

# Chemoenzymatic synthesis and *in situ* application of S-adenosyl-L-methionine analogs

Marie Thomsen, Stine B. Vogensen, Jens Buchardt, Michael D. Burkart\*, and Rasmus P. Clausen\*

## Supporting information

### Content

1	Methods and materials .....	2
1.1	General methods and materials .....	2
1.2	Analytical HPLC .....	2
1.3	Molecular modeling.....	2
1.4	pH dependence of FDAS .....	5
1.5	Cloning, expression and purification of rat PRMT1 .....	6
1.6	Synthesis and purification of the RGG peptide.....	7
1.7	Enzymatic synthesis of SAM and analogs .....	11
1.8	Enzymatic modification of the RGG peptide by rPRMT1 using SAM analogs.....	12
1.9	Stability of SAM, ethyl-SAM, CIDA, adenosine, and SAH .....	12
2	Supporting results .....	13
2.1	Molecular modeling and rational design of mutants .....	13
2.2	pH dependence of FDAS .....	14
2.3	Enzymatic synthesis of SAM analogs .....	15
2.4	Degradation of SAM and ethyl-SAM.....	21
2.5	Enzymatic modification of the RGG peptide by rPRMT1.....	24
2.6	References .....	28

# 1 Methods and materials

## 1.1 General methods and materials

All chemicals were purchased from Sigma-Aldrich unless stated otherwise. Column materials and columns for protein purification were purchased from GE Healthcare, COMPLETE™ protease inhibitor cocktail from Roche Applied Science, and Benzonase and Bugbuster HT Protein Extraction Reagent were from Novagen (Merck). In all cases of concentrating protein samples, Amicon Ultra Centrifugation Filter Units of varying sizes with a cutoff at 10 kDa (Millipore) were used. S-adenosyl-L-methionine (SAM) and restriction enzymes were purchased from New England Biolabs.

The plasmid encoding SalL (pAEM7)<sup>[1]</sup> and FDAS (pFLA\_HT)<sup>[2]</sup> were generously donated by Professor Bradley S. Moore and Professor James H. Naismith, respectively.

LC MS was performed on an Agilent 6230 ESI-TOF LC MS coupled to an Agilent HPLC system series 1200 with an Agilent Poroshell 300SB-C18 (5 μm, 2.1x75 mm) column.

## 1.2 Analytical HPLC

Analytical HPLC was performed on an Agilent 1100 HPLC system using a polar endcapped C18 column: Synergi™ 4 μm Hydro RP 80Å, 150 x 4.6 mm (Phenomenex). For all analytical HPLC, a buffer system consisting of 10 mM ammonium formate (pH 3.2) in A) H<sub>2</sub>O and B) 90% acetonitrile was used. The acetonitrile was of isocratic grade for liquid chromatography (LiChrosolv®, Merck). The following elution gradient was used: 0-5 min. – 0-2%B (F = 1 mL/min.), 5-10 min. – 2-30% B (F = 1 mL/min.), 10-12 min. – 30-100% B (F = 1-1.5 mL/min.).

## 1.3 Molecular modeling

PDB files<sup>[1-6]</sup> were downloaded from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank ([www.rcsb.org](http://www.rcsb.org))<sup>[7]</sup>, whereas electron densities were downloaded from the Electron Density Served (EDS) at Uppsala University (<http://eds.bmc.uu.se>)<sup>[8]</sup>. Note that when specific amino acids or mutants are mentioned, S and F will be subscribed to highlight whether it is from SalL (S) or FDAS (F). Unless stated otherwise, L-Met and analogs were assumed to be a zwitterion in all modeling experiments.

All minimizations and docking experiments were performed in the Maestro suite (Schrödinger) using MacroModel and Glide, respectively, whereas Molecular Interaction Fields (MIFs) were calculated using GRID (Molecular Discovery). All protein structures and MIFs were visualized using Pymol (Schrödinger).

MIFs of water were calculated to evaluate the energetics of the binding of water in the active sites of several structures in complex with L-Met and 5'-chloro or fluoro-5'-deoxyadenosine (CIDA/FDA). For FDAS the wt structure of PDB-ID 1RQR was used, whereas for SalL a structure was modeled from PDB-ID 2Q6I, as the available SalL PDB-files were not consistent with I) the orientation of L-Met, II) the orientation of R243<sub>s</sub> (equivalent to R270<sub>f</sub>) and III) the presence of two water molecules in the active site (henceforth denoted A and B, see section 2.1 for elaboration). For FDAS, these issues were

consistent in nearly all available crystal structures except from the absence of the water molecules in one out of thirteen structures. From an analysis of the available electron densities of the SalL structures, it was however concluded that both L-Met and R243<sub>S</sub> should be oriented as observed in FDAS. L-Met in PDB-ID 2Q6I was, therefore, replaced by L-Met from FDAS. Furthermore, the coordinates of R243<sub>S</sub> were changed to those in the PDB-ID 2Q6L (a Y70T/G131S SalL mutant) which had a similar orientation as observed in FDAS. Finally, although the two water molecules were only present in one out of four PDB-files, it was evaluated from the electron densities that these were actually present in three out of four structures which indicated that they were important constituents of the active site. Thus, they were modeled into the 2Q6I structure from PDB-ID 2Q6O (an Y70T SalL mutant). This modeled SalL structure will henceforth be denoted SalL\*.

In addition to these two wt enzyme structures, four mutants were modeled for each enzyme: T155A<sub>F</sub>/T128A<sub>S</sub>, W217F<sub>F</sub>/W190F<sub>S</sub>, Y266F<sub>F</sub>/Y239F<sub>S</sub>, and a double mutant W217F:Y266F<sub>F</sub>/W190F:Y239F<sub>S</sub>.

Initially, a substructure minimization was performed for all proteins (wt and mutants). Here, a 10 (FDAS) or 15 Å (SalL) shell was selected around L-Met for minimization, whereas the rest of the protein was constrained by a 50 Å shell surrounding the inner shell. Except these changes, default settings were applied. After minimization, all water molecules were removed and these “stripped” structures were used for the GRID calculation. The GRID calculation was performed in a box size of 30 Å centered around L-Met in the active site. Except from adjusting the grid point to 3 (NPLA = 3) and setting the number of hetero atoms to 19 (NETA = 19 = the number of atoms in Cl/FDA), default settings were used.

Furthermore, L-Met, *S*-ethyl-, *S*-propyl-, and *S*-butyl-L-Hcy were docked into four different structures: FDAS and SalL\* with both water molecules (subscribed AB for the two water molecules) and the same two protein structures without the two water molecules (subscribed NW for no water).

### **1.3.1 Site-directed mutagenesis, expression, and purification of SalL and FDAS**

Based on molecular modeling, three amino acids were selected for a mutagenesis study (Table 1.1). The mutants were produced from pAEM7 (for SalL)<sup>[1]</sup> and pFLA\_HT (FDAS)<sup>[2]</sup> using the QuikChange Site-directed Mutagenesis kit (Agilent Technologies) and the primers listed in Table 1.1. The mutations were verified by sequencing in both forward and reverse direction.

All proteins (wt and mutants) were expressed in BL21(DE3) cells. The cells were grown in LB medium with 50 µg/mL kanamycin at 37 °C until the OD<sub>600</sub> was 0.6-0.8 after which the temperature was adjusted to 16 °C. After 0.5-1 h of incubation at this temperature, protein expression was induced by 0.25 mM IPTG and the cells were left at 16 °C for o/n protein expression. Next day, cells were harvested at 6000 xg for 10 min. and resuspended in the following lysis buffer: 1X BugBuster Protein Extraction Reagent (Novagen, Merck), 100 mM phosphate buffer pH 6.8, 2 kU/mL chicken egg white lysozyme (Sigma), 25 U/mL benzonase (Novagen, Merck) and cComplete Protease Inhibitor Cocktail (Roche) supplied in tablets according to manufactures protocol. The cells were lysed by gentle shaking at room temperature for 15-20 min. After lysis, insoluble cell debris was removed by centrifugation at 16.000 xg for 20 min.

**Table 1.1: List of point-mutations and the applied mutant primers for Sall and FDAS. Both forward (F) and reverse (R) primers are shown in the 5'-3' direction and the mutant codons are underlined.**

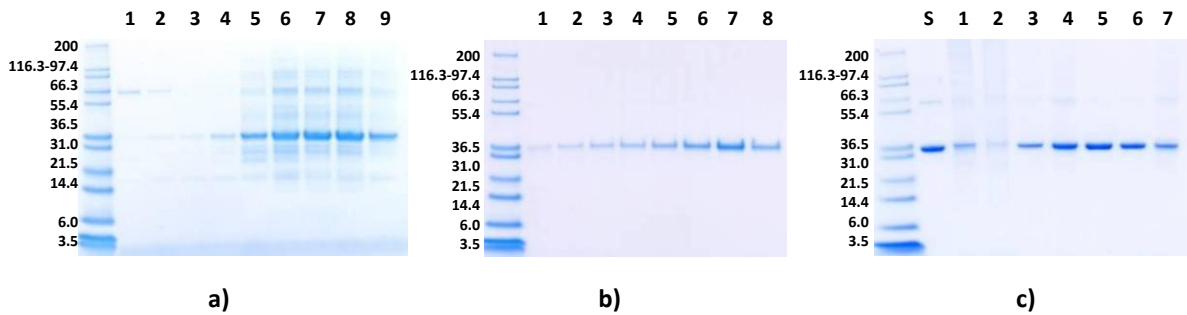
Sall	Mutant Sall primers	FDAS	Mutant FDAS primers
T128S	F: GTGACGCCA <u>AGCT</u> TGGTACGGCAAGG R: CCTTGCCGTACC <u>AGCT</u> TGGCGTCAC	T155S	F: CAGCCCGAACCGA <u>GCT</u> TTCTACAGCCGGGAG R: CTCCCGGCTGTAGA <u>GCT</u> CGGTTCCGGGCTG
T128A	F: GTGACGCCA <u>AGCT</u> TGGTACGGCAAGG R: CCTTGCCGTACC <u>AGCT</u> TGGCGTCAC	T155A	F: CAGCCCGAACCG <u>GCT</u> TTCTACAGCCGGGAG R: CTCCCGGCTGTAGA <u>GCT</u> CGGTTCCGGGCTG
T128G	F: GTGACGCCA <u>GGCT</u> TGGTACGGCAAGG R: CCTTGCCGTACC <u>AGCT</u> TGGCGTCAC	T155G	F: CAGCCCGAACCG <u>GGCT</u> TTCTACAGCCGGGAG R: CTCCCGGCTGTAGA <u>GCT</u> CGGTTCCGGGCTG
W190F	F: CTTTCGGCAACGTAT <u>TTT</u> ACCAACATACCCAC R: GTGGGTATGTTGGT <u>AAAT</u> ACGTTGCCGAAAG	W217F	F: GTTCGGCAACGTG <u>TTT</u> ACCAACATCCACC R: GGTGGATGTTGGT <u>AAAC</u> ACGTTGCCGAAAC
W190H	F: CTTTCGGCAACGTAC <u>CAC</u> ACCAACATACCCAC R: GTGGGTATGTTGGT <u>GTG</u> TACGTTGCCGAAAG	W217H	F: GTTCGGCAACGTG <u>CAC</u> ACCAACATCCACC R: GGTGGATGTTGGT <u>GTG</u> CACGTTGCCGAAAC
W190A	F: CTTTCGGCAACGTAG <u>CG</u> ACCAACATACCCAC R: GTGGGTATGTTGGT <u>CGCT</u> ACGTTGCCGAAAG	W217A	F: GTTCGGCAACGTG <u>CGC</u> ACCAACATCCACC R: GGTGGATGTTGGT <u>CGC</u> ACGTTGCCGAAAC
Y239F	F: CAGCCACTGCTG <u>TTT</u> CTCAACAGTCG R: CGACTGTTGAG <u>AAA</u> CAGCAGTGGCTG	Y266F	F: GCAACATCGCCATC <u>TTT</u> CTCAACAGC R: GCTGTTGAG <u>AAA</u> GATGGCGATGTTGC
Y239H	F: CAGCCACTGCTG <u>CAC</u> CTCAACAGTCG R: CGACTGTTGAGG <u>TG</u> CAGCAGTGGCTG	Y266H	F: GCAACATCGCCATC <u>CAC</u> CTCAACAGC R: GCTGTTGAGG <u>TG</u> GATGGCGATGTTGC
Y239V	F: CAGCCACTGCTG <u>GT</u> CCTCAACAGTCG R: CGACTGTTGAGG <u>ACC</u> CAGCAGTGGCTG	Y266V	F: GCAACATCGCCATC <u>GTC</u> CTCAACAGC R: GCTGTTGAGG <u>AC</u> GATGGCGATGTTG
Y239L	F: CAGCCACTGCTG <u>CTC</u> CTCAACAGTCG R: CGACTGTTGAGG <u>AG</u> CAGCAGTGGCTG	Y266L	F: GCAACATCGCCATC <u>CTC</u> CTCAACAGC R: GCTGTTGAGG <u>AG</u> GATGGCGATGTTGC
Y239I	F: CAGCCACTGCTGAT <u>CT</u> CAACAGTCG R: CGACTGTTGAGG <u>AT</u> CAGCAGTGGCTG	Y266I	F: GCAACATCGCCATC <u>ATC</u> CTCAACAGC R: GCTGTTGAGG <u>AT</u> GATGGCGATGTTGC
Y239A	F: CAGCCACTGCTG <u>CC</u> CTCAACAGTCG R: CGACTGTTGAGG <u>GCC</u> CAGCAGTGGCTG	Y266A	F: GCAACATCGCCATC <u>CC</u> CTCAACAGC R: GCTGTTGAGG <u>GCC</u> GATGGCGATGTTGC

Initial expression experiments in small volumes indicated that five Sall mutants (Y239H-A) did not express at all or at very low level. Attempts to express these mutants in higher yields were unsuccessful, and these were omitted from further analysis.

Although the T128A<sub>5</sub> mutant expressed in low yields it was attempted to purify it in larger yields, but as can be seen in Figure 1.1.c (well 2) only very low yields were obtained, and it was difficult to estimate the purity. It was, however, attempted to test activity for this mutant.

Just prior to loading onto Ni Sepharose<sup>TM</sup> 6 Fast Flow resin, 50 mM imidazole and 150 mM NaCl were added to the lysate to reduce unspecific ionic interactions. Furthermore, 4 mM MgCl<sub>2</sub> was added to prevent Ni<sup>+</sup> stripping of the resin by the EDTA present in the protease inhibitor cocktail.

The supernatant was then loaded onto 1.5 mL resin (pr. 500 mL culture) by gravity flow. After extensive washing with 30 column volumes of wash buffer (50 mM imidazole, 20 mM phosphate, 150 mM NaCl and 10% glycerol), protein was eluted by a step-wise gradient using buffers similar to the wash buffer except from changing the imidazole concentrations (fractions no. in brackets): 4 mL 75 mM (1), 2 x 4 mL 100 mM (2,3), 2 mL 150 mM (4), 2 mL 200 mM (5), 2 x 2 mL 250 mM (6,7), and 5 x 2 mL 500 mM imidazole (8-12) (Figure 1.1.a).



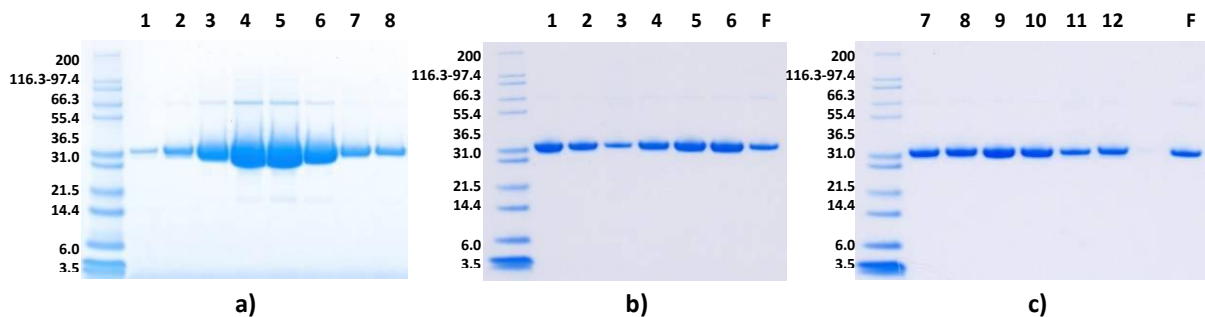
**Figure 1.1: Purification of SalL mutant T128G on Ni Sepharose™ 6 Fast Flow.** The first and second purification step is shown in gel a) and b), respectively, and are representative for all other SalL variants. The numbers above the gels correspond to the fraction numbers given in the text above. C) The purified pools of wt (S) and mutant SalL (1-7, numbers are according to the list in Table 1.1 with 1 being T128S, 2 T128A and so forth, but excluding Y239H-A).

FDAS and mutants expressed at higher levels than SalL and mutants which resulted in differences in purity from this first His-tag purification. FDAS and mutants were >95% pure from just this single purification step, whereas SalL and mutants were still impure (Figure 1.1 and Figure 1.2). SalL and mutants were therefore further purified on Ni Sepharose™ 6 Fast Flow resin. The purified pools from the first purification step were concentrated using spin-filters, diluted to approximately 150 mM imidazole and loaded onto new Ni Sepharose™ 6 Fast Flow resin (1.5 mL for 500 mL culture). After washing with approximately 5 column volumes of 100 mM imidazole, the protein was eluted with (fractions in brackets): 2 x 3mL 150 mM (1,2), 3 mL 175 mM (3), 3 mL 200 mM (4), 3 mL 225 mM (5), 3 mL 250 mM (6), 3 x 3 mL 500 mM imidazole (7,8,9) (Figure 1.1.b). During both washing and elution, protein content in the eluate was accessed by Bradford.

To remove salt and imidazole, the pools of purified protein were concentrated and buffer exchanged into 20 mM phosphate buffer pH 6.8 and 10% glycerol using PD-10 columns. Wt and mutant SalL were then concentrated in this buffer using spin-filters to varying concentration ranging from 40-540  $\mu$ M. Finally, glycerol was added to a final concentration of 50%. Wt and mutant FDAS, on the other hand, were first mixed with glycerol to 50% and then concentrated using the spin-filters. This was done to avoid precipitation of protein during ultrafiltration. Protein concentration was estimated by the Bradford protein assay using a protein standard consisting of BSA and lysozyme in a 1:1 mass ratio. All protein solutions were stored at -20 °C and they were all stable for at least 6 months.

#### 1.4 pH dependence of FDAS

FDAS activity was assayed at five different pH values (4.7, 6.1, 7.0, 7.8, and 9.5) using saturating concentrations of L-Met (50 mM) and CIDA (200  $\mu$ M). Different buffers were used: pH 4.7 was 50 mM sodium acetate buffer, pH 6.1-7.8 were 50 mM phosphate buffer, and pH 9.5 100 mM sodium carbonate buffer. As the L-Met solution contained NaOH, the pH of the buffers with L-Met were estimated and these estimated pH values are given above. 35  $\mu$ M FDAS was mixed with 50 mM L-Met and 200  $\mu$ M CIDA in the respective buffers and incubated at 37 °C for 0, 10, 20, 30, and 40 min. The reactions were quenched with an equal volume of 400 mM sodium formate buffer (pH 3) to eliminate further enzymatic activity and to increase the stability of SAM. Finally, the samples were



**Figure 1.2: Purification of wt and mutants of FDAS. a) Purification of mutant Y260V on Ni Sepharose™ 6 Fast Flow. This is representative for all FDAS variants. The numbers above the gel correspond to the fraction numbers given in the text. b) and c) The purified pools of wt (F) as well as mutants of FDAS (1–12, numbers are according to the list in Table 1.1 with 1 being T155S, 2 T155A and so forth).**

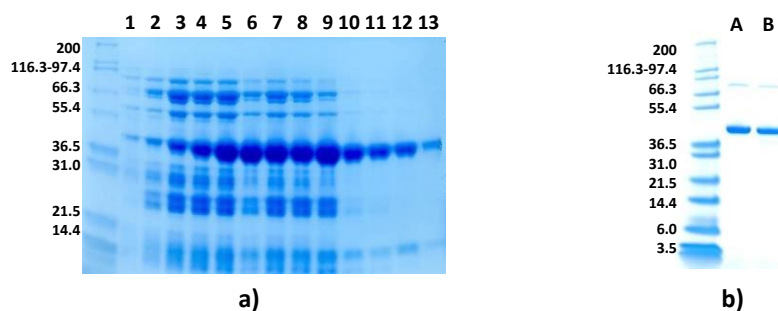
analyzed by HPLC and the FDAS activity was calculated as the initial velocity of the reaction estimated from the area of SAM in the HPLC chromatogram.

## 1.5 Cloning, expression and purification of rat PRMT1

The gene for rat PRMT1 isoform 1 (rPRMT1v1, henceforth denoted rPRMT1) was isolated from a cDNA library from Rat Normal Brain Tissue (BioChain Institute Inc.) using Expand High Fidelity PCR System (Roche) and primers containing an NdeI restriction site in the forward (F) primer and a BamHI restriction site in the reverse (R) primer (restriction sites are underlined): F) GACTATCATATGCGGCAGCCGAGG and R) GACTACTGGATCCTTATTAGCGCATCC. The PCR product was digested with the indicated restriction enzymes and ligated into pET-15b (Novagen, Merck) giving pET15b\_rPRMT1. pET-15b\_rPRMT1 was sequenced in both the forward and reverse direction.

rPRMT1 was expressed in BL21(DE3) cells. The cells were grown at 37 °C and 200 rpm in ZYM-5052r with 100 µg/mL ampicillin until  $OD_{600nm}$  was around 0.6-0.8. ZYM-5052r is a revised formula of ZYM-5052 described by Studier (2005) without lactose and with only half the concentration of glucose.<sup>[9]</sup> The temperature was then adjusted to 16 °C and after 1 h at this temperature expression was induced by 0.2 mM IPTG. After o/n expression, cells were harvested by centrifugation at 6000 xg and resuspended in lysis buffer (30 mL pr. 1.5 L culture): 40 mM phosphate buffer (pH 6.8), 300 mM NaCl, 0.1 mM MgCl<sub>2</sub>, 10% glycerol, 2 kU/mL lysozyme, 2.5 U/mL benzonase and EDTA-free cComplete protease inhibitor cocktail tablet according to manufacturer's instructions. Cells were lysed in Constant Cell Disruption Systems (Constant Systems Ltd) at 20 kpsi and the lysate was then cleared by centrifugation at 27.500 xg for 1 h.

The supernatant was supplemented with 20 mM imidazole, loaded onto pre-equilibrated Ni Sepharose™ 6 Fast Flow resin (1.5 mL resin pr. 1.5 L culture) and washed with 30 column volumes of wash buffer: 20 mM imidazole, 20 mM phosphate buffer (pH 6.8), 300 mM NaCl and 10% glycerol. Elution was achieved by a stepwise gradient of 1 mL fractions containing increasing concentrations of imidazole (fractions numbers in brackets): 1 x 50 mM (1), 1 x 75 mM (2), 1 x 100 mM (3), 1 x 150 mM (4), 14 x 200 mM (5-18) (See Figure 1.3). The buffer was as for the wash buffer except for changing



**Figure 1.3: Purification of rPRMT1. Gel a)** shows the elution from the first purification step on Ni Sepharose™ 6 Fast Flow resin. The numbers listed above the gel indicate the fractions mentioned in the text. **Gel b)** shows two pure (>95%) pools (A and B) of rPRMT1 after the second purification step.

the imidazole concentration. During both washing and elution, protein content in the eluate was accessed by Bradford protein assay.

Since the first purification step did not give pure protein, a second purification was performed. The pool of protein from the first purification was diluted to an imidazole concentration of 50 mM and loaded onto 1.5 mL pre-equilibrated Ni Sepharose™ 6 Fast Flow resin. The resin was then washed with seven column volumes of wash buffer and elution was achieved with 1 mL fractions of increasing imidazole concentrations (fraction numbers in brackets): 2 x 75 mM (1,2), 2 x 100 mM (3,4), 2 x 150 mM (5,6), 6 x 200 mM (7-13) (Figure 1.3). The pool of pure rPRMT1 is shown in Figure 1.3. The protein was concentrated by ultrafiltration to 60  $\mu$ M.

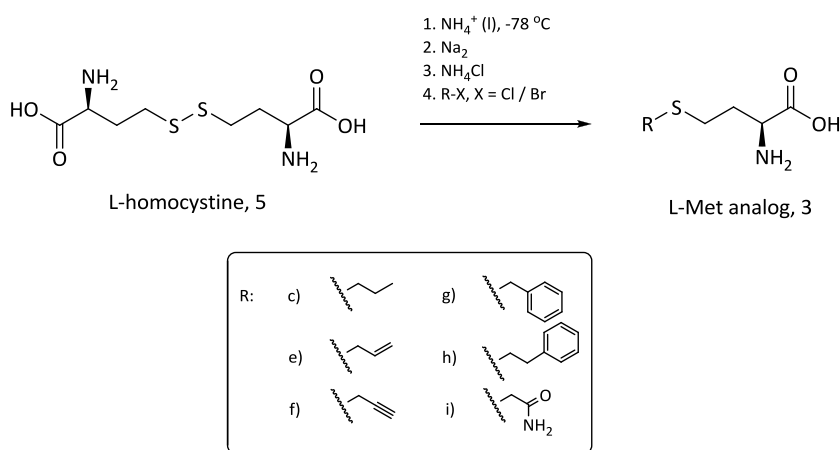
Protein concentration was estimated by the absorbance at 280 nm and using an extinction coefficient of  $54,320 \text{ M}^{-1}\text{cm}^{-1}$  (calculated by the ProtParam Tool at Expasy.org)<sup>[10]</sup>.

## 1.6 Synthesis and purification of the RGG peptide

The RGG peptide ( $\text{GGRGGFGRRGGFGRRGGFG-NH}_2$ ,  $M_{\text{monoisotopic}} = 1667.8137 \text{ g/mol}$ ) was synthesized on Applied Biosystems 433A Peptide Synthesizer using 1 g of Tentagel S RAM resin (RAPP Polymere GmbH) and 1 mmol F-moc protected amino acids (Applied Biosystems). After peptide synthesis, the peptide was cleaved off the resin by incubation in 2.5% triisopropylsilan, 7.5% MilliQ water and 90% TFA for 2.5 h at rt. The peptide was then precipitated by ether and collected by centrifugation. The peptide was purified by preparative reverse phase HPLC using a YMC ODDMS 120 Å 5  $\mu$ m (250 mm x 20 mm) on a Gilson Preparative HPLC instrument (321 pump, 155 UV/VIS detector, GX-271 liquid handler, 831 temperature regulator). Using a buffer system containing 0.1% TFA in water/acetonitrile, the following gradient of acetonitrile was used: 0-5 min. – 0%, 5-45 min. – 0-40% and 45-50 min. – 40-90% (flow: 10 mL/min., column temperature = 40 °C). The peptide eluted at approximately 30% acetonitrile and fractions containing pure peptide were freeze dried.

After purification, peptide purity was checked by analytical HPLC showing one symmetrical peak (data not shown) and by LC MS with:  $[M+H^+]_{\text{monoisotopic}} = 1669.8356$ ,  $[M+H^+]_{\text{measured}} = 1669.8174$ ;  $[M+2H^+]_{\text{monoisotopic}} = 1670.8356$ ,  $[M+2H^+]_{\text{measured}} = 1670.8227$ .

Finally, the peptide was dissolved in water and its final concentration was estimated by N-detection on an Agilent HPLC system series 1200 coupled to an ANTEK 8060 Nitrogen Detector.



**Figure 1.4: Synthesis of L-Met analogs (3) from L-homocystine (5) and alkyl halides. The numerations of the R-groups are according to that given in the paper.**

### 1.6.1 Synthesis of L-Met analogs

L-Met (**3a**, Sigma Aldrich), S-ethyl-L-homocysteine (**3b**, Sigma Aldrich), S-butyl-L-homocysteine (**3d**, Toronto Research Chemicals Inc.) and S-(2-amino-2-carboxyethyl)-L-homocysteine (**3i**, Sigma Aldrich) were bought. All other L-Met analogs were synthesized according to the same protocol illustrated in Figure 1.4.<sup>[11]</sup> The chemicals were purchased from Sigma-Aldrich or Chem-impex (L-homocystine) and were used without further purification. The specific protocols are given below. Furthermore, purification of selected analogs will be described. These analogs were applied in the coupled enzymatic assays and had to be purified to eliminate L-Met contaminations from L-homocystine. For NMR, most products were insolvable in D<sub>2</sub>O but dissolved with the addition of NaOD. NMR spectra were recorded on a 300 MHz Varian Gemini or 300 MHz Varian Mercury spectrometer. Chemical shifts are given in ppm ( $\delta$ ) using dioxane as internal standards, and the coupling constants are given in Hertz. For the NMR spectra interpretation: s = singlet, sex = sextet, dd = double doublet, m = multiplet.

LC-MS were obtained with an Agilent 6410 Triple Quadrupole Mass Spectrometer instrument using electron spray coupled to an Agilent 1200 HPLC system (ESI-LC MS) with a C<sub>18</sub> reverse phase column (Zorbax Eclipse XBD-C<sub>18</sub>, 4.6 mm  $\times$  50 mm), diode array detector, and a Sedere Sedex 85 Light Scattering Detector. Elemental analyses were performed by J. Theiner, Department of Physical Chemistry, University of Vienna, Austria, and the results were within 0.4% of the calculated values unless otherwise stated. Melting points were recorded in open capillary tubes using a OptiMelt instrument from Stanford Research Systems and are uncorrected.

### 1.6.2 S-propyl-L-homocysteine (3c, (S)-2-amino-4-(propylthio)butanoic acid)

L-Homocystine (268 mg, 1 mmol) was dissolved in liquid ammonia (30 mL) at -78°C and sodium was added in small pieces until a dark blue color persisted for 15 min. NH<sub>4</sub>Cl was added until the blue color of the mixture faded. 1-Bromopropane (200  $\mu$ L, 2.2 mmol) was added and the mixture was stirred at -78°C for 2h. The cooling bath was removed and the NH<sub>3</sub> was allowed to evaporate. The white solid was dissolved in water (50 mL) and extracted with ether (2 x 10 mL). The pH of the



aqueous phase was adjusted to pH 6 with 4M HCl and 1M NaOH and the volume was reduced. The solution was left for crystallization to give the product as white crystals (125 mg, 0.71 mmol, 35%). Crystallization of the motherliquid gave a second batch as white crystals (127 mg, 0.72 mmol, 36%), mp. 248-251 °C. <sup>1</sup>H NMR (D<sub>2</sub>O/NaOD): δ 0.96 (t, *J* = 7.5, 3H), 1.60 (sex, *J* = 7.2, 2H), 1.73-1.96 (m, 2H), 2.53-2.61 (m, 4H), 3.31 (dd, *J* = 6.6 and 6.0, 1H). <sup>13</sup>C NMR (D<sub>2</sub>O/NaOD): δ 13.43, 22.90, 28.11, 33.76, 35.37, 55.97, 183.20. LC-ESI-MS: [M+H<sup>+</sup>]<sub>monoisotopic</sub> = 178.09, [M+H<sup>+</sup>]<sub>measured</sub>: 178.00. Anal. (C<sub>7</sub>H<sub>15</sub>NO<sub>2</sub>S·0.2 H<sub>2</sub>O) C: calcd, 46.49; found, 46.05; H: calcd, 8.58; found 8.13; N.

### 1.6.3 S-allyl-L-homocysteine (3e, (S)-2-amino-4-(allylthio)butanoic acid)

L-Homocystine (268 mg, 1 mmol) was dissolved in liquid ammonia (30 mL) at -78°C and sodium was added in small pieces until a dark blue color persisted for 15 min. NH<sub>4</sub>Cl was added until the blue color of the mixture faded. Allyl bromide (186 μL, 2.2 mmol) was added and the mixture was stirred at -78°C for 2h. The cooling bath was removed and the NH<sub>3</sub> was allowed to evaporate. The white solid was dissolved in water (30 mL) and extracted with ether (2 x 10 mL). The pH of the aqueous phase was adjusted to pH 6 with 4M HCl and 1M NaOH and the volume was reduced. The solution was left for crystallization to give the product as white crystals (115 mg, 0.66 mmol, 33%). Crystallization of the motherliquid gave a second batch as white crystals (146 mg, 0.83 mmol, 42%), mp. 244-245 °C. <sup>1</sup>H NMR (D<sub>2</sub>O/NaOD): δ 1.72-1.96 (m, 2H), 2.54(t, *J* = 7.4, 2H), 3.20 (d, *J* = 7.2, 2H), 3.31(dd, *J* = 7.2 and 5.5, 1H), 5.11-5.20 (m, 2H), 5.75-5.90 (m, 1H). <sup>13</sup>C NMR (D<sub>2</sub>O/NaOD): δ 26.94, 34.18, 35.05, 55.97, 118.11, 134.46, 183.25. LC-ESI-MS: [M+H<sup>+</sup>]<sub>monoisotopic</sub> = 176.08, [M+H<sup>+</sup>]<sub>measured</sub>: 176.00. Anal. (C<sub>7</sub>H<sub>13</sub>NO<sub>2</sub>S·0.1 H<sub>2</sub>O) C, H, N.

### 1.6.4 S-propargyl-L-homocysteine (3f, (S)-2-amino-4-(prop-2-yn-1-yl-thio)butanoic acid)

L-Homocystine (804 mg, 3 mmol) was dissolved in liquid ammonia (30 mL) at -78°C and sodium was added in small pieces until a dark blue color persisted for 20 min. NH<sub>4</sub>Cl was added until the blue color of the mixture faded. Propargyl bromide (80% in toluene, 710 μL, 2.2 mmol) was added and the mixture was stirred at -78°C for 2h. The cooling bath was removed and the NH<sub>3</sub> was allowed to evaporate. The white solid was dissolved in water (50 mL) and extracted with ether (2 x 10 mL). The pH of the aqueous phase was adjusted to pH 6 with 4M HCl and 1M NaOH and the volume was reduced. The solution was left for crystallization to give the product as reddish crystals (330 mg, 1.9 mmol, 64%), mp. 218-225 °C. <sup>1</sup>H NMR (D<sub>2</sub>O): δ 2.09-2.30 (m, 2H), 2.66 (t, *J* = 2.6 Hz, 2H), 2.83 (t, *J* = 6.9 Hz, 2H), 3.38 (d, *J* = 2.6 Hz, 1H), 3.86 (dd, *J* = 6.9 and 5.5 Hz, 1H). <sup>13</sup>C NMR (D<sub>2</sub>O): δ 18.74, 27.15, 30.48, 54.42, 72.79, 80.87, 174.49. LC-ESI-MS: [M+H<sup>+</sup>]<sub>monoisotopic</sub> = 174.06, [M+H<sup>+</sup>]<sub>measured</sub>: 174.10. Anal. (C<sub>7</sub>H<sub>11</sub>NO<sub>2</sub>S·0.8 H<sub>2</sub>O) C: calcd, 44.81; found, 43.84; H: calcd, 6.77; found 5.61; N.

When dissolving this analog in 250 mM NaOH it was observed that a degradation product with a mass of +44 was formed. This mass difference is similar to that observed for Propargyl-SAM.<sup>[13-15]</sup>

### 1.6.5 S-benzyl-L-homocysteine (3g, (S)-2-amino-4-(benzylthio)butanoic acid)

L-Homocystine (268 mg, 1 mmol) was dissolved in liquid ammonia (30 mL) at -78°C and sodium was added in small pieces until a dark blue color persisted for 15 min. NH<sub>4</sub>Cl was added until the blue color of the mixture faded. Benzyl chloride (253 μL, 2.2 mmol) was added and the mixture was stirred at -78°C for 2h. The cooling bath was removed and the NH<sub>3</sub> was allowed to evaporate. The white

solid was dissolved in water (50 mL) and extracted with ether (2 x 10 mL). The pH of the aqueous phase was adjusted to pH 6 with 4M HCl and 1M NaOH which resulted in heavy precipitation. The product was isolated as white crystals (350 mg, 1.56 mmol, 78%), mp. 228-232 °C. <sup>1</sup>H NMR (D<sub>2</sub>O/NaOD): δ 1.70-1.96 (m, 2H), 2.50 (t, *J* = 7.2, 2H), 3.28 (dd, *J* = 7.4 and 5.5, 1H), 7.29-7.43 (m, 5H). <sup>13</sup>C NMR (D<sub>2</sub>O/NaOD): δ 27.74, 35.00, 35.59, 55.92, 127.75, 129.33, 129.45, 139.05, 183.15. LC-ESI-MS: [M+H<sup>+</sup>]<sub>monoisotopic</sub> = 226.09, [M+H<sup>+</sup>]<sub>measured</sub>: 226.00. Anal. (C<sub>7</sub>H<sub>11</sub>NO<sub>2</sub>S) C, H, N.

#### 1.6.6 S-phenethyl-L-homocysteine (3h, (S)-2-amino-4-(phenethylthio)butanoic acid)

L-Homocystine (268 mg, 1 mmol) was dissolved in liquid ammonia (30 mL) at -78°C and sodium was added in small pieces until a dark blue color persisted for 15 min. NH<sub>4</sub>Cl was added until the blue color of the mixture faded. (2-Bromoethyl)benzene (298 μL, 2.2 mmol) was added and the mixture was stirred at -78°C for 2h. The cooling bath was removed and the NH<sub>3</sub> was allowed to evaporate. The white solid was dissolved in water (50 mL) and extracted with ether (2 x 10 mL). The pH of the aqueous phase was adjusted to pH 6 with 4M HCl and 1M NaOH which resulted in heavy precipitation. The product was isolated as white crystals (353 mg, 1.47 mmol, 67%), 207-211 °C. <sup>1</sup>H NMR (D<sub>2</sub>O/NaOD): δ 1.70-1.95 (m, 2H), 2.57 (t, *J* = 7.7, 2H), 2.87-2.94 (m, 4H), 3.29 (dd, *J* = 7.2 and 5.5, 1H), 7.25-7.40 (m, 5H). <sup>13</sup>C NMR (D<sub>2</sub>O/NaOD): δ 28.14, 33.09, 35.25, 35.36, 55.93, 127.01, 129.16, 129.22, 141.11, 180.10. LC-ESI-MS: [M+H<sup>+</sup>]<sub>monoisotopic</sub> = 240.11, [M+H<sup>+</sup>]<sub>measured</sub>: 240.00. Anal. (C<sub>12</sub>H<sub>17</sub>NO<sub>2</sub>S·0.1 H<sub>2</sub>O) C, H, N.

#### 1.6.7 S-(2-amino-2-oxoethyl) -L-homocysteine (3i, (S)-2-amino-4-((2-amino-2-oxoethyl)thio)-butanoic acid)

L-Homocystine (804 mg, 3 mmol) was dissolved in liquid ammonia (40 mL) at -78°C and sodium was added in small pieces until a dark blue color persisted for 20 min. NH<sub>4</sub>Cl was added until the blue color of the mixture faded. 2-Chloroacetamide (617 mg, 6.6 mmol) was added and the mixture was stirred at -78°C for 2h. The cooling bath was removed and the NH<sub>3</sub> was allowed to evaporate. The white solid was dissolved in water (50 mL) and washed with ether (2x20 mL). The pH of the aqueous phase was adjusted to pH 6 with 4M HCl and 1M NaOH. The volume was reduced resulting in white precipitation. Recrystallization (H<sub>2</sub>O) gave the product as white crystals (302 mg, 1.57 mmol, 52%), mp. 207-211 °C. <sup>1</sup>H NMR (D<sub>2</sub>O/NaOD): δ 2.06-2.27 (m, 2H), 2.72 (s, 2H), 3.33 (s, 2H), 3.84 (dd, *J* = 7.7 and 5.8 Hz), 1H). <sup>13</sup>C NMR (D<sub>2</sub>O/NaOD): δ 29.31, 30.69, 34.99, 54.37, 174.43, 175.91. LC-ESI-MS: [M+H<sup>+</sup>]<sub>monoisotopic</sub> = 193.07, [M+H<sup>+</sup>]<sub>measured</sub>: 193.09. Anal. (C<sub>6</sub>H<sub>12</sub>N<sub>2</sub>O<sub>3</sub>S) C, H, N.

#### 1.6.8 Purification of selected L-Met analogs

S-ethyl-, S-propyl-, S-butyl-, and S-allyl-L-homocysteine were purified in small amounts by analytical HPLC. The analogs were dissolved in 250 mM NaOH and purified using analytical HPLC with fractionation. They were purified according to the method described in section 1.1 using program I. The pools of purified L-Met analog were freeze-dried and analyzed by LC MS confirming the presence the specific analogs.

### 1.6.9 Dissolution of L-Met analogs

The estimated  $K_M$  values of L-Met for SalL and FDAS are  $2.6^1$  and  $4.5^5$  mM, respectively. The  $K_M$  values are not known for the L-Met analogs but assuming that they were in the same concentration range, it would be preferable to perform the enzymatic reactions with approximately 15-25 mM L-Met analog to obtain maximal activity. The idea was to dissolve L-Met and analogs in 10X stocks. However, only L-Met and the smaller analogs (*S*-ethyl- and *S*-propyl-L-Hcy) could be dissolved in 200-250 mM concentrations in water or other neutral pH solutions. For *S*-butyl-L-Hcy addition of HCl or NaOH was needed whereas the benzyl analog could not be dissolved in that high concentration. The use of HCl was unfavorable due to the fact that  $\text{Cl}^-$  is a product of the halogenase catalyzed SAM formation. Several different methods were probed for dissolving *S*-benzyl-L-Hcy using e.g. DMSO, acetonitrile, alcohols, Tween-20 and -80, or (2-hydroxypropyl)- $\beta$ -cyclodextrin in different aqueous solutions of phosphate buffer (pH 7.8) or NaOH. However, a maximum concentration of 150 mM was obtained in 250 mM NaOH. All analogs were therefore dissolved as 150 mM stocks in 250 mM NaOH.

When diluting 10 fold into the neutral reaction buffer, *S*-benzyl- and *S*-phenethyl-L-Hcy, however, precipitated. Again various additives such as alcohols, tween-20 and 80 as well as organic solvents (DMSO and acetonitrile) were probed for their ability for increase the solubility, but all of these tests were unsuccessful. It was found that the maximum soluble concentration of these analogs were approximately 2 mM in the reaction buffer. The stock solutions of *S*-benzyl- and *S*-phenethyl-L-Hcy were therefore diluted 10 fold with 250 mM NaOH to 15 mM. The final concentration in the reaction mixture was then 1.5 mM.

### 1.7 Enzymatic synthesis of SAM and analogs

Based on the pH dependences of SalL<sup>[12]</sup> and FDAS (section 2.2) showing that these enzymes have maximum activities around neutral pH (pH 6-8), all enzymatic assays were performed in 100 mM phosphate buffer (pH 6.8). However, as L-Met and analogs were dissolved in 250 mM NaOH, the pH of the buffer changed after addition of L-Met and analogs. To all assays, the addition of L-Met or analog corresponded to 10% of the total reaction volume resulting in a final pH of 7.4.

The enzymatic assays were performed using 200  $\mu\text{M}$  5'-chloro-5'-deoxyadenosine (CIDA) and 15 mM L-Met and analogs, except for the benzyl and phenethyl analogs, which due to solubility issues could only be assessed at 1.5 mM. Initially, the reactions were performed using 3  $\mu\text{M}$  SalL enzyme (wt or mutant) or 100  $\mu\text{M}$  FDAS enzyme (wt or mutant), but due to reduced kinetics when using mutants and L-Met analogs, the enzyme concentrations were in some assays increased to 60  $\mu\text{M}$  and 200  $\mu\text{M}$  for SalL and FDAS enzymes, respectively.

In general, buffer, enzyme, and CIDA were mixed before the final addition of L-Met and analogs. This sequence of mixing was performed to avoid precipitation of the L-Met analogs in the neutral pH buffer. The reaction was incubated at 37 °C for varying times (e.g. 0, 1.5, 3 and 5 h) and immediately quenched by the addition of an equal volume of 400 mM sodium formate buffer (pH 3.0). Both SalL and FDAS were inactive at this pH value (data not shown). Samples were then stored at 4 °C until analysis by analytical HPLC as described in 1.2.

The initial velocity of the enzymatic reaction was determined from the ratio of the area of product against the total area of product and unreacted substrate (CIDA) at a time point at which the reaction

did still proceed at a constant initial velocity. This velocity was similar to velocities determined by linear regression of the substrate consumption over time. In many cases, the reaction was very slow and the last time point at 5 h was assumed to be within the range of the initial velocity. For comparison, all initial velocities were divided with the enzyme concentration giving a relative initial velocity.

Control reactions without enzyme were also performed to verify that SAM analog formation was enzyme catalyzed.

All assays were performed in at least duplicate.

## **1.8 Enzymatic modification of the RGG peptide by rPRMT1 using SAM analogs**

Initially, the activity of rPRMT1 was assessed using commercially available SAM. Here 2  $\mu\text{M}$  rPRMT1 was incubated with 200  $\mu\text{M}$  SAM and 10  $\mu\text{M}$  RGG peptide in 100 mM phosphate buffer (pH 6.8) at 37  $^{\circ}\text{C}$  for 12 h after which the reaction was quenched with an equal volume of 400 mM sodium formate buffer (pH 3). The product was assessed by LC MS.

The *in situ* synthesis of SAM coupled to methylation of RGG was then probed by incubating 2  $\mu\text{M}$  rPRMT1, 1  $\mu\text{M}$  SalL, 200  $\mu\text{M}$  CIDA, 15 mM L-Met, and 10  $\mu\text{M}$  RGG peptide in phosphate buffer (pH 6.8) at 37  $^{\circ}\text{C}$  for 12 h. The reaction was quenched and analyzed as described above.

Modification of the RGG peptide with larger chemical groups than the methyl group from SAM was subsequently assessed in a similar manner using 15 mM L-Met analog (for all except the benzyl and phenethyl analog of which only 1.5 mM was used due to solubility issues), 200  $\mu\text{M}$  CIDA and either wt or W190A SalL in varying concentrations: S-ethyl-L-homocysteine – 15  $\mu\text{M}$  wt SalL; S-propyl-L-homocysteine, S-allyl-L-homocysteine, S-benzyl-L-homocysteine – 45  $\mu\text{M}$  wt SalL; S-butyl-L-homocysteine – 45  $\mu\text{M}$  W190A SalL. These coupled assays were performed as described for the coupled assay with L-Met and the product was analyzed by LC MS.

As a negative control, all above mentioned assays were also performed without addition of rPRMT1.

All assays were performed in duplicate.

## **1.9 Stability of SAM, ethyl-SAM, CIDA, adenosine, and SAH**

To obtain a fresh solution of SAM and ethyl-SAM with the lowest amount of degradation as possible, 200  $\mu\text{M}$  CIDA was incubated with 3 or 60  $\mu\text{M}$  SalL and 15 mM L-Met or S-ethyl-L-homocysteine, respectively, in 100 mM phosphate buffer (pH 7.4) for 30 minutes at 37  $^{\circ}\text{C}$ . This resulted in 100% conversion.

The degradation of SAM and ethyl-SAM was then tested in two separated assays. In one assay, SAM or ethyl-SAM (100  $\mu\text{M}$ ) was incubated at 70  $^{\circ}\text{C}$  for 3 hours in the same phosphate buffer as above. The degradation products were fractionated by HPLC and analyzed by LC MS. In the other assay, SAM or ethyl-SAM (100  $\mu\text{M}$ ) was incubated at 37  $^{\circ}\text{C}$  for 0, 0.5, 1, 1.5, 2, 3.5, and 5 h in either 100 mM phosphate buffer (pH 7.4) or 200 mM sodium formate buffer (pH 3.0). These buffer conditions are equivalent with the reaction buffer and the quenching buffer, respectively. Samples were mixed with an equal volume of 400 mM sodium formate buffer (pH 3) and stored at 4  $^{\circ}\text{C}$  until analysis by analytical HPLC. From the HPLC chromatogram, the areas of the peaks containing SAM or ethyl-SAM

and their degradation products were estimated and used to calculate the concentration of each compound by assuming that all compounds have the same extinction coefficient. This assay was performed in duplicate.

The stabilities of CIDA, adenosine and the MTase product SAH were explored in a similar assay as the latter described above.

## 2 Supporting results

### 2.1 Molecular modeling and rational design of mutants

In 12 out of the 13 available FDAS crystal structures two water molecules (A and B) were observed in a cavity into which the L-Met analogs were assumed to propagate (henceforth denoted water cavity). These were also observed in three of the four SalL structures. If these water molecules could not be displaced from the active sites upon binding of the L-Met analogs, they might hinder the use of the larger L-Met analogs.

The affinities of these waters in the active sites of FDAS and SalL\* were explored by estimating the MIFs of water in the active sites of FDAS and SalL\* (Figure 2.2). These MIFs clearly indicated that water did have high affinity for the exact locations of the two above mentioned water molecules as large interaction fields were present at -10.0 kcal/mol. Docking experiments indicated that these two water molecules, indeed, affected the binding of L-Met analogs larger than *S*-ethyl-L-Hcy. When docking into structures containing both water molecules, the sulfurs in the propyl and butyl analogs were shifted away from the location in L-Met (analyses not shown). However, when docking into a structure without these water molecules, the shift was significantly reduced and it was clear that the analogs propagated into the water cavity (Figure 2.1). From analyses of the wt enzymes it was speculated that O4' of C1/FDA, F186<sub>S</sub>/F213<sub>F</sub> and the carbonyl of T155<sub>F</sub> were important for catalysis by transiently donating electrons to sulfur. This would increase the nucleophilicity of sulfur in L-Met and stabilize the positive charge in SAM. Thus, a shift of sulfur could be significant for enzyme activity. Based on these analyses mutants were designed in which the amino acids hydrogen bonding to these

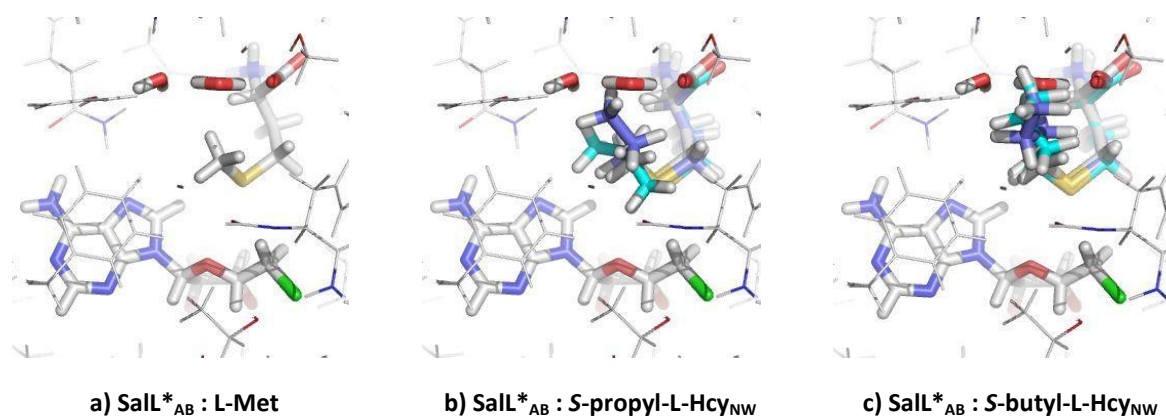
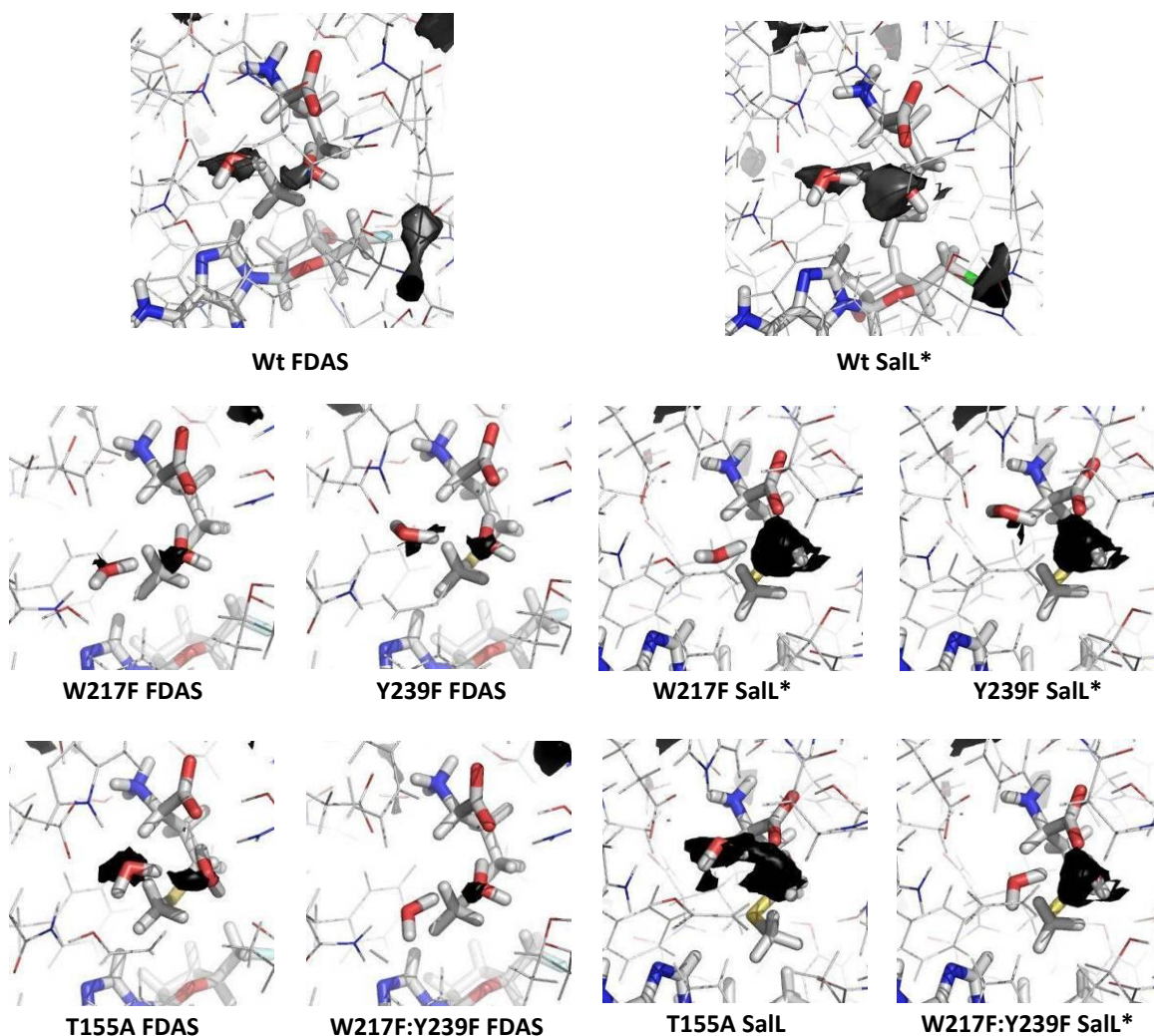


Figure 2.1: SalL\*<sub>AB</sub> in complex with a) L-Met as well as the two best docking poses of b) *S*-propyl- and c) *S*-butyl-L-Hcy from docking into SalL\*<sub>NW</sub>. This illustrated that especially *S*-butyl-L-Hcy would propagate into the water cavity if it was unoccupied. The same trend was observed for FDAS.



**Figure 2.2:** Illustration of the MIFs of water at  $-10$  kcal/mol in the active sites of wt FDAS and wt Sall\* as well as selected mutants. The MIFs are illustrated as dark grey surfaces. The two water molecules, L-Met and F/CIDA are highlighted as bold sticks. The water molecules are denoted A and B from left to right.

water molecules were exchanged. As illustrated in the paper, water molecule A hydrogen bonds to the side chains of W217<sub>F</sub>/W190<sub>S</sub> and Y270<sub>F</sub>/Y239<sub>S</sub>, whereas water molecule B interacts with the carboxylate group of L-Met, the side chain hydroxyl group of T155<sub>F</sub>/T128<sub>S</sub> as well as the backbone amide of S269<sub>F</sub>/S242<sub>S</sub>. The three amino acids in which the side chains were involved (Trp, Tyr and Thr) were therefore selected for a mutagenesis study. The aim was to enlarge the active sites of Sall and FDAS by destabilizing the water molecules and introducing smaller amino acids. MIFs of selected mutants indicated reduced affinity for water in the water cavity (Figure 2.2). All mutants are listed in Table 1.1.

## 2.2 pH dependence of FDAS

In Lipson *et al* it was shown that Sall is most active in the neutral pH interval from 7-8.<sup>12</sup> The pH dependence of FDAS was assayed in a similar manner, and it was observed that FDAS had maximal activity at slightly lower pH values, pH 5 to 7 (Figure 2.3). However, to ease the performance of the assays it was decided to use pH 7.4 for both FDAS and Sall although this would lead to approximately 10% loss of FDAS activity.

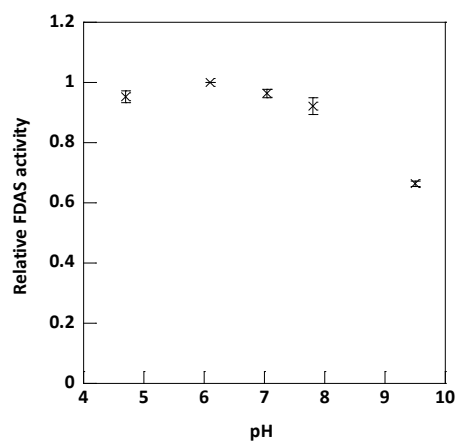


Figure 2.3: The pH dependence of the formation of SAM catalyzed by FDAS from CIDA and L-Met.

### 2.3 Enzymatic synthesis of SAM analogs

Assays with halogenases (either wt or mutant) were mixed as described in section 1.7 and representative chromatograms showing the formation of SAM, ethyl-, propyl-, butyl-, allyl- and benzyl-SAM are shown in Figure 2.5 to Figure 2.7. For the negative controls no formation of SAM or analogs were detected verifying that their formation was catalyzed by the respective enzymes. Using the given reaction condition five hours of incubation resulted in 100% conversion of CIDA to ethyl-SAM, whereas for the other SAM analogs the yield was 1.5-15% due to reduced activities. The severely reduced rate of benzyl-SAM formation could result from use of a 10 fold lower concentration of this analog compared to the other analogs.

As described in the paper and shown in section 2.4, SAM and analogs were readily degraded at neutral pH, and with longer incubation times degradation to adenine and the respective alkyl-thioadenosine was observed (see elaboration on degradation in section 2.4). Racemization could be detected in some but not all assays as the separation of the diastereomers was difficult. Since ethyl-SAM eluted after short time almost without addition of acetonitrile the diastereomers were not separated in all assays. For propyl- and allyl-SAM racemization was clearly observed. Two peaks formed and LC MS analysis of these peaks confirmed the presence of the SAM analogs. As seen in Figure 2.6 and Figure 2.7, racemization was not immediately visible in for butyl- and benzyl-SAM. However, as oppose to the other SAM analogs both of these SAM analogs eluted during a steep gradient of acetonitrile, and the two diastereomers might elute in one peak. This was probed for butyl-SAM using another acetonitrile gradient and results indicated that two diastereomers were indeed present (data not shown). This was not examined for benzyl-SAM due to the formation of only low amounts of this analog. Degradation of SAM and analogs was an issue only in cases where the synthesis of SAM analogs needed longer incubation times. When applying an *in situ* synthesis of SAM analogs coupled to consumption of SAM analogs this would not pose a problem.











When analyzing SAM and analogs by LC MS, degradation was observed, but the extent depended on the individual LC MS instrument. The degradation was similar to that observed in solution, and the typical degradation products of SAM are shown in Figure 2.4. Similar degradation was seen for SAM analogs.

The relative initial velocities of SAM and analog formation for all SalL and FDAS enzymes are shown in Table 2.1 and Table 2.2, respectively. All assays were performed in mixtures of 200  $\mu$ M CIDA and 15 mM L-Met analog except from the *S*-benzyl- and *S*-phenethyl-L-Met analogs which due to solubility issues were tested at a concentration of only 1.5 mM. The  $K_m$  for L-Met has been estimated to be 2.6 and 4.5 mM for SalL<sup>1</sup> and FDAS<sup>5</sup>, respectively. As the applied concentration of L-Met analog did not exceed 5 x  $K_m$  for FDAS and as the  $K_m$  values for the L-Met analogs were unknown, it was decided to estimate enzyme activity as the relative initial velocity and not  $k_{cat}$ . Furthermore, the initial velocities were relative to the enzyme concentration to ease comparison of among enzymes.

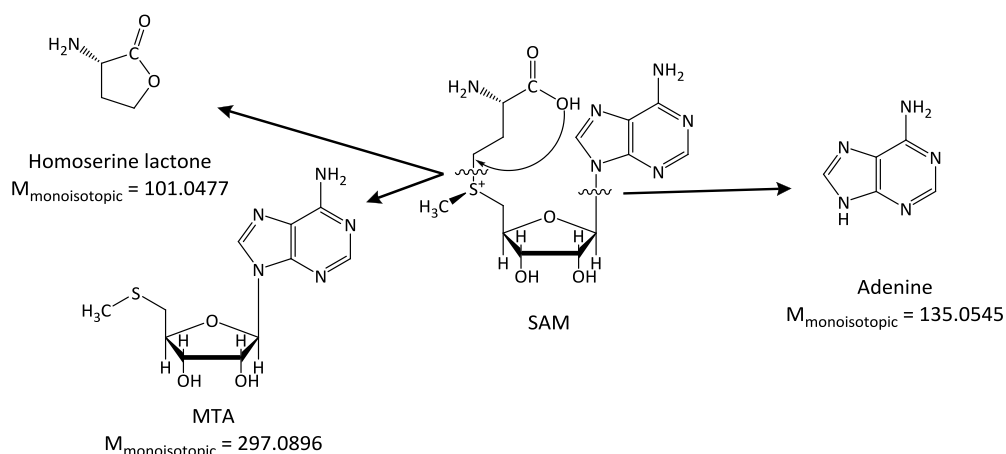
In addition, as degradation of most SAM analogs was observed during the incubation time, the area of product used to calculate the initial velocity was estimated as a summation of the area of SAM analog (both diastereomers) and the degradation products (see section 2.4).

Using L-Met, the activity of wt SalL ( $v_0 = 4,930 \pm 230$  nM/min) was approximately 500 fold higher than the activity of wt FDAS ( $v_0 = 10 \pm 2$  nM/min). These activities were for 1  $\mu$ M enzyme. This difference in activity was in the same range as reported earlier with  $k_{cat}$  values of 12 min.<sup>-1</sup> for the SalL catalyzed reaction of L-Met and CIDA and 0.03 min.<sup>-1</sup> for the FDAS catalyzed reaction of L-Met and FDA.<sup>[1, 5]</sup> This suggested that the applied concentration of L-Met was adequately above  $K_M$  resulting in maximal activities. Assuming that saturating conditions were used for L-Met, the  $k_{cat}$  for SalL and FDAS could be estimated to 4.9 and 0.01 min.<sup>-1</sup> which, although slightly different, were in the range of those reported in literature.

**Table 2.1: The relative initial velocities (nM/min) of the formation of SAM and analogs from CIDA and L-Met analogs (a-j) using different SalL variants (S) at 1  $\mu$ M. The applied denominations for the L-Met analogs (3, R = a-j) are similar to that used in the paper. NA: No activity, – : Not estimated.**

S	a	b	c	d	e	f	g	h	i	j
										
wt	4.93·10 <sup>3</sup> ± 230	453 ± 19	1.01 ± 0.10	0.318 ± 0.006	1.02 ± 0.03	-	0.101 ± 0.010	NA	NA	NA
T128S	356 ± 52	6.23 ± 1.37	NA	NA	NA	-	-	-	NA	NA
T128A	NA	NA	NA	NA	NA	-	-	-	NA	NA
T128G	123 ± 7	8.86 ± 0.14	NA	NA	NA	-	-	-	NA	NA
W190F	843 ± 51	3.37 ± 0.42	0.154 ± 0.011	0.116 ± 0.028	0.336 ± 0.012	-	-	-	NA	NA
W190H	NA	NA	0.095 ± 0.018	0.042 ± 0.012	NA	-	-	-	NA	NA
W190A	30.9 ± 2.8	1.56 ± 0.03	0.904 ± 0.096	0.934 ± 0.010	0.492 ± 0.041	-	0.075 ± 0.001	NA	NA	NA
Y239F	915 ± 294	67.5 ± 1.1	0.357 ± 0.093	NA	0.413 ± 0.082	-	-	-	NA	NA



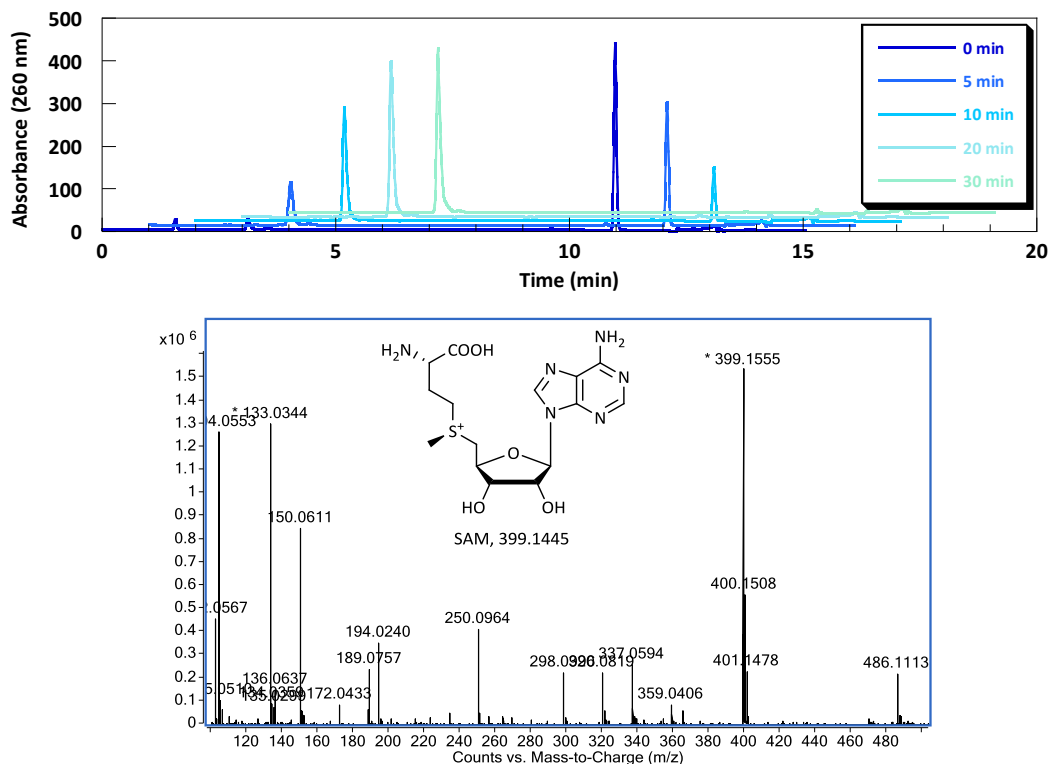


**Figure 2.4: Typical degradation of SAM observed during LC MS analysis. MTA is methylthioadenosine. The same degradation pattern was observed for SAM analogs.**

**Table 2.2: The relative initial velocities (nM/min) of the formation of SAM and analogs from CIDA and L-Met analogs (a-j) using different FDAS variants (F) at 1  $\mu$ M. The applied denomination of L-Met analogs (3, R = a-j) is similar to that used in the paper. AD: Activity detected but below quantification limit, NA: No activity. -: Not estimated.**

F	a	b	c	d	e	f	g	h	i	j
wt	10.6 $\pm 1.5$	0.729 $\pm 0.228$	NA	NA	NA	-	-	-	NA	NA
T155S	18.5 $\pm 0.14$	1.34 $\pm 0.04$	NA	NA	NA	-	-	-	NA	NA
T155A	0.314 $\pm 0.022$	NA	NA	NA	NA	-	-	-	NA	NA
W217F	5.65 $\pm 0.99$	0.033 $\pm 0.016$	NA	NA	NA	-	-	-	NA	NA
W217H	0.683 $\pm 0.050$	0.018 $\pm 0.003$	NA	NA	NA	-	-	-	NA	NA
W217A	1.60 $\pm 0.02$	0.029 $\pm 0.006$	NA	NA	NA	-	-	-	NA	NA
Y266F	15.6 $\pm 0.5$	1.01 $\pm 0.16$	AD	NA	NA	-	-	-	NA	NA
Y266H	5.97 $\pm 1.29$	0.231 $\pm 0.077$	NA	NA	NA	-	-	-	NA	NA
Y266V	2.54 $\pm 0.17$	0.114 $\pm 0.014$	NA	NA	NA	-	-	-	NA	NA
Y266L	4.69 $\pm 0.57$	0.169 $\pm 0.020$	NA	NA	NA	-	-	-	NA	NA
Y266I	3.66 $\pm 0.89$	0.171 $\pm 0.033$	NA	NA	NA	-	-	-	NA	NA
Y266A	12.2 $\pm 1.3$	0.311 $\pm 0.013$	NA	NA	NA	-	-	-	NA	NA

a) SAM



b) Ethyl-SAM

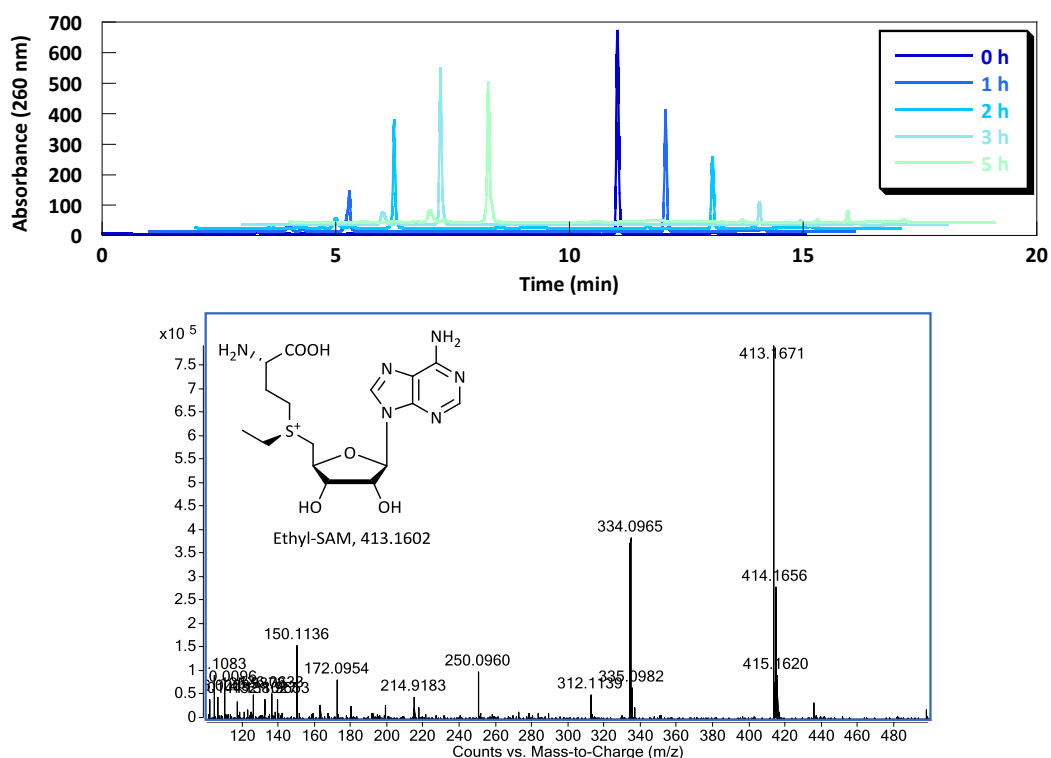
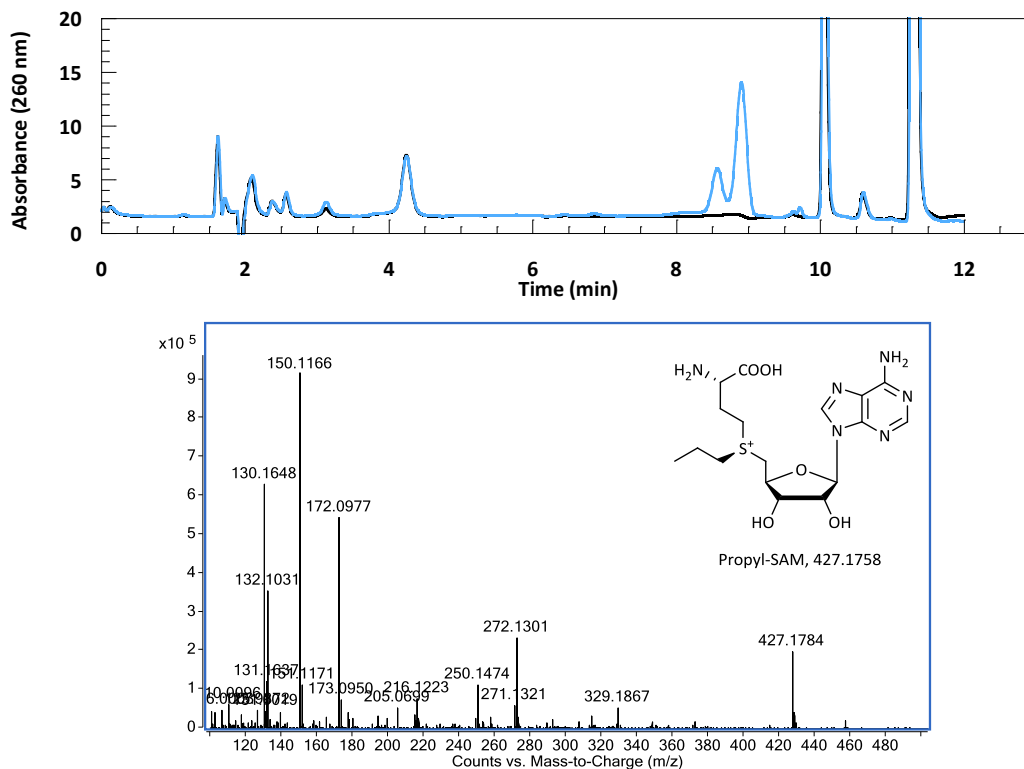


Figure 2.5: Synthesis of a) SAM and b) ethyl-SAM from 200  $\mu$ M CIDA and 15 mM L-Met and S-ethyl-L-homocysteine, respectively, using 3  $\mu$ M Sall. For both compounds, chromatograms of samples at various times (0-30 min. or 0-5 h) visualize the conversion of CIDA ( $R_t = 11$  min.) to SAM ( $R_t = 4$  min.) or ethyl-SAM ( $R_t = 5$  min.). Both reaction proceeded within the shown time span. The LC MS spectra are of the isolated peak of SAM or ethyl-SAM. The monoisotopic masses are shown below the structures. \* Saturated peak

### a) Propyl-SAM



### b) Butyl-SAM

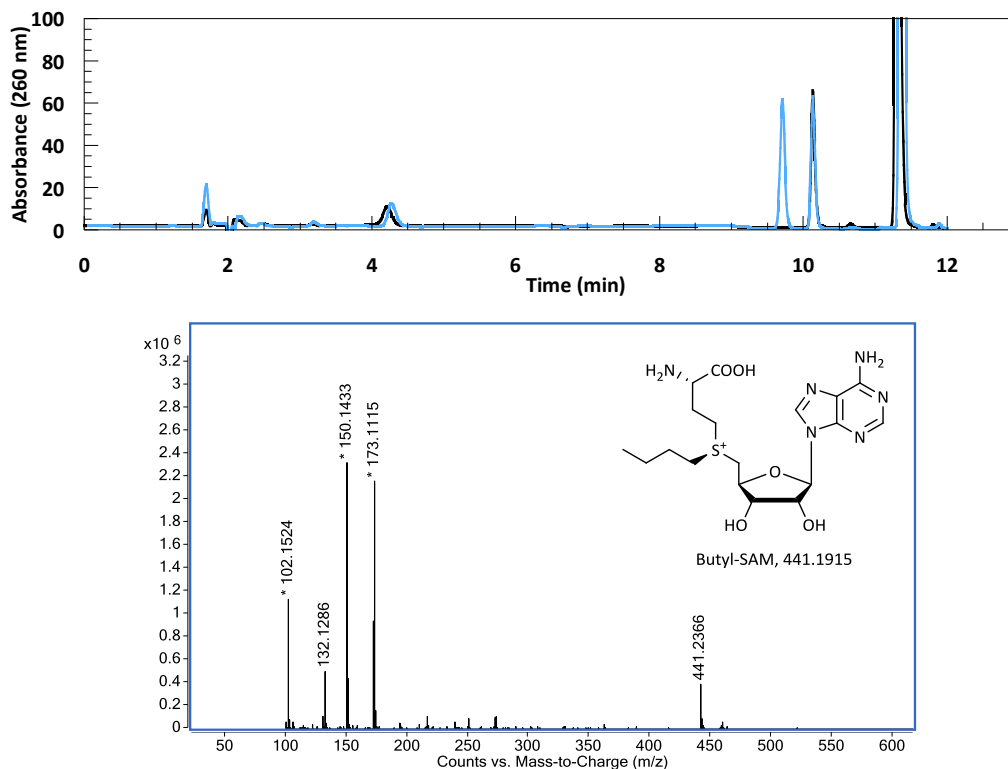
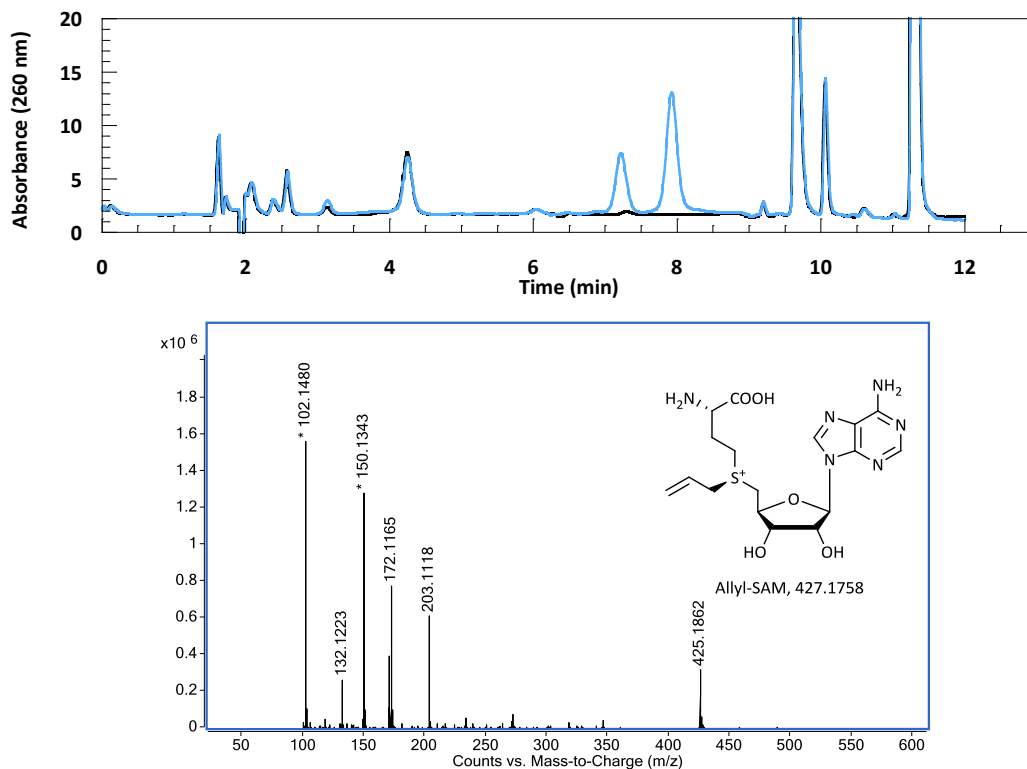


Figure 2.6: Synthesis of a) propyl- and b) butyl-SAM from 200  $\mu$ M CIDA and 15 mM *S*-propyl- and *S*-butyl-L-Hcy, respectively, using 60  $\mu$ M W190A SALL. For both analogs, chromatograms of samples taken at 0 (—) and 5 h (—) are shown to visualize the conversion of CIDA ( $R_t = 11$  min.) to propyl-SAM ( $R_t = 8.5$  and 9 min.) and butyl-SAM ( $R_t = 9.5$  min.). During 5 h's of incubation 9% yield was achieved for both analogs. The LC MS spectra are of the isolated peak with propyl- or butyl-SAM. The monoisotopic masses are shown below the structures.\* Saturated peak

a) Allyl-SAM



b) Benzyl-SAM

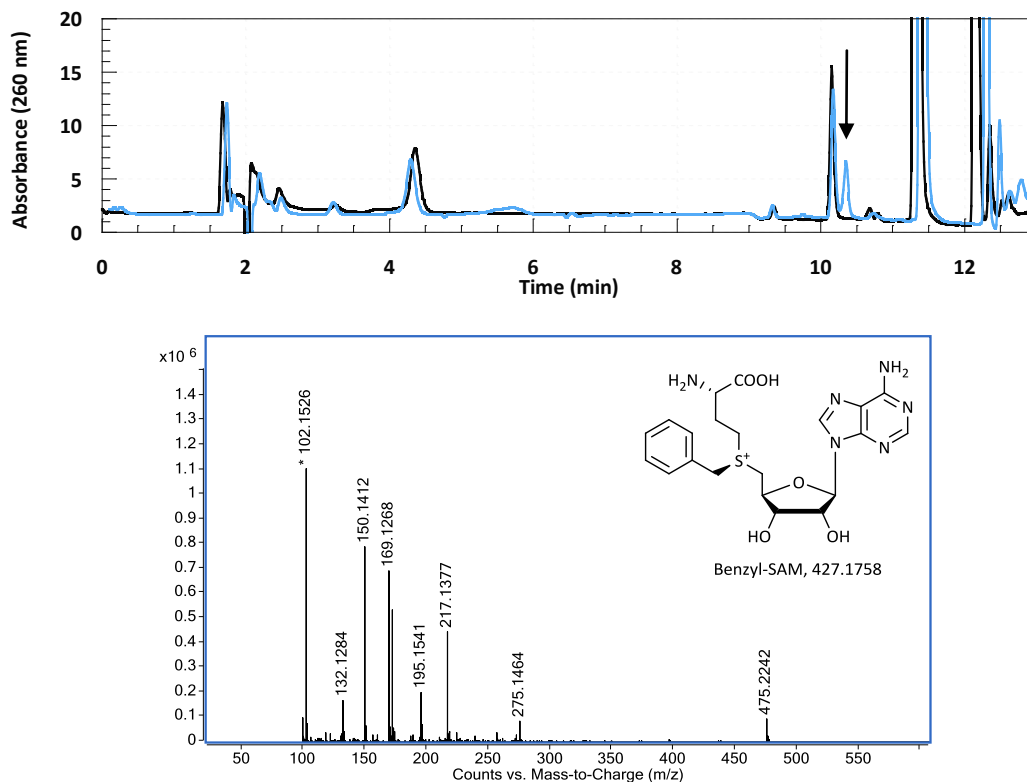


Figure 2.7: Synthesis of a) allyl- and b) benzyl-SAM from 200  $\mu$ M CIDA and 15 mM *S*-allyl-L-Hcy or 1.5 mM *S*-benzyl-L-Hcy, respectively, using 60  $\mu$ M Sall. For both analogs, chromatograms of samples taken at 0 (—) and 5 h (—) are shown to visualize the conversion of CIDA ( $R_t = 11$  min.) to allyl-SAM ( $R_t = 7.2$  and 7.9 min.) and benzyl-SAM ( $R_t = 10.4$  min., marked with  $\downarrow$ ). During 5 h's of incubation 9% and 1.5% yield was achieved for the allyl- and benzyl-SAM, respectively. The LC MS spectra are of the isolated peaks with allyl- or benzyl-SAM. The monoisotopic masses are shown below the structures. \* Saturated peak

## 2.4 Degradation of SAM and ethyl-SAM

It is well known that SAM is degraded by three mechanisms: I) Intermolecular nucleophilic attack by the oxygen of the carboxylate group of the L-Met moiety of SAM at the C<sub>γ</sub> next to the sulfonium center forming 5'-methylthioadenosine (MTA) and homoserine lactone, II) Hydrolysis into adenine and S-pentonylmethionine and III) Racemization at the sulfonium center. Whereas the rate of mechanism I and II increases with increasing pH values, the rate of racemization is independent of pH. However, due to the pH dependence of mechanism I and II, these cease below pH 1.5 and 6, respectively.<sup>[16-18]</sup>

To probe the stability of SAM and ethyl-SAM, they were prepared fresh as described in section 1.9. Chromatograms of the newly prepared samples are shown in Figure 2.8 to visualize the homogeneous samples of the intact molecules. To obtain fully degraded samples, both compounds were exposed to 70 °C in a neutral pH buffer. As can be seen in Figure 2.8, both SAM and ethyl-SAM were completely degraded into two detectable molecules (using absorbance at 260 nm). LC MS analysis of the degradation products verified the breakdown of SAM into methylthioadenosine (MTA) and adenine which was consistent with the reported breakdown mechanisms (Figure 2.9).<sup>[16-18]</sup> For ethyl-SAM, the same degradation pattern was observed (Figure 2.9). Here, ethyl-SAM was degraded into ethylthioadenosine (EtTA) and adenine. In the LC MS spectrum of MTA as well as EtTA, adenine was also detected. However, as adenine and MTA / EtTA were completely separated during the analytical HPLC, this adenine peak was assumed to originate from the degradation of MTA and EtTA during the LC MS run. As noted in section 2.3, this degradation was also seen for SAM and analogs.

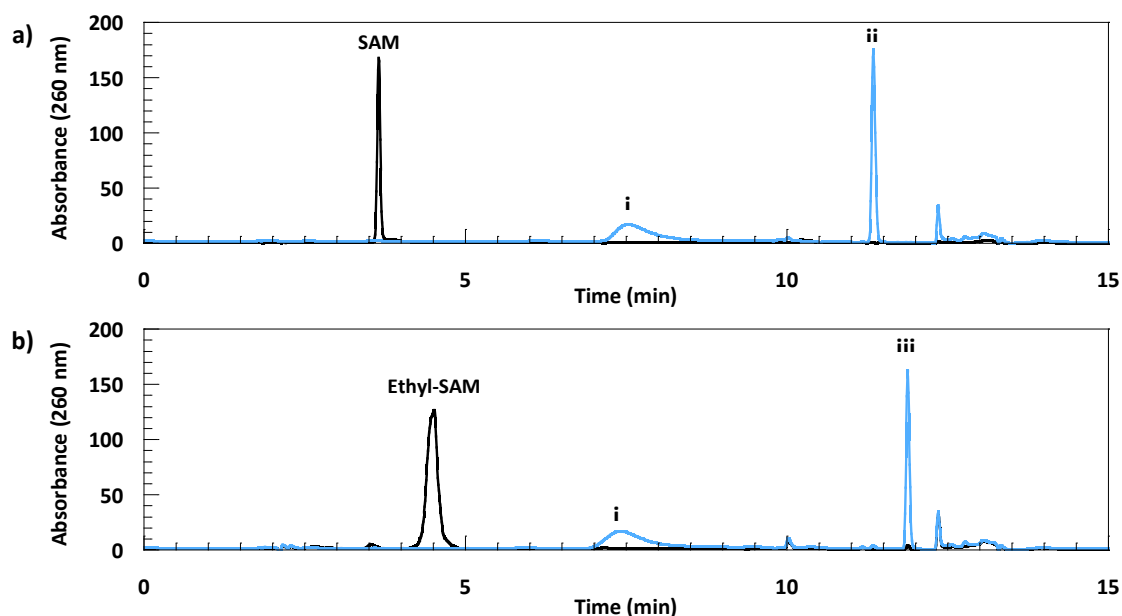


Figure 2.8: HPLC chromatogram of freshly synthesized ( — ) and degraded ( — ) a) SAM and b) ethyl-SAM. SAM and ethyl-SAM eluted at 3.7 min and 4.5 min, respectively. Three degradation products were observed: i) was formed from both SAM and ethyl-SAM, ii) was formed only from SAM and iii) was formed only from ethyl-SAM.

To obtain information about the rate of the degradation under the conditions employed in the enzyme assays, SAM and ethyl-SAM were again prepared fresh from CIDA and L-Met or S-ethyl-L-homocysteine, respectively. In this assay, however, the stability was tested at 37 °C in the reaction buffer (pH 7.4) and in the quenching buffer (pH 3.0). Samples were taken with regular time intervals and degradation products were detected by analytical HPLC (Figure 2.10). For SAM, the rate of racemization could not be accessed as the diastereomers were not properly separated by analytical HPLC using the described conditions. The estimated concentration of SAM was therefore the total concentration of both diastereomers. For SAM, the total degradation (excluding racemization) within

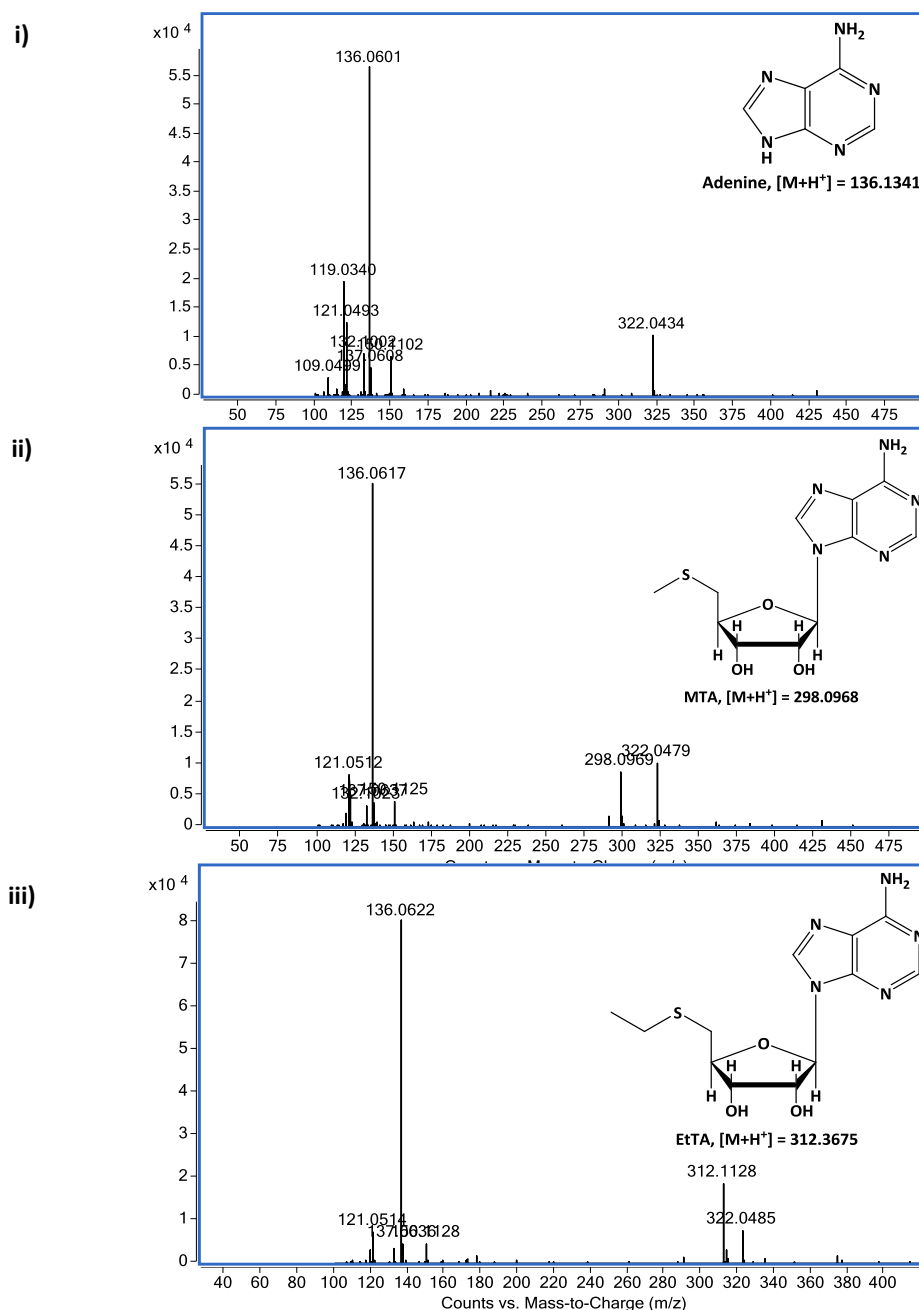


Figure 2.9: LC MS spectra of the breakdown products of SAM and ethyl-SAM (i, ii and iii) observed in the HPLC chromatograms (Figure 2.8).

5 hours of incubation was lowered from 15.3% at pH 7.4 to 3.5% at pH 3.0 due to a 50% reduction in MTA degradation and a complete elimination of adenine degradation.

For ethyl-SAM the same trends were seen. Here the total degradation within five hours of incubation was reduced from 32.4% at pH 7.4 to 21.6% at pH 3.0 (including racemization). The increase in the percentage of total degradation was due to the fact that for ethyl-SAM the racemization could be assessed and the racemization accounts for approximately 13% at pH 7.4 and 19% at pH 3.0. The pH variance was believed to be the result of an increased degradation of both the inactive and active diastereomers at pH 7.4. If SAM racemized with the same rate as observed for ethyl-SAM the degradation pattern was similar for SAM and ethyl-SAM.

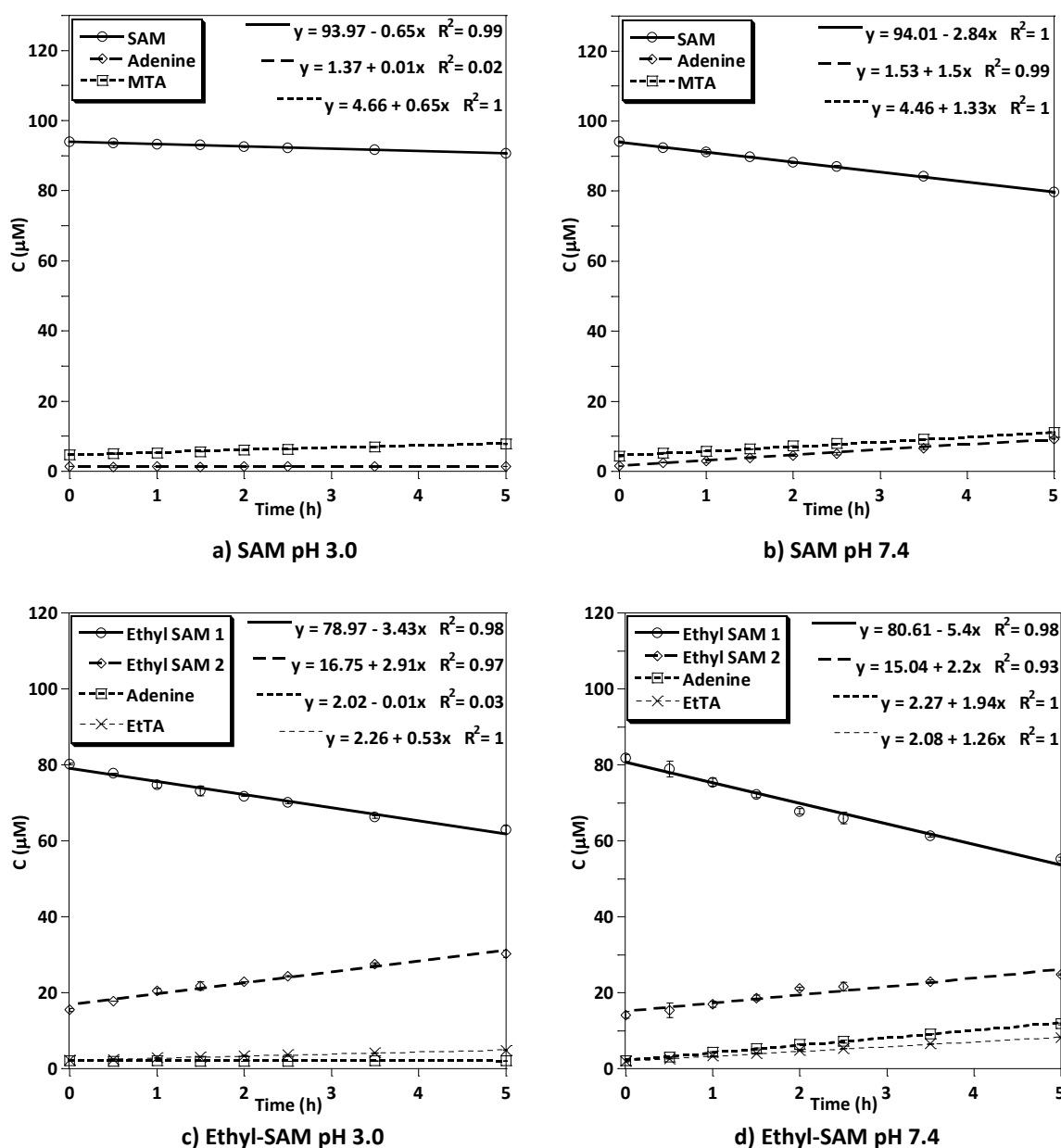


Figure 2.10: The rates of SAM and ethyl-SAM degradation at 37 °C and pH 3.0 or 7.4. The concentrations of each degradation product as well as starting compounds were estimated as described in section 1.9. Standard deviations are shown as bars for all data. Racemization was detected for ethyl-SAM but not SAM due to the lack of a proper separation of the SAM diastereomers on the analytical HPLC.

These data were consistent with a study of Hoffmann (1986) except from the estimated relative rate of racemization.<sup>[16]</sup> Here the rate of racemization was 40% of that of adenine formation at pH 7.5, whereas in this study the rate of racemization was estimated to 120% of the rate of adenine formation. Possibly, the rate of racemization was overestimated due to overlapping peaks in the HPLC chromatogram as this complicated the exact calculation of the areas of each diastereomer.

The same assays were performed for CIDA, adenosine, and SAH. However, for all of these compounds no degradation was observed at either pH 3.0 or 7.4 (data not shown). This confirmed that the inherent instability of SAM and analogs owed to the presence of the positively charged sulfonium ion which was consistent with a study by Borchardt.<sup>[17]</sup>

## 2.5 Enzymatic modification of the RGG peptide by rPRMT1

Application of the SAM analogs in a coupled assay was then probed. Initially, two positive controls were performed – one in which the activity of rPRMT1 was tested by incubation with SAM and one in which the enzymatic SAM formation was coupled to the methylation of the RGG peptide. For both assays, a symmetrical peak of fully hexa-methylated RGG peptide was observed by LC MS indicating that the rPRMT1 was indeed active and that the coupled assay was successful (Figure 2.11 and Figure 2.16).

Then, it was tested if the L-Met analogs were accepted by rPRMT1. Initially, unpurified L-Met analogs were used. However, here methylation activity was observed in all but the sample with the *S*-benzyl-L-Hcy. This was an issue as the masses of ethyl-, propyl- and butyl-modified peptide were equal to di-, tri-, and tetra-methylated peptide. It was concluded that the L-Met impurity came from L-homocystine as a coupled reaction with L-homocystine resulted in methylation (Figure 2.16).

Using purified L-Met analogs no methylation was observed and modification could be assessed for all L-Met analogs. Representatives for all assays are shown in Figure 2.12 to Figure 2.15. Furthermore, for comparison the LC MS of unmodified RGG peptide is shown in Figure 2.13. As noted in the paper, ethyl, allyl, and benzyl were accepted as substrates. Thus, there was a clear difference in the activity of the allyl and propyl analog although they had similar sizes. As both SAM analogs were produced in similar rates, this indicated that the double bond beta to the sulfur accelerated the reaction which was consistent with a study of DNA MTases.<sup>[19]</sup> Concerning the benzyl analog of L-Met, modified RGG peptide was observed in small but reproducible amounts (Figure 2.15).

For most of the LC MS chromatograms shown here, there were some minor impurities eluting at similar retention times as for un- and modified RGG peptide. These originated from the LC MS instruments and enzyme solutions (data not shown).

Please note, that the MS software by default numbers the peak with highest occurrence. This meant that for some MS peaks the numbered peak was not the monoisotopic ion. However, for all peaks which were assigned as double, triple or tetra-protonated peptides, the respective peaks were checked manually for the monoisotopic mass and charge.



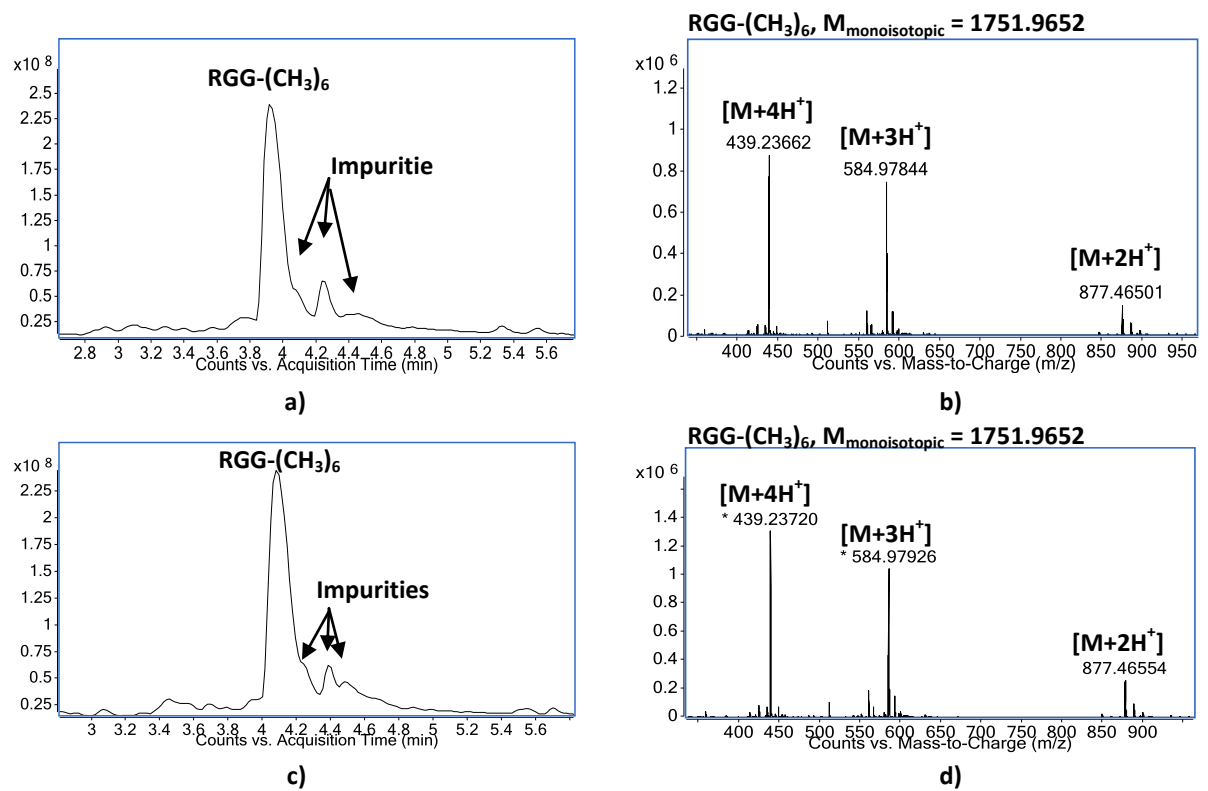


Figure 2.11: Modification of RGG peptide using SAM (a and b) and in a coupled assay with Sall, L-Met, and CIDA (c and d). a and c) LC MS chromatogram of the RGG peptide peak. b and d) MS spectrum of the modified RGG peptide.

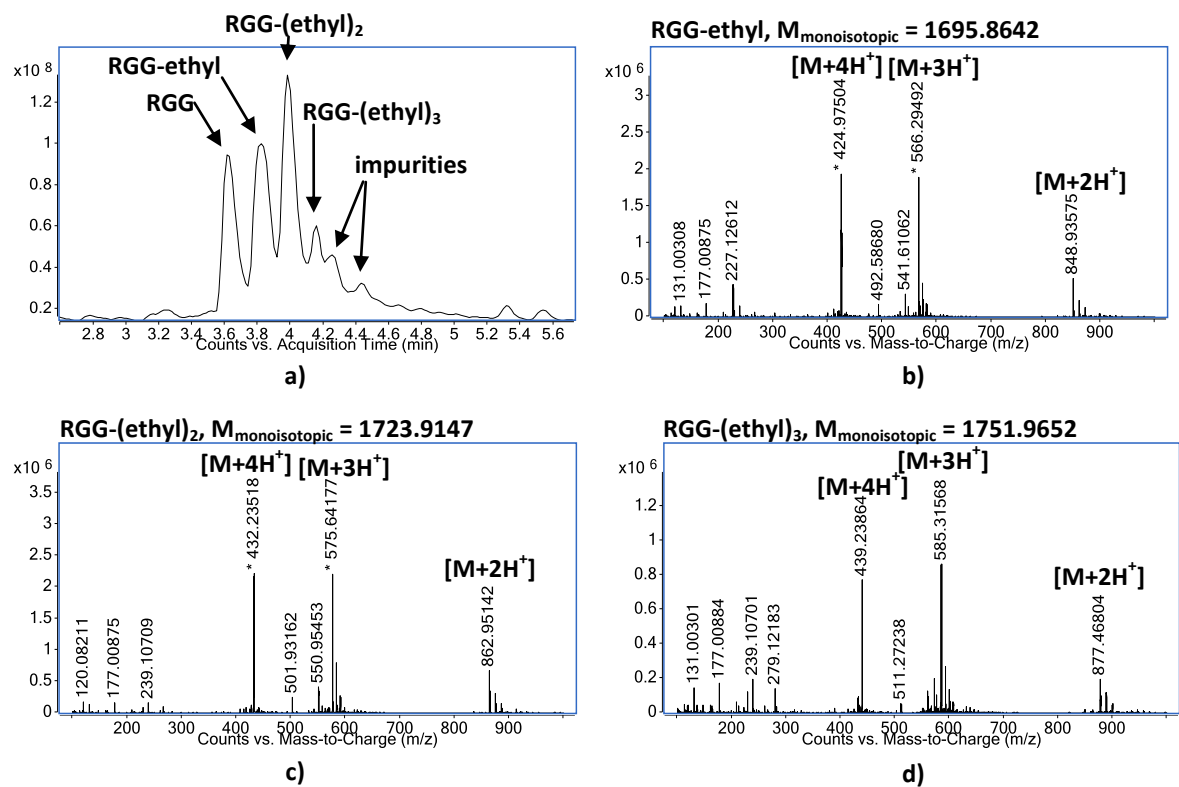
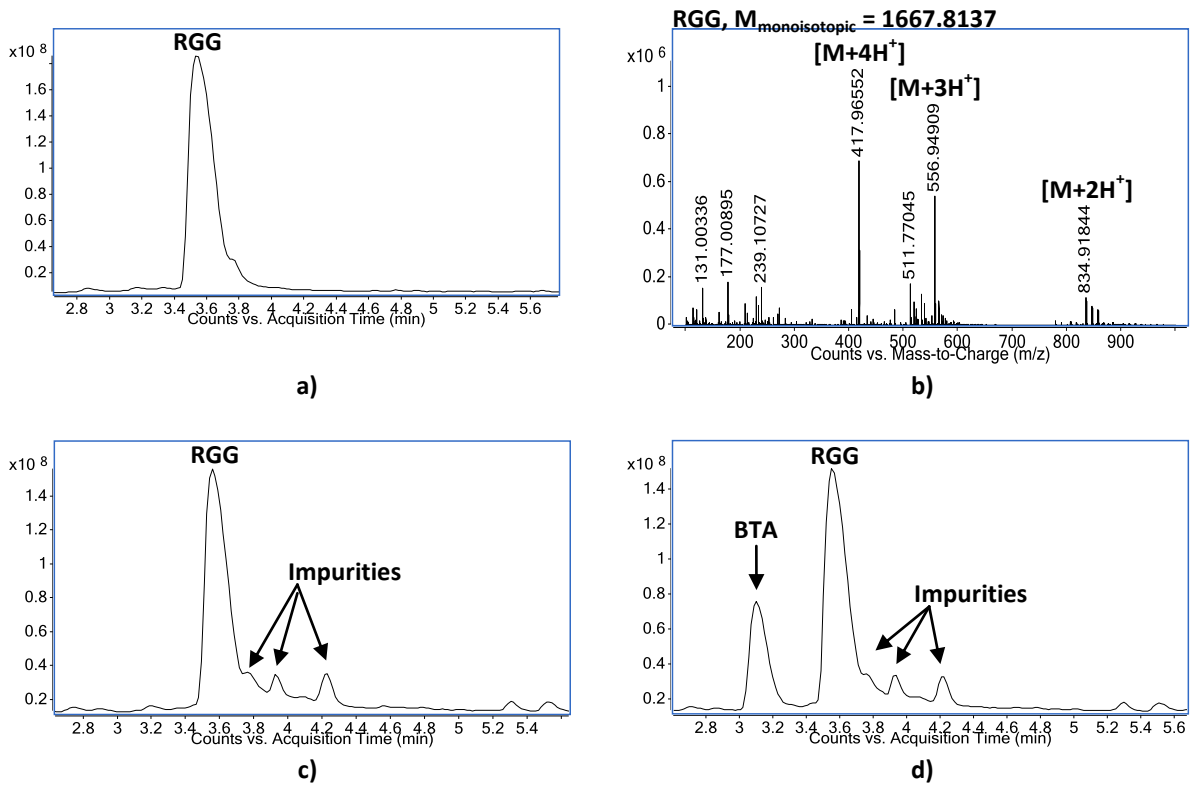
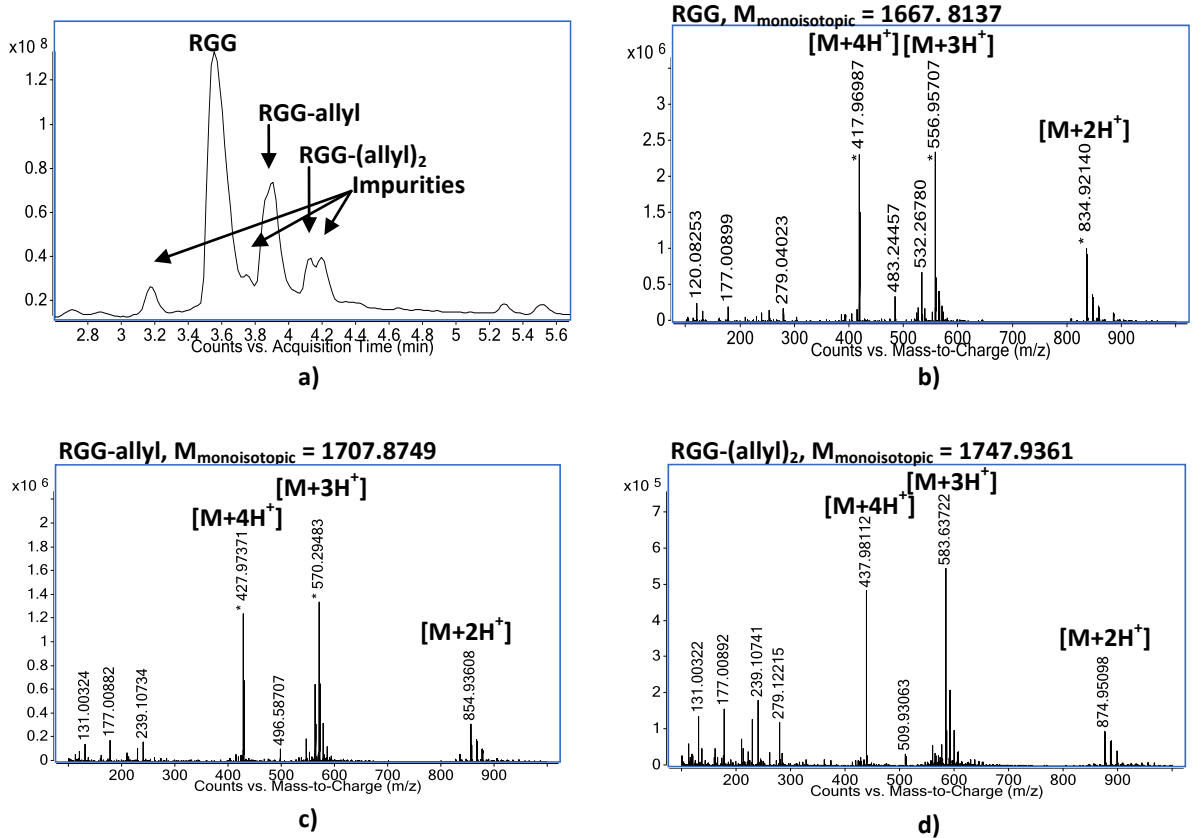


Figure 2.12: Modification of the RGG peptide in a coupled assay with Sall, S-ethyl-L-Hcy, and CIDA. a) LC MS chromatogram. b), c) and d) MS spectra of the mono-, di- and tri-modified RGG peptide, respectively.



**Figure 2.13:** a) and b) LC MS chromatogram and spectrum of unmodified RGG peptide, respectively. c) and d) modification of the RGG peptide in a coupled assay with Sall, CIDA, and S-propyl- or S-butyl-L-Hcy, respectively. No activity was observed in these assays. BTA: Butylthioadenosine.



**Figure 2.14:** Modification of the RGG peptide in a coupled assay with Sall, S-allyl-L-Hcy, and CIDA. a) LC MS chromatogram. b), c) and d) MS spectra of the un-, mono- and di-modified RGG peptide.

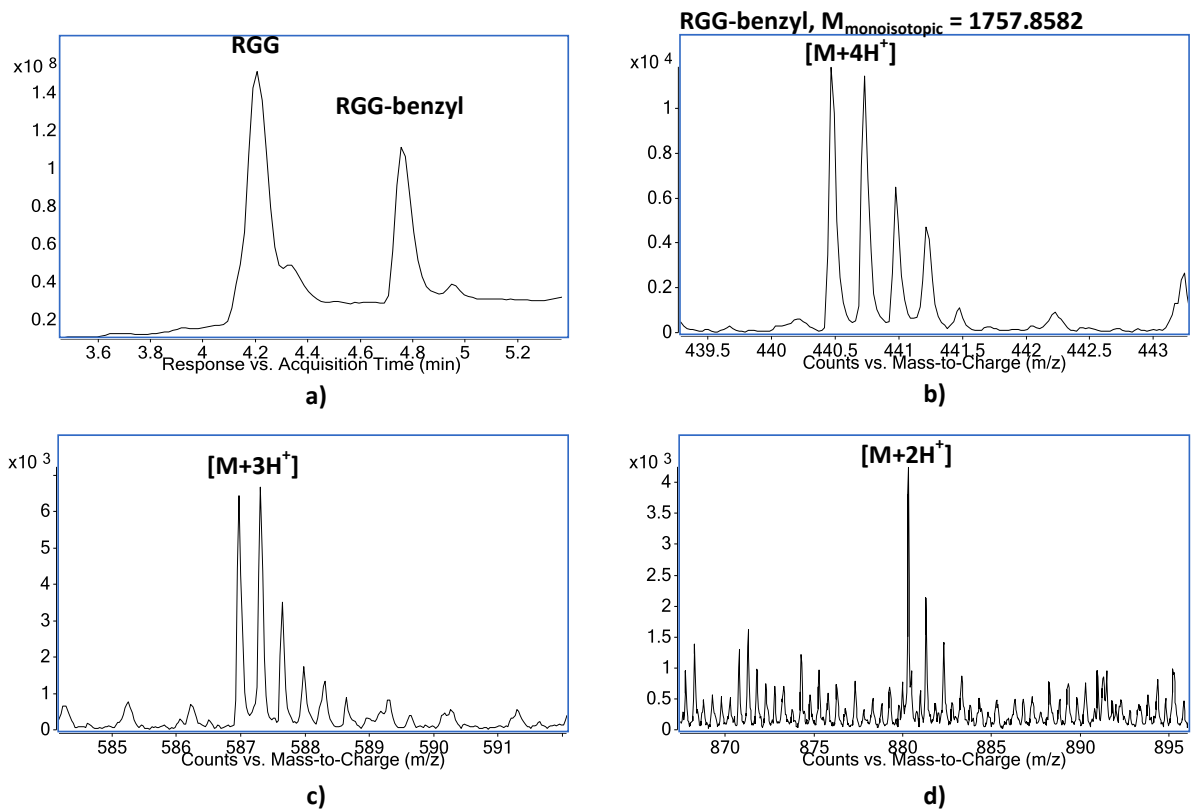


Figure 2.15: Modification of the RGG peptide in a coupled assay with wt SalI, S-benzyl-L-Hcy and CIDA. a) LC MS chromatogram. b-d) Enlargements of the MS spectra of the modified RGG peptide.

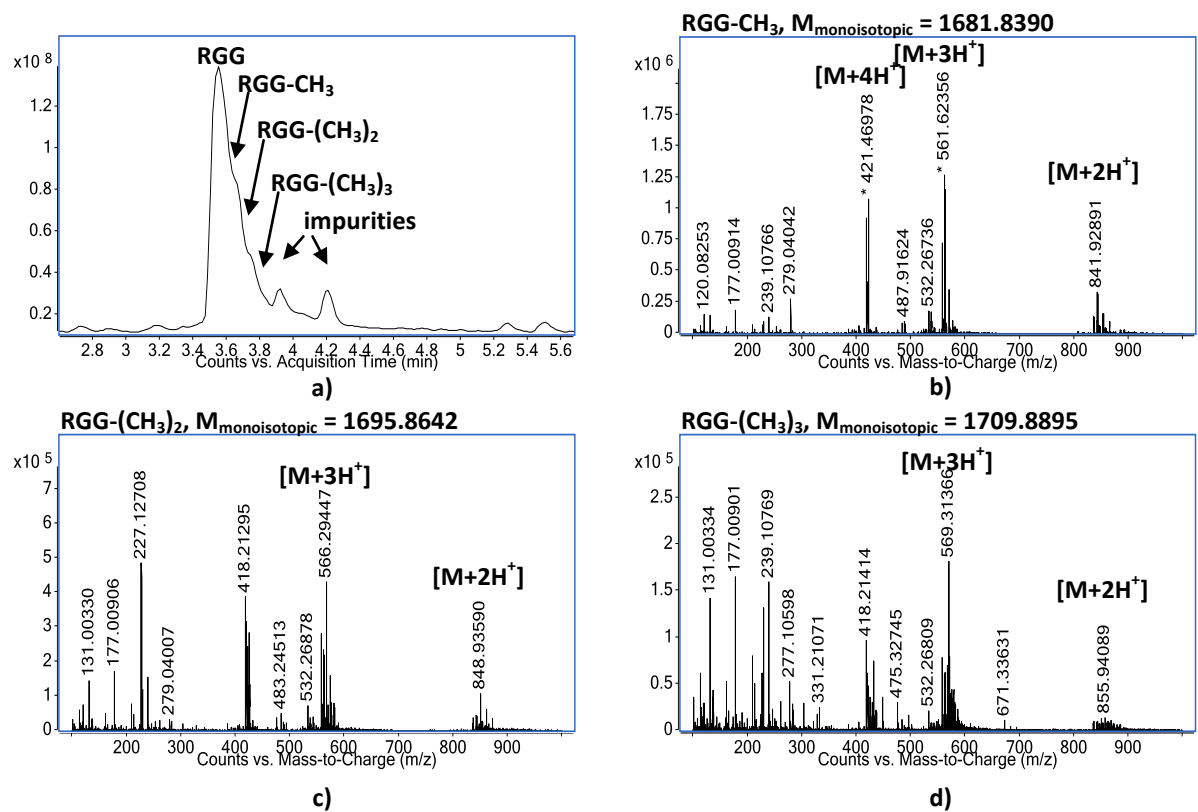


Figure 2.16: Modification of the RGG peptide in a coupled assay with SalI, L-homocysteine, and CIDA. a) LC MS chromatogram. b), c) and d) MS spectra of the mono-, di-, and tri-methylated RGG peptide.

## 2.6 References

- [1] A. S. Eustaquio, F. Pojer, J. P. Noel, B. S. Moore, *Nat Chem Biol* **2008**, *4*, 69.
- [2] X. F. Zhu, D. A. Robinson, A. R. McEwan, D. O'Hagan, J. H. Naismith, *J Am Chem Soc* **2007**, *129*, 14597.
- [3] C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan, J. H. Naismith, *Nature* **2004**, *427*, 561.
- [4] S. L. Cobb, H. Deng, A. R. McEwan, J. H. Naismith, D. O'Hagan, D. A. Robinson, *Org Biomol Chem* **2006**, *4*, 1458.
- [5] H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson, J. B. Spencer, *Angew Chem Int Ed* **2006**, *45*, 759.
- [6] A. R. McEwan, H. Deng, R. P. McGlinchey, D. R. Robinson, D. O'Hagan, J. H. Naismith, *According rcsb.org to be published*.
- [7] P. W. Rose, B. Beran, C. Bi, W. F. Bluhm, D. Dimitropoulos, D. S. Goodsell, A. Prlic, M. Quesada, G. B. Quinn, J. D. Westbrook, J. Young, B. Yukich, C. Zardecki, H. M. Berman, P. E. Bourne, *Nucleic Acids Res* **2011**, *39*, D392.
- [8] G. J. Kleywegt, M. R. Harris, J. Y. Zou, T. C. Taylor, A. Wahlby, T. A. Jones, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2240.
- [9] F. W. Studier, *Protein Expr Purif* **2005**, *41*, 207.
- [10] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel, A. Bairoch, *Nucleic Acids Res* **2003**, *31*, 3784.
- [11] D. Yamashiro, H. L. Aanning, L. A. Branda, W. D. Cash, V. V. Murti, V. du Vigneaud, *J Am Chem Soc* **1968**, *90*, 4141.
- [12] J.M. Lipson, M.Thomsen, B.S. Moore, R.P. Clausen, J.J. La Clair, M.D. Burkart, *ChemBioChem*. DOI: 10.1002/cbic.201300221
- [13] R. Wang, W. Zheng, H. Yu, H. Deng, M. Luo, *J Am Chem Soc* **2011**, *133*, 7648.
- [14] W. Peters, S. Willnow, M. Duisken, H. Kleine, T. Macherey, K. E. Duncan, D. W. Litchfield, B. Luscher, E. Weinhold, *Angew Chem Int Ed* **2010**, *49*, 5170.
- [15] K. Islam, W. Zheng, H. Yu, H. Deng, M. Luo, *ACS Chem Biol* **2011**, *6*, 679.
- [16] J. L. Hoffman, *Biochemistry* **1986**, *25*, 4444.
- [17] R. T. Borchardt, *J Am Chem Soc* **1979**, *101*, 458.
- [18] R. T. Borchardt, Y. Shiong, J. A. Huber, A. F. Wycpalek, *J Med Chem* **1976**, *19*, 1104.
- [19] C. Dalhoff, G. Lukinavicius, S. Klimasauskas, E. Weinhold, *Nat Chem Biol* **2006**, *2*, 31.