The Histidine Methyltransferase Universe

Characterization of METTL9 and CARNMT1 –
Two Novel Protein Histidine Methyltransferases

Thesis for the degree of *Philosophiae Doctor*

by

Lisa Schroer

Department of Biosciences

Faculty of Mathematics and Natural Sciences

University of Oslo

2022

Für Mama & Papa

Für Jannis & Niklas

Für Oma & Opa

Danke.

Table of Contents

Acknowledgements

The work presented in this thesis was carried out at the Department of Biosciences, University of Oslo. The position was financially supported by a PhD fellowship from the University of Oslo. This thesis would not have been possible without all the exceptional people in my life. With these acknowledgements, I would like to thank them.

Pål – You have a brilliant mind and your curiosity and knowledge on methyltransferases inspired me on my 4 year journey as your PhD student. I am truly grateful for the freedom that you have given me during my journey and for letting me be a part of your team. Thank you also for your support for my internship in Mexico and the helpful and very detailed feedback on my thesis. These experiences have helped me become the independent person that I am now. Thank you also for always being kind to me and others.

Erna – Over the last 4 and a half years you have become much more than just a supervisor. You have become one of the most important people in my life. You have been my red thread in my PhD, the daily advice, the voice to talk me down when my mind was flying away, and the wings to lift me up when I was on the floor. I owe this PhD to you, because if it wasn't for you, I would have quit a long time ago. Thank you, you are an amazing supervisor and my dearest friend.

The Falnes lab – Jedrek, Carmen, Mads, Kamilla, Melanie! Thank you for helping out in need, giving me feedback, and mostly thank you for joking around and singing or dancing along to our favorite "radio norgeee" in the lab. Thank you for the fun trips to the conferences, the work parties and the invitations to your home countries (Jedrek & Erni ☺). Thanks for the daily interesting and fun discussions at lunch. Good luck in the future to you all and please someone publish my cool projects! Thanks also to all our wonderful Master students, especially Shabi, Anniken and Ayse, I wish you the best in your (scientific) future!

Marta – When we first met, you just came back from your "boat year" and over time we started to not only share dugnads, gym and lønnignspils together but even crystallography! (Let's pop the bottle for the last one if anything!). Thank you for being not just my colleague (who worked dead hard on the METTL9 structure and manuscript), but mostly for being my lovely friend. I will really miss our daily coffee breaks, but who knows, maybe one day we will work together again?

Hans-Petter – Thank you for being so eager in solving the METTL9 structure even though you were so busy with teaching. You are an amazing teacher and colleague and I wish you all the best!

Kristian – I think you are the most well-read and studied man I know. Your wit and knowledge about literally everything is astonishing and always a welcome chat at lunch. I will miss these stories and you!

Julien – Frenchie! When I think about future professors I think about you. You are the whole package. The passion, the attitude and the skills. Thank you for helping me making beautiful microscope pictures, as everything you do, the quality was excellent and the time spent with you in the lab was not just informative but also fun!

Bernd and **Margarita** – Thank you Bernd for being my co-supervisor and thank you Margarita for being the organized all-rounder at BMB!

Innkjøpsavdeling – **Steinar, Christopher, Gro, Maciej** – You are the heart and the whole cardiovascular system of IBV. Honestly, without you we would all be stuck to writing reviews, no one could do proper lab work. Thank you for being organized and kind when all you're surrounded by is awkward scientists. You deserve the Nobel peace prize for sure, with all the f-ups we do with ordering, you fix it for us. Thank you, "the one and only" Lisa ☺

IT – **Lars, Ola, Bobben** – If innkjøp is the heart and cardiovascular system, then you are the lungs. Thanks for supporting us smart scientists with how to turn a PC on. We need the occasional "have you tried turning it on and off again". Thank you guys, you are awesome. Special thanks to Lars specifically for accompanying me for almost 2 years of my PhD journey and introducing me to your amazing family.

Rest of IBV – Shout out to all the great colleagues at IBV with whom I've not collaborated, but who have enlightened my time here such as **Ada, Adrian, Jens, Nils, Daniel, Per Eugen** and the entire **Ciosk group!** I know I have most likely forgotten a lot of people, so thanks to you as well!

Grünenthal México – Muchísimas gracias por la oportunidad de trabajar con ustedes. Los tres meses en la ciudad que pasé con ustedes fueron una experiencia muy importante en mi vida. He aprendido mucho. Muchísimas gracias por todo. ¡Hasta luego mis amigos!

Daniel Soule – Thank you for correcting my thesis and giving me feedback from a viewpoint outside of academia, from a real professional writer!

Kygo – Your music inspired me to pack my bags and start a new life in Norway, thank you!

Severns family – Emma, Brandon, Kathy and Brian. My international journey started with you. If you hadn't had made my journey to Minnesota so wonderful and exciting, I would never have dared to study in the Netherlands or do this PhD in Norway. Thank you for adopting me and showing me my first foreign and exciting part of the world. I will forever be grateful.

Meine besten Freunde der Welt (in alphabetischer Reihenfolge ;))
Dagmar, Jil, FK, Julia, Lena, Lynn, Neele, Miro, Philip, Pia, Sarah, Torsten und alle anderen tollen Freunde in meinem Leben – Ihr seid über viiiel zu viele Jahre mit viel zu wenig Treffen trotzdem immer an meiner Seite gewesen. Egal ob man mal Monate nicht miteinander redet, ihr seid mein Netzwerk auf das ich zurückfallen kann. Ohne eure Unterstützung hätte ich den PhD schon lange aufgegeben, denn die spannenden Besuche, stundenlange Anrufe, und ermutigenden Nachrichten waren der Grund jeden Tag erneut mit einem Lächeln zu starten. Ich freue mich auf unsere weiteren gemeinsamen Jahre zusammen. Danke für alles.

Of course also shout out to my numerous non-German friends such as my best boxing friends **Ninja, Reza and Jamshid**!☺ Thank you for the fun times to keep my mind off work. Especially thanks to my wonderful roommates **Regine, Ole, Julie, Maddie, Ingrid, Linnea**, and of course **Warner** who have kept me (in)sane before, during, and after Corona isolation.

Meine liebe Familie – Das Beste kommt immer zum Schluss! Hinter all der Arbeit steht immer meine Familie. Am Ende des Tages hätte ich nichts von alledem machen können hätte ich nicht die (finanzielle) Unterstützung von euch bekommen. Vom Umzug, über finanzielle Spritzen oder die mentale Zusicherung. Nichts von meinem Weltbummel hätte ich ohne euch jemals anfangen können, geschweige denn es bis zum Beenden eines PhDs bringen können. Danke für die Freiheit, die Unterstützung (und den billigen Alkohol), dieses Buch ist für euch. Danke.

♡

Thesis Summary

Methyltransferases (MTases) are enzymes that attach methyl (CH_3) groups to a number of substrates, such as nucleic acids, proteins, and various small molecules. In proteins, the most commonly methylated amino acid residues are lysine and arginine. However, recent accumulating evidence indicates that also histidine is frequently methylated. This remarkable amino acid can be methylated on two positions of its imidazole ring - either the proximal (π) or the distal (τ) nitrogen atom. During the work described in this thesis, we characterized two MTases that specifically generate π-methylhistidine in proteins.

The first enzyme that we discovered is **METTL9**, an MTase that generates π-methylhistidine in many human proteins. We showed that it has multiple substrates, which share a common alternating histidine, or "HxH" motif, where the "x" is preferably a small, uncharged amino acid. As for the biological significance of the methylation, we found that METTL9 activity is important for mitochondrial function, and that histidine methylation of a zinc transporter-derived peptide reduces its affinity for zinc.

The second part of the **METTL9** story is the characterization of putative METTL9 orthologues. Here, we investigated nine eukaryotic orthologues, among which are those from the fruit fly (*Drosophila melanogaster*) and the picoplankton *Ostreococcus tauri (O. tauri)*. Interestingly, we discovered that they have substrate specificities that deviated from that of human METTL9. In addition, we solved the protein structure of METTL9 from *O. tauri* which shows a conserved active site, and a model was proposed for how METTL9 binds to its peptide substrate. Our findings present an important contribution to the understanding of substrate specificity and mode of action in METTL9 enzymes.

The second MTase that we investigated is **CARNMT1,** which was reported to methylate a histidine residue on the dipeptide carnosine. We, in turn, found evidence that CARNMT1 is also able to generate π-methylhistidine in a model protein substrate derived from a histidine-rich fragment of a zinc transporter. We also showed that CARNMT1 is able to methylate at least seven, yet unidentified proteins in human cell extracts.

Our discoveries unveil a small part of the growing universe of (π) histidine methylation, and sets the stage for future studies on this fascinating but still unexplored protein modification.

Norsk Sammendrag

Metyltransferaser er enzymer som setter metylgrupper (CH_3) på en rekke substrater som nukleinsyrer, proteiner og ulike småmolekyler. I proteiner er det vanligvis aminosyrene lysin og arginin som blir metylert. Imidlertid indikerer nyere forskning at også aminosyren histidin ofte blir metylert. Denne bemerkelsverdige aminosyren kan bli metylert i to posisjoner på sin imidazol-ring – enten på det proksimale (π) eller på det distale (τ) nitrogenatomet. Gjennom arbeidet som er beskrevet i denne avhandlingen, karakteriserte vi to MTaser som spesifikt genererer π-metylhistidin i proteiner.

Det første enzymet vi oppdaget er METTL9, en metyltransferase som introduserer π-methylhistidine i en rekke menneskeproteiner. Vi viste at det har en rekke substrater, som alle har et alternerende-histidin-, eller "HxH"-motiv, hvor "x" som regel er en liten og ikke-ladet aminosyre. Når det gjelder den biologiske signifikansen av metyleringen, fant vi at METTL9 er viktig for at mitokondriene skal fungere optimalt, og også at metylering av et peptid avledet fra en sink-transportør reduserer dets affinitet for sink.

Den andre delen av **METTL9**-historien er karakteriseringen av antatte METTL9-ortologer. Her undersøkte vi ni eukaryote ortologer, blant dem fra fruktflua *Drosophila melanogaster* og pikoplanktonet *Ostreococcus tauri* (*O. tauri*). Interessant nok oppdaget vi at de har substratspesifisiteter som avvek fra humant METTL9. I tillegg løste vi proteinstrukturen til METTL9 fra *O. tauri*, som viser et konservert aktivt sete, og en modell for hvordan METTL9 binder til et peptid-substrat ble foreslått. Våre funn tilfører et viktig bidrag i forståelsen av substratspesifisitet samt virkningsmåten til METTL9 enzymer.

Den andre MTasen som vi undersøkte er **CARNMT1**, som var rapportert å metylere en histidin i dipeptidet karnosin. Videre fant vi at CARNMT1 også er i stand til å generere π-metylhistidin i et modell-protein-substrat som stammer fra et histidinrikt fragment av en sinktransportør. Vi viste også at CARNMT1 er i stand til å metylere minst syv stadig uidentifiserte proteiner i humane celleekstrakter.

Våre oppdagelser avslører en liten del av det voksende universet av (π) histidin-metylering, og legger grunnlaget for fremtidige studier på denne fascinerende, men fortsatt uutforskede proteinmodifikasjonen.

9

Abbreviations

7BS	Seven-β-strand
ANGST	Alanine, asparagine, glycine, serine or threonine
BAG2	BAG family molecular chaperone regulator 2
CARNMT1	Carnosine methyltransferase
CARNS1	Carnosine synthase
Dm	*Drosophila Melanogaster*
ER	Endoplasmic reticulum
FM	Fibromyalgia syndrome
GAR	glycine-arginine rich
HCC	Hepatocellular carcinoma
hisMeH	histidine methylation
Hpm1	Histidine protein methyltransferase 1
HxH	Alternating histidine (H) sequence, "x" = small, uncharged amino acid
IUPAC	International union of pure and applied chemistry
MAP7	Microtubule-stabilizing protein 7
METTL18	Methyltransferase like 18
METTL9	Methyltransferase like 9
MS	Mass spectrometry
MTase	Methyltransferase
MYLK2	Myosin-light-chain kinase
Ot	*Ostreococcus tauri*
OSBPL1A	Oxysterol-binding protein-related protein 1
p53	tumor protein 53
PTM	Post-translational modification
Rpl3	Ribosomal protein L3
SAM	*S*-adenosyl-methionine
SAH	*S*-adenosyl-homocysteine
SARM1	NAD(+) hydrolase
SET	Suppressor of variegation, Enhancer of zeste, and Trithorax
SGC	Scirrhous gastric cancer

List of Papers

Paper I

The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes

Erna Davydova, Tadahiro Shimazu, Maren Kirstin Schuhmacher, Magnus E. Jakobsson, Hanneke L. D. M. Willemen, Tongri Liu, Anders Moen, Angela Y. Y. Ho1, Jędrzej Małecki, Lisa Schroer, Rita Pinto, Takehiro Suzuki, Ida A. Grønsberg, Yoshihiro Sohtome, Mai Akakabe, Sara Weirich, Masaki Kikuchi, Jesper V. Olsen, Naoshi Dohmae, Takashi Umehara, Mikiko Sodeoka, Valentina Siino, Michael A. McDonough, Niels Eijkelkamp, Christopher J. Schofield, Albert Jeltsch, Yoichi Shinkai & Pål Ø. Falnes

Nature Communications, 2021, https://doi.org/10.1038/s41467-020-20670-7

Manuscript I

METTL9-like histidine methyltransferases from different organisms display distinct substrate specificities

Lisa Schroer, Marta Hammerstad, Sara Weirich, Hans-Petter Hersleth, Ida Andrietta Grønsberg, Lars Hagen, Geir Slupphaug, Albert Jeltsch, Erna Davydova, Pål Ø. Falnes

Manuscript II

The human carnosine Nπ-histidine methyltransferase CARNMT1 methylates proteins

Lisa Schroer, Erna Davydova, Marijke Baltissen, Nelleke Spruijt, Lars Hagen, Geir Slupphaug, Michiel Vermeulen, Pål Ø. Falnes

Introduction

The Fine-Tuning of Proteins: Post Translational Modifications

The human genome contains about 3.2 billion base pairs of information that together form the manuscript for the existence and survival for all of our cells [1]. This DNA, however, is not the executive power of our cells: the double helix first has to be transcribed into RNA which is subsequently translated into proteins, summarizing the central dogma of biology. The resulting proteins are the delicate workers that regulate cell homeostasis and ensure cell survival [2].

To fine-tune a harmonious collaboration, many proteins require chemical modification after translation that help diversify the limitations of the genetic information. These post-translational modifications (PTMs), represent one of the most important mechanisms that diversify and expand the proteome [3]. They make it possible that proteins consist of more than 140 different residues, while DNA encodes merely 20 primary amino acids. PTMs come in a variety of reversible and irreversible chemical changes, ranging from enzymatic cleavage of peptide bonds to the covalent addition of chemical groups [4]. The top five most common PTMs are phosphorylation, acetylation, ubiquitination, succinylation and methylation (Fig. 1) which help to optimize and dynamically regulate the activity of the targeted protein [5]. Methylation is particularly interesting as it is originally associated with histone modification, but more and more non-histone proteins are found to be methylated as well.
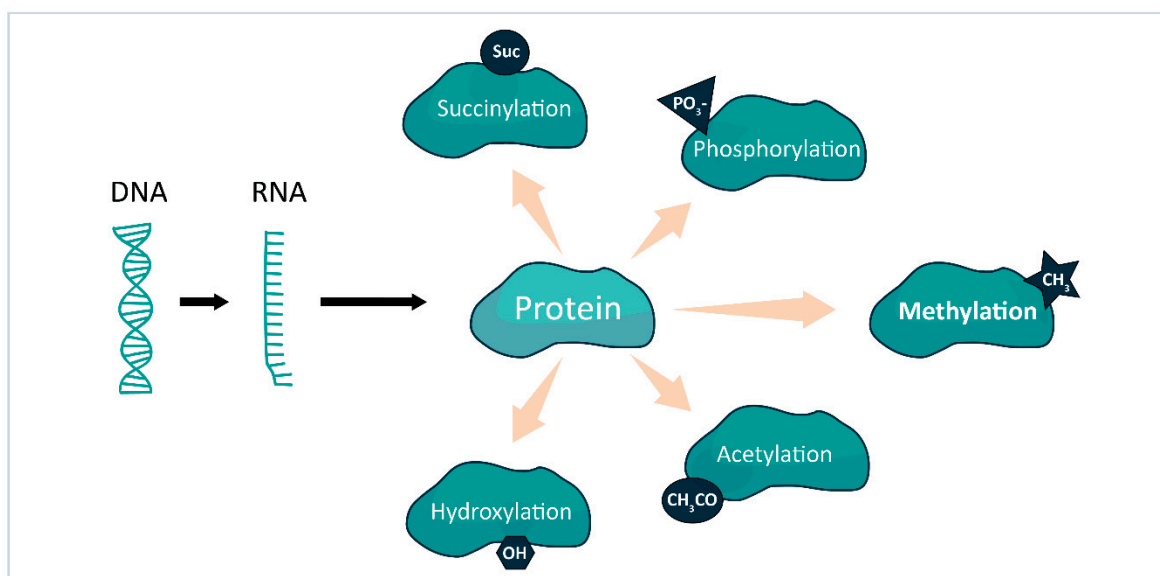

Figure 1: Post-translational modification (PTMs) change the biological function of the protein. The central dogma of biology explains that DNA is transcribed to RNA which is subsequently translated into proteins. However, these proteins can then be modified with PTMs to expand the genetic code and fine-tune the functioning of proteins. The PTM that this thesis will focus on is methylation.

The Extensive Scope of Methylation

Methylation is the chemical addition of a methyl group (-CH$_3$) to a substrate. A group of enzymes called methyltransferases (MTases) catalyze the addition of the methyl group and in some cases demethylase enzymes can remove the methyl group, making this chemical alteration a reversible, dynamic process. The variety of the substrates is immense; they range from small molecules, such as the dipeptide carnosine [6], to DNA [7], RNA [8] and proteins [9].

The epigenetic modification of DNA is one of the most studied methylation reactions. It is, among others, associated with gene repression, genomic imprinting, X-chromosome inactivation, aging and carcinogenesis [10]. The second most widely investigated methylation reactions are histone protein methylations. These can either be associated with gene activation or repression. One example is the mono-, di- or tri-methylation of lysine 4 in histone H3, which are active marks for transcription, as they are detected surrounding transcription start sites. On the other hand, di-methylation of histone H3 at lysine 9 is a mark for transcriptional repression, as it has been associated with heterochromatin formation and gene silencing [11]. Moreover, besides histones, methylation is observed on non-histone proteins, and can influence the substrate's activity, stability, distribution and interaction with other proteins [12]. Disruption of this delicate regulation of proteins has been associated with diseases such as a variety of cancers [13]. The investigations in this field revealed a diverse range of functional output of this PTM, which is why we wanted to explore this exciting section of biology.

SAM-dependent Methyltransferases

While the entire collection of methylation substrates reveals an enormous diversity, the MTases responsible for their modifications are found in a distinct number of structural arrangements, which makes it possible to classify them into superfamilies, among which are the seven β-Strand (7BS), SET domain and SPOUT MTases. The common factor that unites all these superfamilies is their usage of *S*-adenosyl methionine (SAM; in literature also known as AdoMet) as a methyl donor.

The molecule SAM is composed of a methionine attached to an adenosine. During the catalysis of methylation, the MTase aligns the electrophilic methyl group of SAM with the nucleophilic target of the substrate, which is usually an S, N, O or C atom. The MTase then deprotonates the substrate, thereby increasing its nucleophilic character and allowing it to make a nucleophilic attack, during which the methyl group is transferred to the target in a bimolecular nucleophilic substitution (S$_N$2) reaction, leaving behind *S*-adenosyl homocysteine (SAH, in literature also known as AdoHcy) and a methylated substrate (Fig.2) [14].
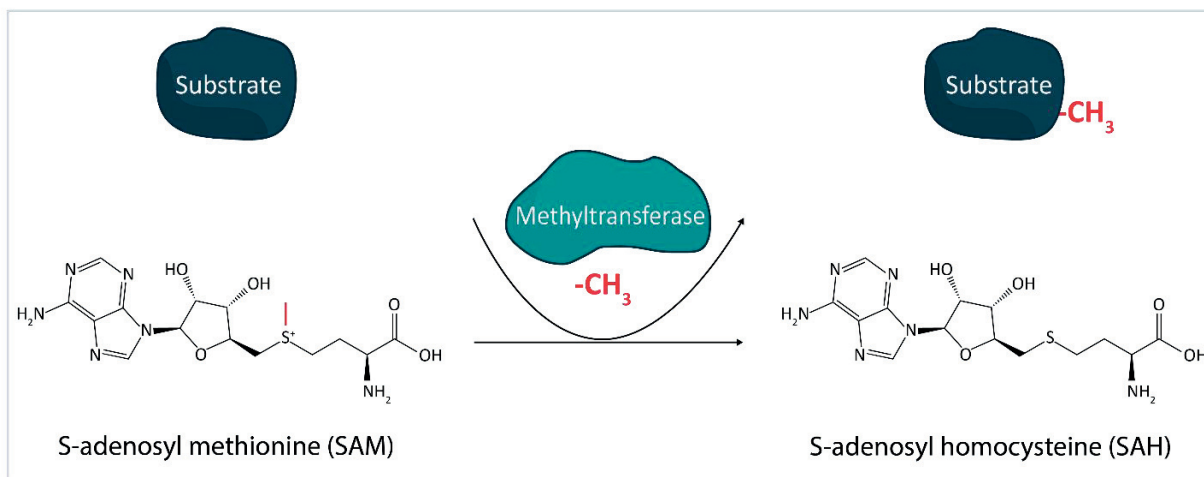
Figure 2: S-adenosyl methionine (SAM) dependent methylation. The methyltransferase takes the methyl group from the donor molecule SAM and incorporates it into a substrate, resulting in S-adenosyl homocysteine (SAH) and the methylated substrate.

Collectively, the human methyltransferasome is predicted to consist of approximately 208 MTases [15]. This makes up almost 1% of the human protein open reading frames, indicating the importance of these enzymes and the methylation they introduce. The great number of MTases in humans is remarkable, however, considering the needed refinement of molecular processes to accomplish such a complex organism, this is not surprising. What seems to be more remarkable is that a large number of the MTases are conserved throughout evolution. For example, the single-celled eukaryote budding yeast (*Saccharomyces cerevisiae*) has 56 7BS MTases of which 40 have been predicted to have human orthologues through bioinformatics analyses [16]. This evolutionary conservation of MTases is a primary indication of the importance of this enzyme group. A composition of the superfamilies in the human and yeast methyltransferasome can be viewed in figure 3, revealing the similarity of these very distant related species.
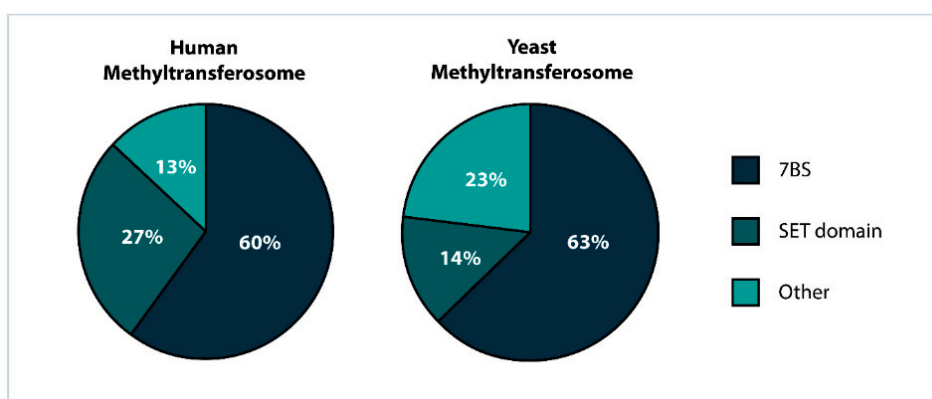

Figure 3: Human and yeast methyltransferasome. Distribution of methyltransferase (MTase) superfamilies within the human and yeast methyltransferasomes, with the seven β-strand (7BS) being the largest superfamily and the SET domain containing MTases the second largest. Figure adapted by Petrossian and Clarke 2011 [16].

The largest class of SAM-depended MTases are the 7BS MTases, contributing to about 60% of the human methyltransferasome (Fig. 3). As the name suggests, the enzymes belonging to this class share a similar structure consisting of seven twisted beta-strands. The first six beta-strands are organized in a parallel way, whereas the seventh beta-strand, closest to the C-terminus, runs in the opposite direction. Additional, a hallmark of the 7BS MTases family are four conserved motifs, which are found at specific positions in the 7BS fold and are called "Motif I", "Motif Post I", "Motif II" and "Motif III" (Fig. 4) Residues in Motif I and Post I together bind and coordinate the methyl-donor SAM [17]. Even though sequence homology of an uncharacterized protein with the four motifs is an indicator for 7BS family affiliation, it should be noted that not all 7BS MTases have all four of these motifs. In addition, different subclasses of 7BS MTases have a conserved Post-II motif that defines their substrate specific binding. One example is the MTase family 16, which share a conserved (D/E)xx(Y/F) motif [17], [18]. This specific subfamily has been investigated extensively by our group in the past, and the histidine MTase like 18 (METTL18), which will be introduced later, is also part of it.

Most of the characterized protein MTases belonging to the 7BS class have been established to methylate lysine or arginine residues. However, recently three 7BS MTases have been identified that methylate histidine. The first one is CARNMT1, which has been attributed to methylate a histidine containing dipeptide [19] and was furthermore investigated in this thesis (Manuscript II) to show its ability to methylate proteins as well. The second one is MTase like 9 (METTL9), a promiscuous MTase, the discovery and characterization of which is summarized in Paper I. The last histidine MTase is METTL18, an enzyme responsible for RPL3 methylation, which also was investigated by our group. Due to its special place in the histidine MTase field it will be introduced in more detail in its own subsection.
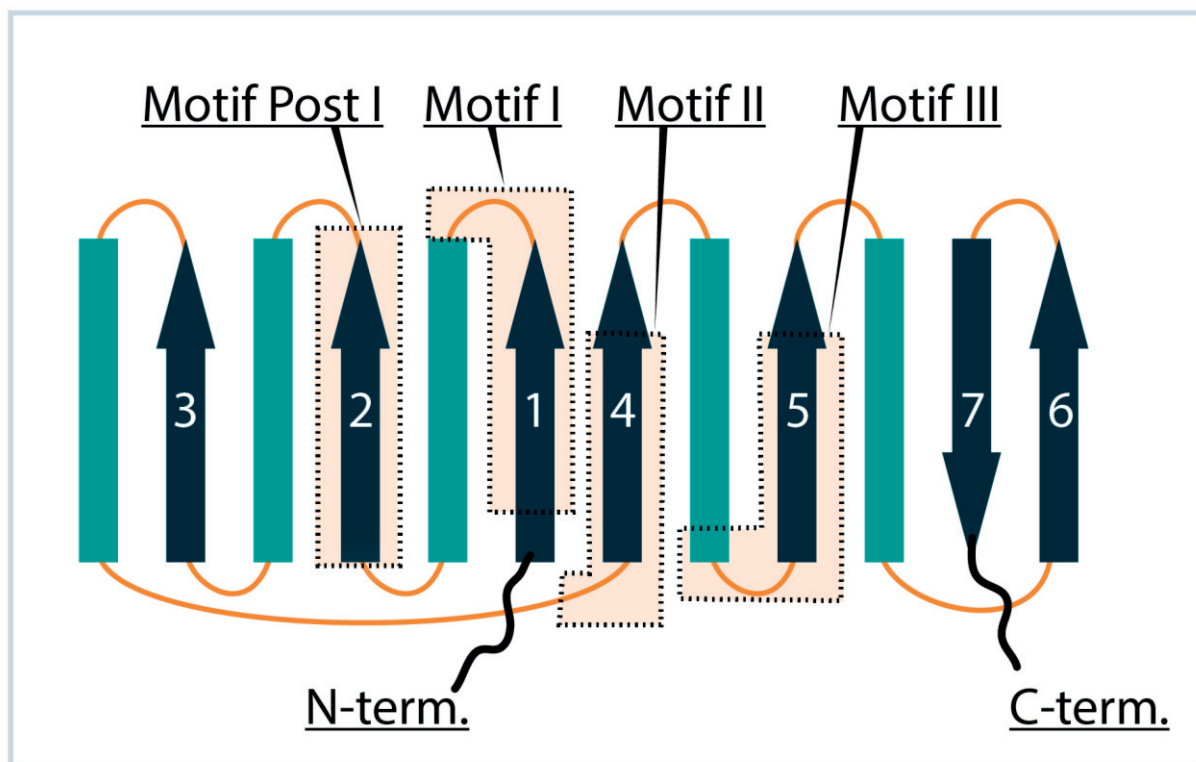
Figure 4: Schematic topology of a typical seven-β-strand methyltransferase structure. The dark blue arrows indicate the seven beta-strands, while the turquoise squares represent the α-helixes. The conserved motifs are indicated (Figure adapted from [20]).

The SET-domain Methyltransferases family

Besides the 7BS MTases, there is a second class of MTases that occupies nearly a third of the human methyltransferasome. In the early 1990s, an approximately 130 amino acids long domain was discovered, which was conserved in several *Drosophila* proteins: *Suppressor of variegation*, *Enhancer of zeste*, and *Trithorax*, and therefore named the SET-domain [21]–[23]. This domain has been shown to contain MTase activity and forms another superfamily of the MTases that is identified in all eukaryotic organisms studied. However, the SET-domain proteins seem to be more abundant in higher eukaryotes, as these MTases make up 27% of the human, but only 14% of the yeast methyltransferasome [16]. The decreased number in the yeast methyltransferasome can be explained by the function of certain SET MTases in humans, which is the methylation of lysine residues in histones. All histone lysine MTases but DOT1L, which is a 7BS MTases, are members of the SET domain superfamily [24] (reviewed in [25]). SET-domain MTases however, do not only methylate lysine residues in histones, but also non-histone proteins. For example, SETD7, in addition to methylation of

histone H3 lysine 4, also methylates the tumor protein 53 (p53), androgen and estrogen receptors, as well as several other proteins (For detailed reviews see [26]).

Until recently, all characterized SET-domain MTases were established to have exclusively lysine-specific activity [9]. However, Wilkinson et al. [27] identified the first protein histidine MTase that belongs to the SET family: SETD3. Due to its unique relevance as a histidine MTase in this thesis, it will be described in its own subsection. Firstly however, we will introduce the remarkable amino acid histidine and give a brief history of the hitherto discovered histidine methylation events and characterized histidine MTases.

The Histidine Methylation & Histidine Methyltransferases Universe

A Remarkable Amino Acid

The human body facilitates a proteome of more than 100,000 unique proteins. And while in nature there are hundreds of difference amino acids found, we only need 20 common amino acids to build up this great variety of proteins. The versatile of these amino acids is histidine. While our bodies are generally able to synthesize histidine by itself, during specific physiological events of growth, for example pregnancy and adolescence growth, the synthesized amount is insufficient, thereby making it a (conditionally) essential amino acid [28].

Histidine, besides the classic backbone of the natural amino acids (the carboxyl- and the amino group), is equipped with an aromatic imidazole ring as a side chain. This ring harbors two nitrogen atoms, one is the proximal (π) position regarding the backbone, and on in the distal (τ) position. Due to this specific structure, histidine is able to exist in two tautomeric forms. Tautomers are structural isomers of chemical compounds that readily interconvert. This reaction commonly results in the relocation of a proton, in this case from one nitrogen to the other. (Fig. 5a). Both of these nitrogen atom may be protonated, yielding neutral or positively charged forms of histidine at physiological pH (Fig. 5b). And either of these nitrogen atoms can be modified, yielding in either π-methylhistidine or τ-methylhistidine (Fig. 5c). This ability of histidine to act as a general acid or base plays major roles in protein interactions and it is therefore often a catalytic residue in enzymatic reactions [29], making histidine arguably the most versatile of the proteinogenic amino acids. Apart from that, histidine can also act as local pH sensor, including coordination of metal cations such as zinc [29], [30].

Out of all the different amino acids, histidine is one of the least frequently modified by PTMs [31]. This is quite surprising, as it is an amino acid with such grand versatility. Therefore, it can be speculated

that the methylation of histidine has the potential to regulate a wide range of o cellular processes and molecular interactions.
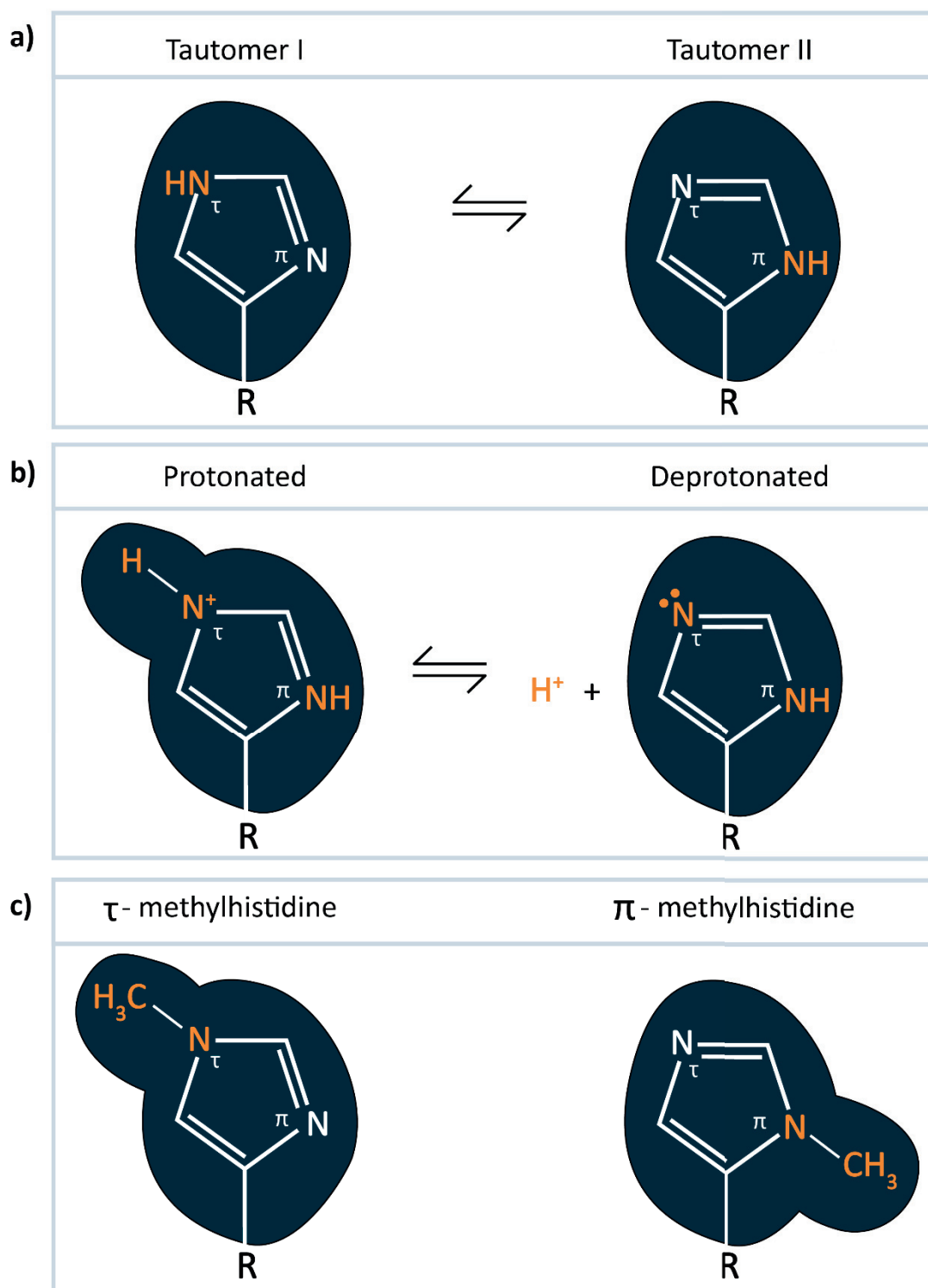

Figure 5: Histidine structure: **a)** Tautomers of histidine protonated on either the distal (τ) or the proximal (π) nitrogen of the imidazole ring. **b)** Protonated and deprotonated histidine (from tautomer II) **c)** τ-methylhistidine and π-methylhistidine with the methyl group on the corresponding nitrogen atoms of the imidazole ring.

Histidine, through its chemical properties associated with the imidazole ring, is an extraordinary amino acid. Even though its ability to become methylated has been known for almost a century, the investigations on the proper mechanism and function of the methylation status have only gained interest in the recent couple of years. This subsection is therefore dedicated to an overview of the research history and the current state of knowledge on histidine methylation. A schematic summary is given in Figure 5.

The first report of histidine methylation dates back almost a century ago. In 1929, the π-methylated form of the dipeptide carnosine, anserine, was observed in goose skeletal muscle [32]. Subsequently, these dipeptides have also been found in high concentration in various mammalian musculatures (summarized in [33]), which implicates the importance of carnosine and anserine in muscle physiology. Much later, the τ-methylated form of carnosine, named balenine (ophidine), was observed in whale meat extract [34].

The next appearance of methylhistidine in scientific literature can be dated back to 1951, where it was found as a free amino acid in human urine samples. Firstly, histidine was found with a methylation on the π-position [35] and later on the τ-position [36], even though the source of the amino acid in the urine samples remained unclear.

Later in the late 1960s and early 1970s, independent studies confirmed the first histidine methylation events in protein context. A histidine in actin was shown to host a methyl group on its τ-nitrogen atom in skeletal muscle of multiple organisms, such as fish [37], bird [37] and mammals [37], [38]. The latter study investigated the origin of the methylation and concluded that the formation of actin τ-methylhistidine has to occur as a PTM, rather than direct incorporation of methylhistidine during synthesis, leading to the hypothesis of the existence of a yet to be identified enzyme that could catalyze such a reaction, later known as a protein histidine MTase. Besides actin, the muscle protein myosin was also found to be methylated on a histidine residue. Myosin τ-methylation level was however found to be varying in different muscle types, unlike actin which was similar in all muscle types. Here again, different species were analyzed, among which rabbit, mouse, pigeon, cow, lobster and crabs [39], [40].

Additional to myosin itself, in 1987, the skeletal muscle myosin-light-chain kinase (MYLK2) has been discovered to also harbor a methyl histidine. Specifically, a π-methylhistidine in MYLK was isolated from rabbit skeletal muscles [41]. The last histidine methylation modification that was found in the 20th century is that on the immunomodulatory protein S100A9. Similarly, it also hosts the methylation on the π-nitrogen, on a histidine that is the last of an alternating histidine (HxH) sequence. In this case the tissue of origin was murine spleen cells. The only low throughput histidine methylation that has

been found in the 21st century, independently of its MTase, was NDUFB3 [42]. This protein is the B12 subunit of mitochondrial respiratory Complex I, and has been shown to contain multiple methylated histidines close to its N-terminus. Notably, the methylation also occurred as part of an HxH motif, similarly to S100A9. Even though the function of protein methylation had not been fully understood, the increasing amount of proteins and the conserved range of species for some proteins harboring this PTM indicates its importance.

There is also some, albeit scarce, evidence that histone proteins may be subject to histidine methylation. Only in the late 20th century, some studies emerged about this issue. The first evidence was found in experiments with duck erythrocytes, a rather specific tissue, which have shown histone histidine methylation. The histone methylation detected was found in histone fraction I (lysine-rich) and V (serine-rich) where the histones were methylated on the τ-nitrogen [43]. Another study characterized the methylation state of HeLa cells and claim that the f3 fraction (arginine-rich) of the histone content the most significant amount of methylated histidine [44]. As histone PTMs have been studied extensively and so far these histidine methylation events could not be confirmed, these original findings of histone histidine methylation are by some experts [45] regarded as potential artifacts, due to e.g. contamination. The original duck erythrocyte experiment [43] however, used multiple steps to control and compensate for potential contamination, even though it cannot be excluded that non-histone methylhistidine-containing proteins were present in the fractions. Additionally, if only a very specific and small fraction of histones bear this histidine methylation, it is possible that this PTM could escape the sophisticated global mass spectrometry (MS) analyses that have so far been conducted. Also, it is difficult to assess histone methylation by MS, due to many different modifications present, usually at low occupancy [46]. Overall, further investigation is needed to decisively conclude whether the histone proteins can be methylated on histidine residues.

Besides the above mentioned experimentally verified histidine methylated proteins, and the introduced histidine MTases and their substrates, there are multiple in vivo and in vitro verified proteins methylated by METTL9. An overview can be found in Paper I Tab. 1.

The Histidine Methylation Universe

Proteomics · Substrate · Methyltransferases

2020

Many substrates: HxH motif (e.g. DNAJB12, ARMC6, ZIP7)

HeLa (299)
HEK293 (80)
A547 (65)
HCT116 (89)
MCF7 (65)
Sy5Y (98)
Prostate tissue (46)
Liver tissue (30)
Colon tissue (47)

HeLa (181)

HEK293 (103)

METTL18 (human)

METTL9

SETD3

CARNMT1

2010

Rpl3

Hpm1p (yeast)

NDUFB3

2000

S100A9 (MRP14)

1990

Myosin Kinase (MYLK2)

1980

1970

Myosin

Actin

Balenine (τ-methylated carnosine)

1960

1950

free meH (in urine)

1940

1930

Anserine (π- methylated carnosine)

Figure 5: The histidine methylation universe. A Schematic overview of the proteomic studies identifying histidine methylation (hisMeH) events; experimentally verified histidine methylations (or substrates); and the discovered histidine methyltransferases (MTases). The stars indicate whether the substrate is methylated on the π- (orange) or τ-nitrogen (turquoise) on the histidine. The responsible MTases are indicated in their respective color and connected to their substrates. MTases of the white substrates are until now, unidentified.

During the previous decade only a hand-full of histidine methylations have been experimentally discovered. Due to rather primitive techniques and high amounts of tissue required it is no surprise that this process had its limitations. Fortunately, with development of more sophisticated experimental tools and the rise of new technology, such as more sensitive MS, there are now new ways of identifying PTMs in a high throughput fashion.

In 2016, Ning *et al.* [15] were the first to identify histidine methylation events in a whole proteome context. In their new approach to probing protein methylation events, they used HEK293 cells and were able to identify 103 methylated histidine residues. Subsequently, a study in 2018 by Wilkinson *et al.* [27], revealed 181 histidine methylation events in HeLa cells. Ultimately, due to the increased interest in the field, histidine methylation got full attention from a proteomics team in 2021 [47]. In this systematic approach, existing data from multiple tissues and cell lines have been analyzed to reveal all methylhistidine-containing sites in the respective cell types and tissues (for details see Fig. 5). Besides identification of novel methylated histidine sites, the study also analyzed the positions and co-occurrence with other important PTMs such as phosphorylation. Together, they concluded that histidine methylation can often be observed in mono-spaced histidine repeat regions (HxH motifs), and in regions of regulatory and functional importance [47]. Concluding, this brief review reveals a drastically growing number of identified histidine methylated residues, thereby strengthening the need to investigate them in the future.

An Overview Histidine Methyltransferases

As the above review has shown, histidine methylation was first explored almost a century ago. The responsible enzymes for these events, however, have only started to be uncovered since 2010. In the next subsection, each of the hitherto identified histidine MTases will be introduced. While SETD3 and METTL18 (and its yeast orthologue Hpm1) will be discussed in detail, METTL9 and CARNMT1 will only

shorty be summarized, as they are the main topics of Paper I and Manuscript I as well as Manuscript II respectively.

CARNMT1 – A dipeptide methyltransferase with a taste for proteins?

Carnosine methyltransferases 1 (CARNMT1) was the first discovered human histidine MTase [19]. Protein methylation was however not attributed to the enzyme, as it methylates the dipeptide carnosine. Carnosine is generated by carnosine synthase (CARNS1) from β-alanine and L-histidine. The latter can be methylated on either the π- or τ-nitrogen, forming the dipeptides called anserine or balenine, respectively. CARNMT1 generates anserine, while the MTase responsible for τ-methylhistidine remains to be discovered. While CARNMT1 is highly conserved throughout evolution, CARNS1 is not, which indicates other functions of the MTase. During the initial characterization of CARNMT1 it was shown that the enzyme methylates not only carnosine but also L-histidine itself and L-histidine-containing di- and tri-peptides in vitro [19]. Moreover, in the structural study of Cao et al., it was confirmed that the MTase dimer hosts enough space to methylate a longer peptide chain or protein, together indicating that CARNMT1 may have proteins as substrates as well [48].

Part of this thesis was to find out if CARNMT1 can methylate proteins (Manuscript II). We first used a model peptide with multiple histidines as a proof of concept and showed the generation of π-methylhistidine. Additionally, we have found seven in vitro substrates in HAP1 cell extracts. While future investigation is required to identify the substrates and reveal the functional significance of the methylation, these results deliver the exciting proof that indeed, CARNMT1 is able to methylate proteins in vitro, adding it to the protein histidine MTase universe.

SETD3 - The long sought actin histidine methyltransferase

As many of the enzymes in the SET domain family are histone lysine MTases [22], it was no surprise that in the early 2010s, a thus far uncharacterized member of the family, SETD3, was reported to be involved in methylation of histone H3 at K4 and K36 [49], [50]. In 2018 however, two independent studies [27], [51], elegantly revealed that the substrate of SETD3 is not a lysine, but a histidine residue. Specifically, it was found that SETD3 is the long sought MTase of H73 in actin, generating τ-methylhistidine. The two research group used a different approach, both with extensive experimental evidence, to reach the same conclusion.

Wilkinson et al. used recombinant wild type or catalytically inactive SETD3 proteins, incubated them with radioactive SAM and mammalian HT1080 cell extract. These in vitro methylation assays were

followed by MS to establish actin as the specific SETD3 substrate [27]. Kwiatkowski et al., on the other hand started with an extensive purification of the hitherto unidentified rat actin MTase enzyme from rat muscle tissue and subsequently identified SETD3 as the candidate for the actin methylation. Results were also confirmed with recombinant SETD3 enzyme and in vitro actin methylation assays [51].

Actin's methylation has been functionally associated with the maintenance of cytoskeleton integrity, promotion of smooth muscle contraction and stabilization of actin filaments [27], [51]. Notably, SETD3 has also been associated with other functions, outside actin homeostasis (summarized by Drozak et al. in Table 1 [52]). SETD3 had originally, as mentioned above, been associated with histone methylation, and was predicted to play a role in gene regulation [49], [50]. The MTase has also been linked to enterovirus pathogenesis [53] and response to hypoxia conditions [54]. An extensive overview of the biological function, activity and structure of SETD3 are summarized in [52], [55].

Furthermore, a study by Cohn et al. identified the interactome of SETD3 [54]. Thereby, 172 interacting proteins have been identified, including actin, and the transcription factor FoxM1 has been observed to be methylated by SETD3. This observation has been experimentally confirmed with an in vitro methylation assay. It should be noted, that the specific residue was not investigated in this publication. One of the two groups that originally identified SETD3 as the actin-specific histidine MTase [51] recently published a conference abstract called "*Preliminary identification of novel protein substrates for actin-histidine N-methyltransferase (SETD3)*" [56]. Their preliminary data suggests quinone oxidoreductase and glyoxalase domain containing protein 4 as potential novel substrates for SETD3 in human cells. As this is preliminary data, we must wait for further investigations on their part for any physiological importance, however this already indicates that exciting new findings will come in the near future.

In conclusion, more than 50 years after the original characterization of H73 actin methylation [37], [38], SETD3 has finally been identified as the responsible enzyme [27], [51], making SETD3 the first human histidine MTase. The pioneer work of the two independent research groups shines light on the histidine methylation universe, which already got multiple groups involved in the functional studies about human histidine MTase, specifically SETD3 [27], [51].

METTL9 – The HxH all star

Our group was the first one to characterize the human 7BS MTase METTL9 (Paper I). We have found that METTL9 is a broad-specific MTase that generates pervasive π-methylhistidine in mammalian proteomes. Specifically, the methylation requires an HxH motif, in which "x" is preferably alanine,

asparagine, glycine, serine or threonine (ANGST) or in other words, a small amino acid. Among its verified in vivo substrates are the immunomodulatory protein S100A9 and the NDUFB3 subunit of mitochondrial respiratory Complex I. The methylation of the latter enhances respiration via Complex I, and the presence of π-methylhistidine in an HxH-containing peptide reduced its zinc binding affinity. METTL9 is conserved throughout most eukaryotes, notably absent from land plant and fungi. We have therefore also investigated METTL9 orthologues among which are the fruit fly *Drosophila* (dm) and the picoplankton *Ostreococcus tauri* (ot)*. We were able to show that they have different substrate preferences compared to human METTL9. Lastly, we have solved the structure from otMETTL9 revealing a conserved active site buried in a binding pocket.

METTL18 − The human orthologue to the yeast RPL3 methyltransferase Hpm1

In 2010, the previously uncharacterized yeast YIL110W was found to methylate the 60S Ribosomal Protein L3 (Rpl3) on τ-histidine and named histidine protein MTase 1 (Hpm1). The position of the methylation is proposed to be the His-243 and its depletion results in an accumulation of 35S and 23S pre-rRNA precursors, highlighting an important role for histidine methylation pre-rRNA processing [57]. Hpm1 is the first discovered protein histidine methyltransferase in eukaryotes. The functional human orthologue of Hpm1 was suggested to be the C1orf156 gene, also known as METTL18 [57]. METTL18 is part of the 7BS MTases, and, more specifically, part of the MTase Family 16, which otherwise included only lysine-specific MTase [17], [58].

To investigate the characteristics of this MTase in humans, an initial study conducted METTL18 pull-down experiments and revealed a strong interaction with human RPL3, analogous to yeast Hpm1 [59]. To confirm and further describe METTL18, my colleagues set out to perform biochemical experiments to fully characterize this enzyme. In this study [58], it was confirmed that METTL18 methylates human RPL3 on the His-245 on its τ-nitrogen. Additionally, it was found that, METTL18 localizes to the nucleus of HeLa cells, where it specifically accumulates in the nucleoli, the location of ribosomal assembly. Finally, and importantly, the functional significance of RPL3 methylation was investigated. Cells lacking METTTL18 showed codon-specific changes in mRNA translation; altered pre-RNA processing, and a decrease in polysome formation. Altogether these findings indicated its importance for ribosome biogenesis and function [58]. A biochemical study published the following year confirmed these results and further stressed the importance of RPL3 methylation in modulation of translation elongation [60].

In an investigative proteome-wide study, Han *et al.* identified potential biomarkers associated with the fibromyalgia syndrome (FM) in women [61]. The definition of its symptoms is explained as follows: "*Fibromyalgia is characterized by chronic, widespread musculoskeletal pain and associated fatigue,*

sleep disturbances, and other cognitive and somatic symptoms" (M.J. Bair and E.E. Krebs, 2020) [62]. As METTL18 was found to be part of the differential serum proteomics pattern of patients with FM, it could potentially serve as a biomarker for diagnosis of FM [61].

Besides it implication in FM, METTL18 has also been associated with hepatocellular carcinoma (HCC). In this study, patients with tumors that showed high expression displayed decreased survival rates [63]. *In vitro*, the knockdown of METTL18 further confirmed these results and revealed suppressed HCC cell proliferation, migration and invasion. Taken together, these results indicate that METTL18 can serve as a biomarker for HCC patients prognosis [63].

The last paper that needs to be mentioned identifies METTL18 as an actin MTase. The suggested evidence is however scarce, as they used immunoprecipitated GFP-tagged METTL18 as enzyme source without stating its purity. Additionally, the methylation assay showed only ~2-4x higher activity as the background. This leaves to speculate if they have co-purified SETD3 and a bona fide methylation activity. Lastly, the true MTase of actin has already been proven with convincing evidence, as the knock-out of SETD3 removes actin methylation [27], [51], so this particular publication never made it out of bioRxiv [64]. This article is yet another one in a string of publications that suggest a new MTase to do the function of a known MTase without very good evidence for it [65], [66] .

After having introduced not only the history and state of the art knowledge of histidine methylation and its MTases, but also the importance and versatility of this PTM, it can be concluded that histidine methylation opens a new and exciting horizon of research. While in the last hundred years there were only a dozen histidine methylated proteins discovered, modern high-throughput methods reveal that hundreds more are yet undescribed. There is therefore a great knowledge gap that needs to be filled: The methylated proteins are in need for further characterizations and the MTase responsible for their methylation need to be discovered to elucidate the functional and biological significance of histidine methylation.

Aims of Study

The aim of the research presented in this thesis was to characterize two human methyltransferases (MTases) - the novel MTase METTL9, which, as we discovered, methylates histidine residues (Paper I), and the previously identified dipeptide histidine MTase CARNMT1 (Manuscript II). In short, our purpose was to identify the protein substrates of the MTases and elucidate the biological significance of their methylation. In addition, for METTL9, we aimed to investigate the activity of putative orthologue enzymes to evaluate evolutionary substrate preferences, as well as to solve the structure of METTL9 to better understand how enzyme structure connects with its function (Manuscript I).

Summary of Papers

Paper I - The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes

In this paper we identified and characterized the first human protein histidine π-methyltransferases METTL9, which is responsible for about half of all π-methylhistidine (or 1-methylhistidine) content in mammalian protein. We have identified multiple in vitro, as well as six in vivo, substrates, among which are DNAJB12 and ARMC6. Notably, we were able to show that the substrates share a common alternating histidine motif (HxH) with "x" preferably being a small uncharged amino acid such as A, N, G, S or T. Finally, we demonstrate that METTL9 activity is important for mitochondrial function, and histidine methylation of a zinc transporter peptide reduces its affinity to zinc.

Manuscript II - METTL9-like histidine methyltransferases from different organisms display distinct substrate specificities

Here, we explored putative METTL9 orthologues from seven eukaryotic organisms and found that most of them could methylate the human METTL9 substrates ARMC6 and DNAJB12. However, we could also observe distinct, more stringent substrate preferences for the orthologues from *Ostreococcus tauri* (*O. tauri*) and *Drosophila melanogaster* (*D. melanogaster*), both on the protein and on the peptide level. Finally, we solved the structure of *O. tauri* METTL9 and found key residues that coordinate the HxH histidines in a model peptide.

Manuscript I - The human carnosine Nπ-histidine methyltransferase CARNMT1 methylates proteins

In this study, we found indications that CARNMT1 is not only able to generate π-methylhistidine on its previously characterized dipeptide substrate carnosine, but also on a model protein substrate, derived from the zinc transporter ZIP7. Additionally, we were able to observe at least seven methylated proteins in protein extracts from a human cell line. Unfortunately, we were unable to identify these substrates. However, we identified 24 proteins interacting with CARNMT1 and analyzed which of them are good candidates for further study as potential methylation substrates.

Discussion

Part I – CARNMT1: A dipeptide Nπ-histidine methyltransferase with a taste for proteins

CARNMT1 is the first the π-methylhistidine generating MTase discovered [19]. However, due to the nature of its dipeptide substrate, it has so far been excluded from the histidine protein methyltransferase universe. Interestingly, we found that CARNMT1 is able to methylate a histidine-rich fragment from ZIP7 in vitro (Manuscript II Fig. 2), which is an in vivo substrate of METTL9 (Paper I Fig. 4d). We have furthermore confirmed that CARNMT1 generates π-methylhistidine on ZIP7 (Manuscript II Fig. 3). Employing in vitro methylation assay of CARNMT1 on HAP1 cell extracts, we were able to see that the enzyme methylates at least seven proteins (Manuscript II Fig. 4b), which we were not able to identify. Lastly, we revealed 24 interacting proteins (Manuscript II Fig. 5) some of which may represent in vivo substrates of CARNMT1. In particular, four proteins in total stood out as the most likely candidates: BAG2, OSBPL1A, SARM1 and MAP7, with the latter three with highest abundance as interactors, and BAG2 displays a high stoichiometry of binding its molecular mass corresponds to that of one of the observed bands in the in vitro methylation assay with HAP1 extracts (Manuscript II Fig. 5b, 4b).

CARNMT1 – function and future perspectives

Due to time constrains, we were not able to identify the substrates that were methylated by CARNMT1 in HAP1 protein extracts. However, efforts were made to find a connection between the 24 interacting proteins from HEK293 cells to evaluate them as potential substrates. Unfortunately, after GO term analysis, there seemed to be no general connection between these proteins. Looking more closely into the top four interactors, which are the BAG family molecular chaperone regulator 2 (BAG2), the NAD(+) hydrolase SARM1, the Oxysterol-binding protein-related protein 1 (OSBPL1A) and the microtubule-stabilizing protein 7 (MAP7) (Manuscript II Fig. 5), we checked for enrichment of these four interactors in the Human Protein Atlas [67] and found some interesting correlations (Fig 7). To make a comparison, we evaluated the top ten cell types in which CARNMT1 RNA is enriched and have found them to be early spermatids, thyroid glandular cells, ductal cells, cardiomyocytes, mitotic cells (heart), breast myoepithelial cells, breast glandular cells, urothelial cells, ascending loop of Henle cells and

keratinocytes (Fig. 7a). Next, we looked for enrichment correlations with the interacting proteins. According to this analysis, our two most promising potential substrates would be BAG2 and MAP7, as they show enrichment in five of the top ten enriched cell types of CARNMT1 (Fig. 7b and c). BAG2 even has four of these in its top five enriched cell types. OSBPL1A and SARM1, on the other hand, have only two and three cell lines enriched respectively (Fig. 7d and e). Overall, even though we see a co-occurrence of BAG2 and MAP7 with CARNMT1 in many cell types, additional investigations are needed to confirm them as biological substrates of this MTase.

Interestingly, CARNS1, the enzyme catalyzing the synthesis of carnosine from β-alanine and L-histidine, is only enriched in two of CARNMT1's top ten cell types (Fig. 7f), namely cardiomyocytes and breast myoepithelial cells. This again indicates that CARNMT1 has more functions than the methylation of carnosine, which is likely not present in the cells that do not contain CARNS1. In addition to CARNS1, in cardiomyocytes, all top four of the CARNMT1-interacting proteins are also enriched, and it would be interesting to find out if CARNMT1 is able to methylate these proteins even in the presence of carnosine. As 99% of carnosine in an organism is present in skeletal muscle [68], and HAP1 cells that we investigated are chronic myeloid leukemia derived, it is possible that CARNMT1 here has the availability to methylate other products than carnosine. This may also explain why other groups, working with muscle tissue or cells, have not found these substrates earlier.

Methylation Motifs for the Methyltransferases?

Once the substrates for CARNMT1 are identified, it will be interesting to see whether they share a similar target motif, like the METTL9 substrates do. Thus far, there are only six METTL9 substrates observed in vivo (Paper I Tab. 1), however we know that the enzyme preferentially targets an HxH motif, with "x" preferably being a small amino acid. This consensus motif can be used to predict additional, yet unidentified substrates in humans, overall simplifying the search for novel substrates. Large-scale proteomics studies focusing on histidine methylation have not found any of the top CARNMT1 interactors to be methylated [15], [47], [69]. However, the analyzed cell lines and tissue biopsies were chosen based on the expression levels of SETD3 and METTL9, and CARNMT1 may not be expressed in these samples. Still, at a glance, there is no obvious consensus motif around any histidines in BAG2, OSBPL1A, SARM1 and MAP7, should they be CARNMT1 substrates and not just interaction partners. Future studies will hopefully reveal the substrates for CARNMT1 and elucidate whether they share a recognition motif.
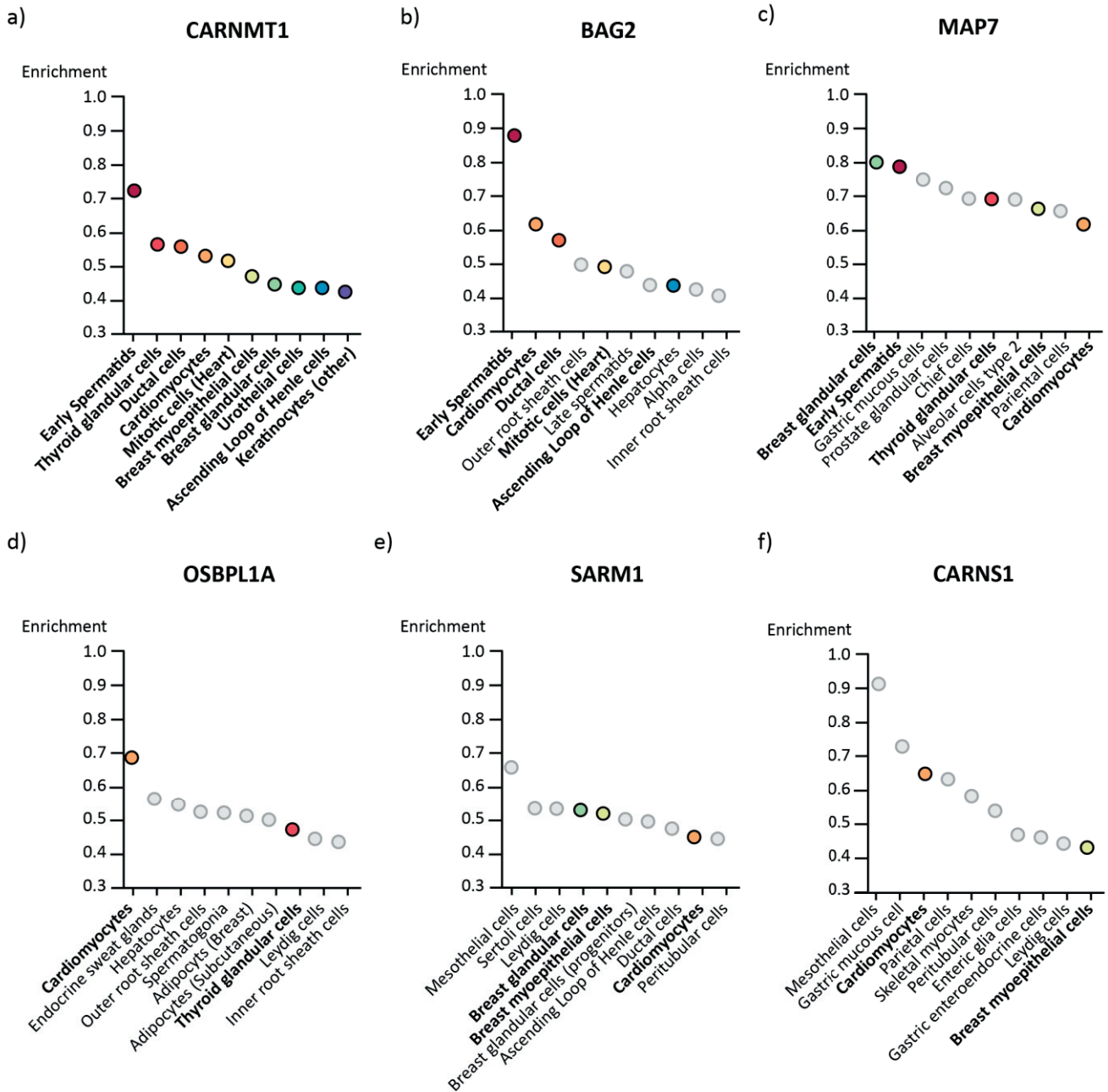
Figure 7. Enrichment of RNA from tissue cells. Ten cell types with greatest enrichment of **a)** CARNMT1 **b)** BAG2 **c)** MAP7 **d)** OSBPL1A **e)** SARM1 and **f)** CARNS1 based on data from the Human Protein Atlas; RNA in tissue cells [67]. Colors are assigned to the top ten cell types form CARNMT1 and used in the other panels.

Histidine methyltransferases have multiple substrates

In accordance with findings on CARNMT1 and METTL9, which generate π-N histidine methylation in multiple protein substrates, the τ-methylhistidine-generating MTase SETD3 has preliminary evidence for its promiscuity as well [56]. While the enzyme has originally been identified to only methylate actin, a poster abstract from a conference this year has indicated that as many as six proteins can be methylated by SETD3 in human cell extracts, as shown by fluorography. After The MS analysis of methylated bands from cell extracts, the authors identified quinone oxidoreductase and the glyoxalase domain containing protein 4 as potential novel substrates for SETD3 [56]. The last hitherto identified τ-methylhistidine-generating MTase METTL18 has thus far only been verified to methylate RPL3. However, during studies of the yeast orthologue of this MTase, it has also been suggested that other, yet undiscovered substrates exist [70]. In conclusion, the two characterized π-methylhistidine MTases, METTL9 and CARNMT1, both have many substrates, and even though we are just starting to discover the histidine methylation universe, we can speculate that the τ-methylhistidine MTases may also be more promiscuous.

Part II – The first human protein π-methylhistidine generating methyltransferase METTL9

METTL9 has many substrates, including the NDUFB3 subunit of mitochondrial respiratory Complex I and the immunomodulatory protein S100A9. Most notably however, these substrates share an HxH motif, where "x" is preferably a small uncharged amino acid. This broad specificity enables METTL9 to generate pervasive π-methylhistidine throughout the human proteome (Paper I). Notably, METTL9 is conserved in the vast majority of eukaryotes; however, we see differences in substrate preference throughout three orthologues that we investigated. We furthermore have solved the structure of METTL9 from the picoplankton *O. tauri*, thereby paving the way for understanding structure-function relationships of histidine MTases (Manuscript I).

METTL9 activity has been validated by independent studies

Following the publication of our pioneering METTL9 work, a study by Lv et al. [71] presented an independent METTL9 characterization. While most of their results are in line with ours, there are some additional and differing results. Validating our results, they found that METTL9 methylates ZIP7 in vitro and in vivo. After overexpression of HA-tagged full length ZIP7 in HEK293 cells and subsequent pulldown and MS analysis, they identified His256 to be methylated in vivo. They have also analyzed ZIP7 fragments in-vitro methylated by METTL9 and found methylation on residues His73, His75, His101 and His103. The inconsistency in detection of methylated sites shows the difficulties of MS analyses of histidine-rich regions. Furthermore, Lv et al. investigated which nitrogen in the histidine ring is methylated and found it to be the π-nitrogen, confirming our results. It has to be noted however, that for this radiolabelled methylation experiment, they used a peptide (residues 66–74 of ZIP7) that was either unmethylated, or methylated either on the τ- or π-nitrogen. Surprisingly, they see methylation not only of the unmethylated but also of the τ-methylated peptide, indicating that METTL9 may be able to also generate dimethylhistidine.

Additionally, Lv et al. were able to see a preference for METTL9 to methylate an xHxH motif, similar to our ANGST-HxH motif results. While they agree that the second "x" is preferably a small amino acid, they found it to be mostly alanine (Ala), cysteine (Cys), glycine (Gly), or serine (Ser) (ACGS). In our initial study of the ANGST motif, Cys and tryptophan (Trp) residues were excluded from the peptide array, as they are prone to non-enzyme-catalyzed methylation (Paper I Fig. 1f). However, in the peptide array with METTL9 orthologues, we have included these residues (Manuscript I Fig. 3c), and observed

methylation of Cys for all three enzymes from *D. melanogaster, O. tauri* and human. However, as Cys is prone to nonenzymatic methylation, we must repeat this experiment with inactive mutants of the enzymes, to show that the methylation is truly generated by METTL9. Additionally, in contrast to the Lv et al study, we found that Asparagine (Asn) and Threonine (Thr) are also preferred as the middle amino acid for human METTL9 (Paper I Fig. 1f), wheras for *D. melanogaster* and *O. tauri*, we observe methylation on "CAGST" amino acids (Manuscript I Fig. 3c). Taken together, this suggests that while METTL9 accepts a variety of amino acids in the "x" position of the HxH motif, it does not tolerate charged (D, E - negatively charged and R, K - positively charged) or aromatic (Y, F, W) residues. This is also in line with the *O. tauri* structure, as the amino acids lining the part of the cleft surrounding the middle residue in the HxH-peptide are relatively hydrophobic in nature, and the shape and size of the putative binding pocket would likely limit the size of the "x" amino acid (Manuscript I Fig. 3).

Lv et al also validated our results on the identity of the amino acid prior to the first histidine (x_N position) in the HxH motif. After conduction of a peptide array with substitutions in this position, they found that this amino acid is preferably not Isoleucine (Ile), Proline (Pro) or Valine (Val).These results are in line with our results as we also analyzed the x_N position, and found that while most amino acid substitutions of were tolerated, Ile, Pro or Val, but also Glutamic Acid (Glu) at the x_N position abolished methylation. Our results on the x_C position, which was not tested by Lv et al, showed that Pro and Val strongly reduced methylation (Paper I Sup. Fig. 5b, c). While peptide arrays from both our and their studies had a Glycine (Gly) flanking the HxH motif at the x_N position, there are differences in the experimental setup. First, Lv et al. used GST-tagged four amino acid long peptides, while we used tag-free peptides which were 15-mers derived from biologically relevant substrates. Additionally, we have collected motif data from three different peptide sequences, while they have only used one peptide. Lastly, Lv et al. used mouse Mettl9, while we worked with human METTL9 recombinant protein. This is very interesting, and could indicate a different substrate preferences for the mouse orthologues, even though it only differs from the human sequence by nine amino acids, eight of which are conservative substitutions, and are not located in conserved motifs (Manuscript I Fig. 2 and 3, [71]).

During our METTL9 investigation, we have generated a KO mouse model which, besides the drastic reduction of π-methylhistidine content in the proteome, had no apparent phenotype. However, we did not test its response to any stress conditions (Paper I). Lv et al., on the other hand, performed additional investigations in mice [71]. They injected WT or METTL9 KO RM-1 murine prostate cancer cells into immunocompetent and immunodeficient mice and observed a significant decrease in tumor growth for the KO cells in both models. During analysis of the tumor immune cells they discovered an increase in CD45[+] as well as CD4[+] and CD8[+] effector T cells, indicating that the influence of METTL9 is

not only on autonomous growth but also, potentially on anti-tumor immunity (Fig. 1 in [71]). Thus, subjecting the KO mice to different stresses, like cancer, inflammation and metal stress is likely to produce a phenotype.

Another difference between the results from our study and that of Lv et al., is the localization of the enzyme. We found that human METTL9 to co-localize with markers for membranous compartments such as the endoplasmic reticulum and mitochondria in HeLa cells (Paper I Sup. Fig. 2b), which is in line with methylation events seen in assays on protein extracts from different cellular compartments (Paper I Fig. 1a and Sup. Fig. 2a). Lv et al. suggest the localization of mouse 3xFLAG-tagged Mettl9 to be in the cytoplasm in RM1 cells. However, the resolution of the microscopy image is rather low and they do not use any compartment-specific markers, which makes it difficult to determine the exact localization of Mettl9 (Sup. Fig. 3b in [71]). This difference may also be due to the choice of cell line or tag. However, in principle, cytoplasmic localization would not prevent METTL9 to methylate zinc transporters, as their histidine-rich loops are mostly found on the cytoplasmic side of the membrane, including His256, which Lv et al found to be methylated in vivo. Also, the mitochondrially-localized METTL9 substrates may conceivably become methylated prior to mitochondrial import.

Another interesting experiment conducted by Lv et al. investigates the ability of METTL9 to regulate cellular zinc concentrations, as zinc transporters are among its most prominent substrates. Here they utilized FluoZin-3, a fluorescent zinc probe, to assess free zinc concentrations in WT and Mettl9 KO RM-1 tumor cell lines. They observed that the zinc level was significantly increased with aggregation in the cytoplasm in Mettl9 KO cells (Fig. 2A in [71]). Additionally, they overexpressed either empty vector, WT or H204/216A or H254/262A ZIP7 mutants in the RM-1 cell lines. Surprisingly, these mutations are not of the methylhistidine sites that they found in either the in vivo or in vitro studies. The empty vector showed no signal for FluoZin-3. For the H204/216A mutant they observed significantly upregulated zinc aggregation in cytoplasm compared to WT ZIP7 (Fig. 2E in [71]), while for the H254/262A mutant they detected slightly downregulated zinc level, as compared to WT ZIP7 (Fig. 2G in [71]). Overall, however, they did not use DAPI as a nuclear marker to control for exposure so this might be an artifact due to different exposure, as it seems like the cells have overall different levels of zinc, rather than just a change in distribution. Additionally, it is also very surprising to see such strong zinc reduction throughout the whole cell as ZIP7 is localized in the membrane of the endoplasmic reticulum (ER), and there are multiple other zinc transporters in the cells that influence zinc homeostasis.

Another group has also confirmed the π-methylhistidine activity of METTL9 on the metal-binding S100A9 protein using siRNA screening [72]. Similarly to our study, they have found that METTL9

localized to the ER in Hela and HEK293 cells. Additionally, they have validated that the enzyme generates π-methylhistidine on His-107 in the mouse proinflammatory protein S100A9, both in vitro and in vivo. Interestingly, His-107 is involved in zinc-binding, and its methylation reduced zinc binding to the corresponding peptide in vitro, similarly to what we saw in our study of a zinc transporter peptide (Paper I Fig. 5). As the biological relevance of the methylation they have excluded any effect on heterodimer formation between S100A9 and S100A8, which forms the calprotectin protein. However, they suggest involvement of METTL9 in the innate immune response to fungal and bacterial infections, as calprotectin exerts an antimicrobial activity [72].

METTL9 has a wide range of substrates

Of the known human protein histidine MTases, METTL9 stands out with its broad substrate specificity. Its preferred target ANGST-HxH sequence is found almost three thousand times in the human proteome, indicating the grand scale on which METTL9 can potentially influence the function of other proteins. A proteome-wide study from 2021 confirmed the widespread presence of histidine methylation in human cell lines, where they reported 299 histidine methylation sites in HeLa cells alone [47]. Interestingly, they specifically found such histidine methylations in the context of an HxH motif, thereby giving more evidence for the abundance of METTL9 methylation events, though these remain to be validated. It has to be noted, that this study did not differentiate between τ- and π-methylhistidine, as these two histidine versions are isobaric and cannot be distinguished in the proteomic data they analyzed. This, as well as analysis of histidine-rich sequences in general, is probably one of the biggest challenges for MS-based discovery of histidine methylation. Additionally, at the present time, there are no commercially available "pan-antibodies" for π- or τ-methylhistidine. Still, the first insights on the scale of METTL9-induced methylation is exciting as it shows similarity to well-characterized arginine MTases, such as PRMT1, which has more than 50 identified non-histone substrates [73]. PRMT1 methylates glycine-arginine rich (GAR) motifs, which are similar to our ANGST HxH motif in that they also contain the target residue intermixed with a small uncharged amino acid, glycine. We will in the future hopefully be able to uncover the depth of the histidine methylation activity of METTL9.

The biological function of METTL9

Even though we are merely beginning to understand the biology of METTL9, it is possible to speculate about its broader biological function. Given the large amount of substrates, it is likely that their

collective methylation "pushes" the cell in the same general "direction". In addition, even though METTL9 is highly conserved throughout most eukaryotes, it is absent from fungi and land plants [74]. Interestingly, the latter two have, in addition to the plasma membrane, a cell wall that consists mainly of polysaccharides, while organisms that express METTL9 do not. This may indicate that methylation of METTL9 substrates, such as zinc transporters localized to the plasma membrane, for example, is not necessary for organisms containing a cell wall, where the transport of zinc may be regulated differently.

In human and mouse, METTL9 is involved in the methylation of metal binding proteins, such as zinc transporters, and the immunomodulatory protein S100A9 (Paper I Fig. 5). Additionally, there was no direct phenotype seen in KO mice, indicating that it is important under specific stress conditions rather than normal living conditions. The latter is supported by in vitro methylation assays, which demonstrated considerable methylation by recombinant METTL9 on WT mouse tissue extracts, indicating hypomethylation of METTL9 substrates in normal mouse tissues. This observation indicates that, rather than being static, METTL9-mediated methylation may be dynamic and of potential regulatory importance, and is not strictly required without certain stresses. The study by Lv et al [71] also found increased levels of CD45$^+$ as well as CD4$^+$ and CD8$^+$ effector T cells of Mettl9 KO tumor cells injected into mice. This evidence strengthens the speculation that METTL9 is important in the immune system of its organism, and that a phenotype can be observed under certain stress conditions.

Mettl9 and disease

Putting methylation into a medical perspective, there have been multiple publication on METTL9 associations with disease. Already prior to its biochemical characterization, a study revealed that METTL9 is part of a distal 16p12.2 microdeletion of three genes (the other two being IGSF6 and OTOA) that is responsible for hearing loss under its homozygous condition, however the causative mutation might be in the IGSF6 gene which is encoded entirely in an intron of METTL9 [75]. Patients may also suffer from other symptoms such as developmental delay and cognitive impairment with uncertain pathogenesis.

Another disease associated with METTL9 is HIV-infection [76]. Specifically, METTL9 was found to be upregulated in HIV-positive patients that maintain undetectable viral load for a long period (elite controllers) and, as well as those who have normal CD4+ count in spite of circulating viral load (viremic nonprogressors). METTL9 has thus an apparent inhibitory effect on HIV-1 viral production. While the role of METTL9 remains to be elucidated, moderate inhibitory effect on HIV-1 production suggest its importance for HIV resistance, or enhanced immunity against it. METTL9 seems to play important roles

in hearing loss and resistance of HIV infection, but has been more strongly implied in cancer, as elaborated in the next paragraph.

METTL9 and cancer

METTL9 [77], as well as METTL18 [63], have been associated with hepatocellular carcinoma (HCC). During evaluation of the human liver cancer cell lines HepG2 and LO$_2$, one study found increased expression of METTL9 in both, even though the HepG2 cell lines showed the bigger increase of more than 100 fold [77]. In another study, the authors investigated expression levels of METTL18 in six HCC cell lines, including HepG2 , however, they found two other cell lines, LM3 and Huh7, to express the highest levels of METTL18 [63]. In tumor samples from HCC patients, high expression of METTL18 was associated with decreased survival rate [63]. These results indicate the diverse role of histidine methylation in cancer, with the need for personalized medicine, including evaluation of individual tumor samples for proper diagnostic.

Elevated METTL9 has also been correlated with metastasis of human scirrhous gastric cancer (SGC) [78]. After analysis of METTL9 expression in samples from healthy donors and patients with or without peritoneal dissemination (metastasis) the authors found that high METTL9 expression was associated with cancer progression. Interestingly, after METTL9 knockdown in metastatic cells from SGC patients, reduced mitochondrial Complex I activity was observed, which decreased cell migration and invasion [78]. This nicely links with our findings that knockout of METTL9 in human HAP1 and HEK293T cell lines also reduced the activity of mitochondrial Complex, possibly due to the lack of methylation on NDUFB3, an accessory subunit of Complex I (Paper I Fig. 5). Furthermore, this shows that METTL9 plays an important role in cancer progression. Taken together, the results of these studies indicate not only the potential value of these MTases and the methylations they introduce, but also suggests that drug targeting of protein histidine MTases may represent novel therapeutic strategies for cancer treatment.

Overall, the evidence that has so far been gathered in characterizing METTL9 indicates that we have barely uncovered the tip of the iceberg. Besides its involvement in immunomodulation and zinc transport, it has also been associated with multiple diseases, such as cancer. Future studies are needed to fully characterize the MTase and exploit that knowledge for the creating of new therapeutic strategies.

Part III – Biochemical studies of protein histidine methyltransferases

Structure biology and biochemistry in symbiosis

In times of algorithms and artificial intelligence not only predicting the structure of proteins but also modelling protein complexes with increasing precision, one can impulsively judge the future of experimental structure biology and biochemistry to be rather dark. However, we have observed, that while the general structure prediction from AlphaFold was correct, the flexible area in the N-terminus of METTL9 from *O. tauri* was predicted to have a more ordered structure than we saw in the structure from the crystal (Manuscript I Sup. Fig. 4). Additionally, structure predictions are thus far not able to give in depth structure accuracy for side chain orientations

In our study about CARNMT1 we have also seen an interesting discrepancy between activity of the D316A mutant used by Cao et al. [48], which in their study almost completely disrupted enzymatic activity of CARNMT1 on the peptide carnosine. After creation of the D316A mutant in our group, we however were not able to reduce activity on protein level, it had however similar activity on the ZIP7 fragment as WT CARNMT1. This may have some alternative explanations such as human error in the generation of the mutant. This is in line with the importance of the residue, as the analogous aspartic acid residue in our METTL9 study as crucial for histidine stabilization, and the human METTL9 D213A mutation was inactive (Manuscript I Sup. Fig. 2). If the observed activity of the D316A mutant is real, it can lead to other suggestions. One is that the mutation, as described by the study is holding the histidine in place so that specifically π-methylhistidine is created. A mutation could then lead to the creation of τ-methylhistidine, which they did not measure as they measured activity with anserine generation, and not balenine. Another hypothesis is that the mode of binding carnosine, which is an unconventional peptide, is different, so that D316 is important for binding carnosine, but not longer conventional polypeptide substrates. Once CARNMT1 contacts with ZIP7 and possible resulting conformational changes allowed this mutant of the enzyme to retain its activity. Additionally, we can see differences in the substrate preferences of different METTL9 orthologues, even though they mostly have slight changes in the motifs.

Another interesting example is a SETD3 study, in which the histidine target of actin was mutated to methionine or lysine to inhibit activity [79]. Surprisingly however, they saw, even though reduced, activity of SETD3 on both methionine and lysine in vitro. In a similar attempt, they have methylated the MTase itself to change substrate specificity, leading to SETD3 methylating lysine residue [80]. Altogether we can see that small changes in either the substrate or the amino acid sequence of the enzyme can alter the activity. The fact that MTases can change its function based on single amino acid

mutations makes it challenging for use bioinformatics (on structure or sequence) to predict the biochemical function of protein MTases. Thorough biochemical studies are needed to verify any bioinformatical predictions, otherwise it becomes easy to falsely assign substrates to protein MTases, as multiple examples have shown.

The false assignment of substrates to methyltransferases

As the results of this thesis indicate, there is a great need for biochemistry to study the function of enzymes, to identify their *bona fide* substrates and through that elucidate their biological function. There are multiple examples in the literature describing MTases that have wrongly been assigned with substrates that are incorrect. One example is SETD3, which was first been described as a human H3K36 lysine-specific MTase in 2011 [50]. However, this study based their conclusion on MS data in which mass shifts were inconsistent with *bona fide* methylation events [27]. Still, the histone lysine activity has been accepted as truth and for seven years used as a basis for considerable research. An impressive number of 10 studies have been published that incorporated this incorrect knowledge into their design of their study [50], [54], [81]–[84] or interpretation of results [85]–[88], until finally the true activity of SETD3 was published by two independent biochemical studies [27], [51]. To address the problem of false MTase substrate assignments, Kudithipudi and Jeltsch published guidelines for the identification of substrates of protein lysine MTases [66]. These seven basic biochemical rules can easily be translated to substrates of protein histidine MTases and include: "(1) include positive controls; (2) use target [histidine] mutations of substrate proteins as negative controls; (3) use inactive enzyme variants as negative controls; (4) report quantitative methylation data; (5) consider [MTase] specificity; (6) validate [methyl-specific] antibodies; and (7) connect cellular and in vitro results." [66]. These rules need to be followed to enable reproducible research with *bona fide* substrate identification.

A comment on the nomenclature of methylated histidine

The inconsistencies in the naming of the nitrogen atom in the imidazole ring have led to confusion during literature research for this thesis and even in communication between collaborators. Several parallel nomenclatures are currently used to denote the two nitrogen atoms in the imidazole group of histidine. The first is based on Greek letters/language, and the N-atoms are denoted according to their distance to the backbone: pros ('near', abbreviated π) and tele ('far', abbreviated τ) (Fig. 8a). In addition, two different and opposite ways of numerical naming exist. The international union of pure and applied chemistry (IUPAC) numbering denotes the τ-nitrogen as the 1st position, while the π-nitrogen is counted as the 3rd position of the imidazole ring (Fig. 8b) [89]. While this numbering is widely used in organic chemistry, biochemists use a system that results in the opposite numbering, where the π-nitrogen is the 1st position and the τ-nitrogen is the 3rd position (Fig. 8c). Biochemical studies, e.g. on protein methylation, have a history of publications using "1-MH" for π-methylation and "3-MH" for τ-methylation, even though the Greek lettering is often mentioned throughout the publication as well. This problem is not new, and was first described in 1962, due to discrepancies in the naming of the methylated variants of carnosine [34], and has been addressed in reviews of the field (for example [68]), however, it has still not been solved. A correction of biochemical counting to the IUPAC system will cause too much confusion retrospectively for the already published papers. Therefore, a better solution would be to use the π- and τ- nomenclature in addressing the methylhistidine numbering in the future, similarly to how it is done in this thesis.
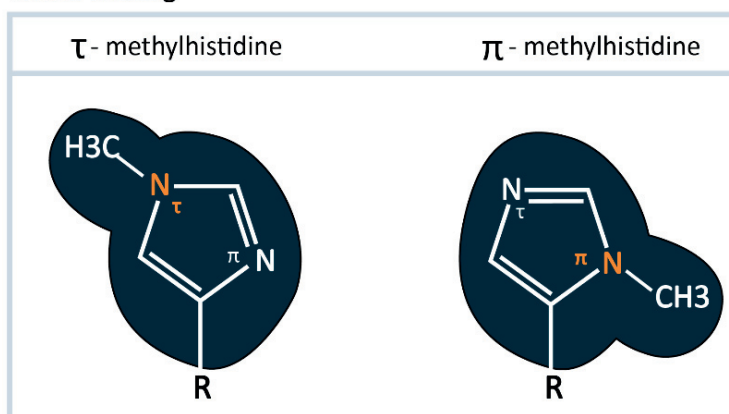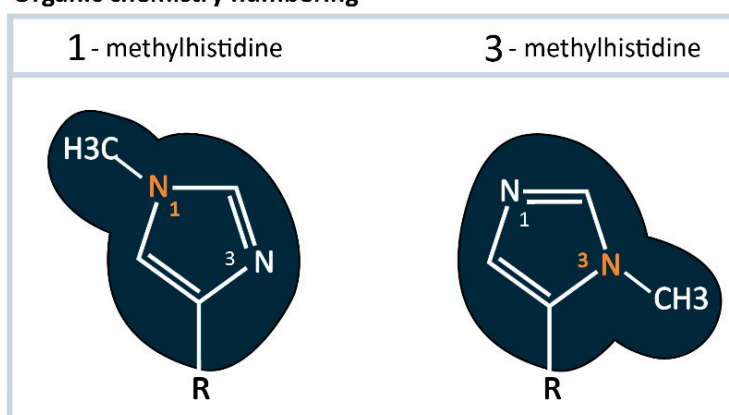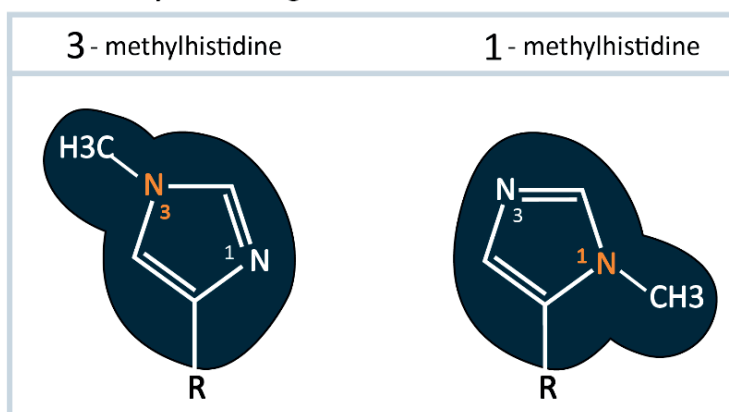
Figure 8: Histidine methylation nomenclature. Differences in counting of the nitrogen atoms have resulted in confusing throughout the disciplines. **a)** In biochemistry the proximal (π) nitrogen is counted as the first, while the distal (τ) nitrogen is counted as the third. **b)** In organic chemistry the IUPAC numbering results in the opposite, the proximal (π) nitrogen is counted as the third, while the distal (τ) nitrogen is counted as the first. **c)** Suggested is the naming according to Greek letters with the proximal (π) or distal (τ) nitrogen atom.

Conclusion and Future Perspective

As this thesis describes, protein histidine methylation is an important molecular event that helps fine-tune the function of its substrates. The immediate future steps for our studies are to identify of additional substrates of METTL9 and CARNMT1 and to look deeper into the biological significance of histidine methylation. For METTL9, the investigation of its involvement in zinc homeostasis and metal biology would be particularly interesting, as METTL9 methylates a great number of zinc transporters and other metal-binding proteins.

To learn more about the evolution of enzymes and structure-function relationships, it will be of essence to solve the structure of multiple orthologues of METTL9 and compare the structural difference and substrate preferences. It would be also interesting to identify the MTase responsible for the τ-methylhistidine generation in carnosine, forming balenine, and compare its structures with that of CARNMT1 to see the molecular basis needed for the minute switch between the generation of either π- or τ-methylhistidine in the same peptide substrate. In the big picture, it will be interesting to investigate how the expression and activity of both of these MTases are regulated, as well as to identify whether any reader proteins or demethylase enzymes exist for the π-methylhistidine modification. This may eventually allow us to generate therapeutic strategies against any disease in which the histidine MTase are involved.

In conclusion, we characterized the first π-methylhistidine generating protein MTase, METTL9, and solved its structure. We studied the activity of METTL9 orthologues and discovered that some of these enzymes have different target sequence preferences. Additionally, we showed that CARNMT1 in vitro methylates not only di- or tri-peptides as previously shown, but also proteins. Altogether, this thesis is a small but significant contribution to the growing histidine methyltransferase universe.

Bibliography

[1]   T. A. Brown, *Genomes, 2nd ed.* 2002.

[2]   H. Lodish, "Molecular Cell Biology 8th ed.," *W.H. Freeman*. 2016.

[3]   T. W. Nilsen and B. R. Graveley, "Expansion of the eukaryotic proteome by alternative splicing," *Nature*. 2010.

[4]   V.N. Uversky, *Edition), Brenner's Encyclopedia of Genetics (Second*, Second Edi. Academic Press, 2013.

[5]   S. Ramazi and J. Zahiri, "Post-translational modifications in proteins: Resources, tools and prediction methods," *Database*, vol. 2021. 2021.

[6]   S. Kwiatkowski, A. Kiersztan, and J. Drozak, "Biosynthesis of Carnosine and Related Dipeptides in Vertebrates," *Curr. Protein Pept. Sci.*, 2018.

[7]   M. V. C. Greenberg and D. Bourc'his, "The diverse roles of DNA methylation in mammalian development and disease," *Nature Reviews Molecular Cell Biology*. 2019.

[8]   Y. Zhou *et al.*, "Principles of RNA methylation and their implications for biology and medicine," *Biomedicine and Pharmacotherapy*. 2020.

[9]   S. G. Clarke, "Protein methylation at the surface and buried deep: Thinking outside the histone box," *Trends in Biochemical Sciences*. 2013.

[10]   M. Kulis and M. Esteller, *DNA Methylation and Cancer*. 2010.

[11]   A. Barski *et al.*, "High-Resolution Profiling of Histone Methylations in the Human Genome," *Cell*, 2007.

[12]   A. Dhayalan and A. Jeltsch, "Special Issue 'Structure, Activity and Function of Protein Methyltransferases,'" *Life*. 2022.

[13]   R. Di Blasi, O. Blyuss, J. F. Timms, D. Conole, F. Ceroni, and H. J. Whitwell, "Non-Histone Protein Methylation: Biological Significance and Bioengineering Potential," *ACS Chemical Biology*, vol. 16, no. 2. pp. 238–250, 2021.

[14]   J. L. Martin and F. M. McMillan, "SAM (dependent) I AM: The S-adenosylmethionine-dependent methyltransferase fold," *Current Opinion in Structural Biology*. 2002.

[15]   Z. Ning *et al.*, "A charge-suppressing strategy for probing protein methylation," *Chem. Commun.*, 2016.

[16]   T. C. Petrossian and S. G. Clarke, "Uncovering the human methyltransferasome," *Mol. Cell. Proteomics*, 2011.

[17]   P. Ø. Falnes, M. E. Jakobsson, E. Davydova, A. Ho, and J. Małecki, "Protein lysine methylation by seven-β-strand methyltransferases," *Biochem. J.*, 2016.

[18]   J. drzej Małecki, M. E. Jakobsson, A. Y. Y. Ho, A. Moen, A. C. Rustan, and P. Falnes, "Uncovering human METTL12 as a mitochondrial methyltransferase that modulates citrate synthase activity through metabolite-sensitive lysine methylation," *J. Biol. Chem.*, 2017.

[19]   J. Drozak *et al.*, "UPF0586 protein C9orf41 homolog is anserine-producing methyltransferase," *J. Biol. Chem.*, 2015.

[20] P. Falnes, M. E. Jakobsson, E. Davydova, A. Ho, and J. Malecki, "Protein lysine methylation by seven-β-strand methyltransferases," *Biochemical Journal*. 2016.

[21] R. S. Jones and W. M. Gelbart, "The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax," *Mol. Cell. Biol.*, 1993.

[22] H. M. Herz, A. Garruss, and A. Shilatifard, "SET for life: Biochemical activities and biological functions of SET domain-containing proteins," *Trends in Biochemical Sciences*. 2013.

[23] B. Tschiersch, A. Hofmann, V. Krauss, R. Dorn, G. Korge, and G. Reuter, "The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes," *EMBO J.*, 1994.

[24] Q. Feng *et al.*, "Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain," *Curr. Biol.*, 2002.

[25] S. C. Dillon, X. Zhang, R. C. Trievel, and X. Cheng, "The SET-domain protein superfamily: Protein lysine methyltransferases," *Genome Biology*. 2005.

[26] S. Pradhan, H. G. Chin, P. O. Estève, and S. E. Jacobsen, "SET7/9 mediated methylation of non-histone proteins in mammalian cells," *Epigenetics*. 2009.

[27] A. W. Wilkinson *et al.*, "SETD3 is an actin histidine methyltransferase that prevents primary dystocia," *Nature*, 2018.

[28] M. J. Lopez and S. S. Mohiuddin, *Biochemistry, Essential Amino Acids*. 2020.

[29] S. M. Liao, Q. S. Du, J. Z. Meng, Z. W. Pang, and R. B. Huang, "The multiple roles of histidine in protein interactions," *Chem. Cent. J.*, vol. 7, no. 1, 2013.

[30] R. Zhu *et al.*, "Allosteric histidine switch for regulation of intracellular zinc(II) fluctuation," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 114, no. 52, 2017.

[31] Z. Li *et al.*, "DbPTM in 2022: An updated database for exploring regulatory networks and functional associations of protein post-translational modifications," *Nucleic Acids Res.*, vol. 50, no. D1, 2022.

[32] D. Ackermann, O. Timpe, and K. Poller, "Über das Anserin, einen neuen Bestandteil der Vogelmuskulatur," *Hoppe. Seylers. Z. Physiol. Chem.*, vol. 183, no. 1–2, 1929.

[33] J. A. Zapp and D. W. Wilson, "QUANTITATIVE STUDIES OF CARNOSINE AND ANSERINE IN MAMMALIAN MUSCLE," *J. Biol. Chem.*, vol. 126, no. 1, 1938.

[34] F. Pocchiari, L. Tentori, and G. Vivaldi, "The presence of the dipeptide β-alanyl-3-methylhistidine in whale meat extract," *Sci Rept Ist Super Sanita*, vol. 2, p. 188, 1962.

[35] J. M. SEARLE and R. G. WESTALL, "The occurrence of free methylhistidine in urine.," *Biochem. J.*, vol. 48, no. 4, 1951.

[36] H. H. TALLAN, W. H. STEIN, and S. MOORE, "3-Methylhistidine, a new amino acid from human urine.," *J. Biol. Chem.*, vol. 206, no. 2, 1954.

[37] P. Johnson, C. I. Harris, and S. V. Perry, "3-methylhistidine in actin and other muscle proteins.," *Biochem. J.*, vol. 105, no. 1, 1967.

[38] A. M. Asatoor and M. D. Armstrong, "3-Methylhistidine, a component of actin," *Biochem. Biophys. Res. Commun.*, vol. 26, no. 2, 1967.

[39] P. Johnson and S. V. Perry, "Biological activity and the 3-methylhistidine content of actin and

myosin.," *Biochem. J.*, vol. 119, no. 2, 1970.

[40] E. M., "Amino acid sequence aroung 3-methylhistidine in rabbit skeletal muscle actin," *Biochemistry*, vol. 10(2):224-, 1971.

[41] H. E. Meyer and G. W. Mayr, "N pi-methylhistidine in myosin-light-chain kinase.," *Biol. Chem. Hoppe. Seyler.*, 1987.

[42] J. Carroll *et al.*, "The Post-translational Modifications of the Nuclear Encoded Subunits of Complex I from Bovine Heart Mitochondria," *Mol. &amp; Cell. Proteomics*, vol. 4, no. 5, pp. 693 LP – 699, May 2005.

[43] E. L. Gershey, G. W. Haslett, G. Vidali, and V. G. Allfrey, "Chemical studies of histone methylation. Evidence for the occurrence of 3-methylhistidine in avian erythrocyte histone fractions.," *J. Biol. Chem.*, vol. 244, no. 18, 1969.

[44] T. W. Borun, D. Pearson, and W. K. Paik, "Studies of histone methylation during the HeLa S-3 cell cycle.," *J. Biol. Chem.*, 1972.

[45] M. E. Jakobsson, "Enzymology and significance of protein histidine methylation," *Journal of Biological Chemistry*, vol. 297, no. 4. 2021.

[46] Z. F. Yuan *et al.*, "Epiprofile quantifies histone peptides with modifications by extracting retention time and intensity in high-resolution mass spectra," *Mol. Cell. Proteomics*, vol. 14, no. 6, 2015.

[47] S. Kapell and M. E. Jakobsson, "Large-scale identification of protein histidine methylation in human cells," *NAR Genomics Bioinforma.*, vol. 3, no. 2, 2021.

[48] R. Cao, X. Zhang, X. Liu, Y. Li, and H. Li, "Molecular basis for histidine N1 position-specific methylation by CARNMT1," *Cell Research*. 2018.

[49] D. W. Kim, K. B. Kim, J. Y. Kim, and S. B. Seo, "Characterization of a novel histone H3K36 methyltransferase setd3 in zebrafish," *Biosci. Biotechnol. Biochem.*, vol. 75, no. 2, 2011.

[50] G. H. Eom *et al.*, "Histone methyltransferase SETD3 regulates muscle differentiation," *J. Biol. Chem.*, vol. 286, no. 40, 2011.

[51] S. Kwiatkowski *et al.*, "SETD3 protein is the actin-specific histidine N-methyltransferase," *Elife*, 2018.

[52] A. Witecka, S. Kwiatkowski, T. Ishikawa, and J. Drozak, "The structure, activity, and function of the setd3 protein histidine methyltransferase," *Life*, vol. 11, no. 10. 2021.

[53] J. Diep *et al.*, "Enterovirus pathogenesis requires the host methyltransferase SETD3," *Nat. Microbiol.*, vol. 4, no. 12, 2019.

[54] O. Cohn, M. Feldman, L. Weil, M. Kublanovsky, and D. Levy, "Chromatin associated SETD3 negatively regulates VEGF expression," *Sci. Rep.*, vol. 6, 2016.

[55] W. J. Shu and H. N. Du, "The methyltransferase SETD3-mediated histidine methylation: Biological functions and potential implications in cancers," *Biochimica et Biophysica Acta - Reviews on Cancer*, vol. 1875, no. 1. 2021.

[56] A. Witecka, M. Zarod, T. Ishikawa, and J. Drozak, "Preliminary identification of novel protein substrates for actin-histidine N-methyltransferase (SETD3)," *FEBS Open Bio*, vol. 12 MA-P, pp. 251-252 WE-Science Citation Index Expanded (SCI, 2022.

[57] K. J. Webb *et al.*, "A novel 3-methylhistidine modification of yeast ribosomal protein Rpl3 is

dependent upon the YIL110W methyltransferase," *J. Biol. Chem.*, vol. 285, no. 48, 2010.

[58] J. M. Małecki *et al.*, "Human METTL18 is a histidine-specific methyltransferase that targets RPL3 and affects ribosome biogenesis and function," *Nucleic Acids Res.*, vol. 49, no. 6, 2021.

[59] P. Cloutier, M. Lavallée-Adam, D. Faubert, M. Blanchette, and B. Coulombe, "A Newly Uncovered Group of Distantly Related Lysine Methyltransferases Preferentially Interact with Molecular Chaperones to Regulate Their Activity," *PLoS Genet.*, vol. 9, no. 1, 2013.

[60] E. Matsuura-Suzuki *et al.*, "METTL18-mediated histidine methylation of RPL3 modulates translation elongation for proteostasis maintenance," *Elife*, vol. 11, 2022.

[61] C. L. Han, Y. C. Sheng, S. Y. Wang, Y. H. Chen, and J. H. Kang, "Serum proteome profiles revealed dysregulated proteins and mechanisms associated with fibromyalgia syndrome in women," *Sci. Rep.*, vol. 10, no. 1, 2020.

[62] M. J. Bair and E. E. Krebs, "In the clinic®: fibromyalgia," *Ann. Intern. Med.*, vol. 172, no. 5, 2020.

[63] T. H. Li *et al.*, "Identification METTL18 as a Potential Prognosis Biomarker and Associated With Immune Infiltrates in Hepatocellular Carcinoma," *Front. Oncol.*, vol. 11, 2021.

[64] A. E. Hammerstrom, D. H. Cauley, B. J. Atkinson, and P. Sharma, "Cancer immunotherapy: Sipuleucel-T and beyond," *Pharmacotherapy*, vol. 31, no. 8, pp. 813–28, 2011.

[65] M. E. Jakobsson, A. Moen, and P. Falnes, "Correspondence: On the enzymology and significance of HSPA1 lysine methylation," *Nature Communications*, vol. 7. 2016.

[66] S. Kudithipudi and A. Jeltsch, "Approaches and Guidelines for the Identification of Novel Substrates of Protein Lysine Methyltransferases," *Cell Chemical Biology*, vol. 23, no. 9. 2016.

[67] M. Uhlén *et al.*, "Tissue-based map of the human proteome," *Science (80-.).*, vol. 347, no. 6220, 2015.

[68] A. A. Boldyrev, G. Aldini, and W. Derave, "Physiology and pathophysiology of carnosine," *Physiological Reviews*, vol. 93, no. 4. 2013.

[69] A. W. Wilkinson *et al.*, "SETD3 is an actin histidine methyltransferase that prevents primary dystocia," *Nature*. 2019.

[70] Q. Al-Hadid, K. Roy, G. Chanfreau, and S. G. Clarke, "Methylation of yeast ribosomal protein Rpl3 promotes translational elongation fidelity," *RNA*, vol. 22, no. 4, 2016.

[71] M. Lv *et al.*, "METTL9 mediated N1-histidine methylation of zinc transporters is required for tumor growth," *Protein and Cell*, vol. 12, no. 12. 2021.

[72] H. Daitoku *et al.*, "siRNA screening identifies METTL9 as a histidine Nπ-methyltransferase that targets the proinflammatory protein S100A9," *J. Biol. Chem.*, vol. 297, no. 5, 2021.

[73] C. Thiebaut, L. Eve, C. Poulard, and M. Le Romancer, "Structure, activity, and function of prmt1," *Life*, vol. 11, no. 11. 2021.

[74] E. Davydova *et al.*, "The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes," *Nat. Commun.*, vol. 12, no. 1, 2021.

[75] E. Tassano *et al.*, "'Distal 16p12.2 microdeletion' in a patient with autosomal recessive deafness-22," *J. Genet.*, vol. 98, no. 2, 2019.

[76] J. Ding *et al.*, "An integrative genomic analysis of transcriptional profiles identifies

characteristic genes and patterns in HIV-infected long-term non-progressors and elite controllers," *J. Transl. Med.*, vol. 17, no. 1, 2019.

[77]   Q. Zheng *et al.*, "Immune signature-based hepatocellular carcinoma subtypes may provide novel insights into therapy and prognosis predictions," *Cancer Cell Int.*, vol. 21, no. 1, 2021.

[78]   T. Hara, Y. Tominaga, K. Ueda, K. Mihara, K. Yanagihara, and Y. Takei, "Elevated METTL9 is associated with peritoneal dissemination in human scirrhous gastric cancers.," *Biochem. Biophys. reports*, vol. 30, p. 101255, Jul. 2022.

[79]   S. Dai *et al.*, "Characterization of SETD3 methyltransferase-mediated protein methionine methylation," *J. Biol. Chem.*, vol. 295, no. 32, 2020.

[80]   S. Dai, J. R. Horton, A. W. Wilkinson, O. Gozani, X. Zhang, and X. Cheng, "An engineered variant of SETD3 methyltransferase alters target specificity from histidine to lysine methylation," *J. Biol. Chem.*, vol. 295, no. 9, 2020.

[81]   Z. Chen, C. T. Yan, Y. Dou, S. S. Viboolsittiseri, and J. H. Wang, "The role of a newly identified SET domain-containing protein, SETD3, in oncogenesis," *Haematologica*, vol. 98, no. 5, 2013.

[82]   A. S. Pires-Luís *et al.*, "Expression of histone methyltransferases as novel biomarkers for renal cell tumor diagnosis and prognostication," *Epigenetics*, vol. 10, no. 11, 2015.

[83]   X. Cheng *et al.*, "Cell cycle-dependent degradation of the methyltransferase SETD3 attenuates cell proliferation and liver tumorigenesis," *J. Biol. Chem.*, vol. 292, no. 22, 2017.

[84]   M. Tiebe, M. Lutz, D. Levy, and A. A. Teleman, "Phenotypic characterization of SETD3 knockout Drosophila," *PLoS One*, vol. 13, no. 8, 2018.

[85]   S. E. Cooper, E. Hodimont, and C. M. Green, "A fluorescent bimolecular complementation screen reveals MAF1, RNF7 and SETD3 as PCNAassociated proteins in human cells," *Cell Cycle*, vol. 14, no. 15, 2015.

[86]   M. J. Ferreira *et al.*, "SETDB2 and RIOX2 are differentially expressed among renal cell tumor subtypes, associating with prognosis and metastization," *Epigenetics*, vol. 12, no. 12, 2017.

[87]   R. A. Seaborne *et al.*, "Human Skeletal Muscle Possesses an Epigenetic Memory of Hypertrophy," *Sci. Rep.*, vol. 8, no. 1, 2018.

[88]   X. Jiang, T. Li, J. Sun, J. Liu, and H. Wu, "SETD3 negatively regulates VEGF expression during hypoxic pulmonary hypertension in rats," *Hypertens. Res.*, vol. 41, no. 9, 2018.

[89]   IUPAC, "Compendium of Chemical Terminology 2nd ed. (the 'Gold Book')," *Blackwell Sci. Publ. Oxford*, vol. 80, no. 2, 2014.

Paper and Manuscripts
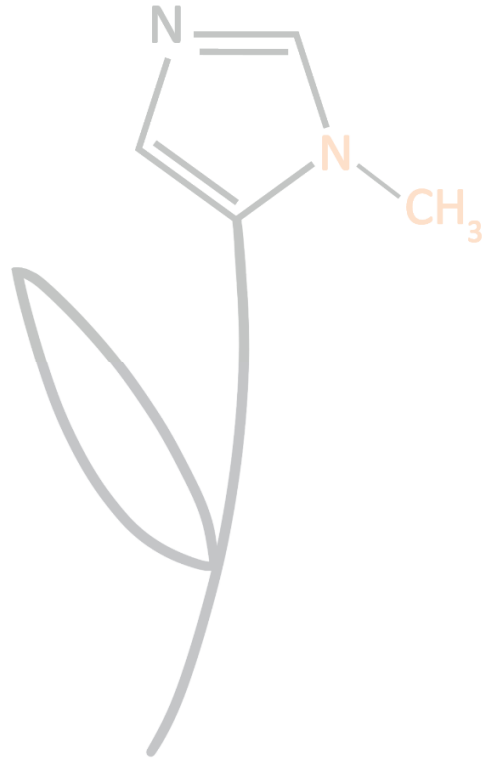
ARTICLE

The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes

Erna Davydova [1,13], Tadahiro Shimazu[2,13], Maren Kirstin Schuhmacher [3], Magnus E. Jakobsson [4,5], Hanneke L. D. M. Willemen [6], Tongri Liu[7], Anders Moen[1], Angela Y. Y. Ho[1], Jędrzej Małecki [1], Lisa Schroer [1], Rita Pinto[1,12], Takehiro Suzuki[8], Ida A. Grønsberg[1], Yoshihiro Sohtome [9,10], Mai Akakabe[9], Sara Weirich[3], Masaki Kikuchi [11], Jesper V. Olsen [4], Naoshi Dohmae[8], Takashi Umehara [11], Mikiko Sodeoka [9,10], Valentina Siino[5], Michael A. McDonough [7], Niels Eijkelkamp [6], Christopher J. Schofield [7], Albert Jeltsch [3,14✉], Yoichi Shinkai [2,14✉] & Pål Ø. Falnes [1,14✉]

Post-translational methylation plays a crucial role in regulating and optimizing protein function. Protein histidine methylation, occurring as the two isomers 1- and 3-methylhistidine (1MH and 3MH), was first reported five decades ago, but remains largely unexplored. Here we report that METTL9 is a broad-specificity methyltransferase that mediates the formation of the majority of 1MH present in mouse and human proteomes. METTL9-catalyzed methylation requires a His-x-His (HxH) motif, where "x" is preferably a small amino acid, allowing METTL9 to methylate a number of HxH-containing proteins, including the immunomodulatory protein S100A9 and the NDUFB3 subunit of mitochondrial respiratory Complex I. Notably, METTL9-mediated methylation enhances respiration via Complex I, and the presence of 1MH in an HxH-containing peptide reduced its zinc binding affinity. Our results establish METTL9-mediated 1MH as a pervasive protein modification, thus setting the stage for further functional studies on protein histidine methylation.

[1] Department of Biosciences, Faculty of Mathematics and Natural Sciences, University of Oslo, 0316 Oslo, Norway. [2] Cellular Memory Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama, Japan. [3] Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany. [4] Proteomics Program, Faculty of Health and Medical Sciences, Novo Nordisk Foundation Center for Protein Research (NNF-CPR), University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark. [5] Department of Immunotechnology, Lund University, Medicon Village, 22100 Lund, Sweden. [6] Center for Translational Immunology (CTI), University Medical Center Utrecht, Utrecht University, 3584 Utrecht, EA, The Netherlands. [7] Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Oxford, UK. [8] Biomolecular Characterization Unit, Technology Platform Division, RIKEN Center for Sustainable Resource Science, Wako, Saitama, Japan. [9] Synthetic Organic Chemistry Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama, Japan. [10] RIKEN Center for Sustainable Resource Science, Wako, Saitama, Japan. [11] Laboratory for Epigenetics Drug Discovery, RIKEN Center for Biosystems Dynamics Research, Yokohama, Japan. [12] Present address: Department of Molecular Oncology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway. [13] These authors contributed equally: Erna Davydova, Tadahiro Shimazu. [14] These authors jointly supervised this work: Albert Jeltsch, Yoichi Shinkai, Pål Ø. Falnes. ✉email: albert.jeltsch@ibtb.uni-stuttgart.de; yshinkai@riken.jp; pal.falnes@ibv.uio.no

Proteins are frequently modified by post-translational methylation, which is primarily found on lysines and arginines, but can also occur at other residues, such as glutamine and histidine[1–4]. The functional significance of protein methylation has been most intensively studied for lysine methylation of histone proteins, playing important roles in the regulation of gene expression and chromatin state[5]. In humans, protein methylation is mediated by a number of S-adenosylmethionine (AdoMet)-dependent methyltransferases (MTases), belonging to two distinct classes, the SET-domain and the seven-β-strand (7BS) MTases[3,6–9]. The SET-domain MTases mainly encompass lysine-specific histone MTases, whereas the 7BS MTases, collectively, target a wide range of substrates[6,8,9].

Histidine can be methylated at either the N1 or N3 position of its imidazole ring, yielding the isomers 1-methylhistidine (1MH; also referred to as π-methylhistidine) or 3-methylhistidine (3MH; τ-methylhistidine), respectively[1]. Histidine methylation was first described five decades ago, when it was shown that actin and myosin from muscle contain 3MH[1,10,11]. Since then, histidine methylation has been firmly established for only a few additional mammalian proteins. These are myosin light chain kinase (MYLK2) from rabbit skeletal muscle and the S100A9 subunit of the dimeric antimicrobial and immunomodulatory protein calprotectin from mouse spleen[12,13], both of which carry a 1MH-modified histidine. In addition, the B12 subunit of mitochondrial respiratory Complex I (NDUFB3) has been shown to contain multiple methylated histidines close to its N-terminus, but the exact chemical nature of the modification (1MH or 3MH) was not investigated[14]. Moreover, recent high-throughput proteomics analyses have indicated the presence of several hundred histidine-methylated proteins in humans, implying that this modification is widespread[15,16]. Importantly, studies on mammalian histidine methylation and its biological significance have so far been hampered by the absence of knowledge about the responsible MTases. However, it was recently shown that the SET-domain MTase SETD3 is the long-sought enzyme responsible for the introduction of 3MH at His-73 in actin, and that SETD3-mediated methylation plays an important role in regulating actin function and muscle contractility[16,17].

Histidine is arguably the most versatile of the proteinogenic amino acids. Both of the nitrogen atoms in its imidazole ring may become protonated, yielding neutral, and positively charged forms of histidine at physiological pH. Histidine is therefore a common catalytic residue, acting as either a general base or acid. It can also be a sensor of local pH, and is involved in a variety of other interactions, frequently including coordination of metal cations[18,19]. Against this background, histidine methylation has the potential to regulate a wide range of molecular interactions and cellular processes.

In the present work, we investigate the function of the previously uncharacterized mammalian 7BS MTase METTL9. We find that METTL9 introduces 1MH at His-x-His (HxH; where x is optimally a small residue) motifs in several different proteins from human and mouse. Amino acid analysis of total protein hydrolysates shows that 1MH is a relatively abundant modification and that METTL9 introduces the majority of 1MH in human and mouse proteomes. Moreover, the mitochondrial Complex I protein NDUFB3 is found to be an in vivo target of METTL9, and, in accordance with this, the enzymatic activity of METTL9 promotes Complex I-mediated respiration. Finally, the presence of 1MH modifications in an HxH-containing peptide is found to diminish its affinity towards zinc.

Results

METTL9 is a protein histidine MTase. As part of efforts to investigate uncharacterized human 7BS MTases, we set out to study METTL9. Beyond containing the classical 7BS MTase hallmark motifs, METTL9 shows no significant sequence homology to other human 7BS MTases. However, putative METTL9 orthologues are widespread throughout eukaryotes, though notably absent from fungi and land plants (Supplementary Fig. 1).

To determine whether METTL9 is a protein MTase, we investigated the ability of a recombinant fusion protein between human METTL9 and glutathione-S-transferase (GST-hMETTL9) to methylate proteins in fractionated human cell extracts in the presence of [3H]AdoMet, and detected protein methylation by fluorography. Both *METTL9* knockout (KO) and wild-type (WT) cells were used, and an MTase inactive mutant (E174A; Supplementary Fig. 1) enzyme was included as a negative control. WT GST-hMETTL9 methylated several distinct proteins in the different cellular fractions (Fig. 1a and Supplementary Fig. 2a). This was particularly striking for the membrane fraction, where several proteins were specifically methylated by the WT enzyme and showed a stronger signal in the KO extracts (Fig. 1a), indicating that these proteins represent *bona fide* cellular substrates of METTL9. In agreement with this, a fusion protein between METTL9 and green fluorescent protein (hMETTL9-GFP) co-localized with markers for membranous compartments such as the endoplasmic reticulum and mitochondria in HeLa cells (Supplementary Fig. 2b). We also observed similar yet distinct hMETTL9-dependent methylation patterns in extracts from various tissues from WT and *Mettl9* KO mice (Supplementary Fig. 2c). These results demonstrate that METTL9 is a protein MTase responsible for methylating a multitude of substrates from different tissues and subcellular compartments.

Attempts to identify the METTL9 substrates observed in fluorography using fractionation and protein mass spectrometry (MS) were unsuccessful. Thus, we searched for substrates among previously published METTL9 interactants[20–22] by incubating several of the corresponding recombinant His6-tagged proteins with recombinant His6-hMETTL9 in the presence of AdoMet. In this way, we found that the poorly characterized Armadillo repeat-containing protein 6 (ARMC6) was an in vitro substrate of METTL9, and MS analysis showed that a peptide corresponding to residues 248–267 of ARMC6 became partially methylated after treatment with the MTase (Fig. 1b, Supplementary Fig. 3), with the modification either at His-263 or Asn-264 (Supplementary Fig. 4a). To pinpoint the methylation site, we introduced alanine substitution mutations at His-263, Asn-264 and surrounding residues, and tested the ability of the resulting mutant proteins to undergo GST-hMETTL9-mediated methylation using [3H]Ado-Met and fluorography. Wild-type ARMC6 and most of the mutants were efficiently methylated, but replacement of either His-261 or His-263 abolished METTL9-mediated methylation (Fig. 1c). Through MS analysis of tryptic peptides derived from proteins co-immunoprecipitated with hMETTL9-GFP inducibly expressed in HEK-293-derived cells, we found that ARMC6 was methylated at His-263 (Supplementary Fig. 4b), whereas the corresponding unmethylated peptide was undetectable. Thus, this shows cellular methylation of ARMC6 within the METTL9 target sequence. These results strongly indicate that METTL9 is a histidine MTase that methylates ARMC6 both in vitro and in cells.

METTL9 methylates HxH motifs. The methylation site of ARMC6 contains three alternating histidines (HxHxH) and, remarkably, such a pattern is also present in two of the few proteins previously reported to contain methylhistidine, mouse S100A9 and bovine NDUFB3 (Fig. 1d)[13,14]. Moreover, when searching the human proteome for instances of HxHxH, we
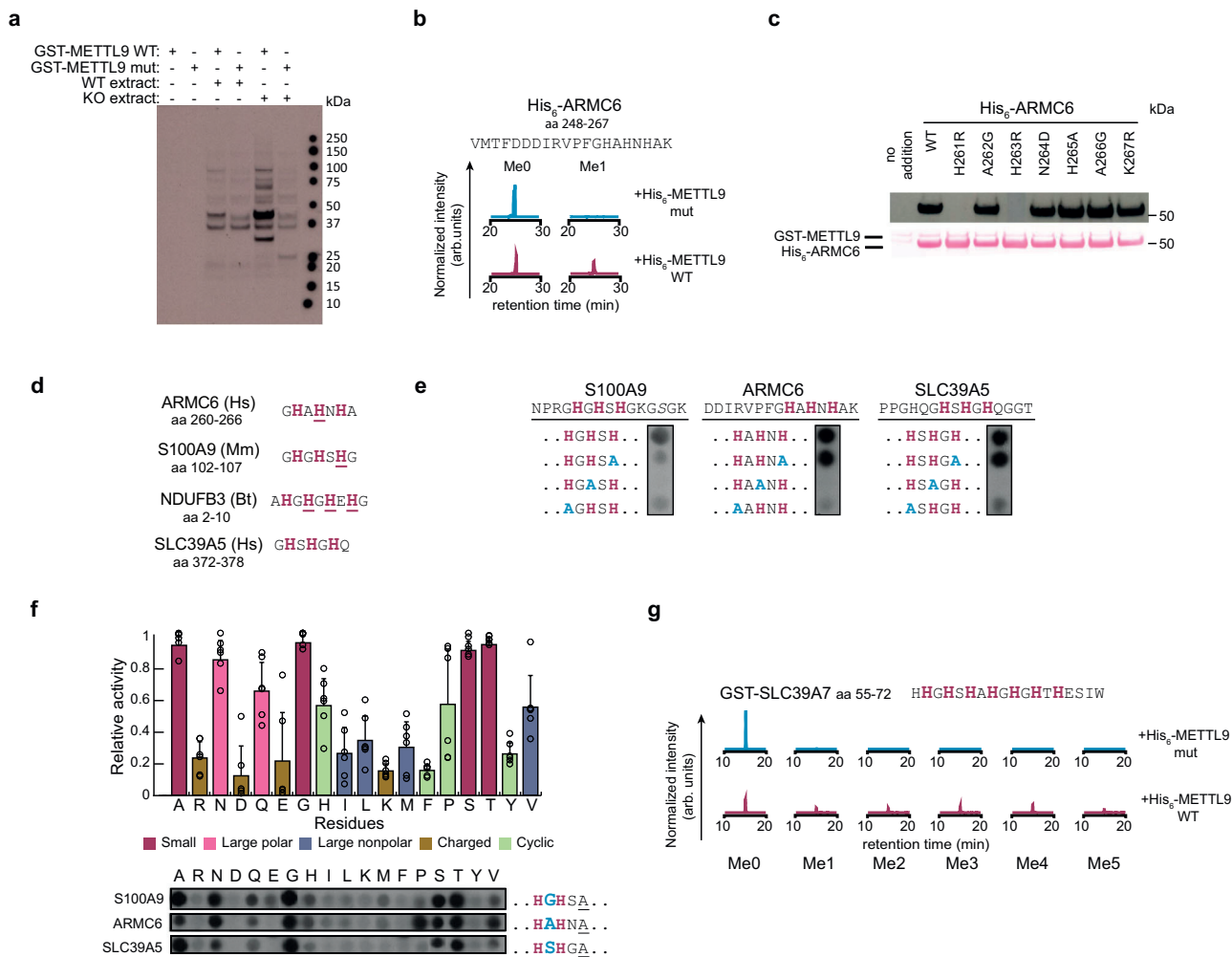
Fig. 1 METTL9 is a protein histidine methyltransferase that methylates the ANGST HxH motif. a Fluorography showing MTase activity, using [³H] AdoMet, of recombinant wild-type (WT) GST-hMETTL9 on proteins in the membrane fraction from WT and *METTL9* KO HAP1 cells (METTL9 mut: inactive mutant E147A). The data are representative of three independent experiments. **b** Enzymatically active METTL9 methylates ARMC6. Recombinant His₆-ARMC6 was treated with His₆-hMETTL9 in the presence of AdoMet in vitro, then digested with trypsin, and the resulting peptides analyzed by LC-MS. Normalized extracted ion chromatograms (XICs) corresponding to the monomethylated (Me1) and unmethylated (Me0) forms of the peptide ARMC6₂₄₈₋₂₆₇ are shown (full XICs are shown in Supplementary Fig. 3). **c** Fluorography showing in vitro activity of GST-METTL9 on WT and mutated recombinant His₆-ARMC6 (top). Ponceau S-stained membrane as loading control (bottom). The data are representative of three independent experiments. **d** HxHxH-containing sequences from human (Hs) ARMC6 and SLC39A5, bovine (Bt) NDUFB3, and mouse (Mm) S100A9 proteins with known methylated His underlined. **e** Activity of His₆-hMETTL9, using [³H]AdoMet, on a peptide array containing WT and His-to-Ala substituted (blue) peptides derived from S100A9, ARMC6 and SLC39A5. Similar results were observed in two independent experiments. **f** Activity of METTL9 on peptides (from **e**) modified (by His-to-Ala mutation; underlined) to contain a single HxH, and where "x" (blue, enlarged) was varied. Mean ± s.d (*n* = 6; i.e. 2 independent experiments for each of the three different peptides). Top, bar graph colored by residue type – small (A, G, S, T; purple), large polar (N, Q; pink), large nonpolar (I, L, M, V; blue), charged (R, D, E, K; brown) and cyclic (H, F, P, Y; green). Bottom, image of one of the two peptide arrays used to generate the bar graph. Note that the original image, which can be found in Supplementary Fig. 5a, has been rearranged so that each amino acid represents a distinct column. **g** Multiple methylation of an SLC39A7-derived fragment by METTL9. GST-SLC39A7₃₁₋₁₃₇ was incubated with recombinant WT or mutant His₆-hMETTL9 in the presence of AdoMet, and digested with chymotrypsin. Normalized LC-MS XICs corresponding to different methylation states of the peptide SLC39A7₅₅₋₇₂ are shown (full XICs are found in Supplementary Fig. 6). Source data are provided as a Source Data file.

found that such motifs were particularly abundant in a number of zinc transporters (SLC39A and SLC30A proteins), and notably, several of these proteins had previously been reported as METTL9 interactants[20–22]. We reasoned that HxHxH could be a recognition sequence for METTL9, and that the recombinant enzyme may be active on peptides containing this motif. To investigate this, peptides were synthesized on an array, followed by incubation with [³H]AdoMet and METTL9, and methylation was detected by fluorography. For these experiments, we used hexahistidine-tagged METTL9 (His₆-hMETTL9) which showed

considerably higher activity than GST-hMETTL9, but was unsuitable for fluorography experiments due to automethylation of the tag. Indeed, we found that His₆-hMETTL9 was able to methylate peptides derived from ARMC6, S100A9, and the zinc transporter SLC39A5 (Fig. 1e) (but no activity was observed on an NDUFB3-derived peptide). To further explore the sequence requirements for METTL9-mediated methylation, we analyzed various substituted peptides. Alanine substitutions of single histidines in the HxHxH motifs revealed that METTL9-mediated methylation was abolished on complete removal of both HxH

motifs via substitution of the middle histidine (HxAxH) (Fig. 1e). Methylation was observed for all the peptides containing a single intact HxH (AxHxH or HxHxA), albeit in some cases at a reduced level compared with the corresponding HxHxH peptide. This indicates that a single HxH is the minimal sequence motif for METTL9-mediated methylation.

Based on the above results, we set out to investigate how the identity of the middle residue, x, as well as the N- and C-flanking residues, denoted x_N and x_C, influence methylation of the METTL9 recognition sequence (x_NHxHx_C). We investigated all three test sequences from above (Fig. 1e), modified to contain only a single instance of HxH (Fig. 1f, Supplementary Fig. 5a), and observed some clear trends. For the x_N and x_C positions, most amino acid substitutions were tolerated, but Pro and Val largely abolished methylation, and Ile or Glu at the x_N position strongly reduced methylation (Supplementary Fig. 5b, c). For the middle position, x, replacement with aromatic or charged residues had a drastic negative effect on methylation, whereas the highest activity was observed with the small and uncharged residues Ala, Gly, Thr, Ser, and Asn (Fig. 1f, Supplementary Fig. 5a), very much in agreement with the presence of such residues in the identified METTL9 substrates. Based on this sequence preference (A,N,G,S,T), we refer, in the following, to histidines alternating with such residues as ANGST HxH motifs.

NDUFB3 was reported to contain three methylated histidines within the sequence HGHGHEH (Fig. 1d)[14], and we therefore set out to explore the ability of METTL9 to introduce multiple methylations in vitro. As we were unable to detect METTL9-mediated methylation of a corresponding NDUFB3-derived peptide in vitro, we instead used as substrate a sequence derived from a histidine-rich loop (amino acids 31–137) of the zinc transporter SLC39A7, since it contains a high number of ANGST HxH motifs. This sequence (as a recombinant GST fusion protein, GST-SLC39A7$_{31-137}$) was incubated with recombinant His$_6$-hMETTL9 and AdoMet, then digested with chymotrypsin, and the methylation status of the resulting peptides analyzed by MS. In general, we found the MS analysis of histidine-rich METTL9 substrates challenging, but for one of the tryptic fragments (amino acids 55–72) which contained seven alternating histidines forming six ANGST HxH motifs, extensive methylation was detected. We observed up to five methylations, with an average of ~3 methylations (Fig. 1g, Supplementary Fig. 6), as well as a considerable pool of unmethylated substrate, possibly resulting from suboptimal activity of the recombinant enzyme produced in E. coli. The above results clearly establish the ANGST HxH motif as a target sequence for METTL9, and demonstrate that METTL9 can introduce multiple methylations in stretches of alternating histidines.

METTL9 catalyzes 1MH formation in vitro and in vivo. Methylation of histidine can occur on the proximal or the distal nitrogen of its imidazole ring, leading to the formation of 1MH or 3MH, respectively (Fig. 2a). Thus, we set out to determine which isomer is generated by METTL9. We methylated three different HxHxH-containing peptides with recombinant His$_6$-hMETTL9 in vitro, followed by acid hydrolysis and analysis of the resulting amino acids by liquid chromatography coupled to MS (LC-MS). Indeed, 1MH was formed in all three peptides, and no 3MH was detected (Fig. 2b and Supplementary Fig. 7), in agreement with the reported presence of 1MH in the mouse S100A9 protein[13].

We reasoned that METTL9, being a broad-specificity histidine MTase, could be responsible for a substantial portion of the 1MH in the human proteome. Thus, we examined whether the total content of 1MH in cellular proteins was affected by the disruption of the METTL9 gene. We collected protein extracts from WT and

METTL9 KO human HAP1 and HEK293T cells, as well as from various tissues and embryonic fibroblasts derived from WT and Mettl9 KO mice, and analyzed the amount of methylhistidine, relative to total histidine, by LC-MS/MS. Interestingly, whereas 3MH was unaffected, 1MH was reduced by ~50% in the human METTL9 KO cells, and showed an even more notable reduction in the Mettl9 KO mouse fibroblasts (Fig. 2c, d). When WT METTL9 was reintroduced into the KO HAP1 and HEK293T cells, the 1MH content was restored back up to, or even above, the WT levels (Fig. 2c). Importantly, no such effect was observed with enzymatically inactive METTL9, demonstrating that the enzymatic activity of METTL9 is required for 1MH formation in vivo. We also found a remarkable reduction of proteinaceous 1MH, but not 3MH, in various tissues from Mettl9 KO mice (relative to WT), although the amount of METTL9-dependent 1MH varied considerably between the tissues (Fig. 2d). Taken together, this indicates that METTL9-mediated methylation occurs in a wide range of mammalian tissues, and that METTL9 is the major enzyme generating 1MH in mouse and human proteins.

Defining a catalog of METTL9 targets. To identify additional METTL9 substrates, we employed a proteomics approach involving the AdoMet analogue propargylic Se-adenosyl-L-selenomethionine (ProSeAM), which allows MTase-mediated transfer to relevant substrates of an alkyne moiety that can be conjugated to biotin by click chemistry[23]. Thus, we incubated protein extracts from Mettl9 KO mouse embryonic fibroblasts (MEFs) with mouse METTL9 (mMETTL9) in the presence of ProSeAM, and pulled down biotinylated proteins with streptavidin (Fig. 3a). To allow specific identification of METTL9 targets, we used a SILAC approach where proteins pulled down from a METTL9-treated heavy isotope-labeled extract were compared with those pulled down from an untreated light isotope-labeled extract (Fig. 3a). Two replicate experiments were performed, and the eight proteins identified in Experiment 1 represented a subset of the 16 proteins identified in Experiment 2. (Fig. 3b). Intriguingly, while an ANGST HxH motif is found in only ~11% of all mouse proteins (number obtained by using ScanProsite[24]), the eight proteins that were identified in both experiments all contain such a motif, and this was also the case for two of the eight additional proteins identified in Experiment 2 only. Reassuringly, NDUFB3 was among the eight proteins identified in both replicates, and so were zinc transporters from the SLC30A and SLC39A families. Taken together, these experiments identified new candidate METTL9 substrates, and further corroborated METTL9 as a broad-specificity enzyme with a preference for ANGST HxH motifs.

To further investigate candidate METTL9 substrates, corresponding peptide sequences were spotted on an array and His$_6$-hMETTL9-mediated methylation assessed using [^{3}H]AdoMet and fluorography. 56 peptides were investigated, representing (ANGST and non-ANGST) HxH-containing sequences from the following categories of proteins: (1) hits from the ProSeAM experiments; (2) METTL9 interactants[20–22]; (3) other proteins, such as reported and putative His-methylated proteins, and proteins of particular biological interest (Fig. 3c, Supplementary Table 1). As negative controls were included corresponding peptides with HxH motifs disrupted by His-to-Ala replacements. About half of the peptides were methylated, to varying extents, by METTL9, and the control peptides were not methylated in any case (Fig. 3d, e, Supplementary Fig. 8). The five peptides that showed the strongest activity stem from proteins involved in a wide range of cellular processes, i.e. chaperone DNAJB12, unconventional myosin MYO18A, two zinc transporters
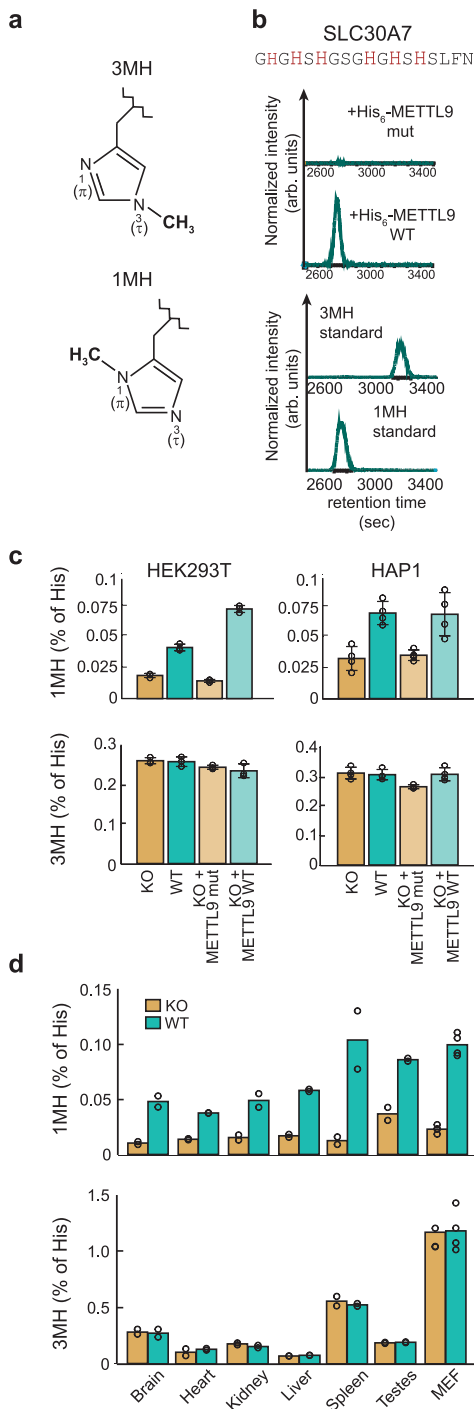
Fig. 2 METTL9 generates pervasive 1-methylhistidine in vitro and in vivo. **a** 3-Methylhistidine (3MH, top) and 1-methylhistidine (1MH, bottom). Polypeptide backbone indicated by jagged line. **b** Methylhistidine analysis of METTL9-treated SLC30A7 peptide. Shown are normalized LC/MS chromatograms of His$_6$-hMETTL9-treated samples subjected to acid hydrolysis, as well as of methylhistidine standards. **c** Proteinaceous 3MH and 1MH in HEK293T and HAP1 cells. WT (teal bars), *METTL9* KO (KO, orange bars), KO complemented with WT (KO + METTL9 WT, light teal) or mutant METTL9 (KO + METTL9 mut, light orange). Note that for HAP1 cells, the complementation is with hMETTL9-3xFLAG and the inactivating mutation is E174A, whereas for HEK293 cells, the complementation is with mMETTL9-HA and the inactivating mutation is D151K/G153R. Mean ± s.d ($n = 3$ (HEK293) or 4 (HAP1) biologically independent samples). **d** Proteinaceous 3MH and 1MH in WT or *Mettl9* KO mouse tissues and MEF cells. Mean with individual values of duplicates (tissues) or triplicates (MEF). Source data are provided as a Source Data file.

we generated the E. coli-expressed recombinant GST-tagged protein, as well as corresponding mutants where the relevant HxH motifs had been disrupted. Similar to what we already observed for His$_6$-ARMC6, all three proteins were methylated by GST-hMETTL9 in vitro, assessed using [^{3}H]AdoMet and fluorography (Fig. 4a). Importantly, methylation was abolished by the HxH-disrupting mutations in all cases (Fig. 4a). In conclusion, we have identified a total of 30 different human proteins that are targeted by METTL9 in vitro in the context of peptides and/or full-length recombinant proteins (summarized in Table 1).

Identification of in vivo substrates of METTL9. Prior to our work, HxH sequences from NDUFB3 and S100A9 were already reported to be histidine-methylated in vivo, and we have here shown that these sequences are methylated by METTL9 in vitro. Taken together with the high levels of METTL9-dependent 1MH in mouse and human proteomes, this strongly indicates that extensive METTL9-mediated methylation of HxH motifs occurs in vivo. To firmly establish this, we set out to directly demonstrate cellular METTL9-mediated methylation of specific in vitro substrates.

We generally found it challenging to assess the methylation status of the histidine-rich METTL9 substrates by protein MS, and relatively large amounts of pure material were required. However, we were able to isolate S100A9 from mouse peritoneal exudate neutrophils and, by matrix-assisted laser desorption-ionization (MALDI) MS, assess the methylation status of a peptide corresponding to the METTL9 target sequence (Fig. 4b, Supplementary Fig. 9a). Importantly, we found that this HGHSH-containing (Fig. 1d) peptide was completely unmethylated in *Mettl9* KO neutrophils, whereas it appeared exclusively in the mono- and dimethylated forms in WT neutrophils (Fig. 4b, Supplementary Fig. 9a). In the case of NDUFB3, over-expression of the FLAG-tagged protein in HEK293T cells enabled us to assess its in vivo methylation status. Importantly, the peptide corresponding to the N-terminal HxH-rich portion of the protein appeared as a mixture of mono-, di- and non-methylated forms in WT cells, whereas it was found exclusively in the unmethylated state in *METTL9* KO cells (Fig. 4c, Supplementary Fig. 9b). These results firmly establish that METTL9-mediated methylation of HxH-motifs occurs in vivo.

To obtain evidence for in vivo methylation of additional METTL9 substrates, we also attempted both targeted proteomics (using parallel reaction monitoring) on total cellular protein and

SLC30A1 and SLC39A6, and cyclin CCNT1, and they show a large variation in sequence and the number of HxH motifs present (from 1 to 4) (Fig. 3f). However, they all contained at least one ANGST HxH motif, whereas most of the peptides that were not methylated lacked such motifs (Fig. 3d and Supplementary Table 1).

Since we had mainly used peptide substrates when studying the in vitro activity of METTL9, we also set out to study methylation in the context of corresponding recombinant proteins. This was particularly relevant in the case of NDUFB3, for which HxH-methylation was reported in vivo, but where we failed to observe METTL9-mediated methylation of the corresponding peptide. For three proteins, i.e. NDUFB3, as well as CCNT1 and DNAJB12, both of which were efficiently methylated as peptides,
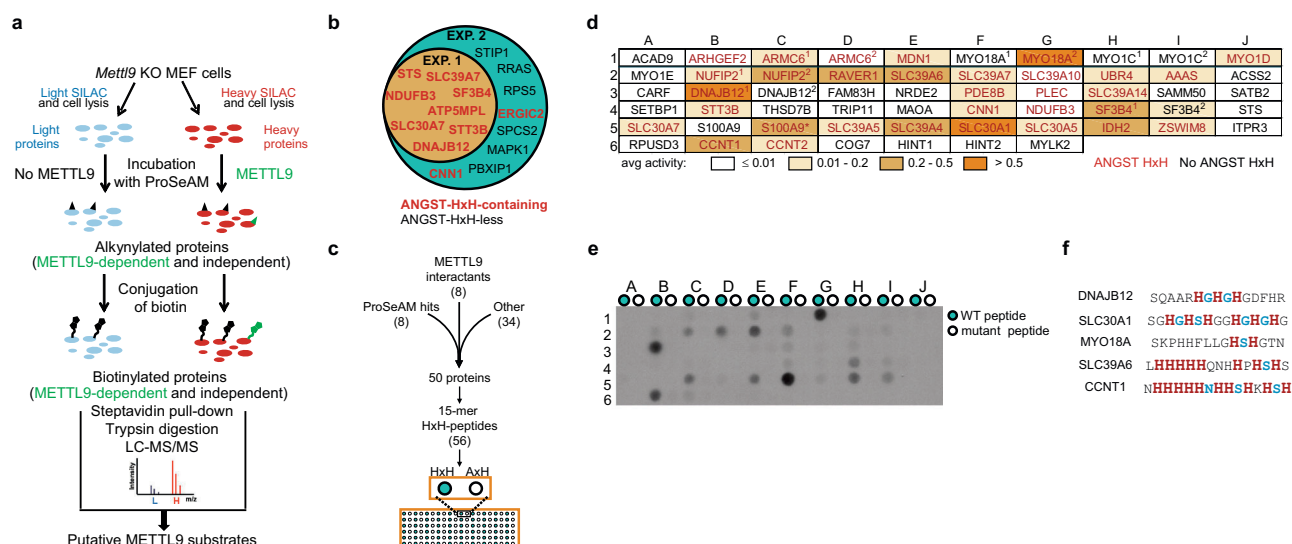
Fig. 3 Identification of new cellular and peptide substrates of METTL9. a Outline of the ProSeAM (propargylic Se-adenosyl-L-selenomethionine) assay for identification of novel METTL9 substrates in cell extracts. **b** Venn diagram of mMETTL9 substrates identified in the two ProSeAM experiments. HxH-containing proteins are in red. **c** Schematic of array design for METTL9 substrate assessment. **d** Relative average MTase activity (from two independent arrays) presented in shades of orange, with gene names of relevant peptides indicated (ANGST HxH-containing in red). Superscripts denote multiple peptides derived from the same protein. All peptides are derived from human proteins, except mouse S100A9 (asterisk). Note that the array also contained an additional, HxH-less, peptide, based on the previously reported histidine-methylated sequence from MYLK2. **e** Representative result of METTL9-mediated methylation of a peptide array designed as in **c** and with each coordinate (e.g. A1) representing WT (HxH-containing; teal) and control (HxH-less; white) versions of a unique sequence (see also Supplementary Table 1). The experiment was performed twice with similar results. **f** Sequences of the top five METTL9 substrates from the peptide array, indicating ANGST residues (blue) within HxH motifs. Source data are provided as a Source Data file.

purification of select substrates by immunoprecipitation, but unfortunately these efforts were largely unsuccessful. However, one exception was DNAJB12, which was identified as a METTL9 target in both ProSeAM experiments, and found to be an excellent in vitro substrate (Fig. 3, Fig. 4a). We detected dimethylation of the relevant HSHSH-sequence in DNAJB12 immunoprecipitated from WT HEK293 cells (Supplementary Fig. 9c, d), but were, unfortunately, unable to detect the corresponding peptide in the KO cells. Still, this result suggests that DNAJB12 is an in vivo target of METTL9.

As we showed that *METTL9* KO substantially diminished global 1MH levels in mammalian proteomes (Fig. 2c, d), we also set out to directly demonstrate the presence of METTL9-dependent 1MH in a single cellular protein. For this purpose, we selected the zinc transporter SLC39A7, which carries a high number (32) of ANGST HxH motifs, and is relatively abundant. SLC39A7 was immunoprecipitated from HEK293T cells, and then subjected to acid hydrolysis and amino acid analysis. Strikingly, 4.4% of total histidine was found as 1MH in the immunoprecipitated SLC39A7 (Fig. 4d), i.e. a level ~100 fold higher than that found in the total HEK293T proteome, and corresponding, on average, to ~2.5 1MH modifications per SLC39A7, which contains 57 histidines. Importantly, 1MH in SLC39A7 was virtually absent in *METTL9* KO cells, but complementation of the KO cells with enzymatically active mMETTL9 restored 1MH to levels even higher than those observed in SLC39A7 from WT cells (Fig. 4d).

The above results directly demonstrate METTL9-dependent in vivo methylation for three proteins, namely S100A9, NDUFB3, and SLC39A7. In addition, we showed that METTL9 target sequences in ARMC6 and DNAJB12 are methylated in vivo, and high-throughput analysis of histidine methylation provided similar evidence for CCNT2 (shown to be a peptide substrate in Fig. 3d, e) and SLC39A7[16]. In summary, our data provide clear evidence that METTL9 is a broad-specificity MTase that

introduces 1MH at ANGST HxH motifs in numerous mammalian proteins, and the underlying in vitro (MTase assays) and cellular (MS detection of modification) findings have been summarized in Table 1.

Effects of METTL9-mediated methylation on mitochondrial respiration and zinc binding. NDUFB3 is an accessory subunit of the mitochondrial Complex I, and contains several methyl-histidines within a stretch of alternating histidines in its N-terminal region (Fig. 5a)[14]. We found that the presence of alternating histidines close to the N-terminus was a conserved feature of vertebrate NDUFB3, and that the number of HxH motifs showed an interesting variation between different organisms (Fig. 5a), indicating that methylation of these alternating histidines is functionally important. Thus, we investigated mitochondrial respiration in isolated mitochondria from HAP1 and in permeabilized HEK293T cells; in both cases ablation of METTL9 led to a reduction in oxygen consumption (Supplementary Fig. 10). The difference was particularly notable for so-called State 3 respiration, i.e. when respiration fuels ATP synthesis. A difference was only observed when respiration was driven by Complex I (of which NDUFB3 is a component), but not via Complex II (Supplementary Fig. 10). In addition, when WT, but not mutant, mMETTL9 was reintroduced into the KO HEK293T cells, respiration was restored to WT levels, showing that METTL9 enzymatic activity is required for optimal Complex I function (Fig. 5b). Taken together, these results indicate that METTL9-mediated methylation of NDUFB3 enhances Complex I-mediated mitochondrial respiration.

Many of the His residues targeted by METTL9 are involved in the binding of zinc and other metals (e.g. in S100A9 and zinc transporters)[13,25], suggesting that methylation may modulate metal binding to histidines. To investigate this, we used isothermal titration calorimetry (ITC) to measure zinc binding
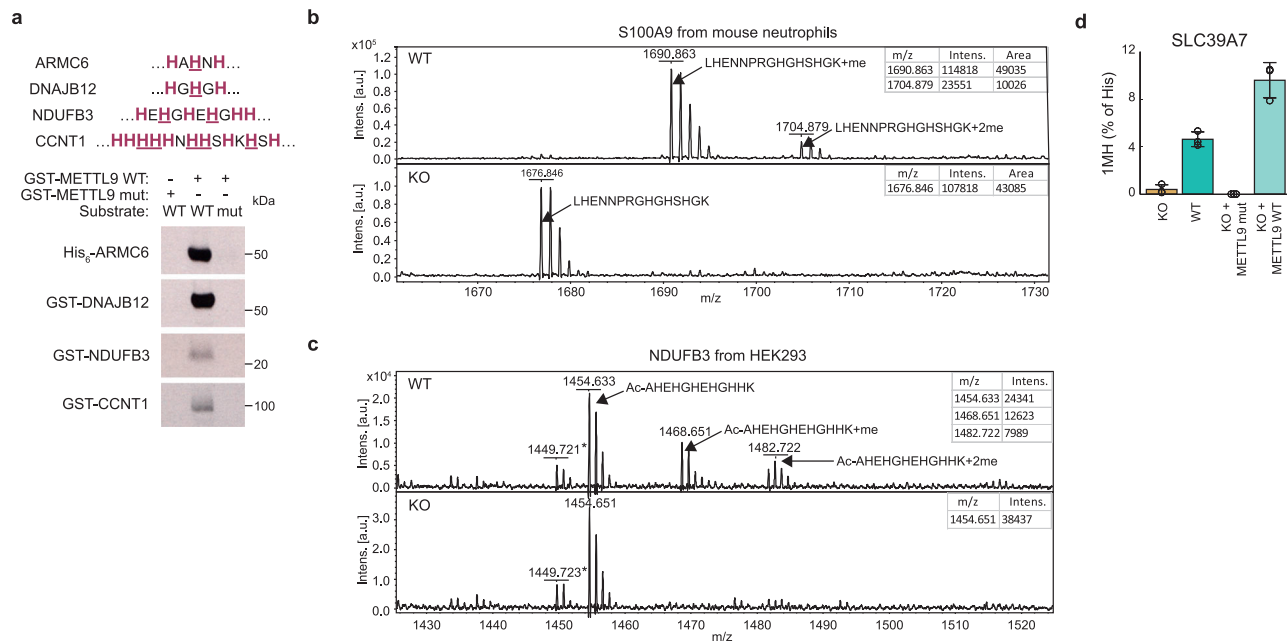
Fig. 4 Confirmation of METTL9-dependent protein histidine methylation in vitro and in vivo. a HxH-containing sequences of protein substrates with histidines marked in red and histidines mutated to arginines in the mutant substrates underlined (top). Fluorography showing in vitro activity of GST-hMETTL9 WT or E174A mutant (METTL9 mut) on the indicated proteins, either WT or with all HxH motifs removed by His-to-Arg mutations (mut) (bottom). 4 days exposure for ARMC6 and DNAJB12, 10 days exposure for CCNT1 and NDUFB3. **b** MALDI MS analysis of S100A9-derived peptides from WT and *Mettl9* KO mouse neutrophils. **c** MALDI MS analysis of peptides from NDUFB3-FLAG immunoprecipitated from WT and *METTL9* KO HEK293T cells. **d** Cellular SLC39A7 contains 1MH. The 1MH content of SLC39A7 immunoprecipitated from various HEK293-derived cells was determined by amino acid analysis. The following cells were used: WT, *METTL9* KO (KO), KO complemented with WT C-terminally HA-tagged mMETTL9 (KO + METTL9 WT) or KO complemented with the corresponding D151K/G153R mutant (KO + METTL9 mut). Mean ± s.d (*n* = 3 biologically independent samples). Source data are provided as a Source Data file. For the MALDI MS experiments (**b** and **c**), a table showing theoretical vs experimental mass of relevant peaks (and ppm deviation), as well as MASCOT results showing peptide coverage, can be found as Supplementary Data 1 and 2, respectively.

Table 1 Identified METTL9 substrates.

Substrate	In vitro activity		Observed in vivo
	On rec. protein	On peptide	
AAAS		X	
ARMC6	X	X	X[a]
CCNT1	X	X	
CCNT2		X	X[16]
CNN1		X	
DNAJB12	X	X	X[a]
IDH2		X	
MDN1		X	
MYO18A		X	
MYO1D		X	
NDUFB3	X	X	X[14,a]
NUFIP2		X	
PDE8B		X	
RAVER1		X	
S100A9[b]		X	X[13,a,c]
SF3B4		X	
SLC30A1		X	
SLC30A5		X	
SLC30A7		X	
SLC39A14		X	
SLC39A4		X	
SLC39A5		X	
SLC39A6		X	
SLC39A7		X	X[16,a,c]
STT3B		X	
UBR4		X	
ZSWIM8		X	

[a]In the present study.
[b]Mouse protein.
[c]METTL9-dependent, on endogenous protein.

to an SLC39A7-derived peptide containing six consecutive ANGST HxH motifs, as well as to a fully methylated (except the first His, in accordance with the presumed product specificity of METTL9; see "Discussion") version of this peptide. Interestingly, methylation significantly reduced zinc binding (Fig. 5c, Supplementary Fig. 11), suggesting that METTL9-mediated methylation modulates the binding of zinc and other metals to its target proteins.

Discussion

Our results define METTL9 as a protein 1MH MTase. We show that METTL9 methylates ANGST HxH motifs in numerous proteins/peptides, and that METTL9-mediated 1MH formation is abundant in mammalian cells and tissues. Moreover, we find that NDUFB3, a component of mitochondrial Complex I, is methylated by METTL9, and that the enzymatic activity of METTL9 promotes Complex I-mediated respiration.

Strikingly, we observed a substantial reduction in 1MH content in total protein from all nine *METTL9* KO cells/tissues tested (six tissues and one MEF cell line from mice, as well as two human cell lines), and METTL9-dependent 1MH was found to account for up to ~0.1% of total histidine (Fig. 2c, d). Based on this, one may perform a simple calculation to further illustrate the potential extent of METTL9-mediated 1MH. The human proteome consists of ~10 million residues (~20,500 unique proteins with an average size of ~460 residues), and ~2.5% of these are histidines[26,27]. Assuming that at least half the proteome (10,000 proteins) is expressed in a typical human cell or tissue, this corresponds to ~125,000 histidine residues[28]. If one also assumes that the frequency of METTL9-mediated 1MH is independent of
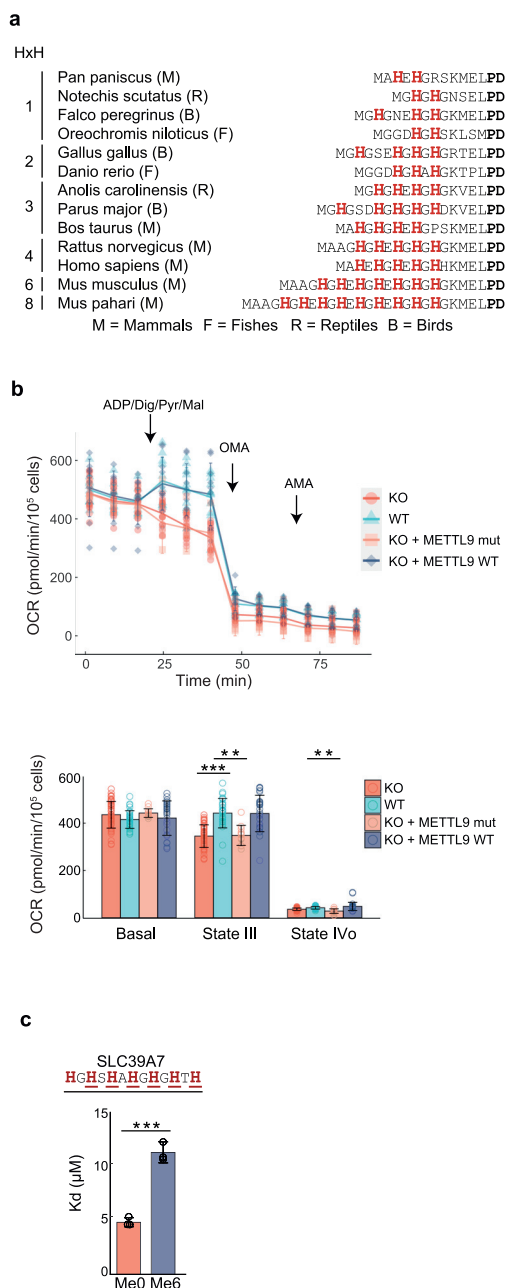
a

b

c

Fig. 5 METTL9-mediated methylation affects mitochondrial respiration and reduces the binding affinity of a zinc transporter peptide to Zn^{2+}. a Sequence alignment of the N-terminal end of NDUFB3 orthologues from various vertebrate species. Histidines are shown in red and the number of HxH indicated. Residues "PD" (in black, bold) are conserved in all vertebrates. **b** Oxygen consumption rate (OCR) of digitonin (Dig)-permeabilized WT (blue), *METTL9* KO (KO, red), *METTL9* KO complemented with WT (KO + METTL9 WT, purple) or D151K/G153R mutant mMETTL9-HA (KO + METTL9 mut, orange) HEK293 cells. Respiration mediated by Complex I was assessed under basal conditions, and after sequential addition of ADP/pyruvate (Pyr)/malate (Mal), oligomycin (OMA), and antimycin A (AMA). Shown are OCR traces (top) and quantification of respiratory states (bottom). Mean ± s.d. ($n = 17$ (KO), 11 (WT), 10 (KO + METTL9 WT), or 4 (KO + METTL9 mut) biologically independent samples). ***$P < 0.001$; **$P < 0.005$ (state III: $P = 0.00372$, state IV: $P = 0.00157$); two-sided Dunnett's multiple comparison test. **c** Dissociation constant (K_d) between Zn^{2+} and a synthetic SLC39A7 peptide that contains either none (Me0) or six 1MH (Me6) residues in an HxH context (methylated histidines underlined). Mean ± s.d. ($n = 3$ independent experiments). ***$P = 0.000451$; two-tailed Student's *t* test. Source data are provided as a Source Data file.

methylation in all tested substrates (28 peptides and 4 recombinant proteins). Moreover, we established, using three different peptide substrates, that a small residue (ANGST) is preferred as the middle residue, x, and this is also supported by other observations. First, ANGST HxH-motifs were present in all the proteins identified as METTL9 substrates in both of the two ProSeAM experiments (Fig. 3b), whereas this motif is found in only ~11% of mouse/human proteins[24]. Second, when HxH-containing candidate sequences were tested as METTL9 substrates (Fig. 3c–e) activity was observed for 27 out of the 32 ANGST HxH-containing peptides, but merely for one out of the 24 sequences that exclusively contained non-ANGST HxH motifs.

We demonstrated that an ANGST HxH motif is sufficient for METTL9-mediated methylation in most, but not all sequence contexts, as we failed to observe in vitro methylation of some (5 out of 32) of the tested ANGST HxH-containing peptide sequences. This was further corroborated by the observation that substitution of residues flanking the HxH motif in some cases abolished methylation (Supplementary Fig. 5). However, this flanking residue effect varied somewhat between the three sequences tested, and not all of the 5 ANGST HxH peptides that failed to be methylated contained a non-compatible flanking residue. Thus, additional sequence features also influence METTL9-mediated methylation of ANGST HxH motifs. In one case, NDUFB3, we observed METTL9-dependent methylation of the full-length protein both in vitro and in cells, but no in vitro methylation of the corresponding peptide. Conceivably, a structuring of the relevant NDUFB3 sequence, e.g. through interaction with other residues, may promote methylation in the context of the full-length protein.

We found that an HxH motif was invariably present in all METTL9 substrates, but many of the identified METTL9 substrates (e.g. ARMC6, S100A9, and NDUFB3) contained several consecutive HxH-motifs, typically HxHxH, and mutating an HxHxH-motif into HxH severely diminished methylation in several cases (Fig. 1e). This suggests that a stretch of more than one HxH may be a biologically more relevant substrate for METTL9. Interestingly, we have not observed in our MS/MS data the methylation of the first His residue in a stretch of alternating histidines, in agreement with previously reported methylation patterns of NDUFB3 and S100A9 (Fig. 1d)[13,14]. This may suggest that METTL9 can only methylate the second

protein abundance, one can calculate that a 1MH frequency of 0.1% corresponds to ~125 stoichiometrically (100%) 1MH-modified histidine residues, and potentially a much higher number of sub-stoichiometrically modified residues. This agrees well with METTL9's relaxed sequence specificity (the human proteome contains 2807 instances of ANGST HxH[24]), and its ability to methylate the majority (~80%) of tested ANGST HxH-containing peptide substrates in vitro (Fig. 3d, e). So far, in vivo methylation of ANGST HxH motifs has been reported, by us and others, for six mammalian proteins (S100A9, NDUFB3, ARMC6, SLC39A7, DNAJB12, and CCNT2; Table 1) and we formally demonstrated METTL9-dependence for three of these (NDUFB3, S100A9, and SLC39A7). Future studies will likely focus on expanding the catalog of the METTL9's in vivo substrates.

Our work has firmly established that an HxH motif is an absolute requirement for METTL9-mediated methylation, as we found that disruption of the HxH motif invariably abrogated

histidine of the HxH motif, i.e. that all but the first residue in a stretch of alternating histidines have the potential to become methylated. In this context, zinc transporters of the SLC30A and SLC39A families are a very interesting group of METTL9 substrates. They are responsible for transporting zinc and other metals across cellular membranes, but many of the 23 human proteins remain poorly characterized[25]. They contain a high number of metal-binding, ANGST-enriched alternating histidines: eleven of these proteins collectively account for 24% of the instances of ANGST HxHxH-motifs in the human proteome (but for only 6.5% of total HxHxH[24]). Interestingly, the METTL9-targeted residues in S100A9 are also involved in metal coordination[13]. Indeed, we showed that 1MH modification of a SLC39A7-derived stretch of alternating histidines caused a reduction in zinc binding (Fig. 5c), and it may be of interest to further investigate the potential role of METTL9 in regulating metal ion homeostasis.

The *Mettl9* KO mice were without apparent phenotype, and this somewhat contrasts with the dramatic reduction in 1MH observed proteome-wide. However, it is not unprecedented that a KO mouse devoid of a modification enzyme shows a strong molecular phenotype but no other, easily observable, phenotype. For example, mouse KO of the tRNA modification enzyme Alkbh8 caused mismodification of several distinct tRNA iso-acceptors at the critical wobble uridine position, but gave no overt phenotype[29,30]. However, later studies showed defects in seleno-protein synthesis and stress tolerance in the *Alkbh8* KO mice[31,32], and it was recently reported that homozygous inactivating mutations in the human *ALKBH8* gene causes intellectual dis-ability and general developmental delay[33]. Likewise, more detailed studies, e.g. on metal biology and mitochondrial meta-bolism, will be required to elucidate the biological function of METTL9.

Our in vitro methylation experiments (Fig. 1a, Supplementary Fig. 2a, c) demonstrated considerable methylation by recombi-nant METTL9 in WT extracts, indicating substantial hypo-methylation of METTL9 substrates. This observation indicates that, rather than being static, METTL9-mediated methylation is dynamic and of potential regulatory importance. It would therefore be intriguing to explore the dynamics of methylation, and whether it can be reversed or modulated by specific deme-thylases and reader proteins, in a manner analogous to lysine methylation on histones[34,35]. Moreover, we also observed sub-stantial METTL9-independent 1MH in all examined cells and tissues, and the corresponding MTases remain to be identified, *e. g.* the one introducing 1MH (in a non-HxH context) in MYLK2.

Note: Since no other enzymes have previously been reported to generate 1MH in proteins, METTL9 cannot readily be given any EC (Enzyme Commission) number, see e.g. the ENZYME data-base[36]. "EC 2.1.1.85 Protein-histidine N-methyltransferase" already exists, but only encompasses enzymes that generate 3MH. Thus, we recommend that a new entry is made for "EC 2.1.1.xxx Protein-histidine N1-methyltransferase", and that EC 2.1.1.85, also for the future, is restricted to protein-histidine N3-methyltransferases.

Methods

Cloning and mutagenesis. ORFs were amplified with Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) from HeLa or HEK293 cDNA. Primers were designed to generate PCR products flanked by sequences com-patible with the appropriate vectors for In-Fusion Cloning (Takara Bio). Mutagenesis was performed using mutagenic primers designed with the PrimerX program and the QuickChangeII Site-directed mutagenesis kit (Agilent) or by overlap extension PCR. All constructs were sequence-verified. Human NDUFB3 cDNA was amplified from a HEK293T cDNA library, and the full length DNA fragment was inserted into either pQCXIH or pcDNA3 vector with C-terminal FLAG tag.

Mouse Mettl9 cDNA was amplified from Fantom clone (AK075586), and the full length DNA fragment was inserted into pcDNA3 vector with C-terminal FLAG or HA tag (pcDNA3-mouse (m)METTL9-FLAG, pcDNA3-mMETTL9-HA). For baculovirus expression system, N-terminal truncated mMETTL9 (residues 22–318) was cloned into the pFastBac HT vector, which encodes a tobacco etch virus (TEV) protease recognition sequence inserted after an N-terminal polyhistidine tag. The oligonucleotides used for cloning, mutagenesis, and for CRISPR-Cas9-mediated gene knock-out are listed in Supplementary Table 2.

Expression and purification of recombinant proteins. GST-tagged proteins GST-hMETTL9, GST-hMETTL9 E174A, GST-SLC39A7 (31-137aa), GST-DNAJB12 (1-243 aa), GST-DNAJB12 H185R, GST-NDUFB3, GST-NDUFB3 H5R/H9R, GST-CCNT1 and GST-CCNT1 H519R/H520R/H523R/H524R/H528R were expressed in *Escherichia coli* and purified using Glutathione Sepharose 4B (Sigma-Aldrich). His$_6$-tagged proteins His$_6$-hMETTL9, His$_6$-hMETTL9 E174A, His$_6$-ARMC6, His$_6$-ARMC6 H261R, His$_6$-ARMC6 A262G, His$_6$-ARMC6 H263R, His$_6$-ARMC6 N264D, His$_6$-ARMC6 H265A, His$_6$-ARMC6 A266G, and His$_6$-ARMC6 K267R were expressed in *E. coli* and purified using Ni-NTA-agarose (Qiagen). In sum-mary, the appropriate pGEX-6P-2 or pET28a constructs were transformed into the chemically competent BL21- CodonPlus (DE3)-RIPL (Agilent Technologies) *E. coli* strain. The bacteria were cultured overnight, inoculated into a larger volume of TB medium with appropriate antibiotics, and after the OD600 reached ~0.8, protein expression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), at 18 °C, 250 rpm, followed by overnight growth. For His-tagged proteins, bacteria were lysed in His-Lysis Buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 5% (v/v) glycerol and 1% (v/v) Triton X-100) supplemented with 1 mg/ml lyso-zyme, 5 mM 2-mercaptoethonol, 10 mM imidazole, and 1x cOmplete (EDTA-free) Protease Inhibitor Cocktail (Roche) and purified using Ni-NTA-agarose (Qiagen), according to the manufacturer's instructions. For GST-tagged proteins, bacteria were lysed in GST-Lysis Buffer (50 mM Tris-HCl pH 7.6, 500 mM NaCl and 0.5% Triton X-100) supplemented with 1 mg/ml lysozyme, 1 mM DTT, and 1x cOm-plete Protease Inhibitor Cocktail (Roche) and purified using Glutathione-Sepharose 4 B (Sigma-Aldrich). Eluted proteins were buffer exchanged to Storage Buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5% glycerol and 1 mM DTT) using cen-trifugal concentrators with a molecular weight cut-off of 10 kDa (Sartorius) and stored at −80 °C.

For the ProSeAM assay, mMETTL9 was expressed in baculovirus-infected Sf9 insect cells. Baculovirus was generated using the Bac-to-Bac baculovirus expression system (Invitrogen) according to the manufacturer's instruction. The baculovirus-infected cells were harvested 48 h after transfection, and cell pellets were stored at −80 °C until purification. Frozen Sf9 cells were resuspended in 20 mM Tris-HCl buffer (pH 8.0) containing 500 mM NaCl, 0.1% NP-40, 10% glycerol, 20 mM imidazole, DNase I (Sigma Aldrich) and cOmplete (EDTA-free) Protease Inhibitor Cocktail (Roche). Cells were lysed by sonication and clarified by centrifugation. The cell lysate was loaded onto a HisTrap HP column (GE Healthcare), and eluted with 50 mM Tris-HCl buffer (pH 8.0) containing 500 mM NaCl, 10% glycerol and 500 mM imidazole. After exchanging the buffer using a HiTrap Desalting column (GE Healthcare), the N-terminal polyhistidine tag was cleaved by incubating with TEV protease overnight at 4 °C. Then, the cleaved mMETTL9 protein was reapplied to a HisTrap HP column, and the flow-through fraction was collected. The mMETTL9 protein was further purified by size-exclusion column chromatography using a HiLoad Superdex 200 16/60 (GE Healthcare), and then the buffer was exchanged to 20 mM Tris-HCl buffer (pH 8.0) containing 150 mM NaCl and 20% glycerol using a HiTrap Desalting column. Purified protein was concentrated using an Amicon Ultra-15 centrifugal filter unit (Millipore, 10 kDa MWCO) to 0.4 mg/ml and flash frozen in liquid nitrogen.

In vitro methyltransferase assays on recombinant proteins and protein extracts. In vitro methyltransferase assays were performed using Reaction Buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM EDTA) with 0.64 μM ^{3}H-AdoMet (Perkin–Elmer, specific activity = ~77–78 Ci/mmol) or 1 mM unlabeled AdoMet for MS analysis of samples. For methylation of recombinant His$_6$ ARMC6 WT and mutant proteins or the GST-SLC39A7 (aa 31–137) fragment, reactions con-tained 10 μg substrates and 1 μM WT or E174A mutant His$_6$-hMETTL9. For methylation of subcellular protein extracts, 100 μg protein from HAP1 fractions isolated using the Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Scientific) were used as substrates, with 1 μM WT or E174A mutant GST-hMETTL9 enzymes. For the recombinant substrate panel, 2 μM His$_6$-ARMC6 WT and H263R mutant, GST-DNAJB12(1-243 aa) WT and H185R mutant, or GST-NDUFB3 WT and H5R/H9R mutant, or 1 μM GST-CCNT1 WT and H519R/H520R/H523R/H524R/H528R mutant were used as substrates with 1 μM GST-hMETTL9 WT or E174A mutant enzyme. Reactions were incubated at 36 °C for 1 h, and stopped by the addition of NuPAGE loading buffer. Proteins were separated by SDS-PAGE. For fluorographic analysis, the proteins were then transferred to PVDF membranes, stained with Ponceau S, dried, sprayed with a scintillation solution (57.5% 2-methylnaphthalene, 40% pentylacetate, 2.5% diphenyloxazole) and exposed to Eastman Kodak Co. Bio-Max MS film (Sigma-Aldrich) at −80 °C.

Methylhistidine content analysis of cultured cells, mouse tissues and peptides. HAP1 cells, MEFs, and mouse tissues were lysed in RIPA buffer (Sigma-Aldrich), followed by precipitation with 10% (v/v) aqueous trichloroacetic acid for 20 min on ice and centrifugation ($16,000 \times g$, 15 min, 4 °C). The pellets were washed with cold isopropanol, pelleted again, and heated to 95 °C to remove residual isopropanol. The dry pellets were transferred to vacuum hydrolysis tubes (Thermo Fisher Scientific), and 200 µl Sequencing Grade 6 M HCl (Thermo Fisher Scientific) was added. The pellets were hydrolyzed under vacuum at 110 °C for 48 h. The samples were then desiccated at 60 °C and the pellet resuspended in 800 µl of H_2O. These solutions were filtered through 0.22 µm syringe filters (Millex-GP) to remove insoluble material and the liquid was again evaporated at 70 °C. The LC-MS/MS analysis to quantify methyl-histidine content was performed by BEVITAL AS.

Peptides derived from SLC30A7 (aa 163-180), CCNT1 (aa 515-531), and ARMC6 (aa 253-267) purchased from Peptide2.0 were methylated with 1 µM His_6-hMETTL9 WT or His_6-METTL9 E174A, and 1 mM AdoMet in Reaction Buffer (50 mM Tris pH 7.5, 50 mM NaCl, 5 mM EDTA). All peptides had C-terminal amides. The samples were then subjected to acid hydrolysis. Briefly, liquid samples were dried in a Speed Vacuum (Thermo Fisher) and put under N^2 in glass vials. 1 mL 6 M HCl was added to the bottom of the reaction vials and the samples were heated to 110 °C for 36 h. The samples were then placed in clean reaction vials and heated for a further 5 h, then desiccated under vacuum for 16 h to remove remaining acid. The samples were analyzed using a ACCQ TAG ULTRA C18 1.7 µM column and the Xevo-G2S LC-MS machine (Waters) and analyzed with MassLynx V4.1 software. 1-Methyl-L-histidine (Sigma-Aldrich) and 3-Methyl-L-histidine (Sigma-Aldrich) were used as 3MH and 1MH standards, respectively. Note the different nomenclature for methylhistidine used in biochemistry and chemistry due to different systems of numbering the atoms in the imidazole ring.

For methylhistidine content analysis of HEK293 cells, the cell pellets precipitated with acetone were hydrolyzed with 6 N HCl at 110 °C for 24 h, and the extracted amino acids were dissolved in 25 µL of 5 mM $HCOONH_4$/0.001% formic acid. The Extracted samples were applied to a liquid chromatograph (Vanquish UHPLC; Thermo Fisher Scientific) coupled to a triple quadrupole mass spectrometer (TSQ Vantage EMR; Thermo Fisher Scientific). The amino acids derived from the protein samples were separated on a C18 column (YMC-Triart C18, 2.0×100 mm length, 1.9 µm particle size; YMC.CO., LTD., Kyoto, Japan). Mobile phase A was comprised of 5 mM ammonium formate with 0.001% formic acid, and mobile phase B was comprised of acetonitrile. Following different slopes were used for a gradient elution at a flow rate of 0.3 mL/min: 0/0 – 1.5/0 – 2/95 – 4/95 – 4.1/0 – 7/0 (min/%B). The effluent from the column was directed to an electrospray ion source (HESI-II; Thermo Fisher Scientific) connected to the triple quadrupole mass spectrometer operating in the positive ion multiple reaction monitoring mode, and the intensity of specific MH $+ \rightarrow$ fragment ion transitions were recorded (His m/z 156.1 → 83.3; 93.2; 110.2, 1MH 170.1 → 95.3; 97.3; 109.2 and 3MH m/z 170.1 → 81.3; 83.3; 124.2). The electrospray conditions were; spray voltage of 3000 V, vaporizer temperature of 450 °C, sheath gas pressure 50 arbitrary units, auxiliary gas pressure 15 arbitrary units, and collision gas pressure 1.0 mTorr. With each batch of experimental samples a series of standard samples was simultaneously prepared and run. Calibration curves were constructed for 1MH, 3MH, and His from the data obtained from the standard samples with a range of 1 – 250 nM. The measured concentration and percentage of 1MH and 3MH in each experimental sample was calculated from the calibration curves.

SLC39A7 methylhistidine content analysis. Pelleted HEK293T cells (2 mg) were dissolved in 500 µL of lysis buffer (50 mM pH7.5 Tris-HCl, 400 mM NaCl, 0.1% NP40), and sonicated for 10 s on ice. After centrifugation ($18,000 \times g$ for 10 min), the supernatant was collected in a new tube. To immunoprecipitate SLC39A7, 1 µg of anti-ZIP7 antibody (Protein Tech, #19429-1-AP) was added to the cell lysate, and incubated for 1 h at 4 °C with a gentle agitation. For antibody alone control, 1 µg of anti-ZIP7 antibody was also added to 500 µL of the lysis buffer without cell lysates. Ten µL of Protein A/G PLUS-Agarose (Santa Cruz, SC-2003) was added in the tube, and incubated for 1 h at 4 °C with a gentle agitation. The agarose beads were washed three times with PBS. The bound proteins were separated with SDS-PAGE and transferred to a PDVF membrane. The PVDF membrane was stained with Coomassie Brilliant Blue R250, and the band corresponding to SLC39A7 (around 50 kDa) was excised. The membrane bound proteins were hydrolyzed and quantitated as described above for HEK293 cells. Each amino acid content value was subtracted with that of the antibody alone control, since molecular weight of SLC39A7 and the heavy chain of anti-ZIP7 antibody was almost the same and the excised band around 50 kDa consisted of two species.

Generation of human METTL9 KO cell lines. HAP-1 METTL9 KO cells were generated as a custom project by Horizon Genomics. The METTL9 gene was disrupted by the CRISPR-Cas9 method using a guide RNA targeted towards exon 2, resulting in an 8 bp frameshift mutation. The METTL9 KO cell line is currently commercially available at Horizon, catalog ID HZGHC004343c010. The HAP1 KO cells were complemented with C-terminally 3xFLAG-tagged WT or E174A mutant hMETTL9 by stable transfection with the appropriate p3xFLAG-CMV-14 constructs using the FuGENE6 Transfection Reagent (Roche), followed by selection in medium containing 1 mg/ml geneticin (Gibco). Individual clones were screened by

western blot for the presence of the tag using anti-FLAG antibodies (Sigma-Aldrich, F1804). WT T-Rex Flp-In 293 cell lines inducibly expressing C-terminally GFP-tagged WT METTL9 were generated according to protocol (Thermo Fisher Scientific). METTL9 KO HEK293T cells were generated by deletion of Exon 3 with the CRISPR-Cas9 system. hCas9 and two guide RNA expression vector, PX330-B/B was kindly provided by Dr. T. Hirose[37]. Two guide RNAs targeted to introns of upstream (5′-TAATAAGTGATTATGGTTGT-3′) and downstream (5′-GCAG-TATTTTTCTGGAGCGG-3′) of exon 3 in METTL9 were cloned into BbsI and BsaI site of PX330-B/B vector, respectively. The plasmid (PX330-B/B-gMETTL9) was then transfected into HEK293T cells together with pEGFP-C1. Two days after the transfection, GFP positive cells were sorted and single cells were seeded onto 96 well plate. The KO clones were screened by genomic PCR.

For the generation of HEK293T KO cells complemented with C-terminally HA-tagged WT or D151K/G153R mutant mMETTL9, plasmids for retroviral expression (pQCXIP-mMETTL9-HA or pQCXIP-mMETTL9-DKGR-HA) were transfected into retrovirus packaging cells using PEI transfection reagent (Polysciences, Inc.). 24 h after the transfection, the virus containing culture supernatant were transferred to METTL9 KO HEK293T cells with 4 µg/mL of polybrene. 24 h after the infection, 1 µg/mL puromycin were added, and the cells cultured for 2 weeks.

Generation of Mettl9 KO mice. The Mettl9 KO mice were generated by the CRISPR-Cas9 system. Guide RNAs that target exon 3 of mouse Mettl9 were screened with CRISPR design tools (http://crispr.mit.edu/). The T7 promoter sequence 5′-TTAATACGACTCACTATAGG-3′ and the 20-mer Mettl9 guide RNA sequence 5′-GAGACTGCTTAGAATTAATC-3′ were cloned into the AflII site of the gRNA Cloning Vector, a gift from George Church (Addgene #41824)19. Oligos Insert-F (5′-TTTCTTGGCTTTATATATCTTAATACGACTCACTATAGG AGACTGCTTAGAATTAATC-3′) and Insert-R (5′-GACTAGCCTTATTT-TAACTTGCTATTTCTAGCTCTAAAACGATTAATTCTAAGCAGTCTC-3′) were annealed and extended to construct a 100-bp double-stranded DNA fragment using Phusion polymerase (New England Biolabs, Japan). The gRNA Cloning Vector was linearized with AflII and the inserts were incorporated via Gibson assembly (NEB, Japan, E2611S). To prepare hCas9 mRNA and Mettl9 guide RNAs, an in vitro transcription reaction kit with T7 RNA polymerase (mMESSAGE mMACHINE T7 Ultra Kit, Life Technologies) was used; the transcribed RNAs were purified with the MEGAclear Kit (Life Technologies). The quality of the guide RNA was checked by an in vitro Cas9 cleavage assay with recombinant hCas9 nuclease (NEB) and a 400 bp PCR product around the guide RNA target site. The Cas9 mRNA and guide RNA were injected into fertilized eggs (mouse strain C57BL/6 J), and the mutated Mettl9 gene versions were screened by PCR after BsrI digestion, because the BsrI site is located after the guide RNA target site. One in 8 heterozygous ($+/-$) males carried the expected 16 bp deletion. The heterozygous mouse was crossed with a WT C57BL/6 J mouse, and the male ($+/-$) and female ($+/-$) mice were crossed to obtain Mettl9 KO ($-/-$) mice and WT ($+/+$) mice. Animals were maintained on a 12-h light/dark cycle with access to food and water ad libitum. The temperature and humidity were maintained at 22–23 °C and 50–60%, respectively. Animal health was checked by the animal facility staff 5 times per week. All experiments involving mice complied with all relevant ethical regulations for animal testing and research, and were carried out according to protocols approved by the Animal Experiment Committee of the RIKEN Center for Brain Science.

MS detection of methylation on recombinant proteins. Samples were run on NuPAGE 4–12% Bis-Tris protein gels in MES buffer (ThermoFisher), stained with SimplyBlue (Invitrogen) and destained in water. Proteins were excised from the gel and subject to in-gel digestion by trypsin or chymotrypsin. The dried peptides were dissolved in 10 µL of aqueous 2% acetonitrile containing 1% formic acid and 5 µL of the sample was injected into a Dionex Ultimate 3000 nano-UHPLC system (Sunnyvale, CA, USA) coupled online to a Q Exactive mass spectrometer (ThermoScientific, Bremen, Germany) equipped with a nano-electrospray ion source. For liquid chromatography separation, an Acclaim PepMap 100 column (C18, 3 µm beads, 100 Å, 75 µm inner diameter, 50 cm) was used. A flow rate of 300 nL/min was employed with a solvent gradient of 2–30% B in 60 min. Solvent A was 0.1% formic acid and solvent B was 0.1% formic acid in 90% acetonitrile. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 400 to 2000) were acquired with the resolution $R = 70,000$ at m/z 200, after accumulation to a target of 1e6. The maximum allowed ion accumulation times were 100 ms. The sequential isolation of up to the ten most intense ions, depending on signal intensity (intensity threshold 5.6e3) were considered for fragmentation using higher-energy collision-induced dissociation (HCD) at a target value of 100,000 charges and a resolution $R = 17,500$ with normalized collision energy (NCE) 28. Target ions already selected for MS/MS were dynamically excluded for 30 s. The isolation window was m/z = 2 without offset. Data were analyzed with in-house maintained human protein database using SEQUEST™ and Proteome Discoverer™ (Thermo Fisher Scientific). The mass tolerances of a fragment ion and a parent ion were set as 0.05 Da and 10 ppm, respectively. Methionine oxidation and cysteine carbamido-methylation were selected as variable modifications.

MS detection of ARMC6 methylation in cells. hMETTL9-GFP was over-expressed in HEK-293-derived T-REx-Flp-IN cells and affinity purified using GFP-Trap beads. In brief, cells in a 6-well plate were induced with 1 mg/ml doxycyclin overnight to express the GFP fusion proteins. The cells were harvested and lysed in RIPA buffer containing cOmplete protease inhibitor cocktail, phenylmethylsulfonyl fluoride and sodium orthovanadate (Sigma-Aldrich) on ice for 20 min. After centrifugation (16,000 × g, 5 min, 4 °C), the cleared lysate was diluted in Dilution Buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA). Washed GFP-Trap beads (Chromotek) were incubated with the lysate with overhead rotation for two hours, collected by centrifugation (3000 × g, 3 min, 4 °C), washed two times with 500 μl of the Dilution Buffer and frozen at −80 °C. Proteins were eluted through on-bead proteolysis using trypsin[38], the resulting peptides desalted using StageTips[39] and analyzed by nLC-MS/MS using EASY-nLC 1200 coupled to a Q Exactive HF-X mass spectrometer (Thermo Scientific) operated in the data-dependent acquisition mode[40]. All MS files were analyzed using MaxQuant (v1.6.3.3)[41] using the default settings except for the following modifications: monomethylation (Lys, Arg and His), di-methylation (Lys and Arg), and tri-methylation (Lys). The data were searched against a database comprising the canonical isoforms of human proteins as downloaded from Uniprot (Uniprot Complete proteome: UP_2017_04\Human\UP000005640_9606.fasta).

MS detection of NDUFB3 and S100A9 methylation in cells. For NDUFB3, the plasmids for retrovirus expression of C-terminally FLAG-tagged NDUFB3 (pQCXIH-NDUFB3-FLAG) were transfected into retrovirus packaging cells using PEI transfection reagent (Polysciences, Inc.). 24 h after the transfection, the virus-containing culture supernatants were transferred to WT or *METTL9* KO HEK293T cells with 4 μg/mL of polybrene. 24 h after the infection, 500 μg/mL hygromycin B was added and cells were cultured for 2 weeks. The cells were harvested and lysed in IP buffer (1xPBS, 1% n-Dodecyl-β-D-maltoside, protease inhibitors). M2-agarose beads (Sigma-Aldrich) were added and incubated with the lysate for 1 h at 4 °C, then washed three times with PBS. The bound proteins were separated with SDS-PAGE, and the gel band corresponding to NDUFB3-FLAG was excised and in-gel digested with trypsin.

For S100A9 isolation, peritoneal exudate neutrophils (PEN) were prepared as described[42]. WT and *Mettl9* KO mice were intraperitoneally injected with 3% (w/v) proteose peptone (Difco Laboratories Inc.) (4 ml per mouse) twice (2nd injection was after 12–15 h 1st injection), and sacrificed 2-3 h after the 2nd injection of proteose peptone to collect PEN. PEN-enriched fractions were isolated from the harvested intraperitoneal cells with a Percoll gradient. PEN fraction (2 × 10^5 cells) was mixed with 2x Laemmli SDS-sample buffer, and separated with SDS-PAGE. The bands corresponding to S100a9 (observed as a major band at 14 kDa) were excised, and in-gel digested with *Achrornobacter* protease I (API). The digested peptide fragments were analyzed with MALDI-MS.

MALDI MS analysis of hNDUFB3 and mS100A9. The SDS-PAGE-separated bands were excised and destained. The gel slices were reduced with 50 mM dithiothreitol and 4 M guanidine-HCl at 37 °C for 2 h, followed by alkylation with 100 mM acrylamide at 25 °C for 30 min. The NDUFB3 and S100A9 were digested with trypsin (TPCK-treated, Worthington Biochemical) and Achromobacter protease I (a gift from Dr. T. Masaki, Ibaraki University, Ibaraki), respectively. The reaction was carried out at 37 °C for 12 h in a digestion buffer (10 mM Tris-HCl (pH8.0)/0.03% n-Dodecyl-β-D-maltoside) An α-cyano-4-hydroxycinnamic acid (Brucker Daltonics)–saturated solution in 33% acetonitrile/67% water containing 0.1% trifluoroacetic acid was prepared. A 0.5 μl volume of each sample was mixed with 1.0 μl of the matrix solution and air-dried at room temperature on an MTP AnchorChip Targets plate (Brucker Daltonics). The MS and MS/MS spectra were acquired with a rapifleX MALDI Tissuetyper (Brucker Daltonics). The mass spectrometer was operated in the positive-ion mode and reflector mode the following high voltage conditions (Ion Source1:20.000 kV, PIE: 2.680 kV, Lens: 11.850 kV, Reflector 1: 20.830 kV, Reflector 2: 1.085 kV, Reflector 3: 8.700 kV). The MSMS spectra were obtained using LIFT method the following high voltage conditions (Ion Source1:20.000 kV, PIE: 2.680 kV, Lens: 11.850 kV, Reflector 1: 20.830 kV, Reflector 2: 1.085 kV, Reflector 3: 8.700 kV, Drift tube 1: 14.000 kV, Drift tube 2: 18.700 kV, MS/MS pulse: 2.600 kV).

The acquired data were processed using FlexAnalysis (version 4.0, Brucker Daltonics) and BioTools (3.2 SR5, Brucker Daltonics). The processed data were used to search with MASCOT (version 2.7.0, Matrix Science) against the in-house database including the amino acid sequences of NDUFB3 and S100A9, using the following parameters: enzyme = trypsin (NDUFB3) or Lys-C/P (S100A9); maximum missed cleavages = 1; variable modifications = Acetyl (Protein N-term), Oxidation (M), Methyl (H), Propionamide (C); product mass tolerance = ±50 ppm; product mass tolerance = ±0.5 Da (LIFT mode); instrument type = MALDI-TOF-TOF.

MS detection of DNAJB12 methylation in cells. HEK293 cells grown to 80% confluency in 150 mm plates were harvested and lysed in RIPA buffer containing cOmplete protease inhibitor cocktail and phenylmethylsulfonyl fluoride (Sigma-Aldrich) on ice for 20 min after brief sonication. The lysate was then cleared by centrifugation (16,000 × g, 5 min, 4 °C), diluted two-fold with Dilution Buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA) and incubated with

(0.5 μg/mg lysate) DNAJB12 antibody (16780-1-AP, Proteintech) with overhead rotation for two hours. Washed Protein A Agarose (Merck) was then added and incubated for 2 more hours. The resin was collected by centrifugation (3000 × g, 3 min, 4 °C), washed twice with Dilution Buffer, then twice with water, transferring the resin to a new low-protein binding tube with each wash, and finally frozen at −80 °C. Immunoprecipitated DNAJB12 was separated by SDS-PAGE and in-gel digested with Chymotrypsin (Roche)[43]. The resulting peptides were, without a desalting step, analyzed using a EASY-nLC 1200 ultrahigh-pressure system (Thermo Fisher Scientific) coupled to a Q Exactive HFX Orbitrap mass spectrometer (Thermo Fisher Scientific) operated in a data-dependent acquisition mode[40]. All raw MS files were analyzed with Max-Quant (version 1.6.0.17)[41] and searched against a database composed of the canonical isoforms of human proteins as downloaded from Uniprot in April 2017 (Uniprot Complete proteome: UP_2017_04/Human/UP000005640_9606.fasta) using the default setting with few exceptions. Methylation of lysine (mono, di, and tri), arginine (mono and di), and histidine (mono) were set as variable modifications and missed cleavages were restricted to one.

Ion chromatograms for the different methylated forms of a peptide covering amino acid 173–190 of DNAJB12, DQFGDDKSQAARHGHGHGDF, were extracted using Xcalibur Qual Browser, ver. 4.1 (Thermo Fisher Scientific). Selective ion settings for Asp171-Phe190 ($z = 4$) in Supplementary Fig. 9c were 546.24 (Me0), 549.75 (1xMe) and 553.25 (2xMe), 7 p.p.m.

Generation and methylation of peptide arrays. Peptide arrays of 15-mer peptides were generated using the SPOT method[44,45]. Methylation reactions were conducted by incubating the arrays in Reaction Buffer (50 mM Tris-HCl pH 7.8, 50 mM NaCl, 5 mM EDTA) with 0.76 μM [3H]-AdoMet (PerkinElmer) and 0.5-1 μM of His6-hMETTL9. For His-to-Ala replacements, single histidines from peptides from ARMC6 (aa 253–267), SLC39A5 (aa 367–381), and mouse S100A9 (aa 99-113, note the additional Cys-to-Ser substitution to reduce non-enzyme-catalyzed methylation) were substituted for alanines. For the middle and flanking HxH residue replacements, the peptides as above were additionally modified to contain only one HxH. The relevant residues were mutated to all proteinogenic amino acids except Trp and Cys. Note that for the middle residue replacements, the individual arrays were cut from single images and pasted together to follow the same order of mutations introduced.

For the candidate substrate array, HxH-containing sequences were identified in METTL9-interacting proteins, human proteins homologous to ANGST HxH-containing mouse proteins from the ProSeAM experiments, and putative methylhistidine-containing peptides identified through exploratory and further analysis of previously published HeLa proteomics datasets[28]. In addition, peptides were derived from sequences reported as histidine-methylated in literature, various zinc transporters and other HxH-containing proteins of particular biological interest. Peptide sequence windows were chosen to exclude Cys and Trp residues prone to non-enzyme-catalyzed methylation. As negative controls, all HxH-sequences in WT peptides were disrupted by appropriate His-to-Ala mutations. The negative controls were spotted next to the corresponding HxH-containing sequences. The quantitative analysis of array methylation data was performed using ImageJ[46].

Mitochondrial respiration analyses. HAP1 cells were seeded in three T75 flasks (8 × 10^6 cells/flask) and grown for 48 h at 37 °C. Cells of three flasks were combined after washing the cells with PBS. Mitochondria from WT and *METTL9* KO cells were isolated as previously described[47]. In brief, pelleted HAP1 cells were resuspended in MIB buffer (210 mM D-mannitol, 70 mM sucrose, 5 mM HEPES, 1 mM EGTA, and 0.5% (w/v) fatty acid-free BSA, pH 7.2), transferred into a glass tube, and disrupted by 30 strokes with a homogenizer. After centrifugation (600×g, 10 min, 4 °C), the supernatant was collected into a new tube and centrifuged at 8000 × g (10 min 4 °C). The pellet was washed twice with MIB buffer; the pellet, containing the mitochondria, was then resuspended in MAS buffer (220 mM d-Mannitol, 70 mM sucrose, 10 mM KH2PO4, 5 mM MgCl2, 2 mM HEPES, 1 mM EGTA, and 0.2% (w/v) of fatty acid-free BSA, pH 7.2) supplemented with 10 mM succinate and 2 μM rotenone or 5 mM malate and 10 mM glutamate, to measure Complex II and I respiration, respectively. Mitochondria (15 ug for Complex I- and 5 μg for Complex II-driven respiration) were added in a non-coated XF24 plates and centrifuged. Oxygen consumption rates (OCRs) were measured under basal conditions, and after sequential addition of ADP (2 mM), oligomycin (3.2 μM), FCCP (4 μM), and antimycin A (4 μM). Each assay cycle consisted of 1 min of mixing and 3 min of OCR measurements. For each condition, three cycles were used to determine the average OCR under a given condition.

Mitochondrial respiration analysis in permeabilized HEK293T cells was performed by following an established method[48] with some modifications. HEK293T cells (2 × 10^4 cells/well in 80 μL DMEM) seeded in microplates were incubated overnight at 37 °C under 5% CO2. The culture medium was replaced with MAS buffer (220 mM mannitol, 70 mM sucrose, 10 mM KH2PO4, 5 mM MgCl2, 2 mM HEPES-KOH, and 1 mM ethylene glycol tetra acetic acid, pH 7.2), after which the cells were incubated for 15 min at 37 °C in a CO2-free incubator. Next, three baseline measurements were taken using the Seahorse XFe96 analyzer, after which 25 μL of MAS buffer containing either digitonin (25 μg/mL, final concentration) 1 mM ADP, 1 mM sodium malate and 10 mM sodium pyruvate (for the assessment of Complex I) or digitonin (25 μg/mL), 1 mM ADP, 10 mM

sodium succinate and 0.5 μM rotenone (for Complex II) was injected from port A to start the reaction. This was followed by sequential treatments with oligomycin A (2 μM) from port B and antimycin A (0.5 μM) from port C. Each assay cycle consisted of 2 min of mixing, 2 min of incubation and 3 min of OCR measurements. For each condition, at least three replications were used to determine the average OCR.

ProSeAM assay. ProSeAM were prepared as reported in[49]. ProSeAM substrate screening was carried out as reported[50] with some modifications. *Mettl9* KO MEF cells were isolated from E13.5 mouse embryos, cultured in DMEM containing either light Arg and Lys or heavy isotope labeled Arg ($^{13}C_6$ $^{15}N_4$ L-Arginine) and Lys ($^{13}C_6$ $^{15}N_2$ L-Lysine, Thermo Scientific) for at least six doubling times. The cells were harvested and lysed in a lysis buffer (50 mM Tris-HCl with pH 8.0, 50 mM KCl, 10% Glycerol, 1% *n*-Dodecyl-β-D-maltoside). The cell lysates containing 150 μg of proteins were incubated with 150 μM of ProSeAM with (Heavy) or without (Light) 10 μg of mMETTL9 (protein:enzyme = 15:1) in MTase reaction buffer (50 mM Tris-HCl, pH 8.0) at 20 °C for 2 h. The reaction was stopped by adding 4 volumes of ice-cold acetone. The reaction tube was centrifuged at 15,000×g for 5 min, and precipitates were washed once with ice-cold acetone. The pellet was dissolved in 58.5 μL of 1× PBS and 0.2% SDS, after which 15 μL of 5× click reaction buffer and 1.5 μL of 10 mM Azide-PEG4-Biotin (Click Chemistry Tools) were added; the reaction mixture was incubated for 60 min at room temperature (RT). The click reaction was stopped with 4 volumes of ice-cold acetone. The pellet was resuspended in 75 μL of binding buffer (1× PBS, 0.1% Tween-20, 2% SDS, 20 mM dithiothreitol (DTT)) and sonicated for 10 s. The Light and Heavy samples were mixed in a tube; 450 μL of IP buffer (TBS, 0.1% Tween-20) containing 3 μg of Dynabeads M-280 Streptavidin (Life Technologies Japan Ltd., Minato-ku, Tokyo, Japan) was added to the tube, and it was incubated for 30 min at RT (the final SDS concentration in the reaction was 0.5%). The protein-bound beads were washed three times using wash buffer (1× PBS, 0.1% Tween-20, 0.5% SDS) and twice using 100 mM ammonium bicarbonate (ABC) buffer, and were analyzed by western blotting or mass spectrometry. For SILAC MS/MS analysis, DTT (20 mM) was added to protein-bound Dynabeads in 100 mM ABC buffer, and the mixture was incubated for 30 min at 56 °C. Then, iodoacetamide was added and the mixture was incubated for 30 min at 37 °C in the dark. The protein samples were then digested with 1 μg trypsin (Promega). The protein fragments were applied to a liquid chromatograph (EASY-nLC 1000; Thermo Fisher Scientific, Odense, Denmark) coupled to a Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific, Inc., San Jose, CA, USA), with a nanospray ion source in positive mode. The peptides derived from the protein fragments were separated on a NANO-HPLC C18 capillary column (0.075-mm inner diameter × 150 mm length, 3 mm particle size; Nikkyo Technos, Tokyo, Japan). Mobile phase A was comprised of water with 0.1% formic acid, and mobile phase B was comprised of acetonitrile with 0.1% formic acid. Two different slopes were used for a gradient elution for 120 min at a flow rate of 300 nL/min: 0–30% of phase B in 100 min and 30–65% of phase B in 20 min. The mass spectrometer was operated in the top-10 data-dependent scan mode. The parameters for operating the mass spectrometer were as follows: spray voltage, 2.3 kV; capillary temperature, 275 °C; mass-to-charge ratio, 350–1800; normalized collision energy, 28%. Raw data were acquired using the Xcalibur ver. 4.0 software (Thermo Fisher Scientific). The MS and MS/MS data were searched against the Swiss-Prot database using Proteome Discoverer 1.4 (Thermo Fisher Scientific) using the MASCOT search engine software version no. 2.6.0 (Matrix Science, London, United Kingdom). The peptides were considered identified when their false discovery rates (FDR) were less than 1%. For substrate identification, proteins exhibiting at least a 2-fold increase with at least 5% coverage were defined as positive hit proteins.

Isothermal titration calorimetry. SLC39A7 peptides that contain either no (Me0; HGHSHAHGHGHGTH) or six 1MH (Me6, HGXSXAXGXGXTX; X = 1MH)) residues were synthesized with a peptide synthesizer (ABI, 433 A) using Fmoc-3-methyl-L-histidine (Merck) for the 1MH residue. The peptides were purified with HPLC, and purity determined by MALDI-MS. Dry weights were quantified with an amino acid analyzer (Hitachi, L-8900) for the calculation of molarity. ITC assay was performed as reported previously[51], with a slight modification. Briefly, SLC39A7 peptides and Zn^{2+} solutions were prepared in 20 mM Tris –HCl pH 7.5, 100 mM NaCl. Titration of $ZnCl_2$ (2 mM) into peptide solutions (50 μM) was performed at 30 °C using an ITC titration calorimeter (Malvern Panalytical, MicroCal iTC200). 39 injections of 1 μL were made with an equilibration time of 1 min between injections. Integration of the thermograms after correction for heats of dilution yielded binding isotherms that fit best to a one-set of sites binding model using the ITC data analysis software Origin 7.0 (MicroCal Inc., Piscataway, NJ). A non-linear least-squares curve-fitting algorithm was used to determine the stoichiometric ratio (n), the dissociation constant (K_d), and the change in enthalpy (ΔH) of the interaction. All ITC experiments were performed in triplicate.

Sequence alignments and structural modeling. Multiple protein sequence alignments were performed using the Muscle algorithm embedded in Jalview[52]. The structure of hMETTL9 was modeled using RaptorX[53] using the Ado-Met-dependent methyltransferase Q8PUk2 from *Methanosarcina mazie* as template (PDB ID 3sm3).

Fluorescence microscopy. HeLa cells were transfected with pEGFP-N1 hMETTL9-GFP constructs using FuGENE transfection reagent (Promega) for 24 h and probed with ER-tracker and Mitotracker (ThermoFisher Scientific). Live cells were imaged with PlanApo ×100, numerical aperture 1.1 oil objective (Olympus).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The immunoprecipitation mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[54] partner repository with the dataset identifier PXD016408. The DNAJB12 methylation data have been deposited to ProteomeXchange, ID: PXD020010. The ProSeAm SILAC screen data have been deposited at ProteomeXchange, ID: PXD016823. NDUFB3 and S100A9 methylation data are available via ProteomeXchange with identifier PXD022067. The following publicly available databases were used: UniProt [https://www.uniprot.org/] for obtaining protein sequences, and RCSB Protein Data Bank (RCSB PDB) [https://www.rcsb.org/] for obtaining protein structures.

Source data are provided with this paper.

References
1. Kwiatkowski, S. & Drozak, J. Protein histidine methylation. *Curr. Protein Pept. Sci.* **21**, 675–689 (2020).
2. Moore, K. E. & Gozani, O. An unexpected journey: lysine methylation across the proteome. *Biochim. Biophys. Acta* **1839**, 1395–1403 (2014).
3. Bedford, M. T. Arginine methylation at a glance. *J. Cell Sci.* **120**, 4243–4246 (2007).
4. Figaro, S., Scrima, N., Buckingham, R. H. & Heurgue-Hamard, V. HemK2 protein, encoded on human chromosome 21, methylates translation termination factor eRF1. *FEBS Lett.* **582**, 2352–2356 (2008).
5. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343–357 (2012).
6. Falnes, P. O., Jakobsson, M. E., Davydova, E., Ho, A. Y. & Malecki, J. Protein lysine methylation by seven-β-strand methyltransferases. *Biochem. J.* **473**, 1995–2009 (2016).
7. Herz, H. M., Garruss, A. & Shilatifard, A. SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends Biochem. Sci.* **38**, 621–639 (2013).
8. Petrossian, T. C. & Clarke, S. G. Uncovering the human methyltransferasome. *Mol. Cell Proteom.* **10**, M110 (2011).
9. Schubert, H. L., Blumenthal, R. M. & Cheng, X. Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.* **28**, 329–335 (2003).
10. Huszar, G. & Elzinga, M. Homologous methylated and nonmethylated histidine peptides in skeletal and cardiac myosins. *J. Biol. Chem.* **247**, 745–753 (1972).
11. Johnson, P., Harris, C. I. & Perry, S. V. 3-methylhistidine in actin and other muscle proteins. *Biochem. J.* **105**, 361–370 (1967).
12. Meyer, H. E. & Mayr, G. W. N pi-methylhistidine in myosin-light-chain kinase. *Biol. Chem. Hoppe Seyler* **368**, 1607–1611 (1987).
13. Raftery, M. J., Harrison, C. A., Alewood, P., Jones, A. & Geczy, C. L. Isolation of the murine S100 protein MRP14 (14 kDa migration-inhibitory-factor-related protein) from activated spleen cells: characterization of post-translational modifications and zinc binding. *Biochem. J.* **316**, 285–293 (1996).
14. Carroll, J. et al. The post-translational modifications of the nuclear encoded subunits of complex I from bovine heart mitochondria. *Mol. Cell Proteom.* **4**, 693–699 (2005).
15. Ning, Z. et al. A charge-suppressing strategy for probing protein methylation. *Chem. Commun. (Camb.)* **52**, 5474–5477 (2016).
16. Wilkinson, A. W. et al. SETD3 is an actin histidine methyltransferase that prevents primary dystocia. *Nature* **565**, 372–376 (2019).
17. Kwiatkowski, S. et al. SETD3 protein is the actin-specific histidine N-methyltransferase. *Elife* **7**, e37921 (2018).
18. Liao, S. M., Du, Q. S., Meng, J. Z., Pang, Z. W. & Huang, R. B. The multiple roles of histidine in protein interactions. *Chem. Cent. J.* **7**, 44 (2013).
19. Zhu, R. et al. Allosteric histidine switch for regulation of intracellular zinc(II) fluctuation. *Proc. Natl Acad. Sci. U. S. A* **114**, 13661–13666 (2017).
20. Huttlin, E. L. et al. The BioPlex Network: a systematic exploration of the human interactome. *Cell* **162**, 425–440 (2015).
21. Huttlin, E. L. et al. Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505–509 (2017).

22. Ignatova, V. V., Jansen, P. W. T. C., Baltissen, M. P., Vermeulen, M. & Schneider, R. The interactome of a family of potential methyltransferases in HeLa cells. *Sci. Rep.* **9**, 6584 (2019).

23. Shimazu, T., Barjau, J., Sohtome, Y., Sodeoka, M. & Shinkai, Y. Selenium-based S-adenosylmethionine analog reveals the mammalian seven-beta-strand methyltransferase METTL10 to be an EF1A1 lysine methyltransferase. *PLoS. ONE* **9**, e105394 (2014).

24. de Castro, E. et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* **34**, W362–W365 (2006).

25. Kambe, T., Tsuji, T., Hashimoto, A. & Itsumura, N. The physiological, biochemical, and molecular roles of zinc transporters in zinc homeostasis and metabolism. *Physiol. Rev.* **95**, 749–784 (2015).

26. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

27. Cagney, G., Amiri, S., Premawaradena, T., Lindo, M. & Emili, A. In silico proteome analysis to facilitate proteomics experiments using mass spectrometry. *Proteome Sci.* **1**, 5 (2003).

28. Bekker-Jensen, D. B. et al. An optimized shotgun strategy for the rapid generation of comprehensive human proteomes. *Cell Syst.* **4**, 587–599 (2017).

29. Songe-Moller, L. et al. Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. *Mol. Cell Biol.* **30**, 1814–1827 (2010).

30. van den Born, E. et al. ALKBH8-mediated formation of a novel diastereomeric pair of wobble nucleosides in mammalian tRNA. *Nat. Commun.* **2**, 172 (2011).

31. Leonardi, A. et al. The epitranscriptomic writer ALKBH8 drives tolerance and protects mouse lungs from the environmental pollutant naphthalene. *Epigenetics* **15**, 1121–1138 (2020).

32. Leonardi, A., Evke, S., Lee, M., Melendez, J. A. & Begley, T. J. Epitranscriptomic systems regulate the translation of reactive oxygen species detoxifying and disease linked selenoproteins. *Free Radic. Biol. Med.* **143**, 573–593 (2019).

33. Monies, D., Vagbo, C. B., Al-Owain, M., Alhomaidi, S. & Alkuraya, F. S. Recessive truncating mutations in ALKBH8 cause intellectual disability and severe impairment of wobble uridine modification. *Am. J. Hum. Genet.* **104**, 1202–1209 (2019).

34. Dimitrova, E., Turberfield, A. H. & Klose, R. J. Histone demethylases in chromatin biology and beyond. *EMBO Rep.* **16**, 1620–1639 (2015).

35. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14**, 1025–1040 (2007).

36. Bairoch, A. The ENZYME database in 2000. *Nucleic Acids Res.* **28**, 304–305 (2000).

37. Yamazaki, T. et al. Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. *Mol. Cell* **70**, 1038–1053 (2018).

38. Keilhauer, E. C., Hein, M. Y. & Mann, M. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol. Cell Proteom.* **14**, 120–135 (2015).

39. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).

40. Kelstrup, C. D. et al. Performance evaluation of the Q exactive HF-X for shotgun proteomics. *J. Proteome Res.* **17**, 727–738 (2018).

41. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

42. Hamasaki, A. et al. Accelerated neutrophil apoptosis in mice lacking A1-a, a subtype of the bcl-2-related A1 gene. *J. Exp. Med.* **188**, 1985–1992 (1998).

43. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).

44. Kudithipudi, S., Schuhmacher, M. K., Kebede, A. F. & Jeltsch, A. The SUV39H1 protein lysine methyltransferase methylates chromatin proteins involved in heterochromatin formation and VDJ recombination. *ACS Chem. Biol.* **12**, 958–968 (2017).

45. Kudithipudi, S., Kusevic, D., Weirich, S., & Jeltsch, A. Specificity analysis of protein lysine methyltransferases using SPOT peptide arrays. *J. Vis. Exp.* **93**, e52203 (2014).

46. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

47. Iuso, A., Repp, B., Biagosch, C., Terrile, C. & Prokisch, H. Assessing mitochondrial bioenergetics in isolated mitochondria from various mouse tissues using Seahorse XF96 analyzer. *Methods Mol. Biol.* **1567**, 217–230 (2017).

48. Futamura, Y. et al. Bioenergetic and proteomic profiling to screen small molecule inhibitors that target cancer metabolisms. *Biochim. Biophys. Acta Proteins Proteom.* **1867**, 28–37 (2019).

49. Sohtome, Y. et al. Unveiling epidithiodiketopiperazine as a non-histone arginine methyltransferase inhibitor by chemical protein methylome analyses. *Chem. Commun. (Camb.)* **54**, 9202–9205 (2018).

50. Shimazu, T. et al. Role of METTL20 in regulating beta-oxidation and heat production in mice under fasting or ketogenic conditions. *Sci. Rep.* **8**, 1179 (2018).

51. Kehl-Fie, T. E. et al. Nutrient metal sequestration by calprotectin inhibits bacterial superoxide defense, enhancing neutrophil killing of Staphylococcus aureus. *Cell Host. Microbe* **10**, 158–164 (2011).

52. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

53. Kallberg, M. et al. Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **7**, 1511–1522 (2012).

54. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

Acknowledgements

Author contributions

E.D. and T.Sh. designed and performed experiments, analyzed and presented data and prepared the manuscript. M.Sc. and S.W. performed and interpreted peptide array experiments. M.E.J. designed and performed MS and bioinformatics analyses of the GFP-METTL9 immunoprecipitation experiment and M.E.J. and V.S. performed proteomics analysis of METTL9 KO and WT cells. T.L. performed amino acid analysis. H.L.D.M.W. performed Seahorse experiment and analyses. A.Y.Y.H. performed cellular imaging experiments. R.P., J.M., and L.S. performed biochemical experiments. T. Su., N.D., and A.M. performed MS analysis. I.A.G. performed DNA cloning. Y. So., M.A. and M. So. provided ProSeAM. M.K. and T.U. prepared recombinant His-mMETTL9 from insect cells. J.V.O., M.A.M., N.E., and C.J.S. planned and analyzed experiments. A.J. planned experiments and analyzed and presented data. Y.Sh. and P.Ø.F. supervised the study, planned experiments, interpreted data, and prepared the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-20670-7.

Correspondence and requests for materials should be addressed to A.J., Y.Sh. or P.Ø.F.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
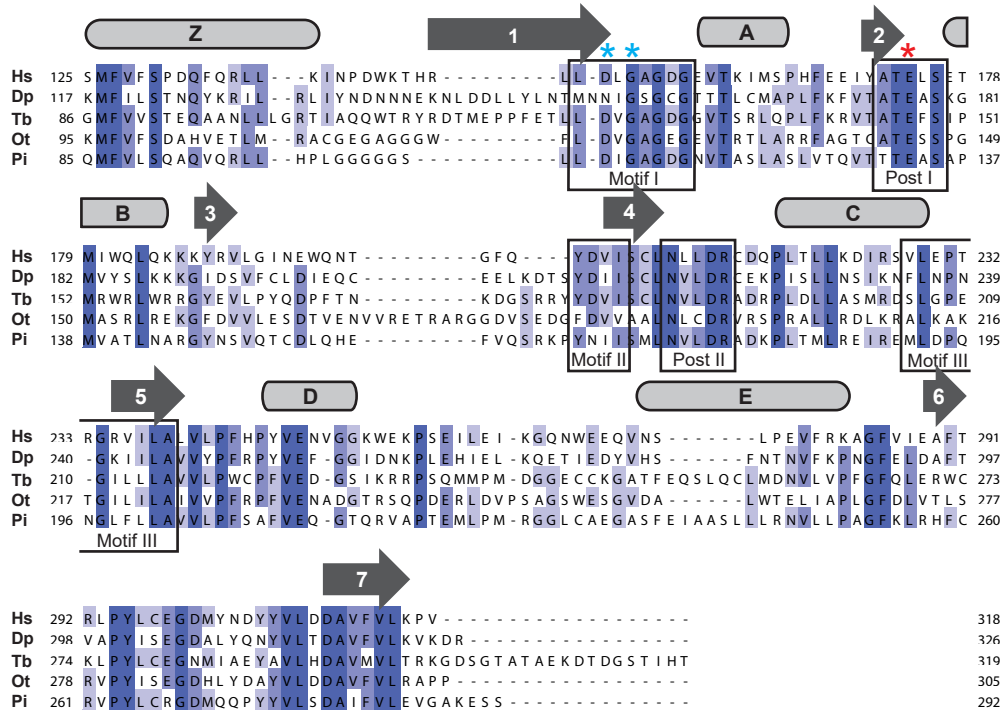
The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes
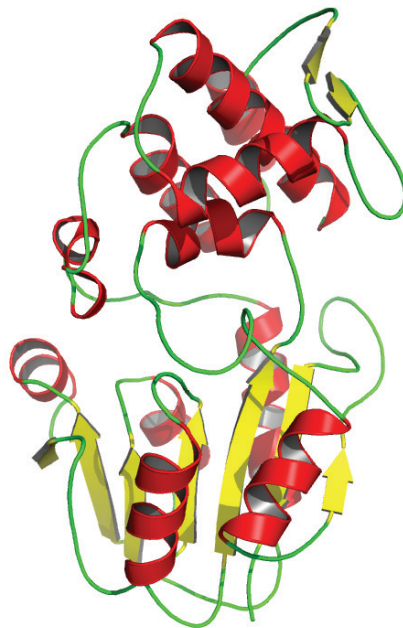
Erna Davydova*, Tadahiro Shimazu* et al. 2020

Supplementary Information

Supplementary Fig. 1: METTL9 is a eukaryotic 7BS methyltransferase.
**a,** Alignment of the core methyltransferase domain of METTL9 orthologues from *Homo sapiens*
(Hs; Q9H1A3-1), *Dictyostelium purpureum* (Dp; XP_003287139.1), *Trypanosoma brucei* (Tb; AAZ11816),
Ostreococcus tauri (Ot; Q01C52), and *Phytophthora infestans* (Pi; PITG_14565T0).
Predicted secondary structure is depicted as rectangles (α-helices) and arrows (β-strands). Hallmark 7BS-MTase
motifs I-III and Post I-II are labelled and boxed. The glutamic acid residue mutated in the E174A hMETTL9 mutant
is marked with a red asterisk, the aspartic acid and glycine mutated in the D151K/G153R mMETTL9 mutant are
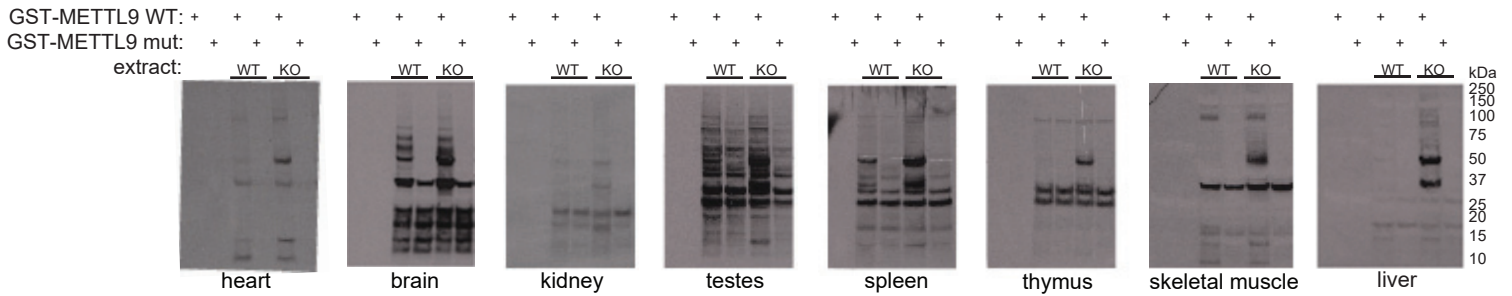marked by blue asterisks. **b,** Predicted structure of human METTL9, α-helices in red and β-strands in yellow.
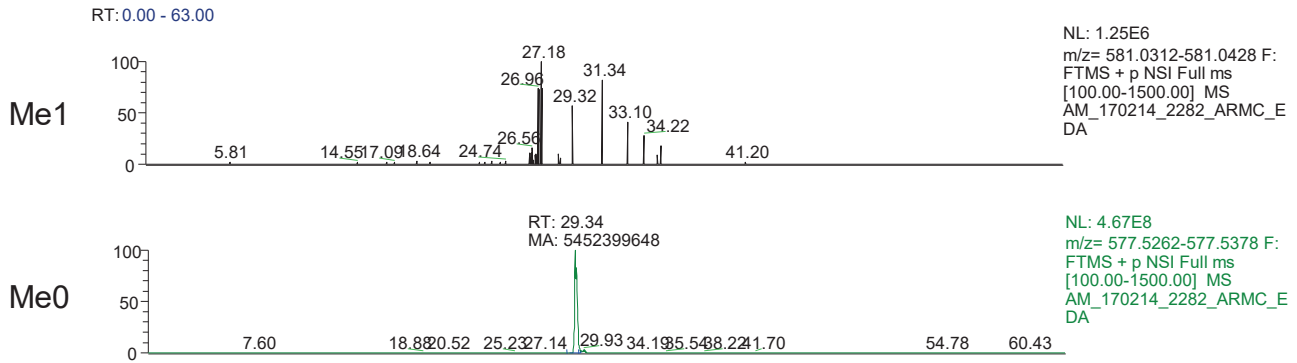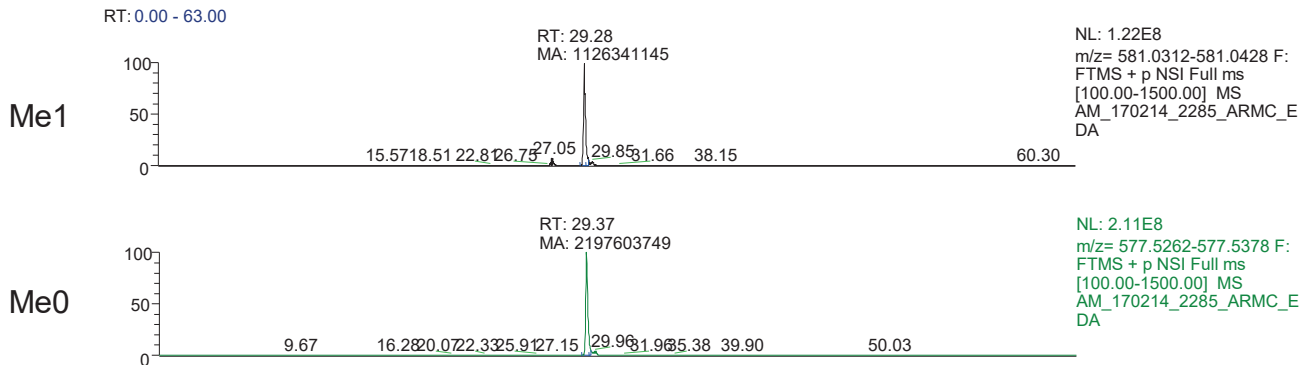
a
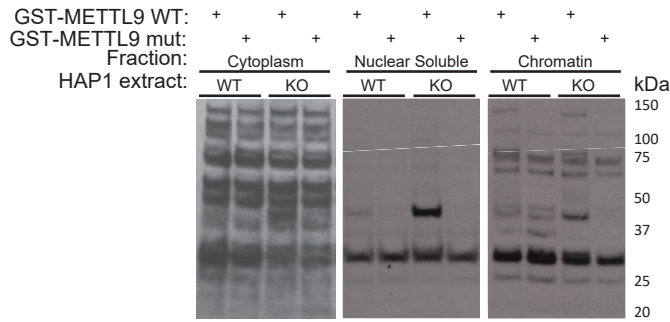
b

c
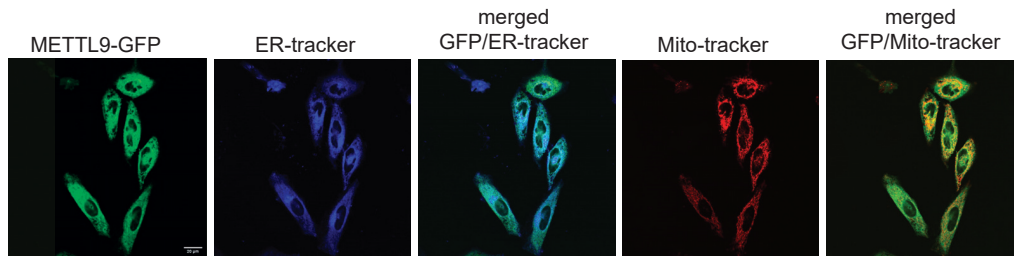
Supplementary Fig. 2: METTL9 is a protein methyltransferase present in different mammalian tissues and subcellular compartments. a, Fluorography showing the activity of wild-type GST-hMETTL9 (METTL9 WT) or the GST-hMETTL9 E174A inactive mutant (METTL9 mut) on cytoplasmic, nuclear soluble and chromatin-bound protein fractions from HAP1 WT or METTL9 knockout (KO) cells. **b,** Subcellular localization of hMETTL9-GFP in HeLa cells. Scale bar: 20 μm. **c,** Fluorography showing the activity of hMETTL9 WT or hMETTL9 mut on protein extracts from different tissues of WT or *Mettl9* KO mice. **a-c** show representative results of at least three independent experiments.

+ METTL9 mut

RT: 0.00 - 63.00

Me1

NL: 1.25E6
m/z= 581.0312-581.0428 F:
FTMS + p NSI Full ms
[100.00-1500.00] MS
AM_170214_2282_ARMC_E
DA

27.18
26.96
31.34
29.32
33.10
34.22
5.81   14.55 17.09 18.64 24.74 26.56   41.20

Me0

RT: 29.34
MA: 5452399648

NL: 4.67E8
m/z= 577.5262-577.5378 F:
FTMS + p NSI Full ms
[100.00-1500.00] MS
AM_170214_2282_ARMC_E
DA

7.60   18.88 20.52 25.23 27.14 29.93 34.19 35.54 38.22 41.70   54.78 60.43

+ METTL9 WT

RT: 0.00 - 63.00

Me1

RT: 29.28
MA: 1126341145

NL: 1.22E8
m/z= 581.0312-581.0428 F:
FTMS + p NSI Full ms
[100.00-1500.00] MS
AM_170214_2285_ARMC_E
DA

15.57 18.51 22.81 26.75   27.05   29.85 31.66 38.15   60.30

Me0

RT: 29.37
MA: 2197603749

NL: 2.11E8
m/z= 577.5262-577.5378 F:
FTMS + p NSI Full ms
[100.00-1500.00] MS
AM_170214_2285_ARMC_E
DA

9.67   16.28 20.07 22.33 25.91 27.15 29.96 31.96 35.38 39.90   50.03

Supplementary Fig. 3: Extracted ion chromatograms (XICs) showing METTL9-mediated methylation of ARMC6 (detailed data for Fig. 1b). Recombinant His_6-ARMC6 was incubated with recombinant WT or mutant (E174A) His_6-hMETTL9 in the presence of AdoMet, and digested with trypsin. The resulting peptides were analysed by high resolution LC-MS. XICs corresponding to the methylated (Me1) and unmethylated (Me0) forms of the peptide VMTFDDDIRVPFGHAHNHAK (ARMC6$_{248-267}$) are shown. z = 4; theoretical m/z: 577.533 (Me0), 581.037 (Me1).
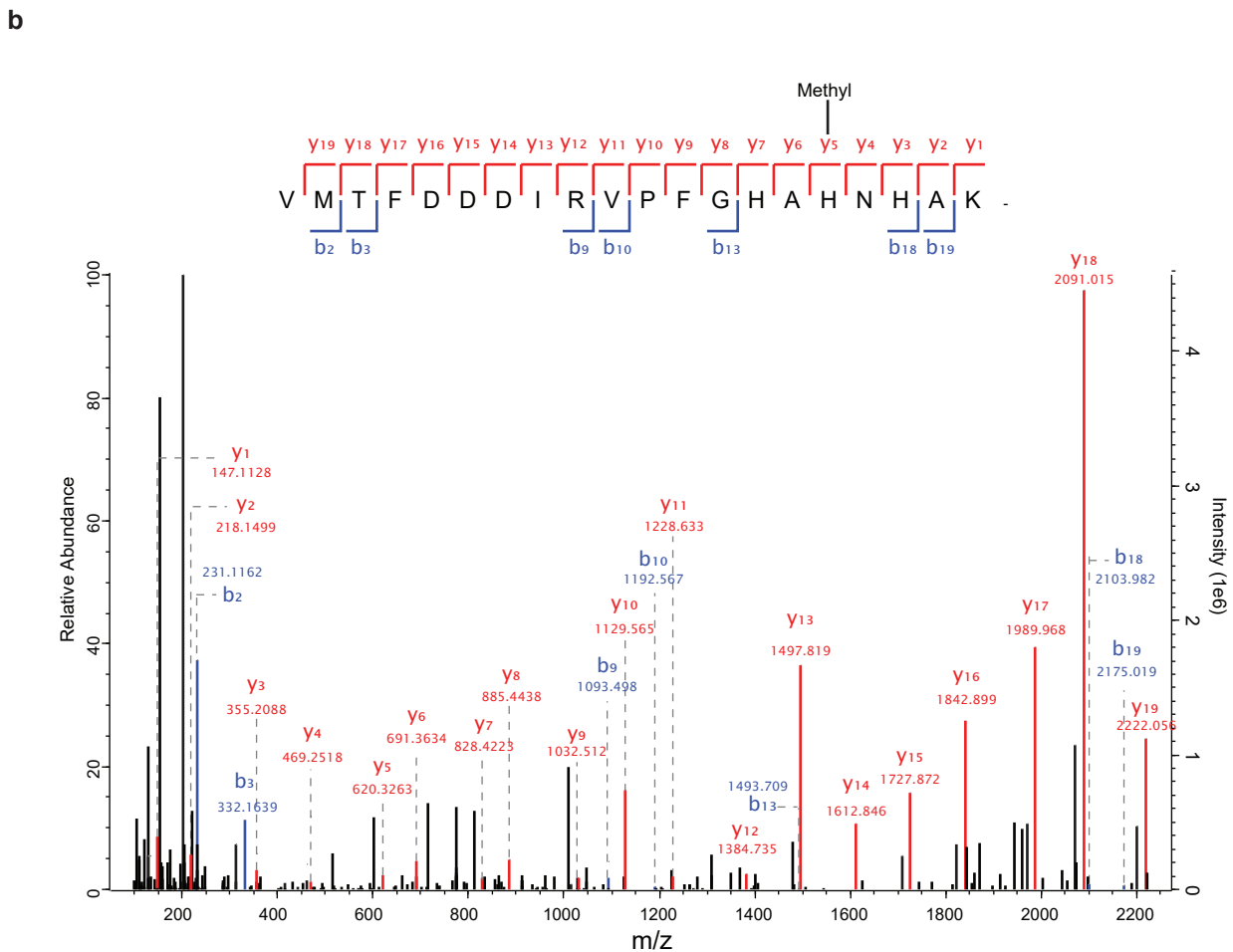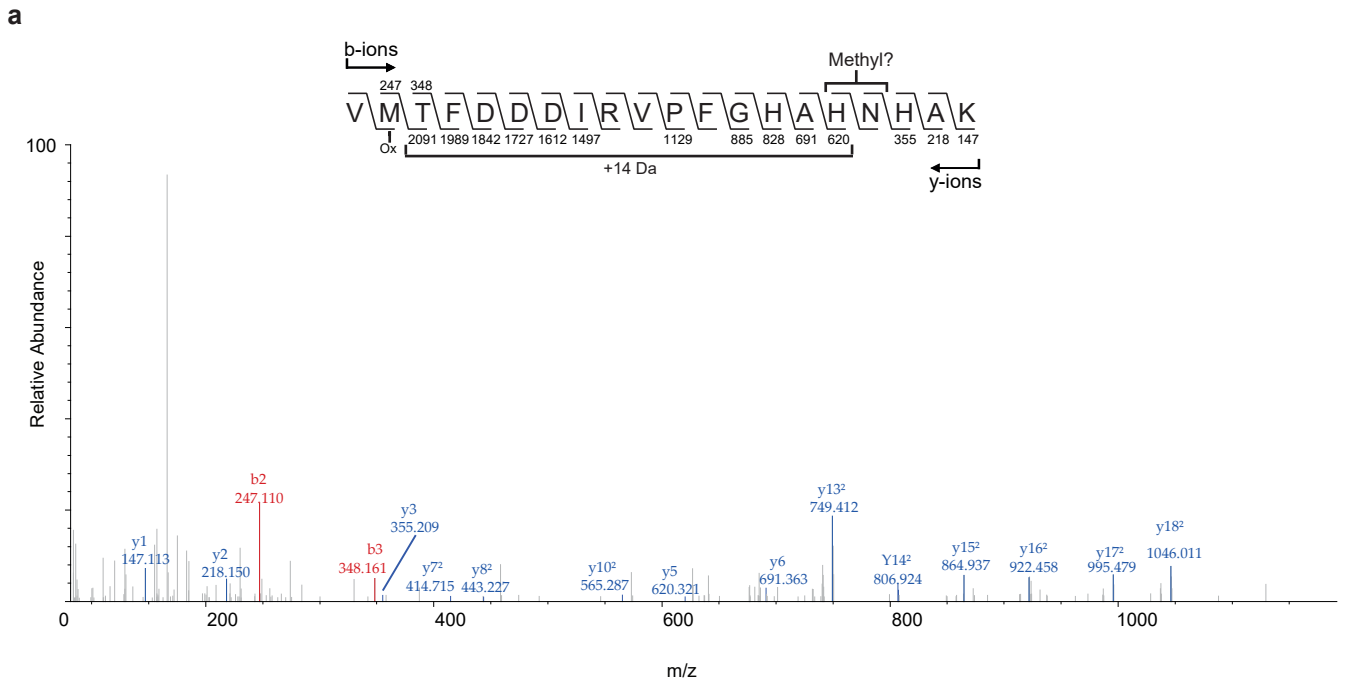
**Supplementary Fig. 4: MS/MS evidence of ARMC6 His263 methylation.**
a, MS/MS fragmentation spectra of a tryptic peptide from His6-ARMC6 in vitro methylated with His6-hMETTL9, indicating monomethylation at either His263 or Asn264 **b,** MS/MS fragmentation spectra of a peptide covering amino acids 248-267 in ARMC6 from Flp-In 293 T-REx cells pulled down with the hMETTL9-GFP bait, demonstrating monomethylation of His263.
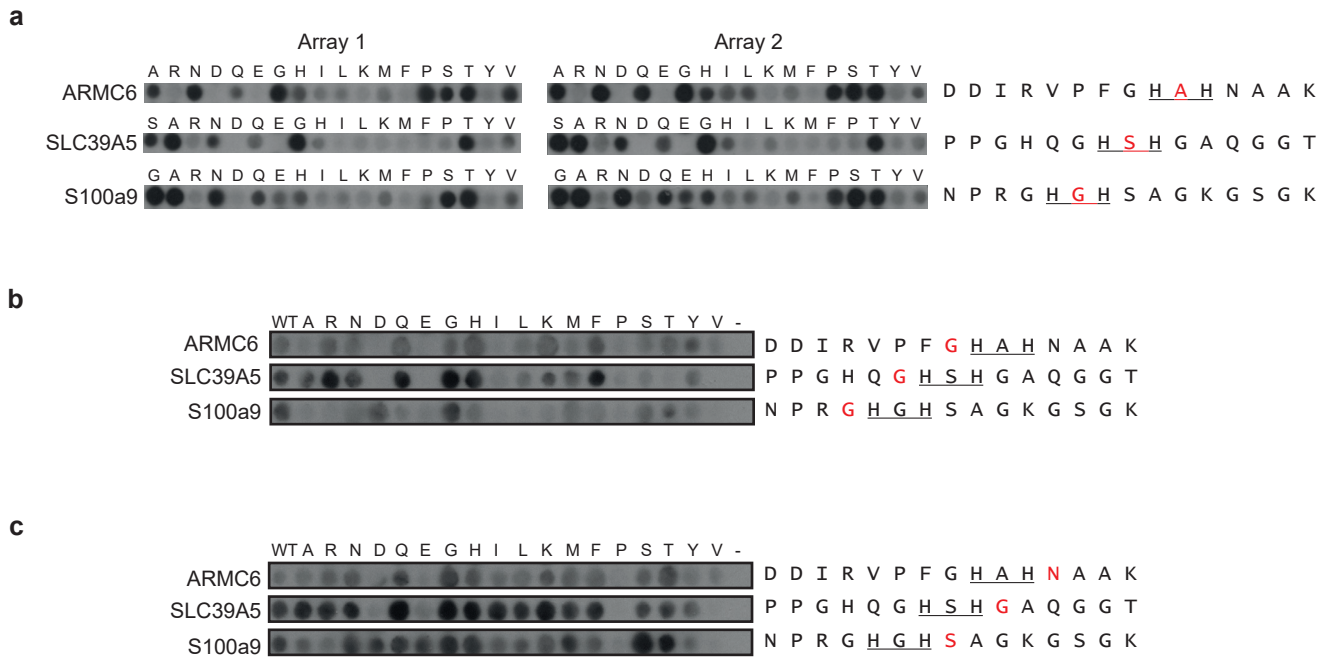
Supplementary Fig. 5: METTL9 activity on HxH peptide arrays.
a, Original peptide arrays used for quantification in Fig. 1f, showing the activity of His$_6$-hMETTL9 with [^{3}H]-AdoMet on peptides derived from ARMC6, SLC39A5 and S100a9 with replacements of the middle residue (red) in the HxH motif (underlined). Note that Array 1 is presented in Fig.1f rearranged to follow the same order of introduced mutations. Two independent peptide array experiments, each assessing three peptides, were performed. **b,** Similar to **a** but single array with replacements of the N-terminally flanking residue (red). **c,** similar to **b** but with replacements of the C-terminally flanking residue (red). The array experiments shown in **b** and **c** were only performed once, but with three different peptides. Thus, the main conclusions are supported by three independent experimental observations.
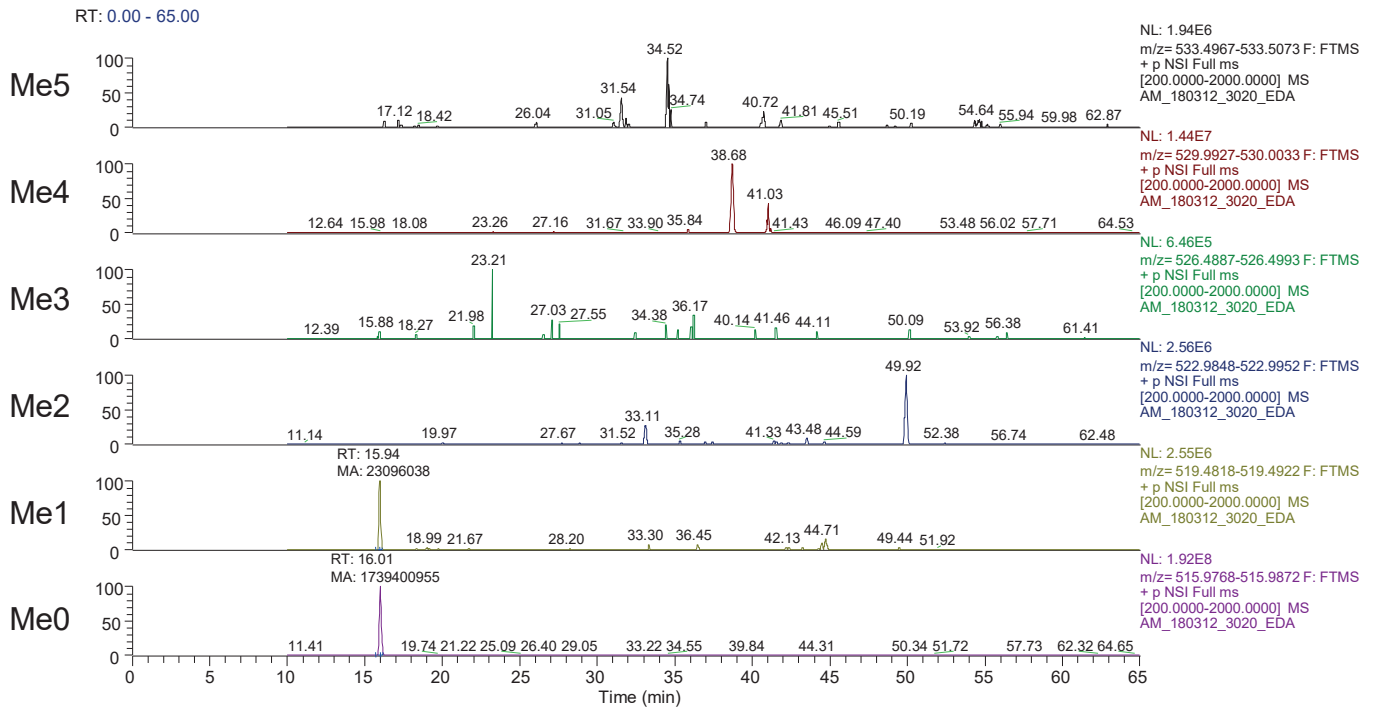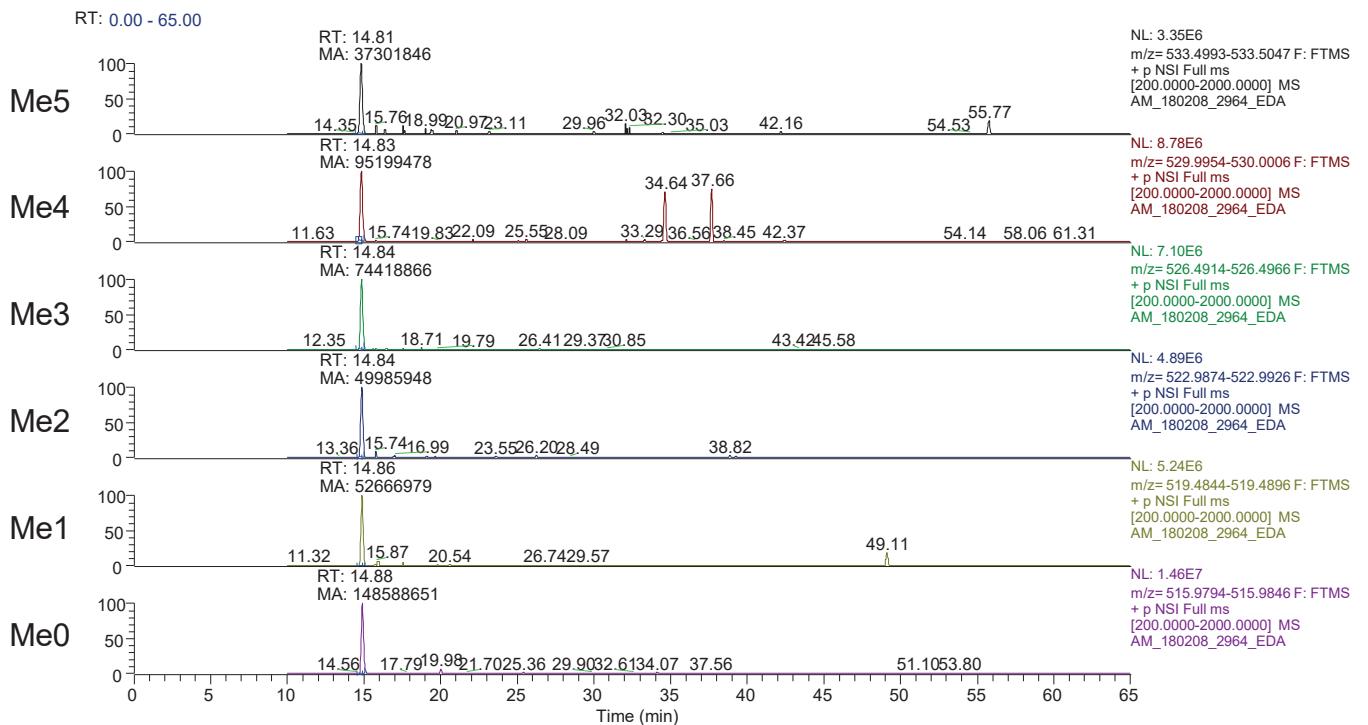
Supplementary Fig. 6: Extracted ion chromatograms (XICs) showing METTL9-mediated introduction of multiple methylations on a SLC39A7-derived fragment in vitro (detailed data for Fig. 1g). A recombinant fusion protein between GST and residues 31-137 of SLC39A7 (GST-SLC39A7$_{31-137}$) was incubated with recombinant WT or mutant His$_6$-hMETTL9 in the presence of AdoMet, and digested with chymotrypsin. The resulting peptides were analysed by high resolution LC-MS. XICs corresponding to different methylation states of the peptide HHGHSHAHGHGHTHESIW (SLC39A7$_{55-72}$) are shown. z=4; Theoretical m/z: 515.983 (Me0), 519.486 (Me1), 522.990 (Me2), 526.494 (Me3), 529.998 (Me4); 533.502 (Me5).
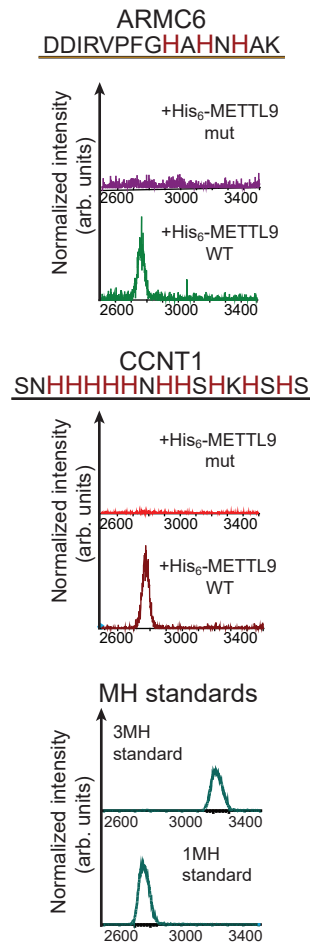
Supplementary Fig. 7: METTL9 forms 1MH in peptides derived from ARMC6 and CCNT1.
Amino acid analysis of ARMC6 and CCNT1 peptides methylated by His_6-hMETTL9 WT or
E174A inactive mutant (mut) as compared to the 3MH- and 1MH-methylhistidine standards.
Shown are the normalized chromatograms with the elution times of the corresponding amino
acid peaks.
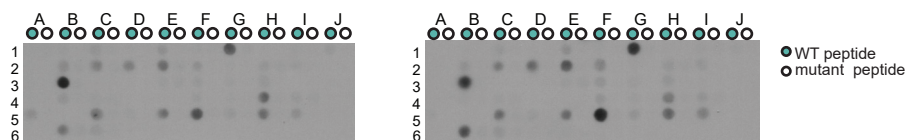
Supplementary Fig. 8: Peptide array both replicates. The peptide array experiments shown in
Fig. 3e was performed twice with very similar results, which were used to generate Fig. 3d. Here,
both arrays are shown (array on the right is presented in Fig. 3e).
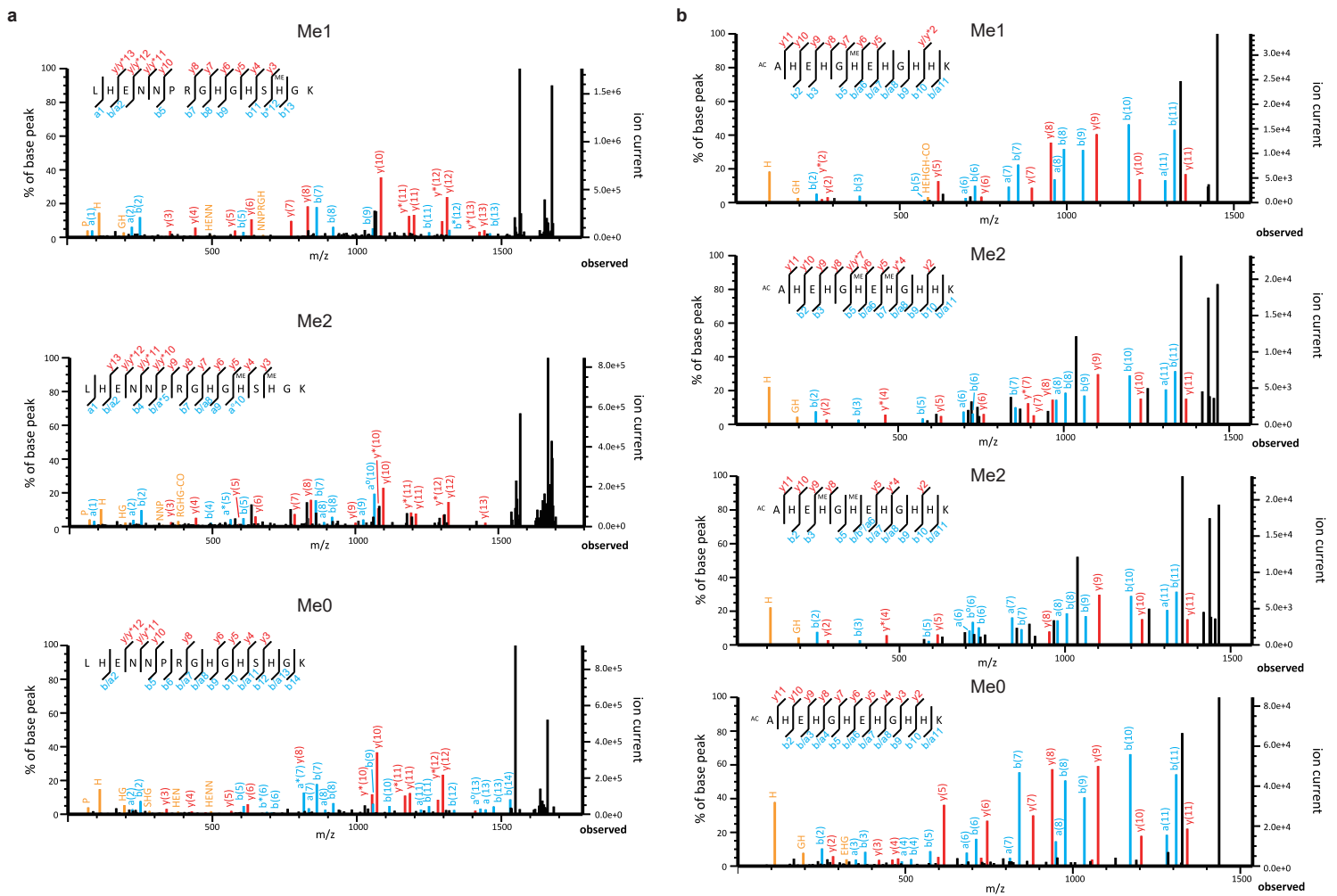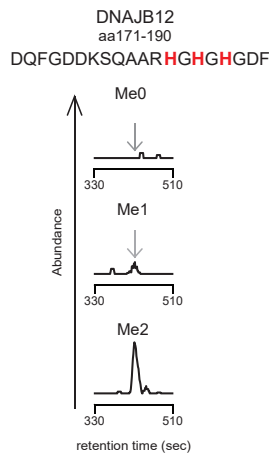
Supplementary Fig. 9: MS support of cellular METTL9-mediated histidine methylation.
a, MS/MS fragmentation spectra of a peptide covering amino acids 94-108 in S100A9 from peritoneal exudate neutrophils (PEN) isolated from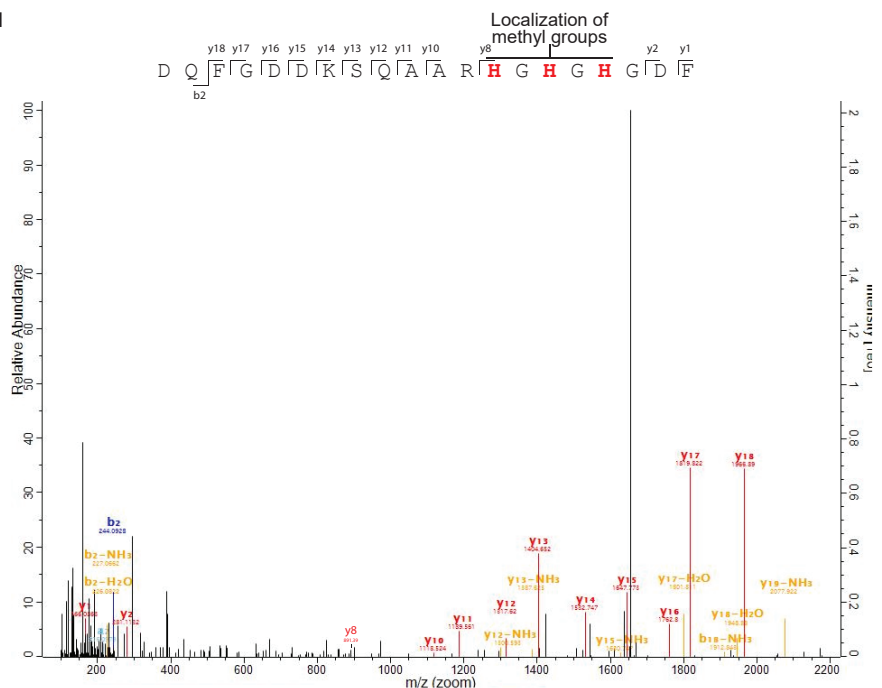 WT mice indicating methylation at His106 (top); and methylation at His104/His106 (middle); or the unmethylated peptide from METTL9 KO mice (bottom). **b,** MS/MS fragmentation spectra of tryptic peptides from immunoprecipitated NDUFB3-FLAG from WT HEK293T cells indicating methylation at His6 (top); methylation at His6/His8 (second from top); methylation at His4/His6 (third from top); and no methylation (bottom). The middle panels show the same MS/MS spectra with different annotation corresponding to methylation at His6/His8 and His4/His6, respectively. **c,** Extracted ion chromatograms of chymotryptic peptides from endogenous DNAJB12 immunoprecipitated from HEK293T cells, with the indicated number of added methyl groups (Me0, Me1, Me2).**d,** MS/MS fragmentation spectra of the dimethylated DNAJB12 peptide suggesting methylation of histidines. Source data are provided as a Source Data file.
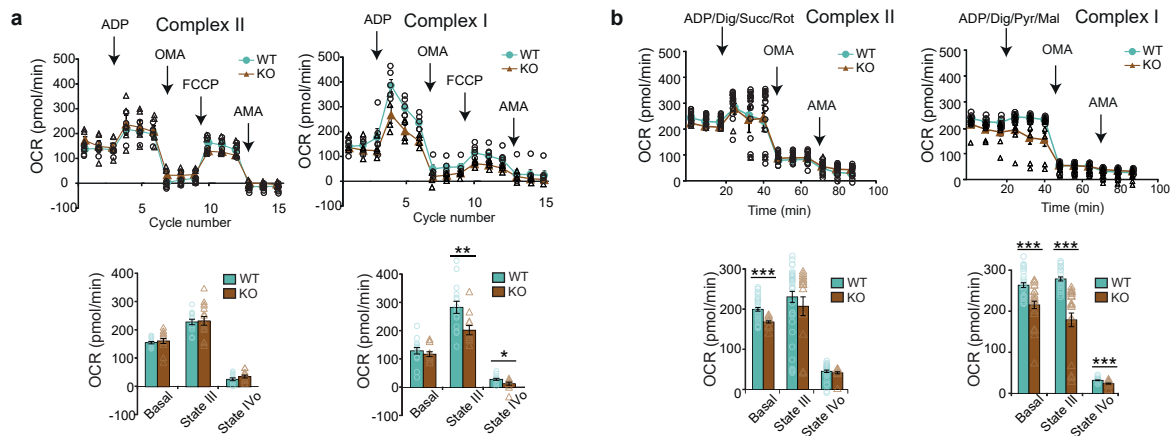
Supplementary Fig. 10: Mitochondrial respiration in WT vs METTL9 KO cells.
a, Oxygen consumption rate (OCR) of isolated mitochondria from WT or METTL9 knockout (KO) HAP1 cells (n = 5 (WT) or 4 (KO) biologically independent samples). Mitochondrial respiration driven by Complex II using succinate, in the presence of rotenone, as electron source (left panels) or Complex I using glutamate and malate (right panels). OCR traces (top) and quantification of respiratory states (bottom). Mean ± SEM. **P<0.01(P=0.0091);*P<0.05(P=0.0279); two-tailed Student's t-test. Source data are provided as a Source Data file. **b,** similar to a, but using digitonin-permeabilized WT or METTL9 KO HEK293T cells, with succinate/rotenone and pyruvate/malate as substrates for respiration mediated by Complex II (left; n = 13 (WT) or 6 (KO) biologically independent samples) and Complex I (right; n = 10 (WT) or 9 (KO) biologically independent samples), respectively. Mean ± SEM. ***P<0.001(Complex II: P=1.630e-04; Complex I: basal P=5.239e-05, state III P=2.564e-06, state IV P=2.710e-06); two-tailed Student's t-test. Source data are provided as a Source Data file.
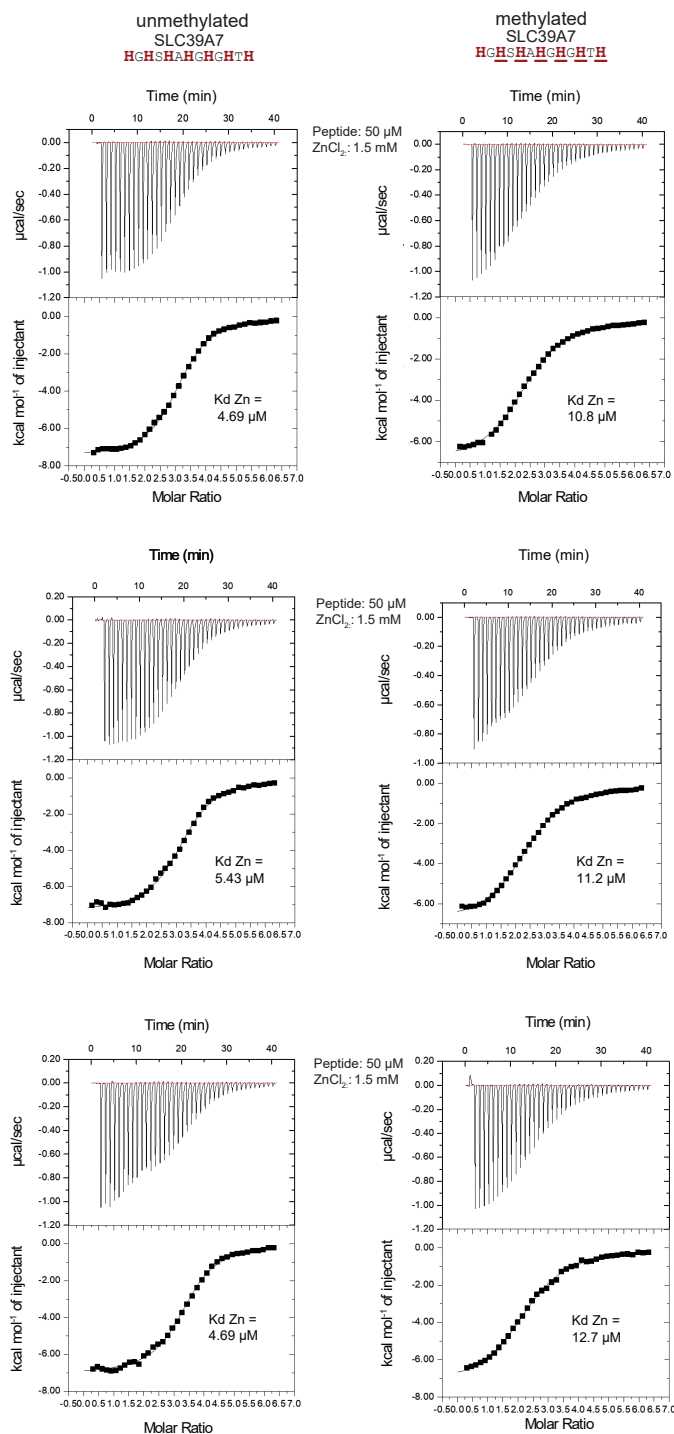
Supplementary Fig. 11: Determination of Zn²⁺ binding to methylated and unmethylated SLC39A7 peptides. Isothermal Titration Calorimetry measurements of the titration of $ZnCl_2$ binding to 50 µM unmethylated (left) or methylated (right, 1MH are underlined) SLC39A7 peptides. Heat changes (upper panels) and fitted data (lower panels) from three independent experiments.

Supplementary Table 1: Peptides used in the candidate substrate peptide array. Indicated are the position in the arrays from Fig. 3d (Pos), Uniprot ID and gene names, WT and control mutant peptide sequences and the corresponding average relative activity of METTL9, as well as the reasoning for inclusion of the peptides (Source) where "other" encompasses sequences reported as histidine-methylated in literature, putative methylhistidine-containing peptides identified through exploratory and further analysis of previously published HeLa proteomics datasets, various zinc transporters and other HxH-containing proteins of particular biological interest. Alternating histidines marked in blue, His-to-Ala mutations in red.

Pos	Uniprot ID	Gene	WT sequence	Mut sequence	WT avg relative activity	Mut avg relative activity	Source
A1	Q9H845	ACAD9	IRIGL RNHDH EVLLA	IRIGL RNADH EVLLA	0	0	Other
B1	Q92974	ARHGEF2	LISRI LQHSH GIEEE	LISRI LQASH GIEEE	0	0	Other
C1	Q6NXE6	ARMC6	RVPFG HAHNH AKMIV	RVPFG HAANH AKMIV	0.03	0	Interactants
D1		ARMC6	LTGAI THHGH HTDVV	LTGAI THAGH HTDVV	0	0	Interactants
E1	Q9NU22	MDN1	KLTQL ASGHS HGTFE	KLTQL ASGAS HGTFE	0.07	0	Interactants
F1	Q92614	MYO18A	LSSAI FKHQH KGGTL	LSSAI FKAQH KGGTL	0	0	Other
G1		MYO18A	SKPHH FLLGH SHGTN	SKPHH FLLGA SHGTN	0.68	0	Other
H1	O00159	MYO1C	VPTGE IIRVV HPHRP	VPTGE IIRVV APHRP	0	0	Other
I1		MYO1C	LEDTV KHHPH FLTHK	LEDTV KHAPH FLTHK	0	0	Other
J1	O94832	MYO1D	LNSKL GKHAH FSSRK	LNSKL GKAAH FSSRK	0.02	0	Other
A2	Q12965	MYO1E	QKLQM QIGSH EHFNS	QKLQM QIGSA EHFNS	0	0	Other
B2	Q7Z417	NUFIP2	YYFYN HSHNH HHHHH	YYFYN HSANH HAAHH	0.09	0	Other
C2		NUFIP2	PQHHH SHHHP HHHPQ	PQHHA SHHAP HHAPQ	0.19	0	Other
D2	Q8IY67	RAVER1	LSPGP NGHSH LLKVR	LSPGP NGASH LLKVR	0.26	0	Interactants
E2	Q13433	SLC39A6	LHHHH HQNHH PHSHS	LHHAA HQNHH PASHS	0.39	0	Interactants
F2	Q92504	SLC39A7	EDFHG HSHRH SHEDF	EDFHG ASHRA SHEDF	0.12	0	ProSeAM, interactants
G2	Q9ULF5	SLC39A10	DHSHS DGLHT IHEHD	DASHS DGLHT IAEHD	0	0	Interactants
H2	Q5T4S7	UBR4	HTSRS AYHSH KDQAL	HTSRS AYASH KDQAL	0.05	0.01	Other
I2	Q9NRG9	AAAS	PGPPP VLPHS PHSHL	PGPPP VLPHS PASHL	0.03	0	Other
J2	Q9NR19	ACSS2	AEAAV VGHPH PVKGE	AEAAV VGAPH PVKGE	0	0	Other
A3	Q8N187	CARF	TVNDI KNHIH EVQKS	TVNDI KNAIH EVQKS	0	0	Other
B3	Q9NXW2	DNAJB12	SQAAR HGHGH GDFHR	SQAAR HGAGH GDFHR	0.87	0	ProSeAM
C3		DNAJB12	SPRPS VGHIH RRVTD	SPRPS VGAIH RRVTD	0	0	ProSeAM
D3	Q6ZRV2	FAM83H	DFRFQ TSHFH RDQLY	DFRFQ TSAFH RDQLY	0	0	Other
E3	Q9H7Z3	NRDE2	EKKKK RKHQH HKKTK	EKKKK RKAQH HKKTK	0.01	0.01	Other
F3	O95263	PDE8B	TKNVH QSHSH LAMPI	TKNVH QSASH LAMPI	0.04	0	Other
G3	Q15149	PLEC	GGIID PVHSH RVPVD	GGIID PVASH RVPVD	0.01	0	Other
H3	Q15043	SLC39A14	QKNEH HHGHS HYASE	QKNEH HAGHS AYASE	0.04	0	Interactants
I3	Q9Y512	SAMM50	NKDVV VQHVH FDGLG	NKDVV VQAVH FDGLG	0	0	Other
J3	Q9UPW6	SATB2	AVSRL LAHQH PQAIN	AVSRL LAAQH PQAIN	0	0	Other
A4	Q9Y6X0	SETBP1	KHKEK QKHQH SEAGH	KHKEK QKAQH SEAGH	0	0	Other
B4	Q8TCJ2	STT3B	MALGN SRHGH HGPGA	MALGN SRAGH HGPGA	0.03	0	ProSeAM
C4	Q9C0I4	THSD7B	TFKHQ SYKAH HHSKS	TFKHQ SYKAA HHSKS	0.01	0	Other
D4	Q15643	TRIP11	TSALQ LEHEH LIKLN	TSALQ LEAEH LIKLN	0	0	Other
E4	P21397	MAOA	SDNII IETLN HEHYE	SDNII IETLN AEHYE	0	0	Other
F4	P51911	CNN1	LGEPA HNHHA HNYYN	LGEPA HNAAA HNYYN	0.04	0	ProSeAM
G4	O43676	NDUFB3	MAHEH GHEHG HHKME	MAHEA GHEAG HHKME	0	0	ProSeAM
H4	Q15427	SF3B4	GHPGH GHSHP HPFPP	GHPGH GASAP HPFPP	0.31	0	ProSeAM
I4		SF3B4	HGPHG LGHPH AGPPG	HGPHG LGAPH AGPPG	0.04	0	ProSeAM
J4	P08842	STS	LLVLS YLHVH TALFS	LLVLS YLAVH TALFS	0	0	ProSeAM
A5	Q8NEW0	SLC30A7	GHGHS HGSGH GHSHS	GHGAS HGSGH GASHS	0.05	0	ProSeAM, interactants
B5	P06702	S100A9	EGDEG PGHHH KPGLG	EGDEG PGAHH KPGLG	0	0	Other
C5	P31725	S100a9*	HENNP RGHGH SHGKG	HENNP RGHGA SHGKG	0.29	0	Other
D5	Q6ZMH5	SLC39A5	PGHQG HSHGH QGGTD	PGHQG HSAGH QGGTD	0.02	0	Interactants
E5	Q6P5W5	SLC39A4	GHSSH SHGGH SHGVS	GHSSA SHGGH SAGVS	0.27	0	Interactants
F5	Q9Y6M5	SLC30A1	SGHGH SHGGH GHGHG	SGHGA SHGGH GAGHG	0.81	0	Other
G5	Q8TAD4	SLC30A5	GHSDH GHGHS HGSAG	GHSDH GAGAS HGSAG	0.05	0	Other
H5	P48735	IDH2	KPITI GRHAH GDQYK	KPITI GRAAH GDQYK	0.33	0	Other
I5	A7E2V4	ZSWIM8	RLSPA HAHNH LRAPA	RLSPA HAANH LRAPA	0.13	0	Other
J5	Q14573	ITPR3	NQALL HKHLH LFLTP	NQALL HKALH LFLTP	0	0	Other
A6	Q6P087	RPUSD3	AQLPL HLHLH RLLLP	AQLPL HLALH RLLLP	0	0	Other
B6	O60563	CCNT1	NHHHH HNHHS HKHSH	NHHAA HNAHS AKHSA	0.38	0.01	Other
C6	O60583	CCNT2	HHTSS HKHSH SHSGS	HHTSS HKASH SASGS	0.06	0	Other
D6	P83436	COG7	MALLP HLHEH NLVKV	MALLP HLAEH NLVKV	0	0	Other
E6	P49773	HINT1	GQSVY HVHLH VLGGR	GQSVY HVALH VLGGR	0	0	Other
F6	Q9BX68	HINT2	AQSVY HLHIH VLGGR	AQSVY HLAIH VLGGR	0	0	Other
G6	Q9H1R3	MYLK2	AAARR GSPAF LHSPS	AAARR GSPAF LASPS	0	0	Other

*mouse protein

Supplementary Table 2: List of primers

Name	Sequence
pcDNA3-mMETTL9-F	AAAGAATTCACCATGAGACTGTTGGCGGGCTG
pcDNA3-mMETTL9-R	AAACTCGAGTACTGGTCTGAGAACAAAG
pET-mMETTL9-(22-)-F	AAACATATGTGGACGCTGCGGAGCCCGCTC
pET-mMETTL9-R	AAACTCGAGTTATACTGGTCTGAGAACAAAG
pQC-mMETTL9-cHA-F	AAAACCGGTACCATGAGACTGTTGGCGGGC
pQC-mMETTL9-cHA-R	AAAGGATCCCTAGATAGCGTAATCTGGAAC
mMETTL9-seq541	TGGCAGCTCCAGAAGAAGAA
mMETTL9-D151K/G153R-F	TTCTTAAGTTACGTGCTGGAGATGGAGAAGTC
mMETTL9-D151K/G153R-R	CCAGCACGTAACTTAAGAAGTCTATGGGTTTTC
pFBHT-Linearized vector-F	TAATGAGCCATGGGATCCGGAATT
pFBHT-Linearized vector-R	GCCCTGAAAATACAGGTTT
pFBHT-mMETTL9-22-F	CTGTATTTTCAGGGCTGGACGCTGCGGAGCCCGCTCT
pFBHT-mMETTL9-318-R	TCCCATGGCTCATTATACTGGTCTGAGAACAAA
pEGFP-N1_hMETTL9_F	GACTCAGATCTCGAGATGAGACTGCTGGCGGGC
pEGFP-N1_hMETTL9_R	GCGACCGGTGGATCCCGTACTGGTTTGAGAACAAAGACAGCGTC
pET28a_hMETTL9_F	GTGCCGCGCGGCAGCCATATGAGACTGCTGGCG
pET28a_hMETTL9_R	GTGGTGGTGGTGGTGCTCGAGTTATACTGGTTTGAGAACAAAGAC
hMETTL9_E174A_F	GAAATCTATGCCACTGCGCTTTCTGAAACTATG
hMETTL9_E174A_R	CATAGTTTCAGAAAGCGCAGTGGCATAGATTTC
pGEX-6P-2_hMETTL9_F	AGTCACGATGCGGCCGCTTATACTGGTTTGAGAACAAAGA
pGEX-6P-2_hMETTL9_R	AATTCCCGGGTCGACTATGAGACTGCTGGCGGGC
p3xFLAG-CMV-14_hMETTL9_F	AAGCTTGCGGCCGCGGCCACCATGAGACTGCTGGCG
p3xFLAG-CMV-14_hMETTL9_R	GTCAGCCCGGGATCCTACTGGTTTGAGAACAAAG
pET28a_ARMC6_F	GTGCCGCGCGGCAGCCATATGAGTGAACGATGTTGCTCTAGATACAG
pET28a_ARMC6_R	GTGGTGGTGGTGGTGCTCGAGTCATGGCGCCAGGTTGC
pET28a_ARMC6_H261R_F	CATGGTTGTGGGCACGGCCAAAGGGCACA
pET28a_ARMC6_H261R_R	TGTGCCCTTTGGCCGTGCCCACAACCATG
pET28a_ARMC6_A262G_F	GGCATGGTTGTGGCCATGGCCAAAGGG
pET28a_ARMC6_A262G_R	CCCTTTGGCCATGGCCACAACCATGCC
pET28a_ARMC6_H263R_F	ATCTTGGCATGGTTGCGGGCATGGCCAAAGG
pET28a_ARMC6_H263R_R	CCTTTGGCCATGCCCGCAACCATGCCAAGAT
pET28a_ARMC6_N264D_F	ATCATCTTGGCATGGTCGTGGGCATGGCCAAAG
pET28a_ARMC6_N264D_R	CTTTGGCCATGCCCACGACCATGCCAAGATGAT
pET28a_ARMC6_H265A_F	TTGGCCATGCCCACAACGCTGCCAAGATGATTGTGC
pET28a_ARMC6_H265A_R	GCACAATCATCTTGGCAGCGTTGTGGGCATGGCCAA
pET28a_ARMC6_A266G_F	CTGCACAATCATCTTGCCATGGTTGTGGGCATG
pET28a_ARMC6_A266G_R	CATGCCCACAACCATGGCAAGATGATTGTGCAG
pET28a_ARMC6_K267R_F	TTCTCCTGCACAATCATCCGGGCATGGTTGTGGGCATG
pET28a_ARMC6_K267R_R	CATGCCCACAACCATGCCCGGATGATTGTGCAGGAGAA
pGEX-6P-2_SLC39A7_31-137_F	GGGCCCCTGGGATCCCATGACGACCTGCACG
pGEX-6P-2_SLC39A7_31-137_R	ATGCGGCCGCTCGAGCTAATAAGCCCAGAGAGTGACAGC
pGEX-6P-2_DNAJB12_1-243_F	GGGCCCCTGGGATCCATGTCATCACTCCGCGC
pGEX-6P-2_DNAJB12_1-243_R	ATGCGGCCGCTCGAGCTACCCGCCATCACCCTG
pGEX-6P-2_DNAJB12_1-243_H185R_F	CCGGCACGGCCGTGGGCATGGGG
pGEX-6P-2_DNAJB12_1-243_H185R_R	CCCCATGCCCACGGCCGTGCCGG
pGEX-6P-2_CCNT1_F	GGGCCCCTGGGATCCATGGAGGGAGAGAGGAAGAAC
pGEX-6P-2_CCNT1_R	ATGCGGCCGCTCGAGTTACTTAGGAAGGGGTGGAAGTG
pGEX-6P-2_CCNT1_H519R/H520R/H523R/H524R/H528R_F	CCATCTAATCATCATCGTCGTCATAATCGCCGCTCACACAAGCGCTCTCATTCCCAACTTC

pGEX-6P-2_CCNT1_H519R/H520R/H523R/H524R/H528R_R	GAAGTTGGGAATGAGAGCGCTTGTGTGAGCGGCGATTATGACGACGATGATGATTAGATGG
pGEX-6P-2_hNDUFB3_F	GGGCCCCTGGGATCCATGGCCCATGAACATG
pGEX-6P-2_hNDUFB3_R	ATGCGGCCGCTCGAGTCAGTGATGCTTCTTATCTTTATTC
pGEX-6P-2_hNDUFB3_H5R/H9R_F	GGGCCCCTGGGATCCATGGCCCATGAACGTGGACATGAGCGTGGACATCATAAAATG
pcDNA3-hNDUFB3-F	AAAGAATTCACCATGGCCCATGAACATGGAC
pcDNA3-hNDUFB3-R	AAAGCGGCCGCCAGTGATGCTTCTTATCTTTA
pQC-hNDUFB3-cFLAG-F	AAAACCGGTACCATGGCCCATGAACATGGAC
pQC-hNDUFB3-cFLAG-R	AAATTAATTAACTAGAGCTTGTCATCGTCGTC
gRNA-hMETTL9-Intron2#1-F	CACCGTAATAAGTGATTATGGTTGT
gRNA-hMETTL9-Intron2#1-R	AAACACAACCATAATCACTTATTAC
gRNA-hMETTL9-Intron3#2-F	CACCGCAGTATTTTTCTGGAGCGG
gRNA-hMETTL9-Intron3#2-R	AAACCCGCTCCAGAAAAATACTGC
hMETTL9-genotype-F	GATCACGAGGGCAGGAGAT
hMETTL9-genotype-R	GTTGTCAGGGGTTACGAGGA