THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 12, pp. 6124–6133, March 18, 2016

Substrate Specificity of the HEMK2 Protein Glutamine Methyltransferase and Identification of Novel Substrates*[S]

Received for publication, December 21, 2015, and in revised form, January 11, 2016 Published, JBC Papers in Press, January 21, 2016, DOI 10.1074/jbc.M115.711952

Denis Kusevic, Srikanth Kudithipudi, and Albert Jeltsch[1]

From the Institute of Biochemistry, Faculty of Chemistry, University Stuttgart, 70569 Stuttgart, Germany

Bacterial HEMK2 homologs initially had been proposed to be involved in heme biogenesis or to function as adenine DNA methyltransferase. Later it was shown that this family of enzymes has protein glutamine methyltransferase activity, and they methylate the glutamine residue in the GGQ motif of ribosomal translation termination factors. The murine HEMK2 enzyme methylates Gln^{185} of the eukaryotic translation termination factor eRF1. We have employed peptide array libraries to investigate the peptide sequence recognition specificity of murine HEMK2. Our data show that HEMK2 requires a GQX_3R motif for methylation activity. In addition, amino acid preferences were observed between the −3 and +7 positions of the peptide substrate (considering the target glutamine as 0), including a preference for Ser, Arg, and Gly at the +1 and a preference for Arg at the +7 position. Based on our specificity profile, we identified several human proteins that contain putative HEMK2 methylation sites and show that HEMK2 methylates 58 novel peptide substrates. After cloning, expression, and purification of the corresponding protein domains, we confirmed methylation for 11 of them at the protein level. Transfected CHD5 (chromodomain helicase DNA-binding protein 5) and NUT (nuclear protein in testis) were also demonstrated to be methylated by HEMK2 in human HEK293 cells. Our data expand the range of proteins potentially subjected to glutamine methylation significantly, but further investigation will be required to understand the function of HEMK2-mediated methylation in proteins other than eRF1.

S-Adenosyl-L-methionine (AdoMet)-dependent[2] methyltransferases (MTases) are known to transfer methyl groups to proteins, nucleic acids, carbohydrates, lipids, and a variety of small molecules (1, 2). A remarkable diversity of substrates exists also within the class of protein MTases that modify different amino acids including lysine, arginine, glutamine, glutamate, histidine, asparagine, aspartate, and cysteine, as well as N-terminal and C-terminal residues (2). In general, protein methylation affects various cellular processes including protein stability, protein-protein interactions, and protein localization. Indirect effects may be mediated by an influence on other post-translational modifications (2–4). Although protein methylation on lysine or arginine residues has been intensively studied (2–7), the investigation of methylation of other amino acids including glutamine is still at the inception. Glutamine methylation was first described almost 40 years ago as a modification present in the *Escherichia coli* ribosomal protein L3 (8). Later, ribosomal polypeptide release factors (RFs), which recognize stop codons in mRNA in the A site of the ribosome were shown to contain the same modification as well (9). Recently, glutamine methylation was identified in the histone H2A protein at Gln^{104} (10).

The HemK enzyme was first identified in *E. coli* in a genetic screen designed to isolate new heme biosynthesis mutants (11). The HemK protein at that time was thought to be involved in the oxidation of protoporphyrinogen to protoporphyrin IX. Despite the mutant phenotype, subsequent biochemical and genetic studies revealed that the *hemK* gene product appeared to have no direct involvement in the heme biosynthetic pathway (12). It is still unclear what exactly caused the defects in heme metabolism observed in the *hemK* mutants. Later, HemK homologs were found in various species including bacteria and lower and higher eukaryotes, and sequence alignments of HemK proteins revealed that all of them share a common (D/N/S)PP(Y/FW) amino acid sequence motif, which is characteristic for adenine-N6 and cytosine-N4 specific DNA MTases (PROSITE accession no. PS00092) (13, 14). This led to the suggestion that HemK was itself an AdoMet-dependent DNA MTase (15), and its classification in data bases as "probable MTase modifying N6-adenine or N4-cytidine in DNA" including the renaming of many entries to N6AMT. However, no evidence could be found that HemK is able to methylate DNA (16, 17). Later it was found that *E. coli* HemK methylates a glutamine residue in the ribosomal release factors (RF1 and RF2) at the universally conserved tripeptide GGQ motif (18, 19), indicating that the alignment based assignment of HemK as an AdoMet dependent MTase was correct although the predicted substrate was wrong. This novel activity triggered renaming of HEMK2 to PrmC (for protein methyltransferase C). Glutamine methylation of the polypeptide release factor stimulates the hydrolysis of the peptidyl-tRNA in the ribosomal P-site, leading to the release of the nascent polypeptide chain and ensuring efficient translational termination (18). The HemK enzyme has two homologs in many eukaryotes, HEMK1 in mitochondria and HEMK2 in the cytosol. The human HEMK2, also called N6AMT1, and its yeast homolog called YDR140w or Mtq2p were shown to methylate the glutamine residue in the GGQ motif of the eukaryotic release factor eRF1

* This work was supported by Deutsche Forschungsgemeinschaft Grant JE 252/25-1. The authors declare that they have no conflicts of interest with the contents of this article.
[S] This article contains supplemental Table S1.
[1] To whom correspondence should be addressed: Inst. of Biochemistry, Faculty of Chemistry, University Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart, Germany. Tel.: 49-711-685-64390; Fax: 49-711-685-64392; E-mail: albert.jeltsch@ibc.uni-stuttgart.de.
[2] The abbreviations used are: AdoMet, *S*-adenosyl-L-methionine; MTase, methyltransferase; RF, release factor.

(17, 20, 21). The human HEMK1 (HMPrmc) and its yeast homolog Mtq1p methylate the corresponding site in the mitochondrial release factor and regulate the mitochondrial translational activity (22).

The first structures of HemK MTases have been determined from *Thermotoga maritima* (23) and *E. coli* (16). Closely similar, the structures revealed two domains: an N-terminal domain with no significant similarity to MTase sequences other than close HemK homologs and a C-terminal domain that contains the (D/N/S)PP(Y/FW) motif and shows strong similarity to members of the family 7 β-strand methyltransferases (also referred as class I methyltransferases). This group of enzymes contains the majority of AdoMet-dependent methyltransferases in most organisms, including DNA adenine and DNA cytosine MTases, RNA MTases, small molecule MTases, and protein arginine and some protein lysine methyltransferases (24, 25). The structure of *E. coli* HemK in complex with the bacterial RF1 and the methyl donor product *S*-adenosyl-L-homocysteine was also solved (26), showing that both the GGQ part and the central region of RF1 interact with the enzyme.

In eukaryotes, HEMK2 alone is not sufficient to methylate eRF1; it requires the binding partner TRM112 to exhibit high methyltransferase activity (21). Structural studies of the yeast homologs, Mtq2 and Trm112, revealed that Mtq2 binds to the zinc binding domain of Trm112 forming a heterodimer. This interaction triggers a conformational change of Mtq2, which increases the binding of AdoMet (27). In addition to that, Trm112 binding masks a hydrophobic region of Mtq2, and this prevents Mtq2 homodimer formation and enhances the expression of the enzyme in soluble form (27). In mice, HEMK2 knock-out heavily impaired postimplantation development of mutant embryos, and it led to early embryonic lethality indicating that HEMK2 is an essential protein in mammals (17). However, the mechanistic basis of the interaction of HEMK2 with its eRF1 substrate is not fully understood. In this context, the structure of the *E. coli* enzyme in complex with its cognate release factor is not of help, because the amino acid sequences of *E. coli* and mammalian release factors are distinct outside of the conserved GGQ motif. Therefore, the aim of this study was to characterize the substrate specificity of murine HEMK2. Based on this, we also examined whether HEMK2 has additional substrates beyond eRF1, similarly to other protein MTases, which often have several methylation substrates. We identified several new substrates of HEMK2 at the peptide and protein level *in vitro* and confirmed methylation of two of them in human calls.

Experimental Procedures

Cloning of Proteins—A bacterial expression pRSF-Duet1 vector that encodes the His-tagged murine HEMK2 (N6AMT1) and untagged TRM112, a pRSET vector encoding human eRF1, and mammalian expression constructs of HA-tagged HEMK2 and Myc-tagged TRM112 were kindly provided by Dr. G. L. Xu (17). The coding sequences of other putative substrate protein domains (see Table 2) were amplified from cDNA prepared from HEK293 cells and cloned into the pGEX-6P-2 vector as GST fusion proteins. The coding sequences of CHD5 (kindly provided by Dr. A. A. Mills), protein NUT (kindly provided by Dr. C. A. French), and eRF1 were subcloned into pEYFP-C1 vector (Clontech).

Expression and Purification of Proteins—For bacterial expression, *E. coli* BL21-CodonPlus (DE3) cells (Novagen) were transformed with the corresponding plasmid, and the cells were grown in Luria-Bertani medium at 37 °C until an A_{600} of 0.6–0.8 was reached. The bacterial cultures were then either shifted to 30 °C and induced with 1 mM isopropyl β-D-thiogalactopyranoside for 4 h or to 20 °C and induced overnight (10–12 h) with 1 mM isopropyl β-D-thiogalactopyranoside. Later, the cells were collected by centrifugation at 5,000 × *g*.

For purification of HEMK2 and TRM112, the cell pellet was resuspended in sonication buffer (30 mM potassium P_i buffer, pH 7.4, 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, 20 mM imidazole, and 10% glycerol), and the cells were disrupted by sonication. The samples were centrifuged at 20,000 × *g* for 90 min, and supernatants were passed through nickel-nitrilotriacetic acid-agarose beads (Qiagen) that were preincubated with the sonication buffer. The beads were washed with sonication buffer. The bound proteins were eluted with the sonication buffer containing 220 mM imidazole. Proteins were dialyzed against low glycerol dialysis 1 buffer (20 mM HEPES, pH 7.2, 200 mM KCl, 0.2 mM DTT, 1 mM EDTA, and 10% glycerol) for 3 h and then overnight against high glycerol dialysis 2 buffer containing 65% glycerol.

For purification of GST-tagged substrate proteins, the cell pellet was resuspended in sonication buffer (50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 1 mM DTT, and 5% glycerol), and the cells were disrupted by sonication. The samples were centrifuged at 20,000 × *g* for 90 min, and supernatants were passed through glutathione-Sepharose 4B resin (GE Healthcare) which were preincubated with the sonication buffer. The beads were washed with wash buffer (50 mM Tris/HCl, pH 8.0, 500 mM NaCl, 1 mM DTT, and 5% glycerol). The bound proteins were eluted with the wash buffer containing 40 mM glutathione. Proteins were dialyzed against low glycerol dialysis 1 buffer (20 mM Tris/HCl, pH 7.4, 100 mM KCl, 0.5 mM DTT, and 10% glycerol) for 3 h and then overnight against high glycerol dialysis 2 buffer (20 mM Tris/HCl, pH 7.4, 100 mM KCl, 0.5 mM DTT, and 60% glycerol).

Cell Culture, Transfection, and Immunoprecipitation—HEK293 cells were grown in Dulbecco's modified Eagle's medium (Sigma) supplemented with 5% fetal bovine serum, penicillin/streptomycin, and L-glutamine (Sigma). The mammalian expression plasmids of HA-HEMK2, myc-TRM112, and YFP-tagged substrate proteins were transfected into cells using polyethylenimine (Promega) according to the manufacturer's instructions. In transfections without HEMK2 and TRM112 expression plasmids, corresponding amounts of empty pcDNA3.1 vector were added. After 72 h of transfections, the cells were washed with PBS buffer and harvested by centrifugation at 500 × *g* for 5 min. The YFP-fused substrate proteins were immunoprecipitated from the cell extract using GFP-Trap A (Chromotek) following the manufacturer's instructions.

Synthesis of Peptide SPOT Arrays—Peptide arrays were synthetized using the SPOT synthesis method using an Autospot peptide array synthesizer (Intavis AG) (28). Each peptide spot had a diameter of 2 mm and contained ~9 nmol of peptide

(Autospot Reference Handbook; Intavis AG). Successful synthesis of peptide arrays were confirmed by bromphenol blue staining of the membranes. Data analysis and calculation of discrimination factors were conducted as described (7, 29).

Peptide Methylation Assays—Peptide SPOT arrays were preincubated for 10 min in methylation buffer containing 10 mM Tris/HCl, pH 7.6, 50 mM KCl, 10 mM $Mg(OAc)_2$, and 1 mM DTT. Afterward the SPOT arrays were incubated for 60 min in methylation buffer containing 1.3 μM HEMK2 and 0.76 μM radioactively labeled [methyl-^{3}H]AdoMet (PerkinElmer Life Sciences) at room temperature (25 °C). The arrays were then washed five times for 5 min with 50 mM NH_4HCO_3 and 1% SDS. Later the SPOT arrays were incubated for 5 min in Amplify NAMP100V solution (GE Healthcare). Finally, the arrays were exposed to Hyperfilm™ high performance autoradiography films (GE Healthcare) in the dark for 1–3 days at −80 °C. The films were developed using an Optimus TR developing machine.

In Vitro Methylation of Protein Domains—Protein methylation reactions were performed in methylation buffer containing 10 mM Tris/HCl, pH 7.6, 50 mM KCl, 10 mM $Mg(OAc)_2$, 1 mM DTT supplemented with 0.76 μM radioactively labeled [methyl-^{3}H]AdoMet and 4,5 μM HEMK2 overnight at 25 °C. The methylation reaction was stopped by incubating the sample in SDS loading buffer at 95 °C for 5 min. Afterward, the samples were separated by SDS-PAGE, and the methylation signal was detected either by autoradiography or using the methylglutamine specific antibody.

Generation of the Methylglutamine Antibody—A methylglutamine antibody was generated by the Biotem (Apprieu, France) following standard procedures. In short, a GQ(me)G tripeptide was synthesized and conjugated to the keyhole limpet hemocyanin protein as carrier. New Zealand White rabbits were immunized with the conjugated protein in four subsequent injections at days 0, 7, 14, and 42. Preimmune serum and serum after days 49 and 63 were collected. The quality of unpurified anti-methylglutamine serum was initially assessed by ELISA using the modified and unmodified peptides using preimmune serum as control. After the initial screening of the serum in ELISA, day 63 serum was collected. Because we still observed some residual binding to unmethylated CHD5 and NUT proteins, it was further purified by passing through GST-Sepharose (GE Healthcare) beads bound with the unmethylated recombinant proteins. The flowthrough was collected, and specific binding to glutamine-methylated eRF1, CHD5, and NUT proteins was validated as described in the text with the methylated and unmethylated proteins. The final dilution of the methylglutamine antibody used in the Western blots was 1:10,000.

Results

Purification and Activity of HEMK2—The His-tagged mouse HEMK2 protein and the mouse TRM112 homolog were coexpressed in *E. coli* BL21-CodonPlus cells and purified with good yield (Fig. 1*A*). Because the eukaryotic translation termination factor eRF1 was the only known HEMK2 substrate, eRF1 peptides were used to test the activity of the purified enzyme. Peptide SPOT arrays were synthesized on cellulose membrane with

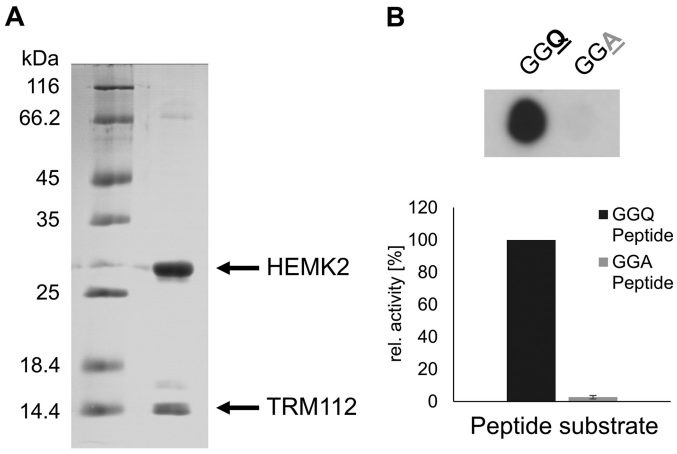

FIGURE 1. **Purification and activity analysis of HEMK2.** *A*, purification of the HEMK2-TRM112 complex. The image shows a Coomassie-stained SDS gel of the purified proteins; the HEMK2 and TRM112 proteins are marked with *arrows*. *B*, activity of the HEMK2-TRM112 complex tested on 15-amino acid-long peptides with eRF1 178–192 sequence (KKHGRGGQSALRFAR) and an eRF1 178–192 Q185A mutant sequence (KKHGRGGASALRFAR). The *upper panel* shows an autoradiography image that indicates the transfer of radioactively labeled methyl groups to the peptides. The bar diagram represents the quantitative analysis of three independent experiments, the *error bar* represents the standard error of the mean.

15-amino acid peptides comprising the eRF1 residues 178–192 surrounding the methylation target site Gln^{185}. Control peptides were used in which Gln^{185} was mutated to alanine. The peptide arrays were then incubated with the HEMK2-TRM112 complex and radioactively labeled [methyl-^{3}H]AdoMet. The transfer of the methyl groups to the peptide was detected by autoradiography. As shown in Fig. 1*B*, a clear methylation signal was observed with the wild type but not with the mutant peptides, which confirms the specific methylation of Gln^{185} in the eRF1 peptide.

Specificity Analysis of HEMK2—To examine the specificity of HEMK2 in peptide recognition, mutational scanning peptide SPOT arrays were synthesized using the sequence of eRF1 (179–192) as a template. In these arrays, each peptide contained an exchange of one amino acid of the original sequence against one of the 20 proteinogenic amino acids, such that all possible single exchange variants of the original sequence are investigated. In total, 300 peptides were synthetized (20 possible amino acids × 14 residues plus one wild type peptide at the beginning of each row). The peptide arrays were then incubated with the HEMK2-TRM112 complex and radioactively labeled [methyl-^{3}H]AdoMet, and the methylation of each peptide was detected by autoradiography as described above (Fig. 2*A*). The advantage of this method is that all peptides are methylated in competition, so that during the linear phase of the methylation reaction, the relative methylation rate of each peptide is proportional to the catalytic single turnover rate constant divided by the dissociation constant (k_{st}/K_D) of the enzyme for the respective peptide substrate (7, 29). As a result, the amount of radioactivity transferred to each peptide spot is directly correlated with the enzymatic activity toward that particular peptide. The experiment was repeated three times, and the results of the individual experiments were normalized and averaged (Fig. 2*B*). Standard deviations indicate that the obtained data are of good quality, because 70% of the peptide substrates exhibited

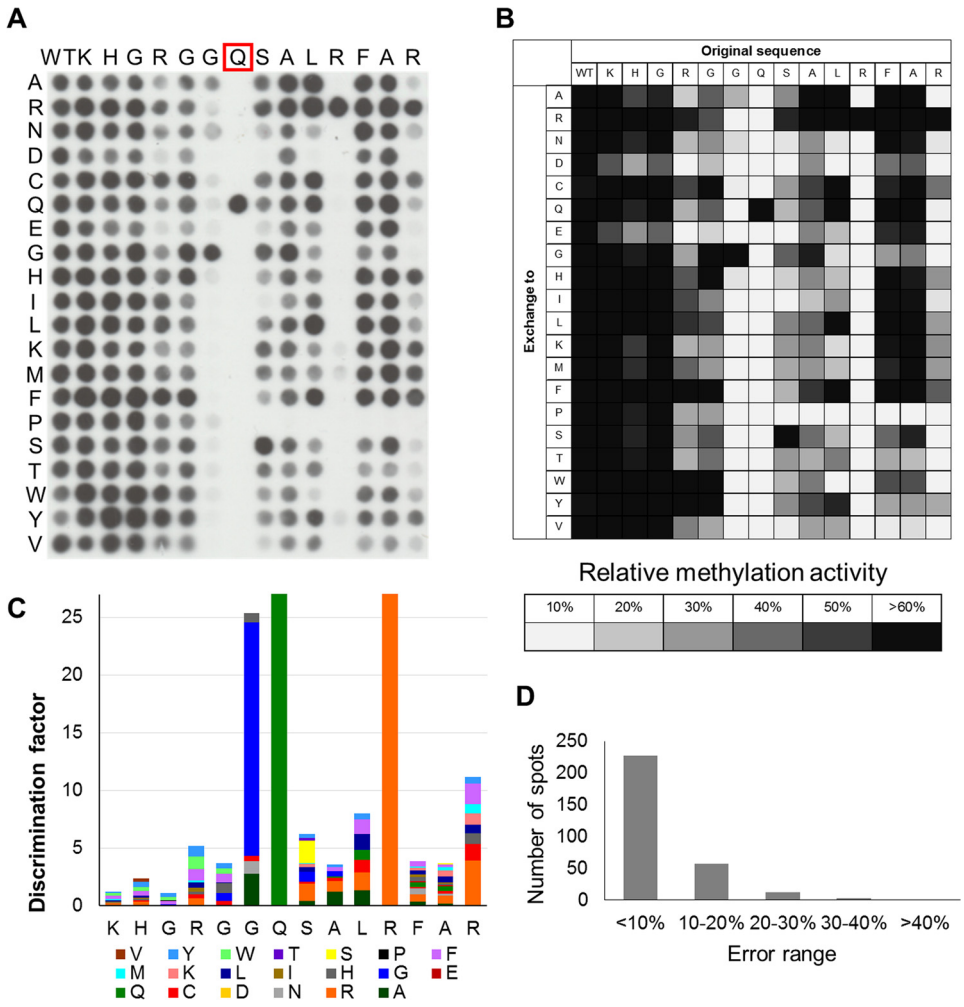

FIGURE 2. **Substrate sequence motif determination of HEMK2.** *A*, example of one complete eRF1 (positions 179–192) peptide array methylated by HEMK2-TRM112. The *horizontal axis* displays the eRF1 sequence, the target Gln^{185} is highlighted with a *red square*. The *vertical axis* represents the residues introduced at the corresponding positions in the original sequence, so that the array contains all possible single amino acid exchanges of the eRF1 sequence. The first column labeled WT contains the original eRF1 (positions 179–192) sequence as reference. The image shows an autoradiography representing the relative efficiency of HEMK2 methylation of each peptide. *B*, compilation of the results of three independent peptide array experiments with HEMK2. The data were averaged after normalizing full activity to 1. Strong to weak methylation is scaled from *black* to *white*. *C*, bar diagram showing the discrimination factors of HEMK2 for the recognition of amino acids at the tested positions. *D*, distribution of the standard errors of the mean of HEMK2 activity on all peptides for the data shown in *B*.

S.D. less than 10%, and more than 95% of the peptides exhibited S.D. less than 20% (Fig. 2*D*). Next, to assess the contribution of each amino acid of the peptide substrate, the discrimination factor for the recognition of each amino acid residue at each position was calculated (Fig. 2*C*) (7, 29). These factors represent the preference of the enzyme for one particular amino acid over all other amino acids at a given site. The results show that HEMK2 has long recognition motif encompassing the residues from Arg^{182} (−3) to Arg^{192} (+7) of eRF1 (defining the target Gln^{185} as position 0). In addition to the target residue Gln^{185}, HEMK2 exhibits absolute specificity at the −1 (Gly^{184}) and +4 positions (Arg^{189}). At these sites, an exchange of the cognate amino acid residue with other amino acids abolished the methylation activity. Hence, the minimal recognition element of HEMK2 is a GQX_3R motif. In addition, stronger preferences were also seen at +1 (Ser^{186}, preference for Ser, Arg, and Gly) and +7 positions (Arg^{192}, preference for Arg). At the −3 (Arg^{182}), −2 (Gly^{183}), +2 (Ala^{187}), and +3 positions (Leu^{188}), some weaker preferences were observed. It is noticeable that proline is not tolerated at positions 184–190. Similarly, Asp and Glu are not preferred in the tested sequence, with exception of position 191.

Methylation of Putative Peptide Substrates—Lysine and arginine methyltransferases were initially identified as histone modifying enzymes, but some of them later were found to methylate other proteins and play an important role in various cellular functions (2–4, 7). The HEMK2 substrate specificity profile shows that it accepts several other residues on both sides of the target glutamine apart from the cognate residues, which indicates that it could also methylate other proteins apart from eRF1. To identify putative candidate substrates, Scansite (30) searches were performed with search profiles as indicated in Table 1. In total, 138 putative novel substrates were identified in the human proteome with substrate specificity profile 1. A second search with a more relaxed specificity profile (profile 2) identified 164 additional substrates candidates. Methylation of these putative substrates was initially tested on peptide arrays. For this, 15-amino acid peptides were synthesized containing

TABLE 1
Substrate specificity profiles used in Scansite searches to identify putative novel HEMK2 substrates

Position	−1	0	+1	+2	+3	+4
Cognate residue	Gly^{184}	Gln^{185}	Ser^{186}	Ala^{187}	Leu^{188}	Arg^{189}
Search profile 1 (stringent)	Gly	Gln	SRYKLG	ARFGLWYCS	LARQCFYT	Arg
Search profile 2 (relaxed)	Gly	Gln	SRYKLGAMTC	ARFGLWYCSQKH	LARCQFYTI	Arg

the target glutamine in the center. Two independent SPOT arrays with the novel substrates from the corresponding profiles were synthesized always including eRF1 and eRF1 Q185A as positive and negative controls. The peptide arrays were methylated with HEMK2-TRM112 as described above (Fig. 3, *A* and *B*). Quantitative analysis of the results revealed that 49 of the putative substrates were methylated with an activity comparable with eRF1, and 76 peptide substrates showed a reduced but clearly detectable activity. The high frequency of methylated peptides in these libraries confirms the validity of the search motifs. Moreover, most of the strongly methylated peptides contained an Arg at the +7 position (corresponding to Arg^{192}), although this was not included in the search parameters. Furthermore, most of the highly methylated peptides contained an Ser, Arg, or Gly at the +1 position, although other residues were allowed at this position in the search profiles. In summary, HEMK2 methylates several novel peptide substrates *in vitro*, and the sequences of the strongly methylated peptides confirmed the specificity data obtained in the mutational scanning approach.

Methylation of Putative Protein Substrates—Although several novel peptide substrates were methylated in par with the eRF1 peptide, this cannot be extrapolated to protein methylation, because the target glutamine may not be accessible to the enzyme in a folded protein. Therefore, we next aimed to investigate the methylation of these substrates at the protein level. 58 proteins were selected that were strongly methylated at the peptide level, and the corresponding protein domains containing the target glutamine residue were cloned as GST fusion proteins. The candidate proteins were overexpressed and purified by affinity chromatography. Finally, 20 protein domains were successfully purified with sufficient yield for a methylation assay. Methylation reactions were performed by incubating the purified protein domains with the HEMK2-TRM112 complex in methylation buffer containing radioactively labeled [methyl-^{3}H]AdoMet. The samples were later separated by polyacrylamide gel electrophoresis, and the transfer of methyl groups to the substrate proteins was detected by autoradiography. As shown in Fig. 4*A*, methylation was detected on 11 of the 20 protein domains (Table 2). Four protein domains (CHD5, NUT, ANKRD34A, and ABCA2) were methylated in par with the eRF1, and six additional domains were methylated at lower level (GHDC, ZSCAN10, RPP1, ASH1L, ARGEF10, and TGFB3). In addition, strong methylation was observed on AMPD2, but it could only be obtained in a highly degraded form. Site-directed mutagenesis was performed to confirm the methylation of the predicted target glutamine. For this, the target glutamine was exchanged to arginine in all 11 protein domains. The mutated protein domains were expressed, purified, and analyzed for methylation using equal amounts of wild type and the target glutamine mutant protein domains (Fig. 4*B*). HEMK2 exhibited no activity on all the target glutamine mutant proteins, confirming methylation at the predicted target glutamine residues.

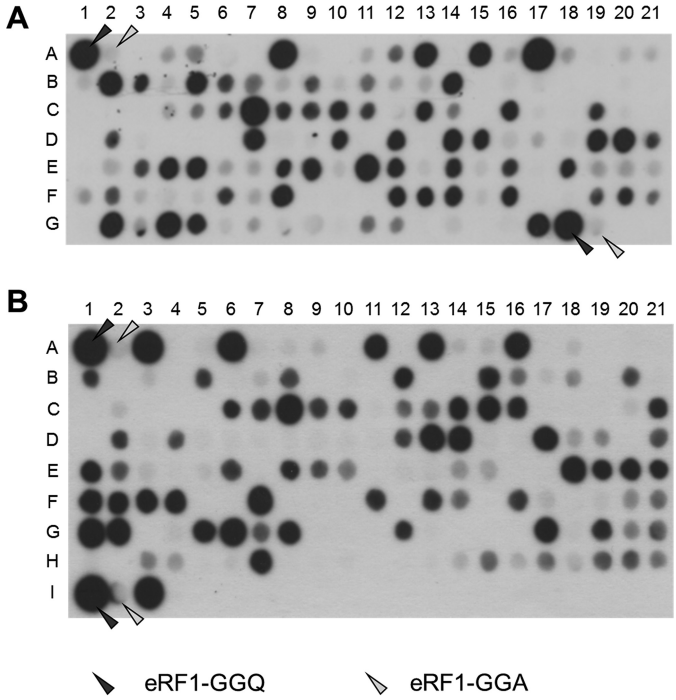

FIGURE 3. **Methylation of novel peptide substrates by HEMK2.** 15-amino acid peptides containing the predicted target glutamine were synthesized on peptide arrays and methylated with HEMK2-TRM112. The names of the proteins and peptide sequences are given in supplemental Table S1). *A*, the potential substrates identified using specificity profile 1. *B*, the potential substrates identified using specificity profile 2 (Table 1). The eRF1 wild type and the Q185A mutant peptides were included as controls and marked with *black* and *gray arrowheads*.

HEMK2 Methylates Novel Protein Substrate in HEK293 Cells—After confirming target glutamine methylation of several protein domains *in vitro*, we next aimed to investigate their methylation in cells. For this purpose, we chose two proteins that were strongly methylated *in vitro* (CHD5 and NUT). For detecting cellular methylation, a methylglutamine specific antibody was generated. Validation of the antibody was performed using *in vitro* methylated and unmethylated proteins and protein domains of eRF1, CHD5, and NUT. Methylated substrate proteins were generated by incubating the proteins with the HEMK2-TRM112 complex in methylation buffer in presence of unlabeled AdoMet. Equal amounts of unmethylated substrate proteins were used as negative control. The proteins were separated on SDS-PAGE, transferred to nitrocellulose membrane, and probed with the antibody. After purification with unmethylated proteins, the antibody specifically recognized the methylated substrate proteins, which indicates that the antibody can be used to detect the methylated glutamine (Fig. 5*A*).

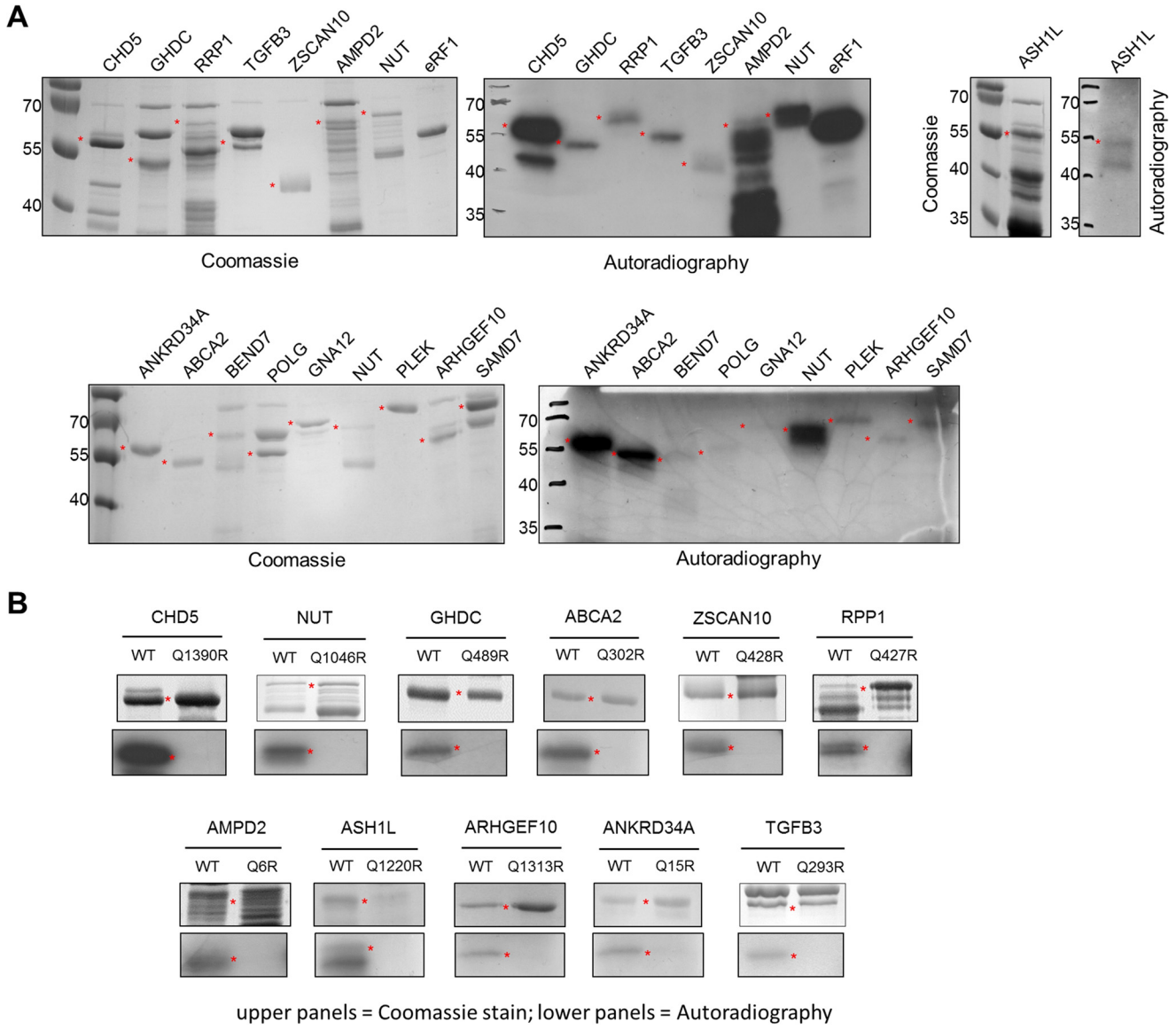

FIGURE 4. **HEMK2 methylates several novel substrates at the protein level.** Details of the substrate proteins are given in Table 2. *A*, detection of protein methylation. Purified GST-fused protein domains were incubated with the HEMK2-TRM112 complex in the presence of radioactively labeled [methyl-^{3}H]AdoMet, and the transfer of methyl groups to the substrate proteins was detected by autoradiography. The Coomassie-stained gel image shows the amount of substrate proteins used in the methylation assays. As a positive control, methylation of eRF1 was included. *B*, confirmation of methylation of the predicted target glutamine. Purified GST-fused WT and target Glu to Arg mutant proteins were incubated with the HEMK2-TRM112 complex in the presence of radioactively labeled [methyl-^{3}H]AdoMet, and the methylation signal was detected by autoradiography. The Coomassie-stained images show the equal loading of wild type and corresponding Gln to Arg mutants.

To investigate the methylation of substrate proteins by HEMK2 in human cells, the YFP-fused substrate protein domains were transiently expressed in HEK293 cells either alone or together with HEMK2 and TRM112. We first confirmed the expression of the HA-tagged HEMK2 (Fig. 5*B*) and Myc-tagged TRM112 (Fig. 5*C*). Next, the YFP fused substrate protein domains were immunoprecipitated from cells using GFP trap beads. Substrate proteins purified from the cells with and without HEMK2-TRM112 were separated on SDS-PAGE and analyzed by Western blot with anti-YFP antibodies to confirm their identity and adjust the protein concentrations of the samples (Fig. 5*D*). Next, methylation of the target proteins was analyzed using the methylglutamine antibody. Increasing concentrations of recombinant unmethylated protein were loaded as negative controls to exclude any cross-reactivity of the antibody with unmethylated proteins in these immunoblots (Fig. 5, *E* and *F*). In both experiments, the methylglutamine antibody specifically bound to the substrate proteins purified from the HEK293 cells, and no binding to recombinant unmethylated substrates was detected even at high protein concentrations. With CHD5, methylation could only be detected after coexpression of HEMK2 and TRM112, indicating that a HEMK2-dependent methylation occurred in the human cells. In the case of NUT, weak methylation was observed in the absence of ectopically expressed HEMK2 and TRM112, which likely was due to the endogenous enzymes. The methylation signal increased strongly after coexpression of NUT with HEMK2 and TRM112, again indicating that a HEMK2-dependent methylation took place.

TABLE 2

List of putative novel HEMK2 substrate proteins, which were selected to test methylation at the protein level

Names, abbreviations, and boundaries of the protein domains used in this study are indicated, as well as the position of the target glutamine. The corresponding peptide sequences are given in supplemental Table S1.

Name	Abbreviation	Swiss Prot no.	Cloned domains	Target Gln
Chromodomain helicase DNA-binding protein 5	CHD5	Q8TDI0	1234–1530	1390
Ribosomal RNA processing protein 1 homolog A	RRP1	P56182	219–461	427
AMP deaminase 2	AMPD2	Q01433	2–135	6
GH3 domain-containing protein	GHDC	Q8N2G8	325–529	489
Protein NUT	NUT	Q86Y26	867–1132	1046
Transforming growth factor β3	TGFB3	P10600	159–405	293
Zinc finger and SCAN domain-containing protein 10	ZSCAN10	Q96SZ4	364–521	428
ATP-binding cassette subfamily A member 2	ABCA2	Q9BZC7	168–403	302
Ankyrin repeat domain-containing protein 34A	ANKRC34A	Q69YU3	5–235	15
Rho guanine nucleotide exchange factor 10	ARHGEF10	O15013	1107–1343	1313
Histone-lysine *N*-methyltransferase ASH1L	ASH1L	Q9NR48	1119–1333	1220
BEN domain-containing protein 7	BEND7	Q8N7W2	9–282	78
DNA polymerase subunit γ1	POLG	P54098	154–387	330
Guanine nucleotide-binding protein subunit α12	GNA12	Q03113	183–320	231
Pleckstrin	PLEK	P08567	2–350	107
Sterile α motif domain-containing protein 7	SAMD7	Q7Z3H4	71–416	179

Discussion

Methylation on lysine and arginine residues of histone proteins has been intensively studied, and its significant role in chromatin regulation has been well established (31, 32), but methylation of these residues also occurs on non-histone proteins, and it has emerged as an important post-translational modification affecting several protein functions like stability, activity, protein-protein interactions (2, 3). However, protein methylation is not limited to lysine and arginine residues; other amino acid residues like cysteine, histidine, and glutamine have also been reported to carry this modification. The first discovered mammalian protein glutamine methyltransferase was HEMK2, which methylates the glutamine residue in the GGQ motif of the eRF1 protein. Recently, Kouzarides and co-workers (10) showed that the Nop1 methyltransferase in yeast, and its human homolog Fibrillarin, methylate H2A protein at Gln^{105} at the 35S ribosomal DNA transcriptional unit. Hence, overall so far only two proteins with a methylglutamine modification are described in the human proteome. Herein, we investigated the substrate specificity profile of the HEMK2 MTase. For this purpose, we used a strategy employing peptide libraries that was highly successful with different protein lysine methyltransferases (33–35). Our results describe the specificity of HEMK2, and we identified several novel protein substrates of this enzyme.

Methylation of peptides with point exchanges in the sequence of eRF1 (179–192) showed that HEMK2 recognizes a long peptide stretch ranging from the −3 to the +7 position. At some sites, clear preferences for the canonical residues could be detected most prominently at the −1 (Gly^{184}) and +4 (Arg^{189}) positions, but also at +1 (Ser^{186}) and +7 (Arg^{192}). Further structural and mechanistic interpretation of these data awaits the successful determination of a HEMK2-TRM112 structure in complex with eRF1 or an eRF1 peptide. However, our data also show that HEMK2 accepts other residues apart from the native sequence at certain positions. Using the derived specificity profile, several potential novel substrates were identified in the human proteome, and we showed that 58 peptides derived from them were efficiently methylated by HEMK2. Most of the highly methylated substrates had an Arg at the +7 position and the preferred residues at the +1 position, although this information was not included in the search filter. This result shows that the sequence motif of HEMK2 remains similar irrespective of the template sequence used for the specificity analysis (eRF1 in our case). This finding indicates that our peptide array data generated by variation of the eRF1 sequence can be extrapolated to other sequences.

We next tested methylation of several of the identified candidate substrates at the protein level; however, many of them failed at different stages of cloning, expression, and purification. Finally, we were able to purify 20 protein domains and detected methylation signals on 11 of them; strong methylation comparable with eRF1 was observed with four proteins (CHD5, NUT, ANKRD34A, and ABCA2); and weaker methylation signals were observed on 7 other proteins. Based on this, one may extrapolate that approximately 8 additional strong methylation substrates should be among the 38 additional highly methylated peptides that could not be tested. Furthermore, we tested the methylation of two proteins (CHD5 and NUT protein) in HEK293 cells and could confirm methylation of both of them, suggesting that strong *in vitro* protein domain methylation is indicative of cellular methylation in this system. Hence, our findings expand the range of human proteins potentially subjected to glutamine methylation considerably. While so far only two human proteins were known to be glutamine-methylated, we now identified 2 more HEMK2 substrate proteins in cellular studies, 9 additional ones *in vitro* at the protein level, and more than 50 additional peptide substrates. Based on the extrapolation described above, our data suggest that HEMK2 has at least approximately 13 strong methylation substrates in human cells (eRF1, 4 candidates identified here at protein level and 8 more candidates expected among the methylated peptides), which suggests that glutamine methylation is more widespread than currently known.

However, the biological significance of the methylation of the novel HEMK2 substrates still needs to be studied. Of note, glutamine methylation is a subtle post-translational modification. In case of eRF1, glutamine methylation changes the properties of the catalytic pocket directly; in case of methylation of H2A, protein-protein interaction is modulated. It is not known at present, if reading domains exist that specifically interact with glutamine-methylated proteins similarly as the well known

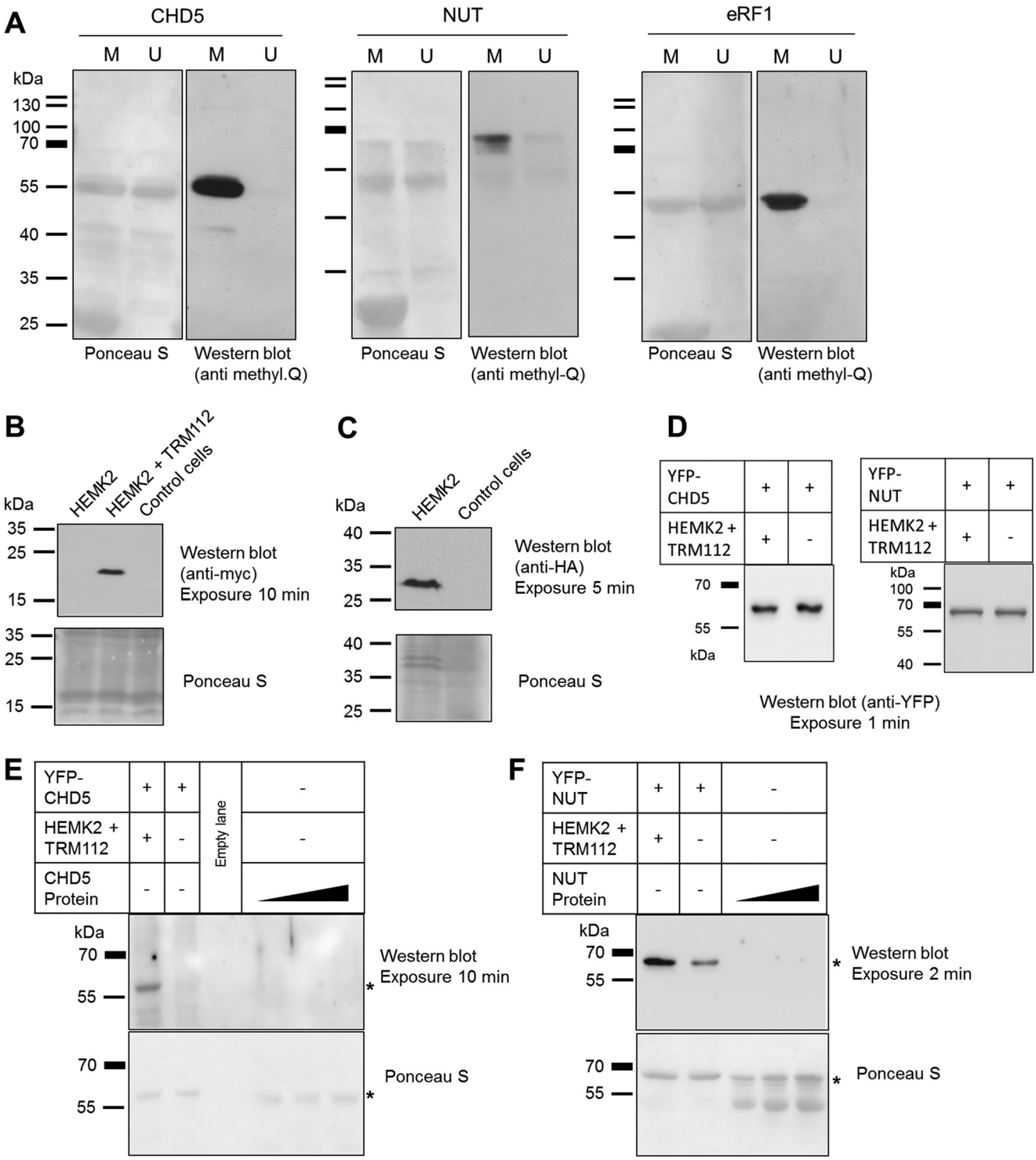

FIGURE 5. **Detection of HEMK2-mediated glutamine methylation of CHD5 and NUT in HEK293 cells.** *A*, validation of the methylglutamine antibody with the *in vitro* methylated CHD5, NUT, and eRF1 proteins, which were prepared by incubation of the proteins with the HEMK2-TRM112 complex in the presence of unlabeled AdoMet. Ponceau S indicates the equal loading of unmethylated (*U*) and methylated (*M*) proteins. Western blots were probed with the methylglutamine specific antibody, showing specific detection of the methylated proteins. *B*, validation of TRM112 expression in HEK293 cells. The cells were transfected with either HEMK2 vector or HEMK2 and TRM112 vector together. Untransfected cells were used as control. Western blot with anti-Myc antibody showed expression of the Myc-tagged TRM112. *C*, validation of HEMK2 expression in HEK293 cells. The cells were either transfected with HEMK2 or untransfected. A Western blot with anti-HA antibody showed expression of the HA-tagged HEMK2. *D*, validation of the expression of YFP-tagged substrate proteins and of the GFP-trap procedure. The cells were transfected with YFP-tagged substrate proteins either alone or together with HEMK2 and TRM112. Cell lysates were harvested, and the target proteins were enriched by GFP trap. The purified target proteins were analyzed by Western blots with anti-YFP antibodies, and the concentrations of the samples obtained with and without HEMK2/TRM112 overexpression were adjusted. *E*, cellular methylation of YFP-tagged CHD5 overexpressed in HEK293 cells in the presence and absence of ectopically expressed HEMK2 and TRM112. The substrate protein was purified using GFP trap as shown in *D*. Protein methylation was tested by Western blot with the methylglutamine antibody. The Ponceau S-stained image indicates that the loading control, recombinant unmethylated CHD5 protein was used as negative control for methylglutamine specificity. *F*, cellular methylation of YFP-tagged NUT overexpressed in HEK293 cells. The experiment was conducted as described in *E*, with the exception that recombinant unmethylated NUT domain protein was used as negative control in the Western blot with the methylglutamine antibody.

domains for readout of methyl-lysine, methylarginine, acetyllysine, and phosphoserine (36). However, it is clear that CHD5 and NUT (which were methylated in cells by HEMK2) both have very important biological functions. They are involved in chromatin regulation and also play important roles in tumor formation. CHD5 belongs to the family of chromatin remodeling proteins, it contains multiple PHD and Chromo domains and functions as tumor suppressor in variety of tumor types (37). NUT has been reported to undergo chromosomal translocations with BRD4, resulting in the formation of a BRD4-NUT fusion oncoprotein in 75% of the NUT midline carcinoma, which is an aggressive human cancer (38). Interestingly, the NLS of NUT resides in the C-terminal sequence (positions 1017–1042) (39), very close to the HEMK2 methylated site Gln^{1046}. However, we did not observe any changes in the localization of methylated and unmethylated YFP-tagged NUT protein in HEK293 cells (data not shown). Nevertheless, it would be interesting to see whether the methylation affects the targeting of BRD4-NUT to chromatin.

Conclusions

Enzymes of the HemK family have been subjected to a long search full of twists and turns to find their biological function. Starting as factors believed to be involved in heme biogenesis, they changed to putative DNA methyltransferases but finally were shown to function as protein glutamine methyltransferases, which methylate translation termination factors, eRF1, in case of the human and mouse HEMK2 enzyme. Here, we identify several novel protein substrates of HEMK2 *in vitro* and *in vivo*, and our data expand the range of proteins potentially subjected to glutamine methylation significantly. The strong phenotype of HEMK2 deletion in mice (17) and the important biological role of the novel HEMK2 substrates described here may suggest that the biological function of HEMK2 is not restricted to its role in translational termination mediated by eRF1. Hence, the search for the biological role of HEMK2 must be continued beyond eRF1 methylation. However, the investigation of the signaling functions of glutamine methylation is complicated by the fact, that the addition of a methyl group on a glutamine residue is a very subtle modification, which only slightly affects its chemical properties. Our results indicate that additional dedicated research will be needed to understand the role of glutamine methylation in signaling networks and other cellular processes.

Author Contributions—S. K. and A. J. devised the research. D. K. and S. K. conducted the experiments. D. K., S. K., and A. J. analyzed the data. All authors contributed to the writing of the manuscript and approved the final version of the manuscript.

Acknowledgments—We sincerely thank Dr. G. L. Xu for providing the HEMK2 and TRM112 bacterial and mammalian expression constructs and also for the bacterial eRF1 expression construct. We thank Dr. A. A. Mills for the CHD5 mammalian expression construct and Dr. C. A. French for the NUT gene.

References

1. Clarke, S. (1993) Protein methylation. *Curr. Opin. Cell Biol.* **5,** 977–983
2. Clarke, S. G. (2013) Protein methylation at the surface and buried deep: thinking outside the histone box. *Trends Biochem. Sci.* **38,** 243–252
3. Biggar, K. K., and Li, S. S. (2015) Non-histone protein methylation as a regulator of cellular signalling and function. *Nat. Rev. Mol. Cell Biol.* **16,** 5–17
4. Hamamoto, R., Saloura, V., and Nakamura, Y. (2015) Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat. Rev. Cancer* **15,** 110–124
5. Lanouette, S., Mongeon, V., Figeys, D., and Couture, J. F. (2014) The functional diversity of protein lysine methylation. *Mol. Syst. Biol.* **10,** 724
6. Bedford, M. T., and Clarke, S. G. (2009) Protein arginine methylation in mammals: who, what, and why. *Mol. Cell* **33,** 1–13
7. Kudithipudi, S., Kusevic, D., Weirich, S., and Jeltsch, A. (2014) Specificity analysis of protein lysine methyltransferases using SPOT peptide arrays. *J. Vis. Exp.* e52203
8. Lhoest, J., and Colson, C. (1977) Genetics of ribosomal protein methylation in *Escherichia coli*: II. a mutant lacking a new type of methylated amino acid, N5-methylglutamine, in protein L3. *Mol. Gen. Genet.* **154,** 175–180
9. Dinçbas-Renqvist, V., Engström, A., Mora, L., Heurgué-Hamard, V., Buckingham, R., and Ehrenberg, M. (2000) A post-translational modification in the GGQ motif of RF2 from *Escherichia coli* stimulates termination of translation. *EMBO J.* **19,** 6900–6907
10. Tessarz, P., Santos-Rosa, H., Robson, S. C., Sylvestersen, K. B., Nelson, C. J., Nielsen, M. L., and Kouzarides, T. (2014) Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* **505,** 564–568
11. Nakayashiki, T., Nishimura, K., and Inokuchi, H. (1995) Cloning and sequencing of a previously unidentified gene that is involved in the biosynthesis of heme in *Escherichia coli*. *Gene* **153,** 67–70
12. Le Guen, L., Santos, R., and Camadro, J. M. (1999) Functional analysis of the hemK gene product involvement in protoporphyrinogen oxidase activity in yeast. *FEMS Microbiol. Lett.* **173,** 175–182
13. Cheng, X. (1995) Structure and function of DNA methyltransferases. *Annu. Rev. Biophys. Biomol. Struct.* **24,** 293–318
14. Jeltsch, A. (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBioChem* **3,** 274–293
15. Bujnicki, J. M., and Radlinska, M. (1999) Is the HemK family of putative *S*-adenosylmethionine-dependent methyltransferases a "missing" ζ subfamily of adenine methyltransferases?: a hypothesis. *IUBMB life* **48,** 247–249
16. Yang, Z., Shipman, L., Zhang, M., Anton, B. P., Roberts, R. J., and Cheng, X. (2004) Structural characterization and comparative phylogenetic analysis of *Escherichia coli* HemK, a protein (N5)-glutamine methyltransferase. *J. Mol. Biol.* **340,** 695–706
17. Liu, P., Nie, S., Li, B., Yang, Z. Q., Xu, Z. M., Fei, J., Lin, C., Zeng, R., and Xu, G. L. (2010) Deficiency in a glutamine-specific methyltransferase for release factor causes mouse embryonic lethality. *Mol. Cell. Biol.* **30,** 4245–4253
18. Heurgué-Hamard, V., Champ, S., Engström, A., Ehrenberg, M., and Buckingham, R. H. (2002) The hemK gene in *Escherichia coli* encodes the N(5)-glutamine methyltransferase that modifies peptide release factors. *EMBO J.* **21,** 769–778
19. Nakahigashi, K., Kubo, N., Narita, S., Shimaoka, T., Goto, S., Oshima, T., Mori, H., Maeda, M., Wada, C., and Inokuchi, H. (2002) HemK, a class of protein methyl transferase with similarity to DNA methyl transferases, methylates polypeptide chain release factors, and hemK knockout induces defects in translational termination. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 1473–1478
20. Heurgué-Hamard, V., Graille, M., Scrima, N., Ulryck, N., Champ, S., van Tilbeurgh, H., and Buckingham, R. H. (2006) The zinc finger protein Ynr046w is plurifunctional and a component of the eRF1 methyltransferase in yeast. *J. Biol. Chem.* **281,** 36140–36148
21. Figaro, S., Scrima, N., Buckingham, R. H., and Heurgué-Hamard, V. (2008) HemK2 protein, encoded on human chromosome 21, methylates translation termination factor eRF1. *FEBS Lett.* **582,** 2352–2356
22. Ishizawa, T., Nozaki, Y., Ueda, T., and Takeuchi, N. (2008) The human mitochondrial translation release factor HMRF1L is methylated in the

GGQ motif by the methyltransferase HMPrmC. *Biochem. Biophys. Res. Commun.* **373,** 99–103
23. Schubert, H. L., Phillips, J. D., and Hill, C. P. (2003) Structures along the catalytic pathway of PrmC/HemK, an N5-glutamine AdoMet-dependent methyltransferase. *Biochemistry* **42,** 5592–5599
24. Schubert, H. L., Blumenthal, R. M., and Cheng, X. (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.* **28,** 329–335
25. Petrossian, T., and Clarke, S. (2009) Bioinformatic identification of novel methyltransferases. *Epigenomics* **1,** 163–175
26. Graille, M., Heurgué-Hamard, V., Champ, S., Mora, L., Scrima, N., Ulryck, N., van Tilbeurgh, H., and Buckingham, R. H. (2005) Molecular basis for bacterial class I release factor methylation by PrmC. *Mol. Cell* **20,** 917–927
27. Liger, D., Mora, L., Lazar, N., Figaro, S., Henri, J., Scrima, N., Buckingham, R. H., van Tilbeurgh, H., Heurgué-Hamard, V., and Graille, M. (2011) Mechanism of activation of methyltransferases involved in translation by the Trm112 "hub" protein. *Nucleic Acids Res.* **39,** 6249–6259
28. Frank, R. (2002) The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports: principles and applications. *J. Immunol. Methods* **267,** 13–26
29. Rathert, P., Zhang, X., Freund, C., Cheng, X., and Jeltsch, A. (2008) Analysis of the substrate specificity of the Dim-5 histone lysine methyltransferase using peptide arrays. *Chem. Biol.* **15,** 5–11
30. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* **31,** 3635–3641
31. Bannister, A. J., and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.* **21,** 381–395
32. Greer, E. L., and Shi, Y. (2012) Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13,** 343–357
33. Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X., and Jeltsch, A. (2008) Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* **4,** 344–346
34. Dhayalan, A., Kudithipudi, S., Rathert, P., and Jeltsch, A. (2011) Specificity analysis-based identification of new methylation targets of the SET7/9 protein lysine methyltransferase. *Chem. Biol.* **18,** 111–120
35. Kudithipudi, S., Lungu, C., Rathert, P., Happel, N., and Jeltsch, A. (2014) Substrate specificity analysis and novel substrates of the protein lysine methyltransferase NSD1. *Chem. Biol.* **21,** 226–237
36. Patel, D. J., and Wang, Z. (2013) Readout of epigenetic modifications. *Annu. Rev. Biochem.* **82,** 81–118
37. Kolla, V., Zhuang, T., Higashi, M., Naraparaju, K., and Brodeur, G. M. (2014) Role of CHD5 in human cancers: 10 years later. *Cancer Res.* **74,** 652–658
38. French, C. A. (2010) NUT midline carcinoma. *Cancer Genet. Cytogenet.* **203,** 16–20
39. French, C. A., Ramirez, C. L., Kolmakova, J., Hickman, T. T., Cameron, M. J., Thyne, M. E., Kutok, J. L., Toretsky, J. A., Tadavarthy, A. K., Kees, U. R., Fletcher, J. A., and Aster, J. C. (2008) BRD-NUT oncoproteins: a family of closely related nuclear proteins that block epithelial differentiation and maintain the growth of carcinoma cells. *Oncogene* **27,** 2237–2242