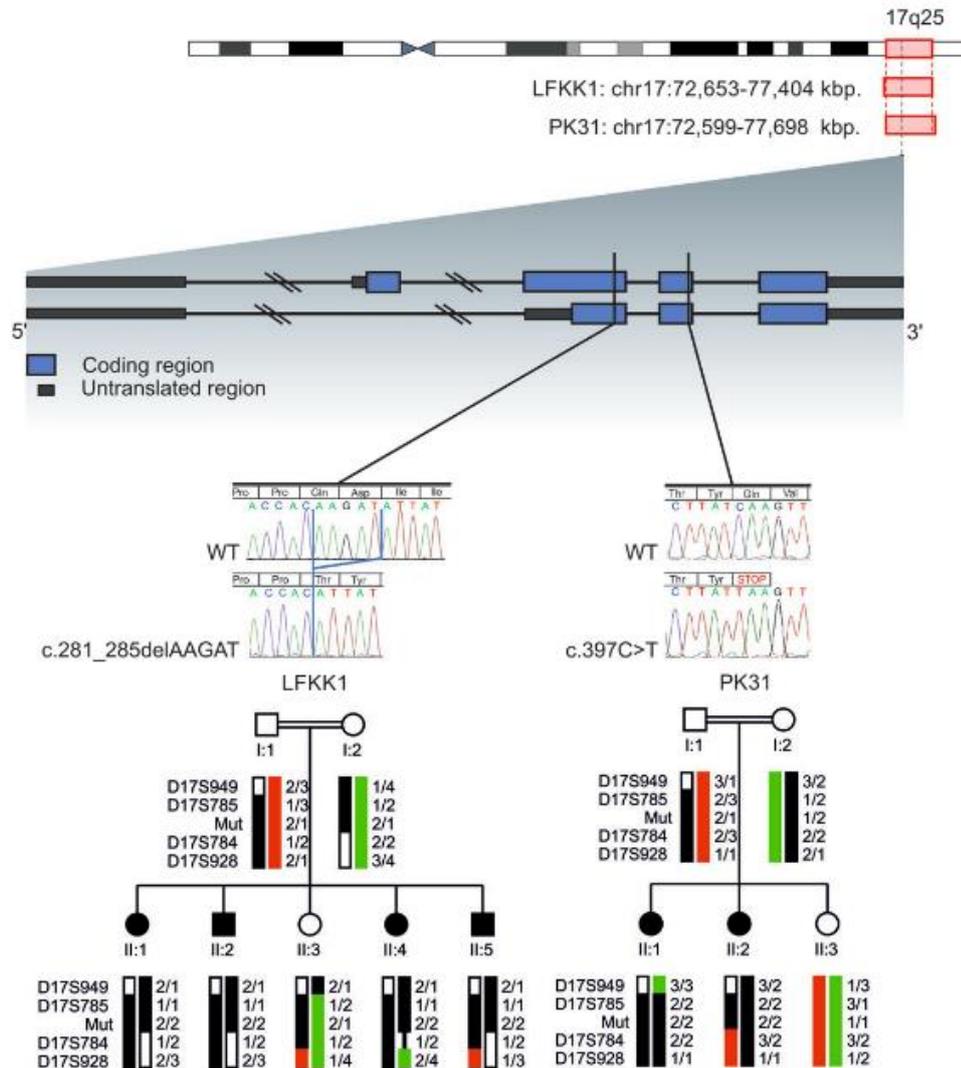
Figure 2. Ideogram of chromosome 17 is shown. Localization, indicated by red rectangles and base pair positions (referring to hg19) defining the homozygous regions of family FLKK1 and PK31 are given. The exon/intron structure of *METTL23* is shown: blue blocks represent coding regions and gray blocks represent untranslated regions, introns are shown as dark lines. Black vertical lines indicate the position of the mutations for LFKK1 and PK31. Electropherograms show the homozygous familial mutations in *METTL23* for LFKK1 and PK31 compared with WT sequences. Haplotypes and simplified pedigrees of LFKK1 and PK31 are shown. The disease haplotype is indicated by black bars. All alleles are recoded. Mutation status (mut) of all tested individuals is indicated by 1 for WT and 2 for mutated.

but not in the patient (Fig. 3C). These observations confirm that the PK31 nonsense mutation does result in the loss of C-terminal METTL23 protein sequence.

Based on sequence comparison, METTL23 appears to be a putative methyltransferase. Accordingly, we modeled the 3D structure of human METTL23. The final model has optimal confidence scores along the entire sequence length (190 residues, i.e. isoform 1) and is shown in Figure 4A. Based on this homology model, possible ligands and their binding sites in METTL23 were predicted using the server 3DLigandSite (14).

In total, 22 similar structures were identified in the PDB, which all contain either *S*-adenosyl-methionine (SAM) or *S*-adenosyl-homocysteine (SAH) as a cofactor. Twenty-five residues of METTL23 were predicted to build up the SAM/SAH-binding site (colored green in Fig. 4A), ~50% of which were completely conserved among this set of SAM/SAH-binding proteins. According to this structural bioinformatics analysis, it is very likely that METTL23 isoform 1 adopts a 3D fold typical for methyltransferases and contains a binding site for the methylation cofactor SAM. Thus, we conclude that METTL23 is
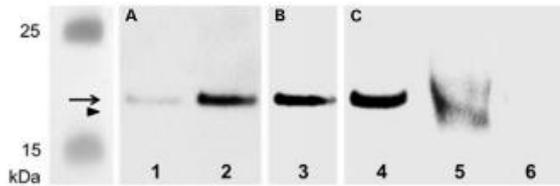
Figure 3. Western blot for METTL23 protein: Purified METTL23-His8-tag protein (expressed in *E. coli*) was loaded in lane 1 (50 ng) and 2–4 (100 ng). METTL23 proteins were immunoprecipitated from lymphoblastoid cells from healthy control (lane 5) and METTL23-mutant patient PK31 II:1 (lane 6) using polyclonal anti-METTL23. Immunoreactive bands were visualized using (**A**) rabbit polyclonal anti-METTL23-His8-tag antiserum, (**B**) rabbit polyclonal anti-His-tag (C-terminal) antibody and (**C**) sequence-specific anti-METTL23 (C17orf95) antibody as primary antibodies. Arrow (METTL23-His8-tag protein, 22.6 kDa); arrowhead (immunoprecipitated METTL23 protein, 21.5 kDa).
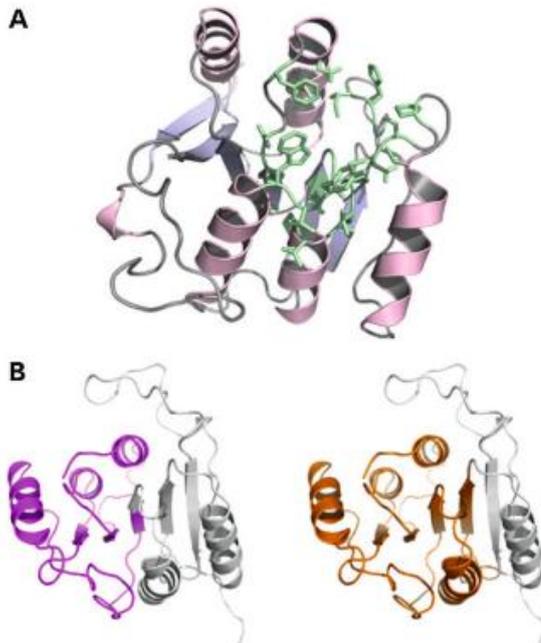

Figure 5. Expression and purification of METTL23 in *E. coli* in the absence and presence of chaperone coexpression. (**A**)The expression of METTL23 in the absence of chaperone coexpression in *E. coli*. METTL23 was found solely in inclusion bodies (left lane), whereas the supernatant was devoid of METTL23 (right lane). (**B**) The purification of METTL23 coexpressed with GroEL-GroES by Ni-NTA affinity and size exclusion chromatography. The middle lane was supernatant of METTL23 coexpressed with GroEL-GroES chaperones after cell lysis. The right lane is a fraction of METTL23 after Ni-NTA affinity and size exclusion chromatography.

Figure 4. (**A**) Homology model of human METTL23 generated using the Phyre2 server (23). Amino acid residues predicted to build an SAM/SAH-binding site are shown in green. (**B**) Cartoon representations of the protein model with purple and orange colored portions representing the extent of the truncation variants (left: 1–98; right: 1–132). The parts missing in these variants are shown in gray.

but was insoluble and formed inclusion bodies (Fig. 5A). We then generated the protein in the presence of chaperones (e.g. GroEL) yielding soluble METTL23. Interestingly, affinity purification of his8-tagged METTL23 and size exclusion chromatography showed that the protein co-eluted with the chaperone (SDS–PAGE: fig. 5B). Attempts to release METTL23 from the chaperone, for example by addition of ATP, were unsuccessful indicating that METTL23 is tightly bound. This finding is in accordance with the suspected association of METTL23 with the endoplasmic reticulum (ER) membrane (as shown in Fig. 6). On the other hand, Cloutier and coworkers (15) recently reported that several human methyltransferases interact with molecular chaperones and regulate their activity by methylation of a conserved lysine residue.

The 5 bp deletion in exon 3 is predicted to lead to a 98 amino acids-protein isoform 1 and 31 amino acids-protein isoform 2, whereas the nonsense mutation in exon 4 is predicted to lead to 132 and 65 amino acids-long gene products, for isoform 1 and 2, respectively. Wild-type-isoform 1 appears to be localized at or in the ER-membrane, whereas isoform 2 is mainly localized in the nucleoplasm in all analyzed cells (Fig. 6). These findings support the predicted transmembrane domain in isoform 1 but not isoform 2. However, this overexpression of the mutant protein leads to the formation of protein aggregates in the cytoplasm in all cells, for both isoform 1 and isoform 2 (Fig. 6).

indeed functioning as a methyltransferase. Both truncation variants, on the other hand, are predicted to lack a significant portion of the core fold (central β-sheet) of the protein (Fig. 4B). Therefore, it is very likely that these variants are not properly folded and thus do not show methyltransferase activity.

To obtain more information on the catalytic properties of METTL23, we heterologously expressed the gene in *E. coli* BL21 host cells. The protein was produced in good yield
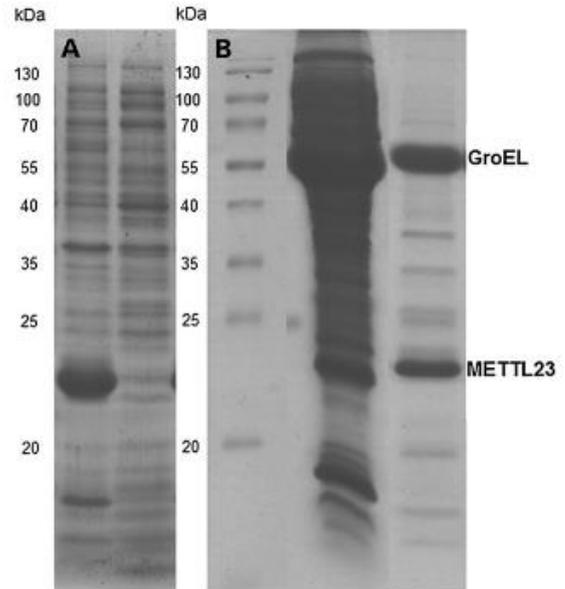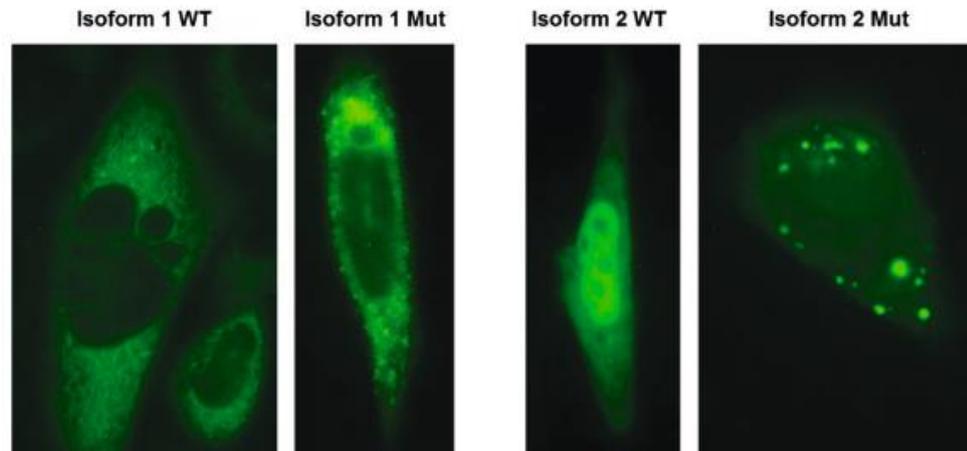
127

Figure 6. Subcellular localization of METTL23-GFP fusion proteins for WT isoform 1, WT isoform 2 and the corresponding 5 bp-deletion mutant proteins are shown in CHO cells after transient transfection. The WT of isoform 1 is predominantly located in the ER and to a lower extent in the nucleus. The corresponding mutant protein shows a similar distribution but additionally forms highly fluorescent protein aggregates. The WT protein of isoform 2 appears to be located in comparable concentrations in the nucleus and in the cytoplasm, whereas the corresponding mutant protein located in the cytoplasm forms numerous concentric aggregates. Green fluorescent signal could also be observed in the nucleus.

DISCUSSION

The findings presented here add support for *METTL23* as an important gene for ARID. Interestingly, a homozygous 4-bp deletion (c.169_172delCACT) in the *METTL23* gene was very recently described in a consanguineous family of Arab origin (16). In contrast to our two families, Reiff *et al.* noted some dysmorphic features in their affected family members. Clinical findings common to all affected individuals included flat occiput, large eyes, depressed nasal bridge, short upturned nose, long philtrum, thin lips and incomplete syndactyly. The cognitive impairment was classified as 'severe, with autistic symptoms' in one individual and as 'moderate' in two other affected family members. Altogether, this suggests that truncating mutations can lead to nonsyndromic ARID as well as ARID with dysmorphic or syndromic features. The cognitive impairment also seems to be milder in families LFKK1 and PK31, as the affected individuals of both families have a relatively high degree of comprehension, including long sentences, and have at least very simple reading, writing and math abilities. There was no evidence of autistic traits. As the three mutations found thus far in *METTL23* are different and present with different severity of ID, a genotype/phenotype correlation seems probable. Importantly, the mutation reported by Reiff *et al.* (16) is located in an exon that is only coding in isoform 1 (and is in a 5′ untranslated region in isoform 2), whereas the mutations in LFKK1 and PK31 are truncating for both isoforms. This perhaps represents something of a paradox, as the symptoms associated with the mutation that is exclusive to isoform 1 are severer.

Localization studies suggest that wild-type (WT) METTL23 isoform 1 is likely to function in the ER, whereas isoform 2 has a nuclear localization. Expression of the mutant (truncated) protein appeared to result in cytoplasmic protein aggregates. It is likely that these aggregates are a secondary effect related to the hydrophobicity and overexpression of the protein, rather than being the primary effect of the mutation. Protein aggregation and inclusion bodies are frequently associated with neurodegeneration; however, there is no evidence of a gain-of-function associated with the two *METTL23* mutations identified as a pathogenic mechanism in these two families. It is most likely that the neurodevelopmental phenotype in our two families is a result of a deficit in the enzymatic function of the methyltransferase.

Methyltransferases catalyze the transfer of a methyl group to diverse substrates, including nucleic acids (DNA and RNA), proteins and lipids. They have been grouped into superfamilies according to their structural properties. Today, 208 proteins in the human proteome, representing ∼0.9% of human gene products, are already characterized as putative methyltransferases, and ∼30% of these have already been linked to diseases (17). By modifying their targets, they essentially influence multiple cellular regulatory mechanisms. Via epigenetic effects, methyltransferases are involved in tissue differentiation and proliferation, and have therefore been associated with a variety of diseases including cancer and neuropsychiatric disorders (18–20).

Many genes encoding demethylases and methyltransferases have now been implicated in syndromic or nonsyndromic forms of ID. These include histone demethylases, histone methyltransferases, tRNA and rRNA methyltransferases (4–8). Epigenetic events influenced by methylases and demethylases have been shown to have a substantial involvement in development and function of the central nervous system, and mutations in the underlying genes may result in cognitive impairment and intellectual disabilities. In the case of METTL23, it appears likely that abnormal methylation may also result in disrupted neuronal development via other (non-epigenetic) pathways. The METTL group of genes is likely to contain many more ID-related genes yet to be discovered.

A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity (15). Our data suggest tight binding of METTL23 to a chaperone protein, GroEL, which is highly homologous to human heat shock protein HSP60 (known to be mutated

in hereditary spastic paraplegia; MIM 605280). Thus, studies of the interaction of METTL23 with various chaperones may eventually lead to the discovery of the cognate substrate(s) of the enzyme. Currently, we are investigating this putative function of METTL23 and developing heterologous expression systems that will yield soluble protein for further characterization of its enzymatic properties.

In summary, we present two different families with *METTL23* mutations, one of them located in Europe, i.e. Austria, the other in Pakistan, which extends the geographical distribution and prevalence of *METTL23* as a gene affecting cognitive function. In fact, our study suggests that the prevalence of mutations in this gene in individuals with ID may be relatively high. Our study adds to the known phenotypic consequences of mutations in *METTL23*. In contrast to the study by Reiff and coworkers (16), our mutations resulted in a nonsyndromic and mild form of ID. We also describe MRI results, which revealed some minor abnormalities in affected individuals. It remains to be determined whether METTL23 isoform 2, as predicted, has a physiological role. Since the mutation reported in Reiff and coworkers (16) only disrupts isoform 1, yet the resulting phenotype is more severe, we conclude that the function of isoform 2 may be restricted. Further functional characterization of the different METTL23 protein isoforms is recommended.

MATERIALS AND METHODS

Family ascertainment

We ascertained an Austrian ID family (LFKK1) with five children, through the Human Genetics Department of the Landes-Frauen und Kinderklinik, Linz. In four siblings—two male and two female—mild, nonsyndromic ID was diagnosed, whereas one sister was unaffected. Consanguinity was not known, however both parents originated from the same village from a mountainous region, which has been geographically isolated throughout previous centuries. We performed a detailed clinical genetic examination of the four affected and the non-affected siblings, and blood was drawn for genetic investigations. Photographs of the individuals were assessed for dysmorphic features by several experienced clinical geneticists and pediatricians, and medical records were studied. The family from Pakistan (PK31) in which the parents were first cousins was recruited through the Lahore Institute of Research and Development, as part of a study researching ID among consanguineous families. The family is located within Lahore, in Punjab Province, and presented with two girls affected with mild ID, and a third female sibling who is unaffected.

The study was approved by the respective ethics committees, and conducted according to the Declaration of Helsinki, and written informed consent were obtained from all included members from the Austrian and Pakistani families. Blood was drawn from family members and DNA extracted by standard procedures.

Single nucleotide polymorphism microarray analysis

For the Austrian family (LLKK1), DNAs from individuals II:1–II:5 were genotyped using Affymetrix 6.0 SNP microarrays, and data were analyzed using Chromosome Analysis Suite (ChAS) version 1.2. For the Pakistani family, PK31, both affected individuals (II:1, II:2) and unaffected sibling (II:3) were genotyped

using Affymetrix 500K NspI microarrays (Affymetrix; Santa Clara, CA, USA). Homozygosity analysis was performed using dCHIP (http://biosun1.harvard.edu/complab/dchip; 21) and HomozygosityMapper (http://www.homozygositymapper.org; 22).

Whole-exome sequencing

Whole-exome sequencing of one ID affected family member (II:5) from family LFKK1 was performed by Macrogen Korea Inc. on the Illumina HiSeq1000 platform using Illumina TruSeq Exome Enrichment. Analysis was performed using the Macrogen exome sequencing standard analysis pipeline, and variants were filtered for homozygous exonic and splice site variants within the 17q HBD region, and excluding known SNPs in the dbSNP131 database. Sequence validation was performed by Sanger sequencing using standard procedures.

Protein 3D modeling

We used the Phyre2 server (23) for modeling the 3D structure of METTL23. The software chose Protein Data Bank (PDB) entries 4LG1, 4LEC and 3BZB as templates. These proteins exhibit 19–26% sequence identity to METTL23 and are annotated as human methyltransferase-like protein 21A (4LEC), and 21D (4LG1) as well as an uncharacterized protein CMQ451C from Cyanidioschyzonmerolae (3BZB). We then used the 3DLigandSite server (14) to predict potential ligands and their binding sites in METTL23.

EBV transformation and culture of lymphoblasts

Patient and control blood samples collected in acid-citrate-dextrose BD Vacutainer® blood tubes (*Becton*, Dickinson and Company) were diluted (1:1) with Roswell Park Memorial Institute medium, and white blood cells separated using ACCUSPIN tubes (Sigma), and transformed with Epstein-Barr virus.

RNA extraction, cDNA synthesis and quantitative RT-PCR

RNA was extracted from lymphoblast cells using the Trizol method. cDNA was synthesized through reverse transcription of 1 μg of RNA using Superscript III™ Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) and random hexamers according to manufacturer's guidelines.

mRNA from Leukocytes was extracted from peripheral blood samples using the Chemagic Magnetic Separation Module I (Perkin Elmer, Baesweiler, Germany). cDNA was synthesized through reverse transcription using M-MLV Reverse Transcriptase (Promega, Madison, WI, USA) and random hexamers according to manufacturer's guidelines. Primers were designed to amplify the coding DNA sequence of the *METTL23* gene, using the primer express software (Applied Biosystems Inc., Foster city, CA, USA) and their details are given in Supplementary Material, Table S1. The Primer-Probe-Sets for *B2M* (Life Technologies, Carlsbad, CA, USA) and *HPRT1* (Life Technologies) were used as an internal reference for all the runs. For *METTL23* isoform-specific qRT-PCR, primers and a FAM-dyed probe were designed. In addition, cDNATotal RNA from a multitissue panel (Clontech) was reverse transcribed into cDNA duplicates and used for qRT-PCR. PCR was performed in quadruplicates on the Roche LightCycler 1.0 System (Roche Diagnostics, Basel, Switzerland)

using the LightCycler FastStart DNA Master HybProbe Assay (Roche) T. The C_t for all reactions was calculated automatically by the LightCycler® Analysis software version 3.5. (Roche Diagnostics). Due to the variability among the internal controls compared with the C_t values for *METTL23* isoform 1 and *METTL23* isoform 2 for the 20 analyzed tissue samples, a relative normalization analysis was not performed.

Immunoprecipitation of WT and mutant METTL23

Lymphoblastoid cells were homogenized in RIPA buffer (50 mmol/l Tris−HCl, pH 7.5, 300 mmol/l NaCl, 5 mmol/l EDTA, 50 mmol/l NaF, 0.1%, v/v, Triton X-100, 0.02%, w/v, NaN$_3$, pH 7.4) containing a protease inhibitor cocktail tablet (Sigma−Aldrich). Cell lysates were centrifuged at 10 000 rpm (4°C, 10 min) to pellet debris and were equilibrated against immunoprecipitation buffer containing (in mmol/l): 50 Tris−HCl (pH 8.0), 10 MgCl$_2$ and 150 NaCl. Lysates containing equal protein concentrations (500 μg) were mixed with 10 μl rabbit polyclonal anti-METTL23 antiserum (raised in the own lab against full-length METTL23 protein (BMWF-2013-10-22T11:53:23) containing a His8-tag, expressed in *E. coli*, see Supplementary Material part) for 2 h at 4°C. Immune complexes were precipitated by mixing 20 μl of protein A/G plus Agarose (Santa Cruz Biotechnology, Heidelberg, Germany) overnight (4°C with shaking). Immune complexes were pelleted by centrifugation at 10 000 rpm (4°C, 1 min). After washing three times with RIPA buffer, immune complexes were resuspended in 40 μl of 4× NuPAGE LDS sample buffer and heated (70°C, 10 min). Western blot analyses were performed as described below.

Western blot analysis

After protein estimation using the Lowry method, indicated concentrations of purified METTL23-His8-tag protein (22.6 kDa) were added to 10 μl of 4× NuPAGE LDS sample buffer containing 2 μl sample reducing agent (Invitrogen, Austria) and heated (70°C, 10 min). Purified METTL23-His8-tag protein and immunoprecipitates were separated by electrophoresis on 12% Bis−Tris gel and transferred to nitrocellulose membranes. Membranes were blocked with 5% (w/v) non-fat milk in TBST (Tris-buffered saline containing Tween 20) (25°C, 2 h) and incubated with either (i) rabbit polyclonal anti-METTL23 antiserum (1:2000 in 5%, w/v, BSA), (ii) rabbit polyclonal anti-His-tag (C-terminal) antibody (1:1000 in 5%, w/v, BSA, Relia Tech, Germany) and (iii) rabbit sequence-specific anti-METTL23 (C17orf95) antibody (1:300 in 5%, w/v, BSA, Abgent, Germany, raised against a synthetic peptide; position 137−166 amino acids) (4°C, overnight). Immunoreactive bands were visualized with either Clean-Blot IP Detection Kit [horseradish peroxidase (HRP)] (1:100 000 in 5%, w/v, non-fat milk in TBST, Thermo scientific, Austria) or HRP-conjugated goat anti-rabbit IgG (1:100 000 in 5%, w/v, non-fat milk in TBST) (25°C, 2 h) followed by Super Signal West Pico Chemiluminescent substrate (Thermo Scientific, Austria) and developed by Bio-Rad ChemiDoc MP Imaging System.

Expression and purification of METTL23

Full-length METTL23 including a C-terminal hexa-histidine tag was synthesized (Life Technology™) and inserted into the *Nde*I/*Xho*I restriction site of pMA-T. One shot® Top10 *E. coli*

harboring the plasmid were cultured in 100 ml LB medium containing 100 μg/ml ampicillin to propagate the plasmid. Extraction and purification of the plasmid was performed according to the protocol of NucleoSpin® Plasmid (Macherey-Nagel, Darmstadt, Germany). The isolated plasmid was digested with *Nde*I and *Xho*I to release METTL23. The insert was then cloned into pET21a (mettl23-pET21a) and used for transformation of competent *E. coli* BL21 (DE3). Expression of the gene was induced by addition of 0.1 mM IPTG at an OD$_{600}$ ∼1 at 20°C for 14 h. Coexpression of GroEL-GroES chaperones was induced by addition of 1 mg/ml of L-arabinose.

Purification of METTL23 was achieved by Ni-NTA affinity (5-ml Ni-Sepharose™ High Performance HisTrap™ HP column) and size exclusion chromatography (Superdex 200 prep grade column). For the latter step, the protein solution was applied to the column after equilibration with 50 mM Tris−HCl buffer, pH 8.0, containing 100 mM NaCl. Protein fractions were pooled according to the absorption at 280 nn. Aliquots of each fraction were analyzed by SDS−PAGE (staining with Coomassie Brilliant Blue R-250) to monitor the progress of purification.

METTL23 wild type and mutant protein localization

To study the cellular localization of METTL23 WT isoform 1, WT isoform 2 and the 5 bp-deletion mutant proteins, we fused the GFP-peptide to the C-terminal end using the GFP Fusion TOPO TA Expression Kit (Invitrogen). For transfection, we used 6.5 μg of DNA from WT and mutant constructs and added OptiMEM I Reduced Serum Medium (Life Technologies) and FuGENE (R) HD Transfection Reagent (Promega) to CHO-K cells. The transfection was performed in duplicates, and 26 h later analyzed using inverted fluorescence microscopy (TE2000, Nikon Corporation).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

REFERENCES

1. Maulik, P.K., Mascarenhas, M.N., Mathers, C.D., Dua, T. and Saxena, S. (2011) Prevalence of intellectual disability: a meta-analysis of population-based studies. *Res. Dev. Disabil.*, **32**, 419–436.
2. Kaufman, L., Ayub, M. and Vincent, J.B. (2010) The genetic basis of non-syndromic intellectual disability: a review. *J. Neurodev. Disord.*, **2**, 182–209.
3. Musante, L. and Ropers, H.H. (2014) Genetics of recessive cognitive disorders. *Trends Genet.*, **30**, 32–39.
4. Khan, M.A., Rafiq, M.A., Noor, A., Hussain, S., Flores, J.V., Rupp, V., Vincent, A.K., Malli, R., Ali, G., Khan, F.S. *et al.* (2012) Mutation in NSUN2, which encodes an RNA methltransferase, causes autosomal-recessive intellectual disability. *Am. J. Hum. Genet.*, **90**, 856–863.
5. Abbasi-Moheb, L., Mertel, S., Gonsior, M., Nouri-Vahid, L., Kahrizi, K., Cirak, S., Wieczorek, D., Motazacker, M.M., Esmaeeli-Nieh, S., Cremer, K. *et al.* (2012) Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am. J. Hum. Genet.*, **90**, 847–855.
6. Kleefstra, T., Brunner, H.G., Amiel, J., Oudakker, A.R., Nillesen, W.M., Magee, A., Genevieve, D., Cormier-Daire, V., van Esch, H., Fryns, J.P. *et al.* (2006) Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am. J. Hum. Genet.*, **79**, 370–377.
7. Freude, K., Hoffmann, K., Jensen, L.-R., Delatycki, M.B., des Portes, V., Moser, B., Hamel, B., van Bokhoven, H., Moraine, C., Fryns, J.-P. *et al.* (2004) Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation. *Am. J. Hum. Genet.*, **75**, 305–309.
8. Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S.S., Chen, W., Hosseini, M., Behjati, F., Haas, S., Jamali, P. *et al.* (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*, **478**, 57–63.
9. Budde, B.S., Namavar, Y., Barth, P.G., Poll-The, B.T., Nürnberg, G., Becker, C., van Ruissen, F., Weterman, M.A., Fluiter, K., te Beek, E.T. *et al.* (2008) tRNA splicing endonuclease mutations cause pontocerebellar hypoplasia. *Nat. Genet.*, **40**, 1113–1118.
10. Barth, P.G. (1993) Pontocerebellar hypoplasias. An overview of a group of inherited neurodegenerative disorders with fetal onset. *Brain Dev.*, **15**, 411–422.
11. Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., DiDonato, N. *et al.* (2012) Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet*, **380**, 1674–1682.
12. Jamra, R.A., Wohlfart, S., Zweier, M., Uebe1, S., Priebe, L., Ekici, A., Giesebrecht, S., Abboud, A., Ayman, M., Khateeb, A., Fakher, M. *et al.* (2011) Homozygosity mapping in 64 Syrian consanguineous families with non-specific ID reveals 11 novel loci and high heterogeneity. *Eur. J. Hum. Genet.*, **19**, 1161–1166.
13. Sobel, E. and Lange, K. (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. *Am. J. Hum.Genet.*, **58**, 1323–1337.
14. Wass, M.N., Kelley, L.A. and Sternberg, M.J. (2010) 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Res*, **38**, 469–473.
15. Cloutier, P., Lavallée-Adam, M., Faubert, D., Blanchette, M. and Coulombe, B. (2013) A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity. *PLoS Genet.*, **9**, e1003210.
16. Reiff, R.E., Ali, B.R., Baron, B., Yu, T.W., Ben-Salem, S., Coulter, M.E., Schubert, C.R., Hill, R.S., Akawi, N.A., Al-Younes, B. *et al.* (2014) METTL23, a transcriptional partner of GABPA, is essential for human cognition. *Hum. Mol. Genet.* [Epub ahead of print].
17. Petrossian, T.C. and Clarke, S.G. (2011) Uncovering the human methyltransferasome. *Mol. Cell. Proteom.*, **10**, M110.000976.
18. Poleshko, A., Einarson, M.B., Shalginskikh, N., Zhang, R., Adams, P.D., Skalka, A.M. and Katz, R.A. (2010) Identification of a functional network of human epigenetic silencing factors. *J. Biol. Chem.*, **285**, 422–433.
19. Albert, H. and Helin, K. (2010) Histone methyltransferases in cancer. *Semin. Cell Dev. Biol.*, **21**, 209–220.
20. Feng, J. and Fan, G. (2009) The role of DNA methylation in the central nervous system and neuropsychiatric disorders. *Int. Rev. Neurobiol.*, **89**, 67–84.
21. Lin, M., Wei, L.J., Sellers, W.R., Lieberfarb, M., Wong, W.H. and Li, C. (2004) dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics*, **20**, 1233–1240.
22. Seelow, D., Schuelke, M., Hildebrandt, F. and Nürnberg, P. (2009) HomozygosityMapper-an interactive approach to homozygosity mapping. *Nucleic Acids Res.*, **37**, W593–W599.
23. Kelley, L.A. and Sternberg, M.J.E. (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat. Protoc.*, **4**, 363–371.

Supplementary information

for

Disruption of the methyltransferase-like 23 gene *METTL23* causes mild autosomal recessive intellectual disability.

Marie Bernkopf[1,†], Gerald Webersinke[1, †], Chanakan Tongsook[2], Chintan N. Koyani[3], Muhammad A. Rafiq[4], Muhammad Ayaz[5], Doris Muller[6], Christian Enzinger[7], Muhammad Aslam[5], Farooq Naeem[5,9], Kurt Schmidt[10], Karl Gruber[11], Michael R. Speicher[8], Ernst Malle[3], Peter Macheroux[2], Muhammad Ayub[5,9], John B. Vincent[4,12,13], Christian Windpassinger[8,*] and Hans-Christoph Duba[6]

†These authors contributed equally

*To whom correspondence should be addressed at Institute of Human Genetics, Medical University of Graz, Harrachgasse 21/III, 8010 Graz, Austria. Tel: +43 (0) 3163804114; Email: christian.windpassinger@medunigraz.at

[1]Laboratory of Molecular Biology and Tumorcytogenetics, Department of Internal Medicine, Krankenhaus der Barmherzigen Schwestern, Linz, Austria, [2]Institute of Biochemistry, Graz University of Technology, Graz, Austria, [3]Institute of Molecular Biology and Biochemistry, Medical University of Graz, Graz, Austria, [4]Molecular Neuropsychiatry and Development (MiND) Lab, The Campbell Family Brain Research Institute, The Centre for Addiction & Mental Health (CAMH), Toronto, ON, Canada, [5]Lahore Institute of Research and Development, Lahore, Punjab Province, Pakistan, [6]Departmen tof Human Genetics, Landes-Frauen und Kinderklinik, Linz, Austria, [7]Department of Neurology and[8] Institute of Human Genetics, Medical University of Graz, Graz, Austria, [9]Division of Developmental Disabilities, Department of Psychiatry, Queen's University, Kingston, ON, Canada, [10]Department of Pharmacology and Toxicology, Karl-Franzens University Graz, Graz, Austria, [11]Institute of Molecular Biosciences, University of Graz, Graz, Austria and [12]Departmentof Psychiatry and [13]Institute of Medical Science, University of Toronto, Toronto, ON, Canada

Supplementary Materials

Table 1. Primers

Amplicon	Primer	Sequence (5' to 3')
Exon 2	METTL23E2-F	GTAATGGATGCCAATGCCTG
	METTL23E2-R	GCTCACAGACCTTACAAACTAGC
Exon 3	METTL23E3-F	CCAGTTACGATACATGGCCC
	METTL23E3-R	AGACCACAATTGGACCTTGG
Exon 4	METTL23E4-F	TTCTTTGAACCAGAAGGTAAGC
	METTL23E4-R	TCATGCCTTTTCTGTTTTGC
Exon 5	METTL23E5-F	TCTACCCTACCCATGCGATAC
	METTL23E5-R	ACCATCTTCAGACCCACTGC
qRT-PCR	Q_Mettl23iso1-F	GGCCATCTTAGAGATTGGAGCT
qRT-PCR	Q_Mettl23iso2-F	CAGTTCTGCGCGTATTGGAGCT
qRT-PCR	Q_Mettl23-R	ATGTTAGTCCTACCACCTGCAG
Probe	Q-Mettl23	FAM-GTAATACTGTCAGACAGCTCAGAACTGCC-TAMRA

Clinical information

Austrian Family (LFKK1)

They are toilet trained, can feed and dress and undress themselves and can help with household chores. Due to lack of such establishments in their surrounding they stayed some years far from home in different facilities for intellectual disabled individuals. Since 2006 all 4 affected sibs live in a sheltered accommodation for intellectual disabled individuals in her hometown. Intellectual disability is mild as they understand everything, including long sentences and have very simply reading, writing and math abilities. All 4 sibs show flatfeet and need orthopaedic insoles. The two girls present with a distinct prominent maxilla, which is also present in one brother and the mother in a milder form. Seizures are not reported. There is no microcephaly.

Individual II:1 (female, born 1960, height 170 cm, weight 67.5 kg, head circumference 55 cm): after an uneventful pregnancy the girl was born at home. Milestones in development were delayed. She shows on the one hand simple skills in literacy and numeracy and on the other hand good fine motor skills (knitting and crochet). Clinical genetic examination revealed a distinct prominent maxilla, which is also present in a mild form in her mother. Additionally she shows a 7.0 mm difference in the length of her lower extremities. At the age of 44/45 there are 3 hospitalizations reported (14, 19 and 14 days respectively). The reason was a S1 radix lesion caused by a discus prolapse L5/S1 which was treated conservatively. As she presented brisk reflexes at one of these hospitalizations a CT of the cerebrum and an EEG were performed, which revealed normal results.

Individual II:2 (male, born 1962, height 185 cm; weight 62 kg, head circumference 55 cm): after an uneventful pregnancy the boy was born at home with complications (a forceps delivery was in discussion but not necessary). After birth he presented as a large cyanotic newborn and showed an icterus neonatorum. Milestones in development were delayed. He was able to walk with 3 years and incontinent until 6 years. At the age of 33 there is one hospitalization for 11 days reported due to a putrid bronchitis, which was effectually treated with Clarithromycin. He shows very simple skills in literacy and numeracy. Clinical genetic examination revealed a laterocollis which is treated with horse riding and massages.

Individual II:5 (male, born 1971, height 180cm, weight 64 kg, head circumference 58 cm): after an uneventful pregnancy the boy was born in a local hospital. He was able to speak single words with 1.5 years. As reported by his mother at this time he was hospitalized for 2 months due to a gastrointestinal infection. During his stay in the hospital he caught measles and had high temperature (41°C). Cerebrospinal fluid was analysed without result.

134

After this long hospitalization it took until the age of five years that speech came back. He shows very simple skills in literacy and numeracy. Clinical genetic examination revealed a mild form of a prominent maxilla as present in the mother.

Individual II:4 (female, born 1975, height 168 cm, weight 62.5 cm, head circumference 57 cm): after an uneventful pregnancy the girl was born in a local hospital. Milestones in development were delayed. After playschool she visited a school for intellectual disabled children. She shows very simple skills in literacy and numeracy. Clinical genetic examination revealed a distinct prominent maxilla, which is also present in a mild form in her mother.

MRI

Visual inspection blinded to clinical information of the scans blinded to clinical information led to the suspicion of increased volume of the subcallosal grey matter in one patient (PK31-II:2). On axial slices, this led to the impression of a decreased delineation of the basal ganglia. These presumably structurally altered structures are commonly attributed to the limbic system and thus phylogenetic old brain areas which are strongly implicated in affect regulation and are in vicinity of other strategically important areas such as the hypothalamus and mamillar bodies. While sulcal atrophy appeared to be present in the 8-year old girl, in both cases we failed to detect the ''molar tooth sign'' or vermal atrophy.

CHAPTER 4

4. Biochemical characterization and kinetic studies on *N,N*-8-demethyl-8-amino-D-riboflavin dimethyltransferase (RosA)

4.1 Structural and kinetic studies on RosA, the enzyme catalysing the methylation of 8-demethyl-8-amino-D-riboflavin to the antibiotic roseoflavin

AUTHOR CONTRIBUTIONS

Structural and kinetic studies on RosA, the enzyme catalysing the methylation of 8-demethyl-8-amino-D-riboflavin to the antibiotic roseoflavin

Chanakan Tongsook[1,†], Michael K. Uhl[2,†], Frank Jankowitsch[3], Matthias Mack[3], Karl Gruber[2] and Peter Macheroux[1]

1 Institute of Biochemistry, Graz University of Technology, Austria
2 Institute of Molecular Biosciences, University of Graz, Austria
3 Department of Biotechnology, Institute for Technical Microbiology, Mannheim University of Applied Sciences, Germany

N,*N*-8-demethyl-8-amino-D-riboflavin dimethyltransferase (RosA) catalyses the final dimethylation of 8-demethyl-8-amino-D-riboflavin (AF) to the antibiotic roseoflavin (RoF) in *Streptomyces davawensis*. In the present study, we solved the X-ray structure of RosA, and determined the binding properties of substrates and products. Moreover, we used steady-state and rapid reaction kinetic studies to obtain detailed information on the reaction mechanism. The structure of RosA was found to be similar to that of previously described *S*-adenosylmethionine (SAM)-dependent methyltransferases, featuring two domains: a mainly α-helical 'orthogonal bundle' and a Rossmann-like domain (α/β twisted open sheet). Bioinformatics studies and molecular modelling enabled us to predict the potential SAM and AF binding sites in RosA, suggesting that both substrates, AF and SAM, bind independently to their respective binding pocket. This finding was confirmed by kinetic experiments that demonstrated a random-order 'bi-bi' reaction mechanism. Furthermore, we determined the dissociation constants for substrates and products by either isothermal titration calorimetry or UV/Vis absorption spectroscopy, revealing that both products, RoF and *S*-adenosylhomocysteine (SAH), bind more tightly to RosA compared with the substrates, AF and SAM. This suggests that RosA may contribute to roseoflavin resistance in *S. davawensis*. The tighter binding of products is also reflected by the results of inhibition experiments, in which RoF and SAH behave as competitive inhibitors for AF and SAM, respectively. We also showed that formation of a ternary complex of RosA, RoF and SAH (or SAM) leads to drastic spectral changes that are indicative of a hydrophobic environment.

Database

Structural data are available in the Protein Data Bank under accession number 4D7K.

138

Introduction

The Gram-positive soil bacteria *Streptomyces davawensis* and *Streptomyces cinnabarinus* produce roseoflavin (RoF), the only known natural riboflavin (vitamin B$_2$) analogue with antibiotic activity. In contrast to most other antibiotics, RoF has multiple cellular targets and negatively affects flavoproteins as well as FMN riboswitches [1–8]. RoF may be considered a natural antimetabolite, and was postulated to be biosynthesized from riboflavin through 8-demethyl-8-amino-D-riboflavin (AF) and 8-demethyl-8-methylamino-riboflavin (MAF) [9,10]. The occurrence of the intermediates AF and MAF was confirmed by the discovery of the first enzyme of RoF biosynthesis, the *S*-adenosylmethionine (SAM)-dependent dimethyltransferase *N*,*N*-8-demethyl-8-amino-D-riboflavin dimethyltransferase (RosA) [11]. This enzyme synthesizes RoF from AF (via MAF) in two consecutive methylation reactions, as shown in Scheme 1. The corresponding gene (*rosA*) is present in a cluster comprising ten genes. The remaining genes of this cluster were found not to be involved in RoF synthesis, and it is presently unclear which genes/enzymes are responsible for synthesis of the key intermediate AF [12].

The important role of RosA for completion of RoF biosynthesis prompted us to determine the three-dimensional structure of this enzyme and to investigate the mechanism of the consecutive methylation reactions in more detail. We show that RosA has a topology similar to other SAM-dependent methyltransferases, containing a mainly α-helical 'orthogonal bundle domain' and a Rossmann-like fold. Determination of the three-dimensional structure also enabled us to locate the binding sites for the substrates AF and SAM. In addition, we performed steady-state and rapid reaction studies, and determined the dissociation constants for the substrates of RosA as well as for the products. It was found that AF and SAM bind independently to RosA according to a random-order mechanism. The products of the RosA reaction, RoF and SAH, exhibit higher affinity for the protein than the substrates AF and SAM do, leading to competitive product inhibition. RosA is a comparably slow enzyme, with a turnover rate (k_{cat}) of 0.06 min^{-1} for the overall dimethylation reaction. As RoF is a more potent antibiotic than AF and MAF, the slow synthesis of RoF and the tight binding of products may contribute to protection of the strains *S. davawensis* and *S. cinnabarinus* from the antibiotic that they produce.

Results

Crystal structure of RosA

Initial attempts to solve the structure of RosA by molecular replacement were unsuccessful despite a thorough search for appropriate template structures using the online tools PHYRE2 [13], FUGUE [14] and CASPR [15]. Therefore, RosA labelled with selenomethionine (SeMet) (seven methionine residues in 353 amino acids) was produced, and the structure was solved by multiple-wavelength anomalous dispersion (MAD) using a single tetragonal crystal (Table 1). The structure (one protomer in the asymmetric unit) was partially refined to a resolution of 3.5 Å, and this model was then used to solve the structure of a triclinic crystal form (space group *P*1) by molecular replacement. This crystal contained six molecules per asymmetric unit, and the structure was refined at a resolution of 2.2 Å, yielding final values of $R = 20\%$ and $R_{free} = 25\%$ (Table 2).

The RosA protomer comprises two domains: a mainly α-helical 'orthogonal bundle domain' and a Rossmann-like domain (α/β twisted open sheet). The orthogonal bundle (five α-helices and one antiparallel β-sheet) consists of residues 1–98. Residues 99–178 form an inter-domain region (five helices) that connects the Rossmann motif to the orthogonal bundle. The Rossmann motif (residues 179–353) comprises a seven-stranded β-sheet core consisting of five parallel and two antiparallel β-strands, which are connected by two pairs of α-helices (Fig. 1A). To classify the domains, we performed a CATH database (http://www.cathdb.info/) search using the CATHEDRAL algorithm [16]. The orthogonal bundle belongs to the CATH superfamily 'winged helix repressor DNA binding domain' (CATH classification: 1.10.10.10; Sequen-
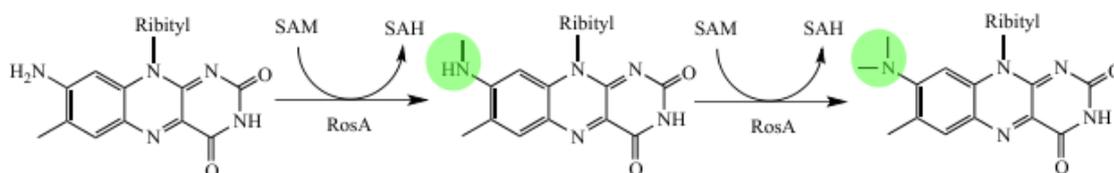
Scheme 1. The catalytic reaction of RosA.

139

Table 1. Statistics for the SeMet MAD datasets. Values in parentheses are for the highest-resolution shell.

	Peak	Inflection	Remote
X-ray source	SLS X06DA-PXIII		
Wavelength (Å)	0.9791	0.9796	0.9714
Temperature (K)	100	100	100
Space group	$I4_122$	$I4_122$	$I4_122$
Cell dimensions			
$a = b, c$ (Å)	112.54, 132.27	112.46, 132.48	112.60, 132.27
Resolution (Å)	47.04–3.35 (3.53–3.35)	47.02–3.54 (3.60–3.54)	47.06–3.42 (3.73–3.42)
Total number of reflections	166 043 (22 355)	141 532 (20 491)	156 742 (21 751)
Unique number of reflections	6400 (905)	5454 (771)	6031 (848)
Multiplicity	25.9 (27.7)	26.0 (26.6)	26.0 (25.6)
Anomalous multiplicity	14.2 (13.1)	14.2 (14.1)	14.2 (13.7)
Completeness (%)	99.9 (99.7)	100.0 (100.0)	99.9 (99.6)
Anomalous completeness (%)	99.9 (99.4)	99.9 (99.4)	99.9 (99.3)
$R_{p.i.m.}$ (%)	1.3 (5.9)	1.0 (5.7)	1.1 (6.3)
R_{meas} (%)	6.4 (29.5)	5.2 (29.8)	5.6 (32.2)
$<I/\sigma I)>$	48.3 (12.8)	51.1 (14.8)	50.0 (12.2)
$CC_{1/2}$	1.000 (0.989)	1.000 (0.987)	1.000 (0.989)
CC*	1.000 (0.995)	1.000 (0.995)	1.000 (0.997)

tial Structure Alignment Program (SSAP) score = 87.6; RMSD = 2.7 Å) and the functional family 'caffeic acid 3-*O*-methyltransferase 1'. The Rossmann-like domain was found to be a member of the CATH superfamily 'Vaccinia Virus protein VP39' (CATH classification: 3.40.50.150; SSAP score = 85.4, RMSD = 4.1 Å). The algorithm treated the inter-domain residues as part of the Rossmann motif.

A PISA analysis [17] predicted that a dimer is the most probable oligomer of RosA in solution. This prediction was confirmed by size exclusion chromatography (Fig. 2). Dimerization occurs in a head-to-head arrangement (Fig. 1B), and is mediated by interactions of adjacent α-helices of the bundle motif (residues 1–36 and residues 60–73), the inter-domain α-helices comprising residues 101–109 and residues 111–126, and one helix from the Rossmann motif (residues 293–305).

The RosA protomers in the asymmetric unit exhibit very similar structures. Pairwise superpositions of the Rossmann domains in the six crystallographically independent protomers yielded RMSD values in the range from 0.2 to 0.9 Å (for approximately 160 aligned Cα atoms). The same analysis yielded mean RMSD values of 0.3 and 0.6 Å for the N-terminal and intermediate regions. The values are slightly higher for superpositions of the entire chains (1.2 Å) and the three dimers (1.6 Å).

To identify protein structures resembling RosA in terms of its tertiary structure, we performed a Dali server analysis [18]. The structures with the highest

Z scores were mitomycin 7-*O*-methyltransferase from *Streptomyces lavendulae* (PDB ID 3GWZ), a probable phenazine-specific methyltransferase from *Pseudomonas aeruginosa* (PDB ID 2IP2), the CALO1 methyltransferase from *Micromonospora echinospora* (PDB ID 3LST), and a caffeic acid *O*-methyltransferase from *Sorghum bicolor* (PDB ID 4PGH), with *Z* scores of 33, 32, 30 and 30, respectively. These proteins share 21–35% sequence identity with RosA. Superposition with the most similar structure (3GWZ, 35% sequence identity, 76% query coverage) resulted in an RMSD of 3.5 Å for the isolated subunit (317 of 324 aligned Cα atoms) and 6.6 Å for the dimers (618 of 650 aligned Cα atoms). Superposition of the Rossmann domains of RosA (residues 179–345) and 3GWZ (residues 179–349) resulted in an RMSD of 0.9 Å (118 of 160 aligned Cα atoms). This indicates that, despite an overall low sequence identity, the structure of RosA, especially the Rossmann domain, closely resembles that of other methyltransferases.

Prediction of the SAM/SAH binding site

Unfortunately, all efforts to co-crystallize RosA with SAM, SAH, AF or combinations of these compounds and cofactors were unsuccessful. Therefore, we used structural bioinformatics methods to locate putative binding sites for the substrate and cofactor. A cavity analysis using the program CASOX [19] yielded several cavities, but only one of those was large enough to

140

Table 2. Data collection and refinement statistics. Values in parentheses are for the highest-resolution shell.

	RosA (native)
Data collection	
X-ray source	ESRF ID23-1
Wavelength (Å)	1.000
Temperature (K)	100 K
Space group	P1
Cell dimensions	
a, b, c (Å)	82.72, 82.76, 96.45
α, β, γ (°)	95.1, 98.7, 114.5
Resolution (Å)	34.77–2.22 (2.26–2.22)
Total number of reflections	168 833 (3484)
Unique number of reflections	86 780 (1838)
Multiplicity	1.9 (1.9)
Completeness (%)	77.4 (33.0)
$R_{p.i.m.}$ (%)	3.1 (39.8)
R_{meas} (%)	4.3 (56.2)
$<I/\sigma(I)>$	12.7 (1.70)
$CC_{1/2}$	0.999 (0.756)
CC^*	1.000 (0.928)
Refinement	
R_{work}	0.1949
R_{free}	0.2485
Number of atoms	15 787
Protein	15 518
Water	269
B-factors (Å^2)	
Protein	66.1
Water	48.9
All atoms	65.8
RMSDs	
Bond lengths (Å)	0.003
Bond angles (°)	0.620
Ramachandran outliers (%)	0.15
PDB ID	4D7K

accommodate AF and SAM, and was located at the interface between the Rossmann domain and the N-terminal domain (Fig. 3A).

In addition, the web tools PROSITE [20], 3DLIGANDSITE [21] and COACH [22] were used to predict and identify ligand binding regions in RosA. PROSITE identified a SAM binding region (residue 235–237) in the C-terminal Rossmann domain. This finding is in line with results obtained using 3DLIGANDSITE, COACH and CaSoX. 3DLIGANDSITE identified 25 crystal structures of proteins structurally related to RosA with bound SAM or SAH, and predicted 22 residues that are putatively involved in cofactor binding to RosA. On the other hand, the program COACH (http://zhanglab.ccmb.me-d.umich.edu/COACH/) found several structures of methyltransferases in complex with various ligands, and clustered them according to a confidence score (C-score between 0 and 1). The best-scored cluster pre-

dicted 13 consensus SAH/SAM binding residues with a C-score of 0.54 and a cluster size (total number of templates in a cluster) of 103. Moreover, COACH generated a structure of RosA with SAM bound in the predicted active site. All other ligand–enzyme complexes had significantly lower C-scores. Among those structures with C-scores between 0.5 and 0.05 were those with ligands such as *N*-acetyl serotonin (C-score 0.10), pisatin (C-score 0.08) and 5-(3,3-dihydroxypropeny)-3-methoxy-benzene-1,2-diol (C-score 0.05). Flavin derivatives such as AF did not occur in these clusters identified by COACH. However, a cluster with an even lower C-score of 0.02 contained structures of methyltransferases with bound SAH and 4-methoxy-E-rhodomycin. 4-methoxy-E-rhodomycin resembles AF with regard to the planar ring system. A comparison of the modelled RosA–SAM complex with the structure of carminomycin 4-*O*-methyltransferase from *Streptomyces peucetius* (PDB ID 1TW2), a representative of the 4-methoxy-E-rhodomycin cluster, showed that, despite the low sequence identity of 33%, the Rossmann domain and especially the β-sheet core appeared to be structurally conserved. Therefore, we superimposed the β-sheet core residues of RosA and 1TW2 (RMSD 0.805 Å; 42 of 42 aligned Cα atoms), and aligned a molecule of AF with 4-methoxy-E-rhodomycin bound to 1TW2 (Fig. 4A). The RosA–AF–SAM complex thus generated was energy-minimized and analysed in terms of the binding interactions of AF and SAM (Figs 3B and 4B).

Determination of dissociation constants

Inspection of the AF and SAM binding sites in the RosA structure suggested that both substrates, AF and SAM, bind independently to their respective pockets. Thus we determined the dissociation constant for AF and SAM by isothermal titration calorimetry (ITC). As shown in Fig. 5A,C, titration with either AF or SAM produced exothermic signals that were fitted to a one-binding-site model, yielding dissociation constants of 10 ± 1 and 22 ± 2 μM, respectively. Surprisingly, both products, i.e. RoF and SAH, bind approximately ten times more tightly than the substrates, displaying dissociation constants of 0.8 ± 0.05 and 2 ± 0.1 μM, respectively (Fig. 5B,D).

Ternary complex formation with substrates and products

The catalytic methylation of AF by RosA was investigated by spectrophotometry using a stopped-flow device and a conventional spectrophotometer,
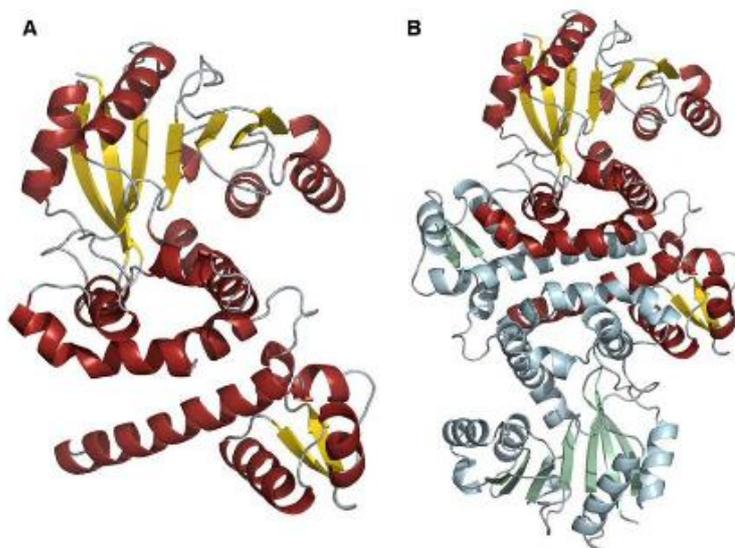
141

Fig. 1. Schematic representation of the structure of RosA. (A) Cartoon representation of the RosA protomer, showing α-helices in red, β-sheets in yellow, and loops in grey. (B) Cartoon representation of the RosA dimer with one protomer shown in red, yellow and grey, and the other in pale blue, green and grey.
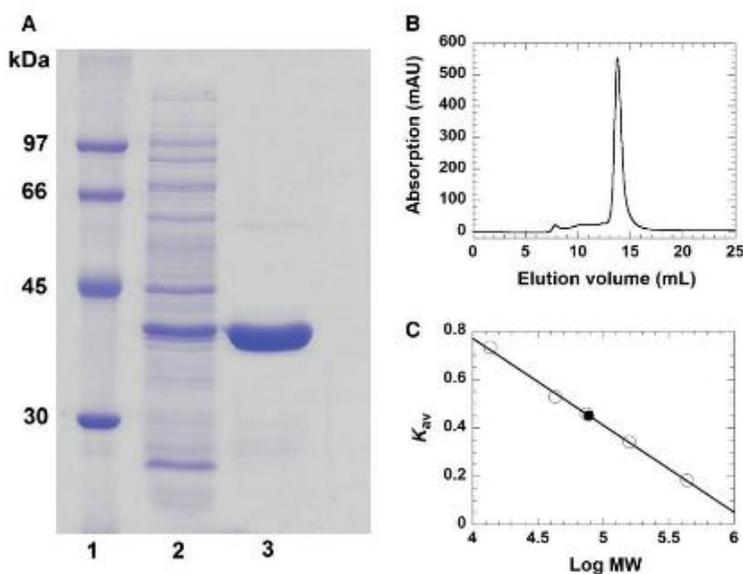

Fig. 2. Determination of the purity and native molecular mass of RosA using SDS/PAGE and FPLC. (A) Determination of the subunit molecular mass of RosA after purification by Ni-Sepharose FPLC was performed by SDS/PAGE (12.5%). Lane 1, low-molecular-mass protein marker; lane 2, crude extract; lane 3, protein fraction after purification by Ni-Sepharose FPLC. The subunit molecular mass of RosA was estimated to be 38 kDa. (B) Determination of the native molecular mass of RosA using Superdex 200 10/300 GL chromatography. (C) Plot of the partition coefficient (K_{av}) against the logarithm of molecular mass of standard proteins (ferritin, 440 kDa; aldolase, 158 kDa; conalbumin, 75 kDa; ovalbumin, 43 kDa; ribonuclease A, 13.7 kDa). The native molecular mass of RosA (filled circle) was estimated to be 77 kDa.

exploiting the spectral changes that occur during methylation [λ_{max} (AF) = 478 nm; λ_{max} (RoF) = 505 nm]. At low concentrations of RosA, the final absorption spectrum features an absorption maximum at 505 nm, which is characteristic of complete formation of RoF (Fig. 6A). However, at high concentrations of RosA the maximum at 505 nm is less pronounced (hypsochromic effect), and additional maxima at a shorter wavelength are observed (Fig. 6B,C). Analysis by HPLC revealed that RoF is the main pro-

duct in both cases, suggesting that the observed spectral changes are due to formation of a complex with RosA and either SAM or SAH. The observed spectral differences suggest considerable changes in the flavin binding site upon formation of a ternary complex, and the UV/Vis absorption spectra of RoF in various solvents suggest that an apolar environment induces similar spectral changes to those observed in our experiments (Fig. 6D). In addition, it is conceivable that steric restriction induced in the ternary complex
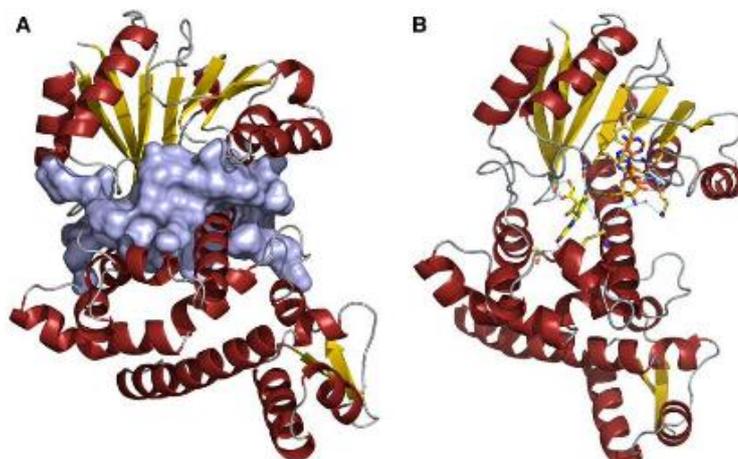
142

Fig. 3. Schematic representation of the predicted binding pocket of RosA. (A) Surface and cartoon representation of the binding pocket of RosA predicted using CASOX [19]. (B) Cartoon representation of RosA with predicted binding of AF (shown as yellow sticks) and SAM (shown as orange sticks). Potential hydrogen bonding interactions are indicated by blue dashed lines. Water molecules are represented as red spheres. The distance between the reactive methyl group of SAM and the amino group of AF is indicated by a green dashed line.
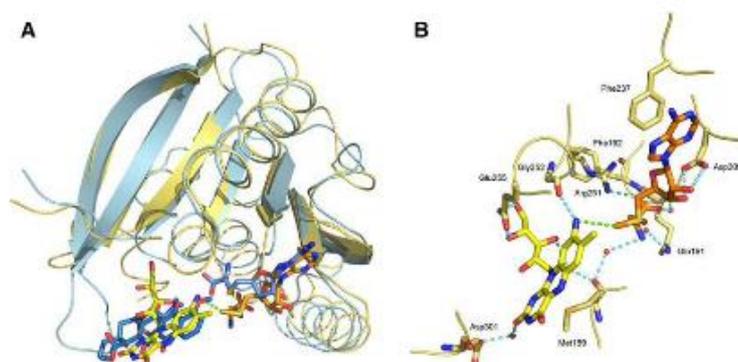
Fig. 4. Structure of the putative active site of RosA. (A) Close-up view of the superimposed structures of RosA (shown in yellow) and carminomycin-4-*O*-methyltransferase (PDB ID 1TW2; shown in blue). Ligands of 1TW2 (*S*-adenosylhomocysteine and 4-methoxy-E-rhodomycin) are shown as blue sticks. AF and SAM are shown in their predicted binding modes as yellow and orange sticks, respectively. (B) Close-up view of the predicted binding pocket of RosA. Potential hydrogen bonding with AF (shown in yellow) and SAM (shown in orange) is indicated by blue dashed lines. Water molecules are represented as red spheres. The distance between the reactive methyl group of SAM and the amino group of AF is indicated by a green dashed line.

prevents de-localization of the electron lone pair into the aromatic system of the isoalloxazine ring, resulting in the observed hypso- and hypochromic effect. Therefore, we assume that formation of the ternary complex generates a more apolar and sterically constrained environment in the RoF binding pocket. In order to reveal the nature of the ternary complex, we also determined the affinity of RosA for SAM and SAH in the presence of RoF. As shown in Fig. 7, the presence of RoF greatly decreases the affinity for SAM, yielding a dissociation constant of 180 ± 20 μM, whereas the affinity of SAH is approximately fourfold higher ($K_d = 0.4 \pm 0.05$ μM). Thus the ternary complex with SAH (product complex) is clearly favoured over a complex with SAM (mixed substrate/product complex).

Next, we performed difference absorption spectroscopy to determine the dissociation constants of RoF from the binary complex of RosA:SAH and RosA:SAM, respectively, as well as the affinity of AF for the RosA:SAH complex. The former experiments, shown in Fig. 8A,B, yielded virtually identical dissociation constants of 14 ± 2 and 15 ± 2 μM, demonstrating that SAH and SAM cause similar decreases in the binding affinity of RoF by a factor of 17–18. In contrast to product binding, AF binding to RosA is not affected by SAH, as the determined dissociation constant of 10 ± 1 μM is almost identical to that observed in the absence of SAH (Fig. 9). Table 3 lists the dissociation constants obtained by ITC or spectrophotometric titrations.

143
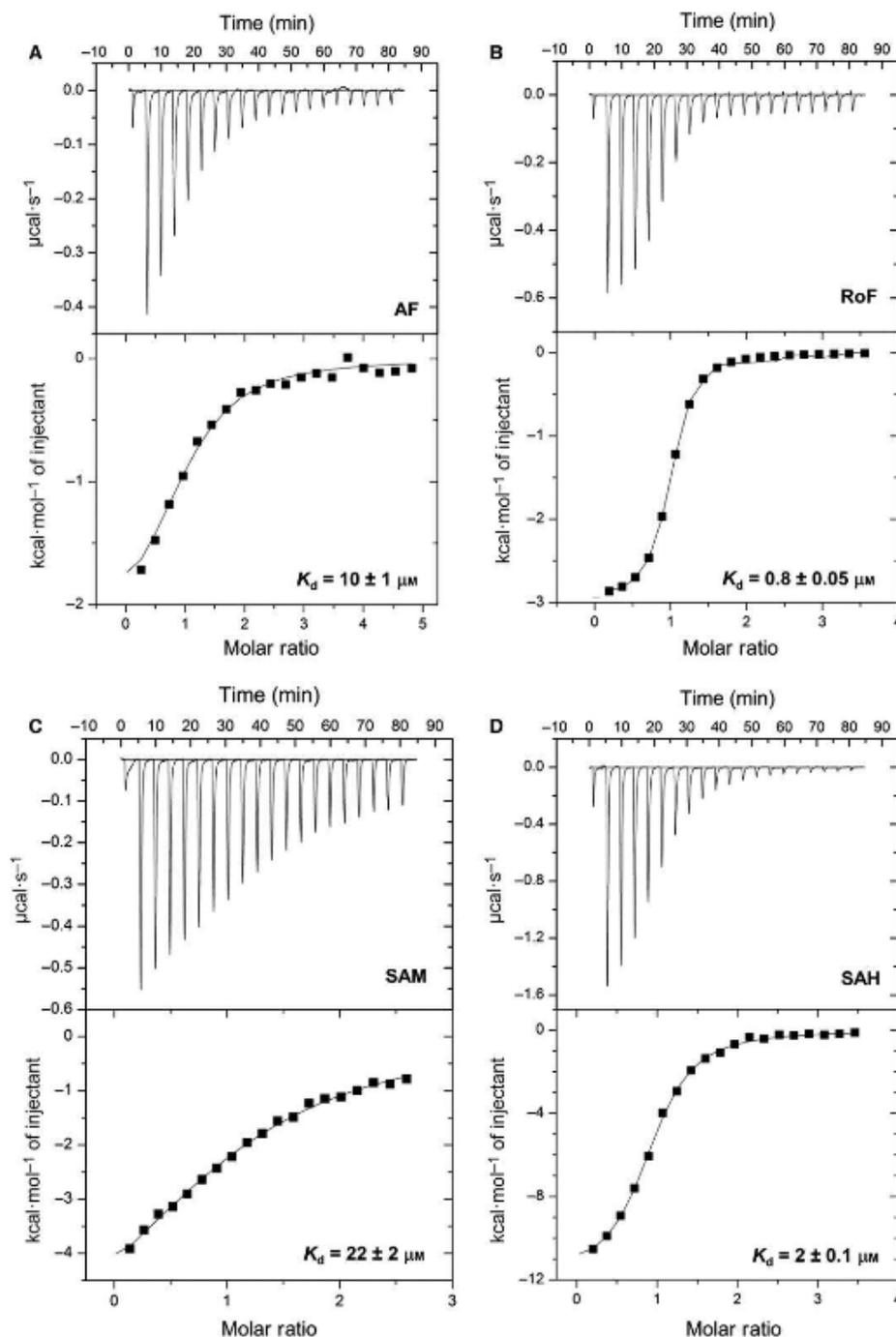
Fig. 5. Binding of AF, RoF, SAM and SAH to RosA. Calorimetric titrations of RosA with AF (A), RoF (B), SAM (C) and SAH (D) were performed in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl at 25 °C using a VP-ITC system (MicroCal). The dissociation constants for binding of AF, RoF, SAM and SAH to RosA were 10 ± 1, 0.8 ± 0.05, 22 ± 2 and 2 ± 0.1 μM, respectively (three independent measurements for each ligand). Data were analysed by non-linear least-squares fitting using ORIGIN version 7.0 (MicroCal).

144

Fig. 6. UV/Vis absorption spectra at various RosA concentrations. (A, B) Spectral changes during the reaction of 5 µM AF and 0.8 mM SAM with 5 and 25 µM RosA, respectively. The reactions were performed in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl, in a single-mix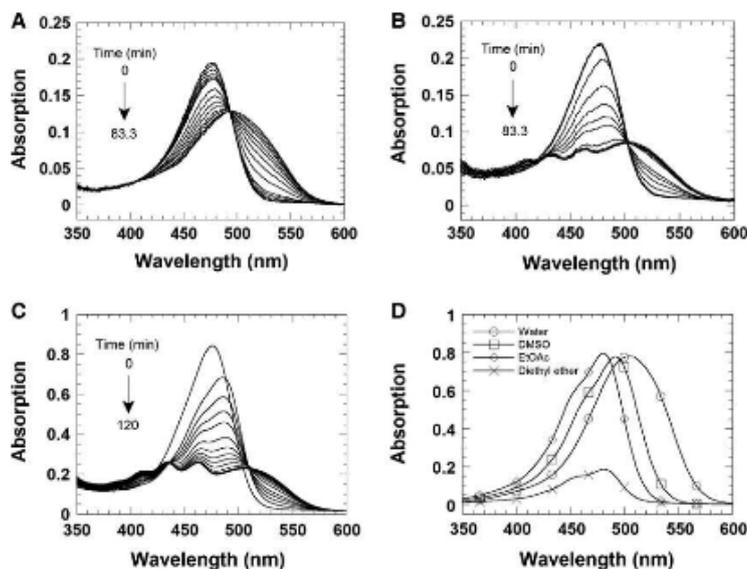ing stopped-flow spectrophotometer. (C) A solution of 20 µM AF and 100 µM RosA was pre-incubated prior to addition of 2 mM SAM. All three reactions were run at 25 °C, and absorption spectra were recorded from 350 to 600 nm using a diode array. (D) Spectra for 25 µM RoF in various solvents [water, dimethylsulfoxide (DMSO) and ethyl acetate] and for 15 µM RoF in diethyl ether.

Mechanistic studies of the dimethylation reaction catalysed by RosA

With regard to the reaction mechanism, the structure of RosA as well as the independent binding of AF and SAM suggest that the enzyme operates by a random-order kinetic mechanism, as shown in Scheme 2. To test this hypothesis, we performed a series of pre-steady-state kinetic measurements in the stopped-flow device. In these experiments, different mixing orders were employed, and the progress of the reaction was monitored at 479 and 530 nm. As shown in Fig. 10A, the rate of reaction was independent of the experimental set-up, i.e. whether RosA was directly mixed with AF and SAM or pre-incubated with either of the substrates. To obtain further insight into the kinetic mechanism, we performed a set of steady-state experiments at various AF and RoF concentrations. As shown in Fig. 10B, a primary reciprocal plot of the reaction velocity against the reciprocal SAM concentration resulted in a set of converging lines in accordance with a random-order mechanism. From a re-plot of the primary data, i.e. slopes and y-axis intercepts versus the reciprocal AF concentration (Fig. 10B, insets), the Michaelis constants of AF and SAM were deduced to be 4 and 70 µM, and the k_{cat} was estimated as 0.06 min^{-1} (Table 4).

As generation of RoF from AF requires two consecutive methylation steps, we also wished to obtain insight into the individual methylation reactions. Towards this aim, we analysed the reaction under single-turnover conditions at two different wavelengths

that represent the consumption of the substrate ($\lambda = 479$ nm) and the generation of the product ($\lambda = 530$ nm). The reaction of RosA (20 µM) in the presence of AF (20 µM) was investigated by mixing with various concentrations of SAM (80–2000 µM). The rates of the observed spectral changes at 479 and 530 nm were biphasic, and therefore fitted to two rate equations (Fig. 11A). The first phase consisted of a decrease in absorption at 479 nm and a concurrent small increase at 530 nm (approximately 25%). This first phase was interpreted to represent monomethylation of AF to MAF, and showed a hyperbolic dependence on the SAM concentration, yielding a limiting rate of 0.5 ± 0.05 min^{-1} and an equilibrium constant of 180 ± 40 µM (Fig. 11B). The second slower phase corresponded to a decrease at 479 nm and a concomitant increase at 530 nm. This phase showed an inverse dependence on the SAM concentration approaching 0.05 ± 0.002 min^{-1} (Fig. 11C). The latter result suggested that monomethylation of AF generates an inhibitory product that prevents binding of SAM. In fact, our binding studies have revealed that SAH binds much more strongly to RosA than SAM does (Table 3). Furthermore, the presence of RoF further weakens SAM binding but the affinity for SAH is increased (Table 3). Thus we reasoned that SAH, which is produced in the first monomethylation reaction, acts as a competitive inhibitor of SAM. To test this hypothesis, we determined the rate of the methylation reaction as a function of SAM and SAH concentrations. The initial velocities were determined as a
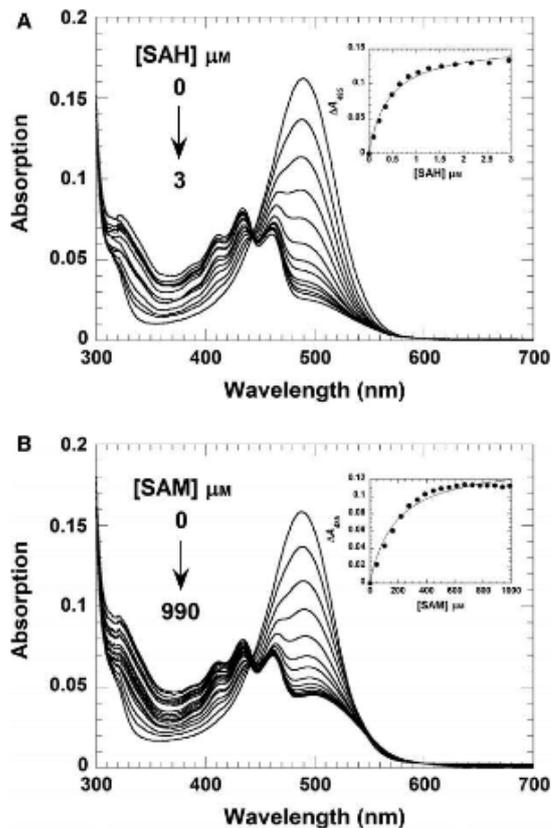
Fig. 7. Formation of a ternary complex comprising the RosA:RoF complex with SAH or SAM. (A) A solution of RosA (25 μM) and RoF (5 μM) was titrated with SAH (0–3 μM). The absorption spectra were recorded from 300 to 700 nm using a UV/Vis spectrophotometer. The inset shows a plot of the absorption change at 485 nm as a function of SAH concentration. A hyperbolic fit to the experimental data yielded a dissociation constant of 0.4 ± 0.05 μM. (B) A solution of RosA (25 μM) and RoF (5 μM) was titrated in tandem cuvettes with SAM (0–990 μM). The inset shows a plot of the absorption change at 485 nm as a function of SAM concentration. A hyperbolic fit to the experimental data yielded a dissociation constant of 180 ± 20 μM. Both titrations were performed in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl.

function of SAM concentration at various fixed concentrations of SAH to obtain information on the type of inhibition. However, this plot did not allow us to discern the inhibition type, i.e. competitive, non-competitive or uncompetitive. As it was shown that SAH binds more tightly to RosA than SAM does, we used the linearized Henderson equation for data analysis [23]. The resulting Henderson plot, as shown in Fig. 12A, yielded a set of lines that converge on the *y* axis. A re-plot of the slopes, as shown in the inset to
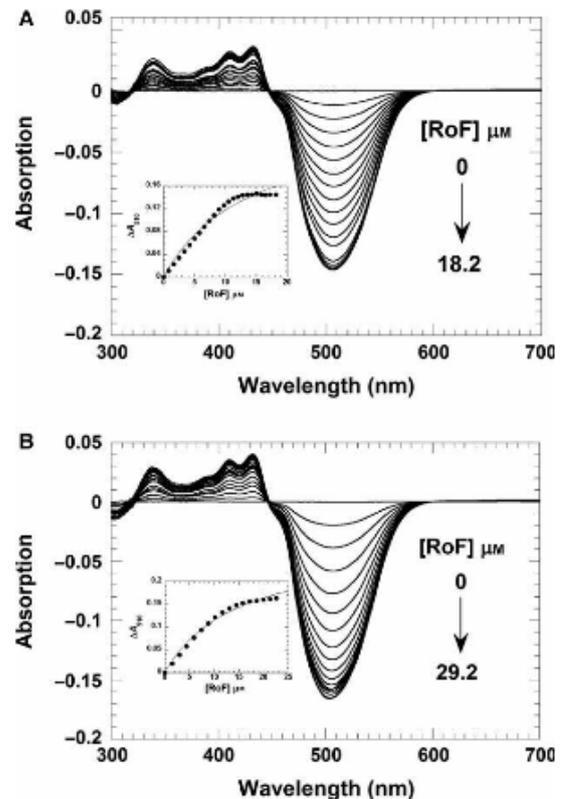
Fig. 8. Formation of a ternary complex comprising the RosA:SAH or RosA:SAM complex with RoF. (A) A solution of RosA (25 μM) and SAH (150 μM) was titrated with RoF (0–18.2 μM). The absorption changes from the UV/Vis difference titrations were recorded from 300 to 700 nm using a UV/Vis spectrophotometer. The inset shows a plot of the absorption change at 510 nm as a function of RoF concentration. A hyperbolic fit to the experimental data yielded a dissociation constant of 14 ± 2 μM. (B) A solution of RosA (30 μM) and SAM (2 mM) was titrated in tandem cuvettes with RoF (0–29.2 μM). The inset shows a plot of the absorption change at 510 nm as a function of the RoF concentration. A hyperbolic fit to the experimental data yielded a dissociation constant of 15 ± 2 μM. Both titrations were performed in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl.

Fig. 12A, allows determination of the inhibition constant of SAH as 7 μM. This value is in the same range as the dissociation constant of SAH in the presence of RoF, suggesting that MAF and RoF have similar effects on SAH binding (Table 3). Thus we conclude that the observed rate decline of the second kinetic phase is due to the competitive effect of SAH, with the limiting value corresponding to k_{off} of SAH.

In a similar set of experiments, it was shown that RoF is a competitive inhibitor of AF, exhibiting a K_i value of 27 μM (Fig. 12B), which represents the disso-
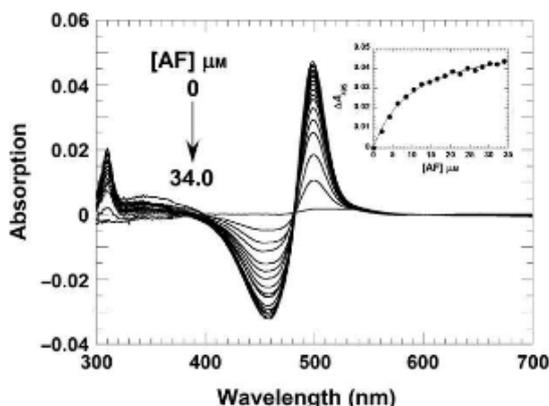
146

Fig. 9. AF binds independently to RosA in the presence of SAH. A solution of RosA (25 μM) and SAH (150 μM) was titrated with AF (0–34.0 μM). The absorption spectra of the titrations were recorded from 300 to 700 nm using a UV/Vis spectrophotometer. The inset shows a plot of the absorption change at 495 nm as a function of SAH concentration. A hyperbolic fit to the experimental data yielded a dissociation constant of 10 ± 1 μM.

ciation constant of RoF in the presence of SAM (compare Tables 3 and 5).

Discussion

Structure analysis revealed that RosA adopts a topology that is characteristic of SAM-dependent methyltransferases. Although attempts to obtain structural information for complexes with AF, SAM and SAH were unsuccessful, further inspection of the active site and the AF and SAM binding pockets provided useful hints for mechanistic considerations that were tested by binding studies and steady-state as well as presteady-state kinetic experiments. The final biosynthesis of RoF from AF by RosA is unusual in the sense that the enzyme performs two consecutive methylations in the same active site, and hence SAH must dissociate to make room for another SAM while the monomethylated intermediate, MAF, may remain bound to the flavin-binding pocket. This complexity of the overall conversion of AF to RoF was analysed by binding studies as well as kinetic experiments to evaluate the interaction of substrates and products in the RosA-catalysed reaction. Much to our surprise, our binding studies revealed that both products, RoF and SAH, bind approximately ten times more tightly to RosA than AF and SAM do (Table 3). Interestingly, the affinity of RoF decreased approximately 20-fold in the presence of SAM and SAH, whereas the affinities of SAM and SAH respond differently to the presence of

Table 3. Dissociation constants (K_d, μM) of RosA for AF, RoF, SAH and SAM, determined by ITC and UV/Vis absorption titrations. Values marked with asterisk were obtained from ITC measurements. Values marked with hash symbols (#) were obtained from UV/Vis absorption titrations.

In the presence of	AF	RoF	SAM	SAH
–	10 ± 1*	0.8 ± 0.05*	22 ± 2*	2 ± 0.1*
RoF			180 ± 20#	0.4 ± 0.05#
SAM		15 ± 2#		
SAH	10 ± 1#	14 ± 2#		

RoF, i.e. the dissociation constant for SAM increased approximately eightfold and that for SAH decreased fourfold. These relative binding affinities clearly suggest that the ternary complex observed in spectrophotometric experiments (Fig. 6) consisted of RoF and SAH bound to RosA. The fact that the tightest binary complex formed is that between RosA and RoF also suggested a potential mechanism to control the concentration of free RoF in *S. davawensis*. It may be assumed that the concentrations of both SAM and SAH are too low to form a ternary complex with RosA:RoF, and thus RoF inhibits binding of AF and further generation of the product. This scenario is also supported by the competitive inhibition of the RosA-catalysed methylation of AF by RoF (Fig. 12). The resistance of *S. davawensis* to RoF appears to be related to the ability of the *ribB* FMN riboswitch to discriminate between FMN and roseoflavin-5'-phosphate [7]. However, it is not clear yet whether other additional mechanisms contribute to resistance in *S. davawensis*. Therefore, it is plausible that the tight binding of RoF to RosA constitutes an additional mechanism to confer resistance towards RoF. Although it is still unknown how RoF is eventually secreted, it is conceivable that RosA also serves as a chaperone to guide RoF to the secretion machinery that is in charge of its translocation from the cell into the environment.

The X-ray crystallographic structure of RosA suggested that both substrates, AF and SAM, bind independently to their respective pockets. This feature was confirmed by kinetic experiments that demonstrated a random-order binding mechanism for the first methylation reaction yielding MAF and SAH (Fig. 10 and Scheme 2). The rate of monomethylation of AF to MAF showed a hyperbolic dependence on the SAM concentration, yielding a limiting rate of 0.5 ± 0.05 min^{-1} (Fig. 11B). Surprisingly, the subsequent reaction phase was characterized by an inverse dependence on the SAM concentration, approaching a
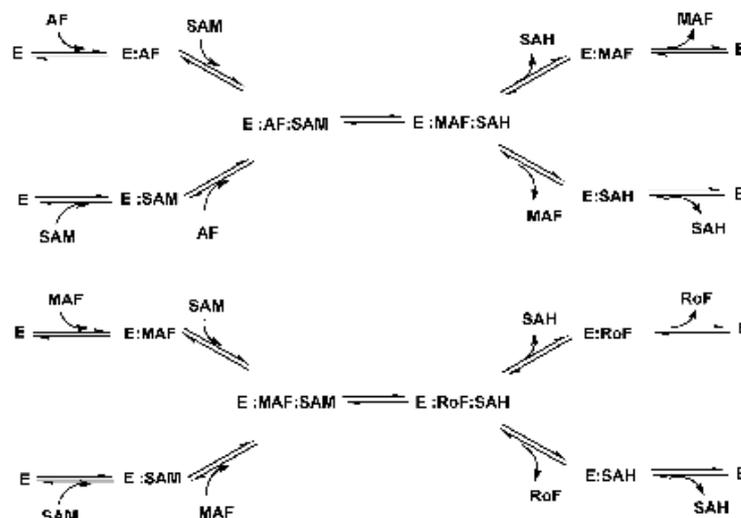
Scheme 2. Scheme of the methylation reactions catalysed by RosA following a random-order 'bi-bi' mechanism.

limiting rate of 0.05 ± 0.002 min^{-1} (Fig. 11C). As noted above, SAH binds tightly to RosA, and its affinity is enhanced in the presence of RoF (Table 3). It is likely that the binding affinity of SAH in the presence of MAF is in the same range, i.e. between 2 ± 0.1 and 0.4 ± 0.05 μM, and thus the second methylation reaction depends on the rate of dissociation of SAH (k_{off}). This k_{off} appears to be the limiting step for conversion of MAF to RoF, and therefore the intrinsic rate for this reaction step is not accessible in our kinetic experiments. A summary of the reaction steps characterized in this study is shown in Scheme 3.

Experimental procedures

Reagents

All chemicals and reagents were of the highest purity commercially available from Sigma-Aldrich (St. Louis, MO, USA) and Merck (Darmstadt, Germany). AF was a generous gift from Sandro Ghisla (Universität Konstanz, Germany). Ni-NTA-agarose was obtained from GE Healthcare (Little Chalfont, UK). RosA was over-produced in *Escherichia coli*, and purified as previously described [11]. The concentrations of the following compounds were determined spectrophotometrically using these extinction coefficients: AF, $\varepsilon_{490} = 42.0$ mM^{-1}·cm^{-1}; RoF, $\varepsilon_{505} = 31.1$ mM^{-1}·cm^{-1} [24]; SAM and SAH, $\varepsilon_{260} = 15.4$ mM^{-1}·cm^{-1} [25]; RosA (based on amino acid sequence) = 27.1 mM^{-1}·cm^{-1}.

Expression and purification

Heterologous production and purification of recombinant hexahistidine-tagged RosA from *S. davawensis* was per-

formed as described previously [11]. *E. coli* Rosetta 2 (DE3) was transformed using pET24a(+)*rosA-His$_6$* (pFJ02) [11]. One of the transformant strains was used to inoculate a pre-culture that was aerobically incubated at 37 °C for 16 h (200 rpm) in lysogeny broth containing 50 μg·mL^{-1} kanamycin. The cells were harvested by centrifugation (5000 *g* at 4 °C for 30 min), and washed three times using PBS (10 min per wash at 4 °C). The cells were then suspended in 30 mL PBS and used as an inoculum for a 15-L bioreactor (Bioengineering AG, Wald, Switzerland). Cell growth was allowed to proceed for 7 h at 750 rpm, 37 °C, with an aeration of 1 L of oxygen per L of culture per minute in a total volume of 10 L until the cells reached an absorption at 600 nm of 0.8. Synthesis of the recombinant protein was induced by addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside. The culture was further incubated for 16 h at 30 °C until an attenuance at 600 nm of 1.2 was reached. Cells were harvested by centrifugation at 5000 *g* (4 °C).

To produce selenomethionine-containing hexahistidine-tagged RosA, the expression protocol for native RosA was slightly altered. *E. coli* BL834 (DE3) was transformed with pFJ02, and incubated for 7 h at 37 °C and 200 rpm in selenomethionine minimal medium [26] supplemented with 50 μg·mL^{-1} kanamycin. The pre-culture was harvested by centrifugation at 5000 *g* (4 °C) and used as an inoculum for a 15-L bioreactor (Bioengineering AG). Synthesis of the recombinant protein was induced at when the cells reached an attenuance at 600 nm of 0.8 by adding 0.5 mM isopropyl-β-D-thiogalactopyranoside. Cells were harvested when they reached an attenuance at 600 nm of 1.2, by centrifugation at 5000 *g* (4 °C). All cell pellets were frozen immediately after harvesting and stored at −80 °C.

The enzyme was purified using an ÄKTApurifier™ system (GE Healthcare). Frozen cell pellets of *E. coli*

148

Table 4. Kinetic parameters for the RosA reaction.

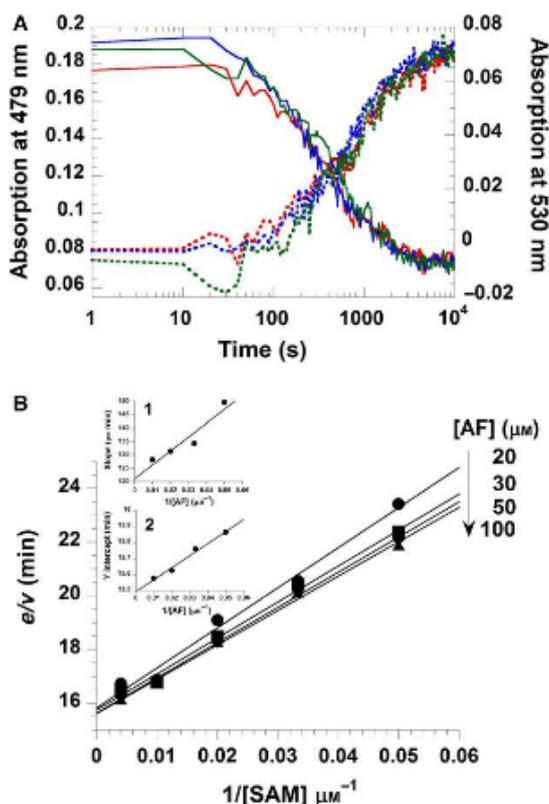| Parameter	Value
K_m^{AF} (μM)	4
K_m^{SAM} (μM)	70
k_{cat} (min^{-1})	0.06

Fig. 10. Steady-state kinetics of RosA and substrate binding sequence of the RosA reaction. (A). The RosA reaction was performed at 25 °C in a stopped-flow spectrophotometer using different mixing orders for the substrates AF and SAM. The traces shown represent AF consumption and RoF formation, respectively, at 479 nm (solid lines) and 530 nm (dotted lines). The red trace represents the reaction of RosA:AF (5 μM RosA and 5 μM AF) with 80 μM SAM. The blue trace represents the reaction of RosA:SAM (5 μM RosA and 80 μM SAM) with 5 μM AF. The green trace represents the reaction of RosA (5 μM) with AF (5 μM) and SAM (80 μM). All solutions were prepared in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl. (B) Double-reciprocal plot of the two-substrate kinetics of the RosA reaction. The enzyme catalytic assay was performed at 25 °C in a spectrophotometer, measuring the increase in RoF formation at 530 nm. The reaction contained 5 μM RosA in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl plus various concentrations of AF (20–100 μM) and SAM (20–250 μM) as indicated. Insets 1 and 2 represent secondary plots of the slope and y-axis intercept obtained from the primary plot against reciprocal AF concentration, resulting in a K_m^{AF} of 4 μM, a K_m^{SAM} of 70 μM, and a k_{cat} of 0.06 min^{-1}.

Rosetta 2 (DE3) and BL834 (DE3) over-producing hexahistidine-tagged RosA and SeMet-RosA, respectively, was resuspended in 50 mL HisTrap binding buffer. This buffer contained 20 mM Na$_2$HPO$_4$, pH 7.4, 500 mM NaCl and

20 mM imidazole, and was supplemented with protease inhibitors (Roche cOmplete™ EDTA-free protease inhibitor cocktail, Roche, Basel, Switzerland). For purification of SeMet-RosA, the HisTrap buffer also contained 2 mM β-mercaptoethanol. The cells were disrupted using a French press (three cycles, 2 × 10^8 Pa, 10 °C). The cell-free extract was cleared by centrifugation at 10 000 *g* for 30 min at 4 °C. The supernatant was again centrifuged at 100 000 *g* for 30 min at 4 °C to remove cell debris and unbroken cells. The cleared lysate was filtered (0.45 μm), equilibrated with loading buffer (20 mM Na$_2$HPO$_4$, pH 7.4, 500 mM NaCl, 20 mM imidazole, plus 2 mM β-mercaptoethanol for SeMET-RosA) and applied to a 5 mL HisTrap column (HisTrap™ HP; GE Healthcare). Gradient elution of the hexahistidine-tagged protein was performed using 0 to 50% elution buffer (20 mM Na$_2$HPO$_4$, pH 7.4, 500 mM NaCl, 500 mM imidazole, plus 2 mM β-mercaptoethanol for SeMET-RosA). Eluted enzyme fractions were pooled and desalted using a HiTrap™ desalting column (GE Healthcare), and equilibrated in a buffer containing 20 mM Tris/HCl, pH 8.0, with 2 mM β-mercaptoethanol for SeMET-RosA only. To achieve enzyme purity appropriate for crystallization experiments, an additional ion exchange chromatography step was performed. The enzyme fraction was applied to a MonoQ 5/50 GL column (GE Healthcare) and eluted using a linear gradient from 0 to 100% elution buffer comprising 20 mM Tris/HCl, pH 8.0, 500 mM NaCl (plus 2 mM β-mercaptoethanol for SeMET-RosA only). Aliquots of the fractions were analysed by SDS/PAGE with staining using Coomassie Brilliant Blue R-250. The homogeneous fractions were tested directly for RosA activity (see below).

Determination of native and subunit molecular masses of RosA

To determine the subunit molecular mass of RosA purified protein was loaded onto SDS/PAGE (12.5%). Protein molecular mass markers used were Low Molecular Weight protein markers (GE Healthcare, Little Chalfont, UK) containing phosphorylase b (97 kDa), albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20 kDa), and α-lactalbumin (14.4 kDa). The gel was stained with Coomassie Brilliant Blue R-250 prior to destaining (10% ethanol, 10% glacial acetic acid in water).

To determine the native molecular mass of RosA gel filtration chromatography at 25 °C was performed. The pro-

tein solution was loaded onto a Superdex 200 10/300 GL column attached to a AKTApurifier™ system (GE Healthcare) and then eluted with elution buffer (50 mM Tris-HCl, pH 8.0 containing 100 mM NaCl) at a flow rate of 0.5 mL·min⁻¹. Protein standard markers used in this experiment were ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (43 kDa), and ribonuclease a (13.7 kDa).

Isothermal titration calorimetry (ITC)

All of the experiments to determine dissociation constants (K_d) for AF, RoF, SAM and SAH to RosA were performed at 25 °C in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl using a VP-ITC system (MicroCal, Northampton, MA, USA). All solutions were degassed before measurements. Titration experiments for AF and RoF binding to RosA (30 μM, 1.42 mL) consisted of 20 injections (15 μL; duration, 30 s; spacing time, 250 s) of a solution containing 0.6 mM AF and 0.48 mM RoF. Determination of the K_d for SAM and SAH was performed by titration (20 injections) of 0.35 mM SAM (15 μL; duration, 30 s; spacing time, 250 s) and 0.48 mM SAH (15 μL; duration, 30 s; spacing time, 250 s) to a RosA solution (30 μM, 1.42 mL). The ITC raw data were analysed using ORIGIN version 7.0 (MicroCal).

Spectrophotometric titration of RosA with AF, RoF, SAM and SAH

UV/Vis absorption spectra were recorded using a Specord 200 Plus spectrophotometer (Analytik Jena, Jena, Germany) at 25 °C. To determine the dissociation constants

for binding of RoF and AF to the RosA:SAH complex by means of difference spectrophotometry, titrations were performed at 25 °C in tandem cuvettes. All solutions were prepared in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl. A solution (0.8 mL) containing RosA (25 μM) and SAH (150 μM) was placed in one of the two chambers of the sample and reference cell. The other chamber of each tandem cuvette was filled with the same volume of buffer only. After recording a baseline, RoF (0–18.2 μM) or AF (0–34.0 μM) were added to the solution containing RosA and SAH in the sample cell and to the buffer in the reference cell at 2 min intervals. Absorption spectra were recorded from 300 to 700 nm after each titration step. Similar titrations of the RosA:SAM complex (30 μM RosA and 2 mM SAM) with RoF were performed by addition of RoF (0–29.2 μM) to the complex in the sample cell and to the buffer in the reference cell. For determination of the dissociation constants for SAH and SAM from the binary RosA:RoF complex (25 μM RosA and 5 μM RoF), SAH and SAM were titrated to final concentrations of 3 and 990 μM, respectively. Absorption changes were recorded from 300 to 700 nm, and were plotted against the concentration of ligands. The dissociation constants were obtained by fitting the data using Levenberg–Marquardt algorithm (Eqn 1) implemented in the KALEIDAGRAPH software (Synergy Software, Reading, PA, USA):

$$\Delta A = \frac{\Delta A_{max}[L]}{K_d + [L]} \tag{1}$$

Steady-state kinetics of RosA

RosA catalyses the *N,N*-dimethylation of AF to yield RoF as the final product. Steady-state kinetic measurements were
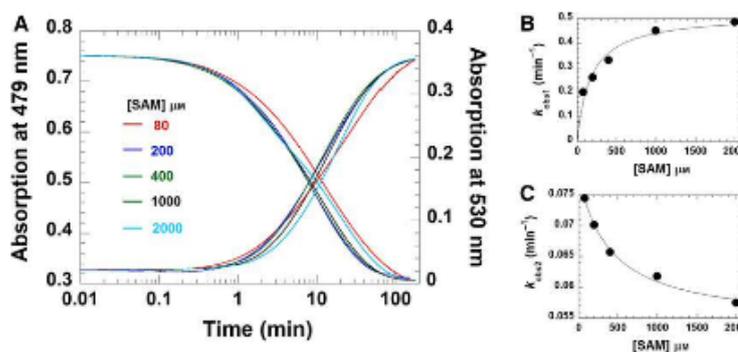
Fig. 11. Kinetic reaction of RosA:AF (1 : 1 ratio) with SAM. (A) Reaction of RosA:AF with SAM. A solution of RosA (20 μM) and AF (20 μM) in 50 mM Tris/HCl, pH 8.0, containing 100 mM NaCl, was mixed with various concentrations of SAM (80–2000 μM), and the reaction was monitored by the absorption change at 25 °C on a spectrophotometer. All concentrations are shown as final concentrations. The absorption decrease at 479 nm (left *y*-axis) and increase at 530 nm (right *y*-axis) are shown. (B) k_{obs} values for the first phase of the reaction representing monomethylation step were plotted as a function of concentration of SAM, resulting a rate constant of 0.5 ± 0.05 min⁻¹ with a K_d of 180 ± 40 μM. (C) Plot of k_{obs} for the second phase against the concentration of SAM. The plot was fitted by non-linear curve fitting, consistent with a rate constant of 0.05 ± 0.002 min⁻¹.

150

performed at 25 °C using a spectrophotometer. The assay reaction contained 5 μM RosA, various concentrations of AF (20–100 μM), and various concentrations of SAM (20–250 μM) in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl. Progress of the RosA reactions was monitored at 530 nm (ε_{530} = 25.2 mM^{-1}·cm^{-1}), at which wavelength RoF formation may be monitored without interference from other chromophores. Initial rates were obtained using KINETIC STUDIO software (TgK Scientific Ltd, Bradford-on-Avon, UK). A double-reciprocal plot of the initial rate and the substrate concentration provides the steady-state kinetic parameters, and also provides information on the mechanism of a two-substrate enzyme reaction according to Dalziel's equation (Eqn 2), in which e/v is the ratio between the enzyme concentration and the initial velocity, [A] and [B] are the concentrations of substrates A and B, respectively, and Φ values are Dalziel's coefficients [27]:

$$\frac{e}{v} = \Phi_0 + \frac{\Phi_A}{[A]} + \frac{\Phi_B}{[B]} + \frac{\Phi_{AB}}{[A][B]} \tag{2}$$

Product inhibition of RosA

Enzymatic activity of RosA in the presence of the reaction products RoF and SAH was studied by measuring the initial rate of the reaction at 25 °C using a spectrophotometer. The assay reaction contained 5 μM RosA, various fixed concentrations of AF (20–100 μM) and various fixed concentrations of SAM (20–100 μM) in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl. Concentrations of RoF in the range of 0–40 μM were used with various concentrations of AF and a fixed concentration of SAM (250 μM). The inhibition by SAH was analysed at various concentrations of SAH and SAM using a fixed concentration of AF (200 μM). Initial rates were obtained from linear data fitting using KINETIC STUDIO software (TgK Scientific Ltd). A Henderson plot of $[I]/(1 - v_i/v_0)$ and the reciprocal of the fractional initial velocity (v_0/v_i) (Eqns 3 and 4) was used to obtain inhibition constants and information on the inhibition mechanism [23]:

$$\frac{[I]}{1 - \frac{v_i}{v_0}} = K_i^{app}\left(\frac{v_0}{v_i}\right) + [E] \tag{3}$$

and for a competitive inhibitor:

$$K_i^{app} = K_i\left(1 + \frac{[S]}{K_m}\right) \tag{4}$$

In Eqns 3 and 4, [I] is the concentration of inhibitor (i.e. RoF and SAH), v_0 and v_i are the initial velocity of a reaction in the absence and presence of a inhibitor, respectively, [E] is the total enzyme concentration, [S] is the variable substrate concentration, K_m is the Michaelis constant, K_i represents the inhibition constant of a tight binding inhibitor and K_i^{app} represents the apparent inhibition constant.

Pre-steady-state reaction studies

The methylation of AF bound to RosA was investigated at various concentrations of SAM at 25 °C using a spectrophotometer. Briefly, a solution of RosA (20 μM) and AF (20 μM) in 50 mM Tris/HCl buffer, pH 8.0, containing 100 mM NaCl was pre-incubated at 25 °C for 3 min, and then the reaction was initiated by adding SAM (final concentrations of 80, 200, 400, 1000 and 2000 μM). Progress of the reaction due to consumption of AF and formation of the final product (RoF) was monitored at 479 and 530 nm, respectively. Data analysis of the kinetic traces was performed using the exponential equations from KINETIC STUDIO software (TgK Scientific Ltd) to obtain the observed rates (k_{obs}). Rate constants were determined from plots of k_{obs} as a function of the SAM concentration using the Levenberg–Marquardt algorithm implemented in the KALEIDAGRAPH software (Synergy Software).

Stopped-flow spectrophotometric studies

All experiments were performed at 25 °C using an SF-61SX2 stopped-flow spectrophotometer (TgK Scientific Ltd, Bradford-on-Avon, UK) with a 1 cm path length. Prior to all experiments, the stopped-flow instrument was flushed with 50 mM Tris/HCl buffer, pH 8, containing 100 mM NaCl several times. To analyse the effect of substrate binding order on the reaction kinetics, experiments with different mixing orders of AF and SAM were performed. Various pre-mixed solutions of RosA, RosA:AF or RosA:SAM complexes were prepared, and then mixed with solutions containing AF and SAM, SAM or AF, respectively. The progress of the reactions was monitored at 479 and 530 nm using a KinetaScan diode array detector (TgK Scientific Ltd).

Product analysis by HPLC

HPLC analysis of AF, MAF and RoF was performed using an Atlantis® dC18 reversed-phase column (5 μm, 4.6 × 250 mm; Waters, Milford, MA, USA) on an Ultamate 3000 HPLC instrument (Dionex, Sunnyvale, CA, USA) at 25 °C. All samples prepared from RosA reactions were loaded onto the column and eluted using a multi-step gradient of 100 mM formic acid and 100 mM ammonium formate (pH 3.7)/methanol (0–3 min, 30% methanol; 3–20 min, 30–75% methanol; 20–22 min, 75% methanol) at a flow rate of 0.8 mL·min^{-1}. Detection of AF, MAF and RoF was performed by measuring the absorption at 479, 490 and 509 nm, respectively.

Crystallization

Crystallization drops containing purified native RosA at a concentration of 20 mg·mL^{-1} in 20 mM Tris/HCl, pH 8.0,
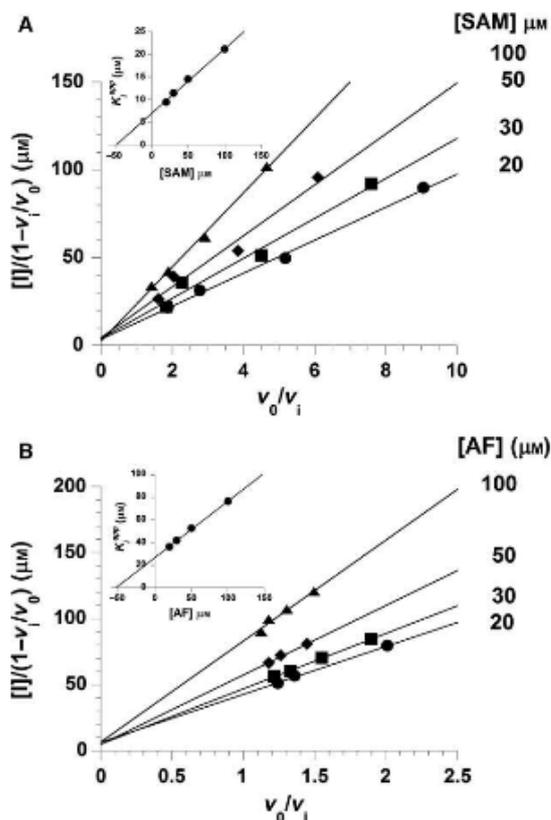
151

Fig. 12. Product inhibition of the RosA reaction. (A) SAH inhibition of the RosA reaction. A Henderson plot for tight-binding inhibition by SAH with respect to SAM was obtained by plotting $[I]/(1 - v_i/v_0)$ as a function of v_0/v_i at various fixed concentrations of SAM as indicated. Initial rates of RosA reactions were measured by the absorption change at 530 nm over time. The reactions contained 5 μM RosA, 200 μM AF, various fixed SAM concentrations (20–100 μM), and various concentrations of SAH (0, 10, 20, 40 and 80 μM). The inset is a secondary plot of K_i^{app}, slopes obtained from the primary plot, against the concentration of SAM yielding a K_i^{app} of 7 μM. (B) RoF inhibition of the RosA reaction. A Henderson plot for tight-binding inhibitor of product inhibition by RoF with respect to AF was obtained by plotting $[I]/(1 - v_i/v_0)$ as a function of v_0/v_i at various fixed concentrations of AF as indicated. Initial rates of RosA reactions were measured by the absorption change at 530 nm by time. The reactions consisted of 5 μM RosA, 250 μM SAM, various fixed AF concentrations (20–100 μM), and various concentrations of RoF (0, 10, 15, 25 and 40 μM). The inset is a secondary plot of K_i^{app}, slopes obtained from the primary plot, against concentration of AF, yielding a K_i^{RoF} of 27 μM.

500 mM NaCl, were set up with commercially available screening solutions using the microbatch method on an Oryx-7 crystallization robot (Douglas Instruments Ltd, Hungerford, UK) at 293 K. Drops were prepared by mixing equal amounts (1.0 μL) of protein and precipitant.

Table 5. Product inhibition pattern and dissociation constants (K_i) of the RosA–product inhibitor complex. Product inhibition was fitted to the tight-binding inhibitor model.

| Product	Type of inhibition	K_i (μM)
RoF	Competitive (AF)	27
SAH	Competitive (SAM)	7

Crystals of native RosA grew within 5 days after mixing the protein solution (in a 1 : 1 ratio) with a solution containing 200 mM sodium chloride, 100 mM Tris/HCl, pH 5.5, and 25% w/v PEG-3350.

The same approach was used to grow crystals of the selenomethionine-labelled derivative. Equal amounts (0.5 μL) of the variant at a concentration of 20 mg·mL^{-1} in 20 mM Tris/HCl, pH 8.0, 500 mM NaCl and 2 mM β-mercaptoethanol were mixed with precipitant containing 100 mM MES/imidazole, pH 6.5, 20 mM D-glucose, D-mannose, D-galactose, L-fructose, D-xylose and N-acetyl-D-glucosamine, 10% w/v PEG-20000 and 20% v/v PEG-MME 550 [28]. SeMet-RosA crystals grew within 7 days after crystallization set-up. For diffraction data collection, crystals were harvested from their mother liquor using Cryo-Loops™ (Hampton Research, Aliso Viejo, CA, USA), and flash-cooled in liquid nitrogen without any additional cryoprotection.

Structure determination and refinement

Diffraction data to a resolution of 2.2 Å were collected for native RosA from a single triclinic crystal (space group *P*1) at beamline ID23-1 of the European Synchrotron Radiation Facility (Grenoble, France). MAD data to a resolution of approximately 3.5 Å were collected at beamline X06DA-PXIII of the Swiss Light Source at the Paul Scherrer Institute (Villigen, Switzerland). Three MAD datasets were collected from a single tetragonal SeMet-RosA crystal (space group $I4_122$) at various wavelengths (peak, inflection, remote) that were determined from an X-ray fluorescence scan recorded around the selenium absorption edge. Data processing and scaling for the native and SeMet MAD datasets were performed using XDS [29], SCALA and AIMLESS [30].

The calculated Matthews coefficient [30,31] of the MAD datasets indicated a 99.9% probability for the presence of one RosA molecule per asymmetric unit of the crystal. To identify selenium sites, the two MAD pipelines AUTO-RICK-SHAW [32,33] and PHENIX AutoSol [34,35] were used. The combination of these programs yielded first phases and a partial structure. These phases were further used as input phases for the automated chain-tracing/building programs PHENIX AUTOBUILD [36] and BUCCANEER [37] to extend the model and to improve its quality. The combined model of both rebuilding programs produced interpretable electron densities, with preliminary values of $R = 37\%$ and

152
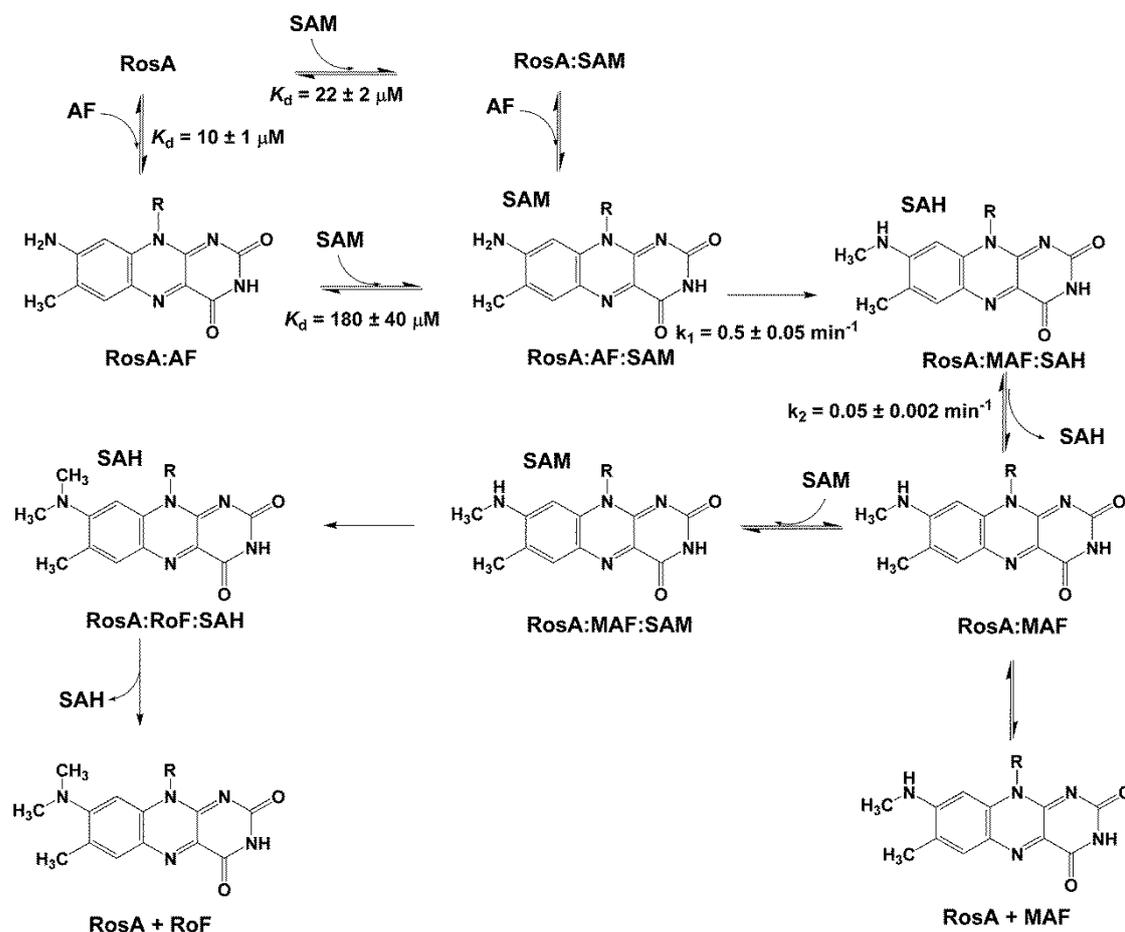
Scheme 3. Suggested reaction scheme for the reactions catalysed by RosA based on our binding studies and kinetic measurements.

$R_{\text{free}} = 49\%$. R_{free} values were computed from a set of randomly chosen reflections (5%), which were not used during refinement [38]. Structure refinement was performed using the refinement programs PHENIX REFINE [34] and REFMAC [39]. Model rebuilding was performed using BUCCANEER [37] and COOT [40] by alternate automated rebuilding and real-space fitting against σ_A-weighted $2 F_o - F_c$ and $F_o - F_c$ electron density maps together with least-squares optimization.

The improved MAD model was used as an initial molecular replacement template for PHASER [41] to determine the structure of native RosA in the triclinic crystal form. Molecular replacement resulted in six molecules in the asymmetric unit. Structure refinement and model building were performed using PHENIX REFINE [34] and COOT [40] by real-space fitting against σ_A-weighted $2 F_o - F_c$ and $F_o - F_c$ electron density maps and least-squares optimizations. Water molecules were placed into the difference

electron density map, and accepted or rejected on the basis of geometry criteria as well as refined B-factors. In later stages of the refinement, four TLS groups per protomer were defined based on an analysis using the TLSMD web server [42]. The final model was refined to $R = 20\%$ and $R_{\text{free}} = 24\%$. Validation of the structure was performed using MOLPROBITY [43], yielding a Ramachandran plot with 97.20% of the residues in favoured regions, 2.65% in allowed regions and 0.15% in disallowed regions. Data statistics and details of structure refinement are given in Table 1 (MAD data) and Table 2 (native data).

Prediction of the SAM/SAH binding site

In order to locate putative binding sites for the substrate AF and the cofactor SAM, the program CASOX [19] together with the web tools PROSITE (sequence-based prediction) [20], 3DLIGANDSITE (sequence- and structure-based

153

prediction) [21] and COACH (sequence- and structure-based prediction) [22] were used. The results of the cavity analysis using CASOX were compared with consensus ligand binding residues predicted by PROSITE, 3DLIGANDSITE and COACH. COACH additionally generated putative enzyme–ligand model complexes and ranked them according to an intrinsic C-score. This C-score represents a confidence score for the prediction. C-scores range between 0 and 1, with a higher score indicating a more reliable prediction. The highest-rated complex, together with other related methyltrans-ferase structures, were superimposed with RosA to identify ligand binding modes. The modelled RosA–ligand complex was further energy-minimized using the program YASARA with the AMBER03 force field, employing the standard optimization protocol [44]. Structure alignments were performed using PYMOL (https://www.pymol.org/).

Acknowledgements

Author contributions

M.M. and P.M. initiated the project; C.T., M.K.U., M.M., K.G. and P.M. designed the experiments and analysed the data; C.T. and F.J. expressed and purified RosA; M.K.U. and K.G. crystallized RosA and determined the crystal structure; C.T. performed analytical and biochemical experiments, and determined dissociation constants as well as kinetic parameters; C.T., M.K.U., M.M., K.G. and P.M. wrote the manuscript.

References

1 Langer S, Hashimoto M, Hobl B, Mathes T & Mack M (2013) Flavoproteins are potential targets for the antibiotic roseoflavin in *Escherichia coli*. *J Bacteriol* **195**, 4037–4045.

2 Langer S, Nakanishi S, Mathes T, Knaus T, Binter A, Macheroux P, Mase T, Miyakawa T, Tanokura M & Mack M (2013) The flavoenzyme azobenzene reductase AzoR from *Escherichia coli* binds roseoflavin mononucleotide (RoFMN) with high affinity and is less active in its RoFMN form. *Biochemistry* **52**, 4288–4295.

3 Pedrolli D, Langer S, Hobl B, Schwarz J, Hashimoto M & Mack M (2015) The ribB FMN riboswitch from *Escherichia coli* operates at the transcriptional and translational level and regulates riboflavin biosynthesis. *FEBS J* **282**, 3230–3242.

4 Pedrolli DB, Jankowitsch F, Schwarz J, Langer S, Nakanishi S, Frei E & Mack M (2013) Riboflavin analogs as antiinfectives: occurrence, mode of action, metabolism and resistance. *Curr Pharm Des* **19**, 2552–2560.

5 Pedrolli DB, Jankowitsch F, Schwarz J, Langer S, Nakanishi S & Mack M (2014) Natural riboflavin analogs. *Methods Mol Biol* **1146**, 41–63.

6 Pedrolli DB & Mack M (2014) Bacterial flavin mononucleotide riboswitches as targets for flavin analogs. *Methods Mol Biol* **1103**, 165–176.

7 Pedrolli DB, Matern A, Wang J, Ester M, Siedler K, Breaker R & Mack M (2012) A highly specialized flavin mononucleotide riboswitch responds differently to similar ligands and confers roseoflavin resistance to *Streptomyces davawensis*. *Nucleic Acids Res* **40**, 8662–8673.

8 Pedrolli DB, Nakanishi S, Barile M, Mansurova M, Carmona EC, Lux A, Gartner W & Mack M (2011) The antibiotics roseoflavin and 8-demethyl-8-amino-riboflavin from *Streptomyces davawensis* are metabolized by human flavokinase and human FAD synthetase. *Biochem Pharmacol* **82**, 1853–1859.

9 Juri N, Kubo Y, Kasai S, Otani S, Kusunose M & Matsui K (1987) Formation of roseoflavin from 8-amino- and 8-methylamino-8-demethyl-D-riboflavin. *J Biochem* **101**, 705–711.

10 Matsui K, Juri N, Kubo Y & Kasai S (1979) Formation of roseoflavin from guanine through riboflavin. *J Biochem* **86**, 167–175.

11 Jankowitsch F, Kuhm C, Kellner R, Kalinowski J, Pelzer S, Macheroux P & Mack M (2011) A novel N, N-8–amino-8–demethyl-d–riboflavin dimethyltransferase (RosA) catalyzing the two terminal steps of roseoflavin biosynthesis in Streptomyces davawensis. *J Biol Chem* **286**, 38275–38285.

12 Jankowitsch F, Schwarz J, Ruckert C, Gust B, Szczepanowski R, Blom J, Pelzer S, Kalinowski J & Mack M (2012) Genome sequence of the bacterium *Streptomyces davawensis* JCM 4913 and heterologous production of the unique antibiotic roseoflavin. *J Bacteriol* **194**, 6818–6827.

13 Kelley LA & Sternberg MJ (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat Protoc* **4**, 363–371.

14 Shi J, Blundell TL & Mizuguchi K (2001) FUGUE: sequence–structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J Mol Biol* **310**, 243–257.

15 Claude J-B, Suhre K, Notredame C, Claverie J-M & Abergel C (2004) CaspR: a web server for automated molecular replacement using homology modelling. *Nucleic Acids Res* **32**, W606–W609.

154

16 Sillitoe I, Cuff AL, Dessailly BH, Dawson NL, Furnham N, Lee D, Lees JG, Lewis TE, Studer RA, Rentzsch R *et al.* (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures. *Nucleic Acids Res* **41**, D490–D498.

17 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.

18 Holm L & Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–W549.

19 Steinkellner G (2012) CASoX – a software tool to visualize and analyze cavities and clefts in protein structures. Presented at BIOCOMP 2012: International Conference on Bioinformatics and Computational Biology, 19 September 2012 to 22 September 2012, Varna, Bulgaria.

20 Sigrist CJA, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L & Xenarios I (2013) New and continuing developments at PROSITE. *Nucleic Acids Res* **41**, D344–D347.

21 Wass MN, Kelley LA & Sternberg MJE (2010) 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Res* **38**, W469–W473.

22 Yang J, Roy A & Zhang Y (2013) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics* **29**, 2588–2595.

23 Henderson PJ (1972) A linear equation that describes the steady-state kinetics of enzymes and subcellular particles interacting with tightly bound inhibitors. *Biochem J* **127**, 321–333.

24 Otani S, Kasai S & Matsui K (1980) Isolation, chemical synthesis, and properties of roseoflavin. *Methods Enzymol* **66**, 235–241.

25 Shapiro SK & Ehninger DJ (1966) Methods for the analysis and preparation of adenosylmethionine and adenosylhomocysteine. *Anal Biochem* **15**, 323–333.

26 Guerrero SA, Hecht HJ, Hofmann B, Biebl H & Singh M (2001) Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. *Appl Microbiol Biotechnol* **56**, 718–723.

27 Dalziel K (1957) Initial steady state velocities in the evaluation of enzyme–coenzyme–substrate mechanisms. *Acta Chem Scand* **11**, 1706–1723.

28 Gorrec F (2009) The MORPHEUS protein crystallization screen. *J Appl Crystallogr* **42**, 1035–1042.

29 Kabsch W (2010) XDS. *Acta Crystallogr* **D66**, 125–132.

30 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A *et al.* (2011) Overview of the CCP4

suite and current developments. *Acta Crystallogr D* **67**, 235–242.

31 Kantardjieff KA & Rupp B (2003) Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Sci* **12**, 1865–1871.

32 Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS & Tucker PA (2005) Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D* **61**, 449–457.

33 Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS & Tucker PA (2009) On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr D* **65**, 1089–1097.

34 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* **66**, 213–221.

35 Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH & Hung L-W (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D* **65**, 582–601.

36 Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung L-W, Read RJ & Adams PD (2007) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D* **64**, 61–69.

37 Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D* **D62**, 1002–1011.

38 Kleywegt GJ (2000) Validation of protein crystal structures. *Acta Crystallogr D* **56**, 249–265.

39 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* **67**, 355–367.

40 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* **66**, 486–501.

41 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.

42 Painter J & Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D* **62**, 439–450.

155

43 Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **66**, 12–21.

44 Krieger E, Darden T, Nabuurs SB, Finkelstein A & Vriend G (2004) Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* **57**, 678–683.

CHAPTER 5

5. Appendix

List of abbreviations

5'-dA·	5'-deoxyadenosyl 5'-radical
ADP	Adenosine diphosphate
AF	8-amino-8-demethyl-D-riboflavin
AFP	8-demethyl-8-amino-riboflavin-5'-phosphate
ATP	Adenosine triphosphate
BHMT	Betaine-homocysteine methyltransferase
CBS	Cystathionine β-synthase
COMT	Catechol O-methyltransferase
CNS	Central nervous system
CSE	Cystathionine γ-lyase
CH$_2$THF	Methylenetetrahydrofolate
DAO	Diamine oxidase
dTMP	2'-Deoxythymidine-5'-monophosphate
dUMP	2'-Deoxyuridine-5'-monophosphate
HCD	Histidine decarboxylase
Hcy	Homocysteine
H$_2$folate	Dihydrofolate
HNMT	Histamine *N*-methyltransferase
HO$_2$C-RP	8-demethyl-8-carboxy-riboflavin-5'-phosphate
MAF	8-methylamino-8-demethyl-D-riboflavin
MAT	Methionine adenosyltransferase
MS	Methionine synthase
MTase	Methyltransferase
MTHF	5-methyl-tetrahydrofolate
MTHFR	Methylenetetrahydrofolate reductase
OHC-RP	8-demethyl-8-formyl-riboflavin-5'-phosphate (or CP1)
PLP	Pyridoxal 5'-phosphate
RoF	Roseoflavin, 8-dimethylamino-8-demethyl-D-riboflavin
RosA	*N*,*N*-8-amino-8-demethyl-D-riboflavin dimethyltransferase
RosB	8-demethyl-8-aminoriboflavin-5'-phosphate synthase
Rubisco	Ribulose-1,5-bisphosphate carboxylase
SAM	*S*-adenosylmethionine (AdoMet)
SAM-MTase	*S*-adenosylmethionine-dependent methyltransferase
SAH	*S*-adenosylhomocysteine
TSase	Thymidylate synthase

List of publications

PUBLICATIONS ON REVIEWED JOURNALS

1. Tongsook, C., Uhl, M.K., Jankowitsch, F., Mack, M., Gruber, K., and Macheroux, P. (2016) Structural and kinetic studies on RosA, the enzyme catalysing the methylation of 8-demethyl-8-amino-d-riboflavin to the antibiotic roseoflavin. FEBS J, 283, 1531-1549

2. Heidari, A., Tongsook, C., Najafipour, R., Musante, L., Vasli, N., Garshasbi, M., Hu, H., Mittal, K., McNaughton, A.J., Sritharan, K., Hudson, M., Stehr, H., Talebi, S., Moradi, M., Darvish, H., Arshad, Rafiq, M., Mozhdehipanah, H., Rashidinejad, A., Samiei, S., Ghadami, M., Windpassinger, C., Gillessen-Kaesbach, G., Tzschach, A., Ahmed, I., Mikhailov, A., Stavropoulos, D.J., Carter, M.T., Keshavarz, S., Ayub, M., Najmabadi, H., Liu, X., Ropers, H.H., Macheroux, P., and Vincent, J.B. (2015) Mutations in the histamine *N*-methyltransferase gene, HNMT, are associated with nonsyndromic autosomal recessive intellectual disability. Hum Mol Genet, 24, 5697-5710

3. Bernkopf, M., Webersinke, G., Tongsook, C., Koyani, C.N., Rafiq, M.A., Ayaz, M., Müller, D., Enzinger, C., Aslam, M., Naeem, F., Schmidt, K., Gruber, K., Speicher, M.R., Malle, E., Macheroux, P., Ayub, M., Vincent, J.B., Windpassinger, C., and Duba, H.C. (2014) Disruption of the methyltransferase-like 23 gene METTL23 causes mild autosomal recessive intellectual disability. Hum Mol Genet, 23, 4015-4023

4. Sucharitakul, J., Tongsook, C., Pakotiprapha, D., van Berkel, W.J., and Chaiyen, P. (2013) The reaction kinetics of 3-hydroxybenzoate 6-hydroxylase from Rhodococcus jostii RHA1 provide an understanding of the para-hydroxylation enzyme catalytic cycle. J Biol Chem, 288, 35210-35221

5. Tongsook, C., Sucharitakul, J., Thotsaporn, K., and Chaiyen, P. (2011) Interactions with the substrate phenolic group are essential for hydroxylation by the oxygenase component of *p*-hydroxyphenylacetate 3-hydroxylase. J Biol Chem, 286, 44491-44502

6. Ruangchan, N., Tongsook, C., Sucharitakul, J., and Chaiyen, P. (2011) pH-dependent studies reveal an efficient hydroxylation mechanism of the oxygenase component of *p*-hydroxyphenylacetate 3-hydroxylase. J Biol Chem, 286, 223-233

ADDITIONAL MANUSCRIPTS IN PREPARATION AND REVISION

1. Tongsook, C., Niederhauser, J., Kronegger, E., Straganz, G., and Macheroux, P. (2016) Leucine 208 in human histamine N-methyltransferase emerges as a hotspot for protein stability rationalizing the role of the L208P variant in intellectual disability

2. Dhammaraj, T., Tongsook, C., and Chaiyen, C. (2016) A single-site mutation at Ser146 expands the reactivity of the oxygenase component of *p*-hydroxyphenylacetate 3-hydroxylase

3. Pornsuwan, S., Tongsook, C., Juttulapa, M., Watthaisong, P., Thotsaporn, K., Nijvipakul, S., Maenpuen, S., and Chaiyen, P. (2016) Biochemical Characterizations of Fe(II)-containing 3,4-Dihydroxyphenylacetate 2,3-dioxygenase from *Pseudomonas aeruginosa*

POSTERS PRESENTATIONS IN INTERNATIONAL CONFERENCES

1. Tongsook, C, Uhl, M.K., Mack, M., Gruber, K., Macheroux, P. (2015) Exploration of the catalytic function and structure of N,N-8-amino-8-demethyl-D-riboflavin dimethyltransferase (RosA), in Enzymes, Coenzymes and Metabolic Pathways, Gordon Research Seminar and Conference, NH, USA

2. Tongsook, C, Uhl, M.K., Mack, M., Gruber, K., Macheroux, P. (2014) Exploration of the catalytic function and structure of N,N-8-amino-8-demethyl-D-riboflavin dimethyltransferase (RosA), in IUBMB Symposium: the 18th International Symposium on Flavins and Flavoproteins, Phetchaburi, Thailand

3. Tongsook, C., Windpassinger, C., Vincent J.B., Macheroux, P. (2014) Effect of a single amino acid exchange in human histamine N-methyltransferase, in Enzyme Mechanisms by Biological Systems, EMBO conference, Manchester, UK

4. Tongsook, C., Bernkopf, M., Windpassinger, C., Macheroux, P. (2013) Exploration of the Function of Human Methyltransferase Like 23 (METTL23), in Mechanisms in Biology, Federation of European Biochemical Societies Congress, St. Petersburg, Russia

ORAL PRESENTATIONS

1. Tongsook, C., Uhl, M.K., Jankowitsch, F., Mack, M., Gruber, K., and Macheroux, P. (2016) Structure and mechanism of RosA, the enzyme catalysing the methylation of 8-demethyl-8-amino-D-riboflavin to the antibiotic roseoflavin, 15th Doc Day NAWI Graz Doctoral School Molecular Biosciences and Biotechnology, Graz

Curriculum vitae

<u>Contact Information and Personal Information</u>

Name	Chanakan TONGSOOK
Address	9/3, Moo.9, Nongaor, Banpong, Ratchaburi province, Thailand, 70110
Mobile	+66/863855956
E-mail	sc_chanakan@yahoo.com & tongsook.c@gmail.com
Date and place of birth	14th December 1984, Kanchanaburi (Thailand)
Citizenship	Thai

<u>Education</u>

1991 – 1997	Elementary School in Ratchaburi, Thailand
1997 – 2000	Junior High School in Ratchaburi, Thailand
2000 – 2003	Senior High School in Nakhonpathom, Thailand
2003 – 2007	Bachelor of Science in Chemistry, Faculty of Science, Silpakorn University, Nakhonpathom, Thailand
04/2007 – 03/2010	Master of Science in Biochemistry, Faculty of Science, Mahidol University, Bangkok, Thailand (Supervisor: Prof. Pimchai Chaiyen)
04/2012 – 07/2016	PhD project at the Institute of Biochemistry, Graz University of Technology, Graz, Austria

<u>Research Activities</u>

03/2002 – 08/2002	Project at the Department of Chemistry, Silpakorn University with Assist.-Prof. Narong Chimpalee: "Determination of nitrogen in fish sauces and soy bean sauces by Flow-injection spectrophotometric method"
02/2006 – 01/2007	Project at the Department of Chemistry, Silpakorn University with Assist.-Prof. Radchada Buntem: "Synthesis and Characterizations of Ethenyl – linked Porphyrin"

162

06/2007 – 02/2010	Master thesis project entitled "Investigation on the Substrate Binding Residues of the Oxygenase Component of *p*-Hydroxyphenylacetate Hydroxylase" (Supervisor: Prof. Pimchai Chaiyen)
04/2010 – 03/2012	Project at the Biochemistry, Faculty of Science, Mahidol University (worked as a research assistant) with Prof. Pimchai Chaiyen: "Biochemical Characterizations of Fe(II)-containing 3,4-Dihydroxyphenylacetate 2,3-dioxygenase from *Pseudomonas aeruginosa*" and "Kinetic studies on 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1"
04/2012 – 07/2016	Worked on PhD project entitled "Biochemical characterization and kinetic studies on *S*-adenosylmethionine-dependent methyltransferases" (Supervisor: Prof. Peter Macheroux)
Further Occupaion activities	
2006 – 2007	Teaching assistant in Analytical Chemistry Laboratory Course, for undergraduate students, Faculty of Science, Silpakorn University, Thailand
2008 – 2010	Teaching assistant in Spectroscopic Technique (SCID508) and Receptor Binding & Enzyme Kinetic Assays (SCID 512), for graduate students, Faculty of Science, Mahidol University, Thailand
04/2010 – 03/2012	Research assistant at the Department of Biochemistry, Faculty of Science, Mahidol University
Other Qualifications	
Language Skills	Thai as mother tongue Fluent in English
Computer Skills	Microsoft Office, Bioinformatics (databases, sequence alignments, Pymol, ect.), Origin 7.0 for ITC data analysis