DOI: 10.1002/cbic.200800605

Fine Tuning the N-Terminal Residue Excision with Methionine Analogues

Birgit Wiltschi, Lars Merkel, and Nediljko Budisa*[a]

This article is dedicated to Prof. Luis Moroder on occasion of his 68th birthday.

Protein biosynthesis starts with methionine (Met) in all living cells.[1–5] In the cytosol of eukaryotes[6] and archaebacteria,[7] protein translation is initiated with Met whereas *N*-formylmethionine (fMet) is used in eubacteria,[8] the mitochondria[9] and chloroplasts.[10] The formyl group is added to the free α-amino group of Met-tRNA^{fMet} by the methionyl-tRNA^{fMet} formyltransferase (EC 2.1.2.9) and fMet-tRNA^{fMet} participates in translation.[2] Met is always the first amino acid incorporated into the N-terminal position of a protein, even if alternative start codons, such as GUG and UUG are used for translation initiation.[11] However, the amino acid sequence of mature proteins seldom contains Met at the first position; the N termini of most proteins are modified by a variety of co-, and post-translational processing events.

Peptide deformylase (EC 3.5.1.88) removes the *N*-formyl group from proteins in prokaryotes and organelles. The unmasked amino terminal Met can subsequently be cleaved from prokaryotic as well as eukaryotic proteins by the Met aminopeptidase (MetAP; EC 3.4.11.18). Additionally, novel N termini are generated by the enzymatic cleavage of leader sequences, or by proteolytic digest at one or more positions within the polypeptide chain. The covalent attachment of diverse chemical functionalities, such as acetyl-, phosphate-, and myristoyl-groups, or ubiquitin represents other N-terminal modifications beyond proteolytic processing.[12,13]

As most N-terminal Met residues are co- or post-translationally removed, their function lies in translation initiation rather than in structure.[14] Depending on the organism, between 55 and 70% of the proteins are subject to N-terminal Met excision (NME) by MetAP.[15] NME is an irreversible cotranslational process that occurs as soon as the first N-terminal residues of the nascent peptide protrude from the ribosomal exit tunnel, before protein folding starts to occur.[12,15,16] In eukaryotic cells, this enzymatic processing coincides with the intracellular localization of MetAP close to the ribosome.[17] Although the localization of bacterial MetAP is still unclear (C. Klein, personal communication), it is reasonable to suppose that it is ribosome associated as well.

MetAP specifically cleaves Met in the first position of the precursor protein.[15] Meinnel and coworkers showed that model peptides with N-terminal Leu and Phe could be processed by *Escherichia coli* MetAP (EcMAP1) in vitro but with considerably lower catalytic efficiency than Met.[18] The cleavage efficiency of Met in the first position largely depends on the bulkiness and nature of the side chain of the second amino acid (Figure 1).[15] In vivo[13] and in vitro[18] data analyses

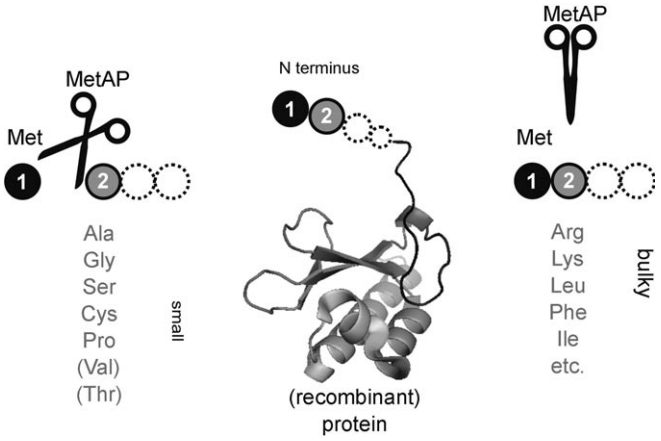

Figure 1. N-terminal Met excision (NME) is an irreversible cotranslational modification that is completed before the nascent polypeptide chain is fully synthesized. The basic rule is that small amino acid residues in the second position after the N-terminal Met facilitate NME, whereas bulky residues inhibit the process. MetAP, Met aminopeptidase.

provided convincing evidence that NME is likely to occur if the second residue of the precursor is Ala, Cys, Gly, Pro or Ser. Met in the first position can also be excised if Thr or Val follows in position 2, but with lower efficiency.[18] Excision of the N-terminal Met with Ile and Asn at the second position is inefficient but may occur if the sequence context (vide infra) supports it.[12] NME is not observed with the other amino acids at the second position.[12,13,18] Although the data were collected predominantly in in vitro studies with EcMAP1 or in vivo with proteins from *E. coli*, MetAPs from diverse organisms all have the same specificity.[13,18] Taken together, smaller amino acids, especially Ala, at the second position optimally support N-terminal Met excision,[18] whereas large side chains at the second position block the action of MetAP (Figure 1).[13]

According to earlier studies, the third position in the precursor protein does not influence NME.[13] However, recently it was shown that not only the amino acid in position 2 but additional residues following the first Met have a major impact on NME efficiency.[18] In vitro excision of the first Met from model peptides with either Ala or Gly in the second position was optimal if Trp, Met, or Ser followed in the third position. In con-

[a] *Dr. B. Wiltschi, L. Merkel, N. Budisa*
Max Planck Institute of Biochemistry, Molecular Biotechnology
Am Klopferspitz 18, 82152 Martinsried (Germany)
Fax: (+49) 89-8578-2815
E-mail: budisa@biochem.mpg.de

trast, Asp, Ile, Thr, Pro, and Glu in the third position had a negative influence on Met cleavage in vitro as well as in vivo. Meinnel and co-workers were able to show that even the amino acid residues in positions 4 and 5 exert certain effects on NME.[18]

Although NME is a rapid and efficient process, the N-terminal Met may erroneously be retained on some proteins that are recombinantly overproduced in *E. coli*, especially when strong promoters like T7 or T5 are used. This phenomenon is most probably due to the saturation of the peptide deformylase and MetAP capacities or targeting of the nascent protein, for example, into inclusion bodies.[13,15]

The nature of the N-terminal residue and the half-life of a protein are directly correlated (N-end rule).[19–22] Apart from rare exceptions,[23] small N-terminal residues stabilize the protein, whereas bulky residues destabilize it. Met occupies a special position since it appears to have a stabilizing effect.[21] Since NME generates new N termini, it determines the metabolic fate of a protein, such as its turnover or its further maturation through other N-terminal modifications. Only recently has NME been recognized as a promising therapeutic target and specific inhibitors for this process are being developed.[15]

Noncanonical Met analogues that are chemically and/or sterically similar to Met have long been used as antagonists to study cellular Met metabolism.[24] Very early it was recognized that the Met analogues norleucine (Nle; Figure 2) and ethionine (Eth; contains an ethyl-thioether instead of the methyl-thioether) can be translated into proteins in *E. coli* in place of Met, though they inhibit bacterial growth.[24] Growth inhibition occurs due to the inability of the analogues to replace Met as the precursor of S-adenosyl Met.[25] Nle and Eth are charged equally onto initiator $tRNA^{fMet}$ and elongator $tRNA^{Met}$ by methionyl-tRNA synthetase, though less efficiently than Met.[26] Norleucyl-$tRNA^{fMet}$ and ethionyl-$tRNA^{fMet}$ are formylated similarly to Met[25–27] and the formylated species efficiently initiate translation.[25,27] Likewise, N-formylnorleucine or N-formylethionine appear to be substrates for the peptide deformylase,[28] and for MetAP.[25,28] This indicates that due to the catalytic promiscuity of the involved enzymes such as the methionyl-tRNA synthetase, methionyl-$tRNA^{fMet}$ formyltransferase, the ribosome, the peptide deformylase, and MetAP, the noncanonical Met analogues can participate in protein translation and subsequent N-terminal modifications.

Recently, Tirrell and co-workers demonstrated nearly quantitative substitution of the Met residues in dihydrofolate reductase by homopropargylglycine (Hpg; Figure 2). The initiator Met was also substituted, but the Hpg in the first position was not excised.[29,30] Similarly, by global replacement of Met with trifluoromethionine in green fluorescent protein, we serendipitously discovered an efficient blockage of the N-terminal excision of trifluoromethionine even though Ser was the second amino acid.[31] In order to study N-terminal excision of Met analogues in more detail, we heterologously expressed human epidermal growth factor (hEGF) in *E. coli* in the presence of Hpg, and azidohomoalanine (Aha; Figure 2). Both Met analogues were incorporated at the N terminus of hEGF and their cleavage rates provided new insights into the rules for N-terminal residue excision (NRE). According to the established NME rules (see above), neither Met nor Aha or Hpg were cleaved from the N terminus of hEGF when Arg was the second amino acid.[32] When the NME supporting amino acid Gly was at the second position, Met > Aha > Hpg were excised in that order.[32] This was the first indication that the excision efficiency depends on the chemical nature of the N-terminal amino acid, even if the second amino acid is small and therefore would facilitate NRE (Figure 2). Similarly, by exploiting NRE, Wang and co-workers sought to preserve Aha or Hpg at the N terminus of recombinantly expressed human β-interferon (IFNβ) for subsequent bioorthogonal modification. They also found that replacing Met in the first position at the N terminus with Aha or Hpg decreased the activity of both the deformylase and MetAP to some extent.[33] These authors were able to detect minor amounts of formylated Hpg,[33] a fact that we did not observe in our protein preparations.[32] Met in the first position of IFNβ was completely cleaved, independent of whether the second amino acid was Ala, Ser, or Gly. Ala in the second position promoted Aha and Hpg excision to nearly completion. However, excision efficiencies of both Met analogues were reduced with either Ser or Gly in the second position. Only when Gln, Glu, or His occupied position 2, was excision of Met as well as Aha and Hpg completely blocked.[33] This observation is in agreement with our findings that NRE is blocked with bulky amino acids in the second posi-

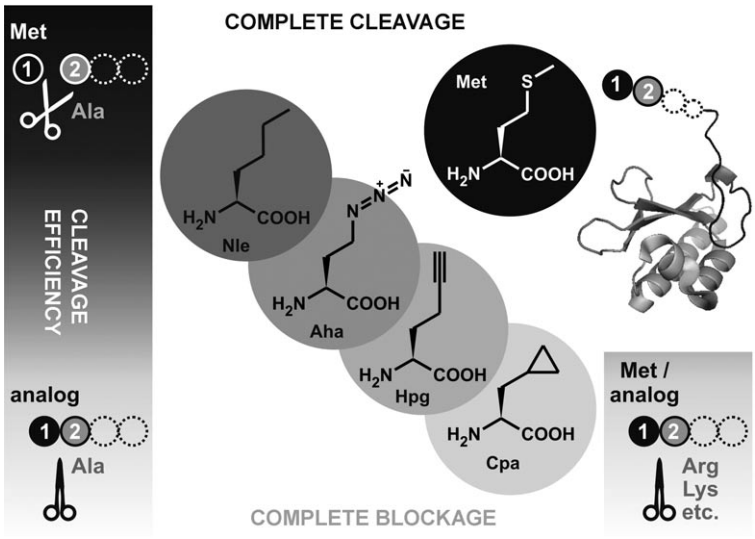

Figure 2. Tuning the N-terminal residue excision from complete cleavage to complete blockage. The chemical nature of the N-terminal amino acid determines its excision efficiency from a recombinant protein with a small amino acid in the second position. Among the tested noncanonical Met analogues, Nle showed excision almost as efficient as Met. Aha was less efficiently excised, whereas Hpg and Cpa were bad substrates for MetAP, and their in vivo cleavage efficiency was dramatically reduced. Note that a bulky residue at position 2 efficiently blocks excision of the N-terminal residue, regardless of its chemical nature. Although Cpa is chemically and structurally not a Met analogue, it is a substrate for methionyl-tRNA synthetase (unpublished results).

tion. However, the excision of Aha and Hpg was found to be less efficient even with Gly occupying this position in hEGF or IFNβ precursor proteins.

In order to further sort out the NRE rules with Met analogues, we conducted in vitro studies with EcMAP1 and model peptides containing the Met analogues Nle, Aha, Hpg and β-cyclopropylalanine (Cpa; Figure 2) in the first and different canonical amino acids in the second and third positions. Cleavage of all Met analogues was efficiently blocked when the second position in the model peptide was occupied by the bulky amino acid arginine.[32] In contrast, Ala at position 2 facilitated the excision of the N-terminal Met analogues except Cpa (cleavage efficiency: Met, Nle, Aha ≫ Cpa; Figure 3A). Thus, the chemical and/or structural nature of the first amino acid influences NRE even if the second amino acid is small and facilitates NRE, such as Ala. Frottin and associates observed a similar trend with the Met analogues, Met sulfoxide (Met(O)), and norvaline (Nva): Met and Nle were more efficiently excised from tripeptides with Ala in the second position than Met(O) and Nva (cleavage efficiency: Met, Nle > Met(O), Nva).[18] We found an altered in vitro cleavage efficiency of Met, Nle, Aha, Hpg and Cpa when Ala (in the second position) was exchanged for Gly (Figures 3B, 3C). In this case, the similarity of the Met analogy became a more important determinant for excision, since the cleavage efficiency declined in the order Met, Nle > Aha ≫ Cpa (unpublished results; Figure 3B) or Met > Aha > Hpg[32] (Figure 3C). The cleavage pattern of Met and its analogues remained unchanged with Ser (unpublished observation) or Gln at the third position (Figures 3A, 3B).

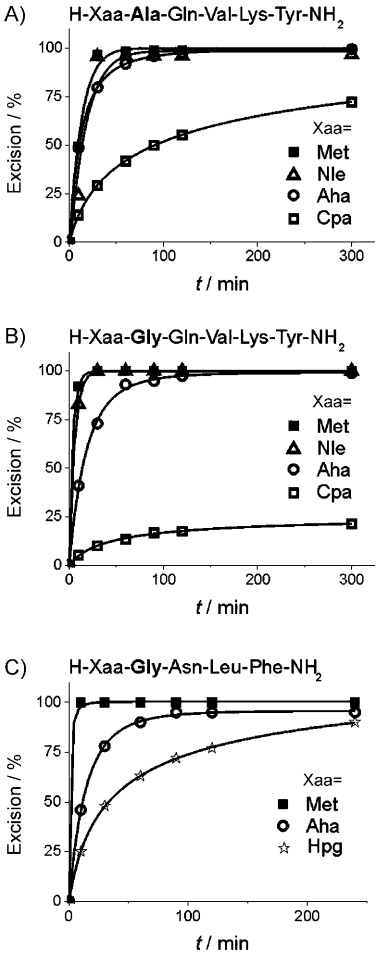

Figure 3. Substrate specificity of *E. coli* MetAP. In vitro excision of N-terminal Met, Nle, Aha and Cpa from the model peptides A) H-Xaa-**Ala**-Gln-Val-Lys-Tyr-NH_2 and B) H-Xaa-**Gly**-Gln-Val-Lys-Tyr-NH_2 by EcMAP1 (our unpublished observations). C) NRE of Met, Aha and Hpg from the model peptide H-Xaa-**Gly**-Asn-Leu-Phe-NH_2.[32] Xaa denotes Met or one of the analogues and the amino acid in position 2 is printed in bold. MetAPs act cotranslationally; their substrates are unstructured, short N-terminal peptides that protrude from the ribosome tunnel. While EcMAP1 cleaves dipeptides extremely inefficiently, tripeptides and larger peptides are suitable substrates for in vitro cleavage assays. In vivo and in vitro data analyses have shown that the kinetics of in vitro excision is directly related to the efficiency of in vivo NRE (unpublished observation).[18,32]

Our findings agree with those reported by Frottin and coworkers. Similar to their in vitro study,[18] we observed that peptides with a noncanonical Met analogue in the first and Ala in the second position were cleaved most efficiently. However, peptides with Gly at position 2 were less efficiently cleaved and the nature of the first amino acid residue became more important. This reflects very well the in vivo situation in the Aha and Hpg variants of hEGF[32] and IFNβ.[33] As for NME, MetAP activity appears to be the limiting factor for NRE. The ribosomal translation of amino acids into polypeptides is a rapid (approximately 10–20 residues per second in *E. coli*)[34] and highly accurate process. Being a cotranslational event, NME must occur with optimal specificity at high rates, as well. Although we did not assess the kinetic parameters for the cleavage of structurally and/or chemically distinct Met analogues by MetAP, it is reasonable to expect that in comparison to Met they are less efficient substrates. This might explain the markedly reduced excision rate of chemically and/or sterically different Met analogues such as Cpa.

As already mentioned above, the MetAPs from different organisms have the same specificity with respect to Met in the first and small or bulky residues in the second position of a polypeptide.[13,18] Interestingly, model peptides with Aha in the first position are processed more efficiently by the *E. coli* MetAP than variant peptides with Hpg. The opposite is true for the MetAP from *Pyrococcus furiosus*; this enzyme cleaves Hpg more efficiently than Aha if Gly is at position 2 (unpublished results). Based on these preliminary findings, one might speculate that the specificities of prokaryotic and archaeal MetAPs towards Met analogues differ.

In vitro and in vivo, Ala promotes NRE, whereas Gly exerts a slightly inhibitory effect. The reason for this inhibitory effect of Gly is unclear but it might be related to the chirality of this residue. Gly is the smallest of all canonical amino acids and it is the only achiral one. Frottin and co-workers observed that both the geometry and the hydrogen environment around the α-C of the second amino acid are critical for EcMAP1 activity.[18] Further in vivo experiments with more model proteins containing noncanonical Met analogues at their N termini may help to gain more insight into this emerging new research area.

From the above-described in vitro and in vivo studies, it is clear that the established NME rules are valid even for processing of noncanonical Met analogues (or in other words, bulky

amino acids in the second position block their excision as they do with Met). Ala in the second position is the best residue for excision (Figure 2). However, N-terminal residue excision by MetAP can be markedly affected by substitution of the natural Met residue with noncanonical analogues. If Gly occupies position 2 in a polypeptide, the chemical and/or structural nature of the Met analogue in the first position becomes critical; the more similar it is to Met, the more likely it is to be excised. Further in vitro and in vivo experiments with peptides and proteins containing noncanonical Met analogues at the N terminus in a defined sequence context will be necessary to fully complete the essential rules of NRE.

Met analogues that more or less efficiently block NRE, such as trifluoromethionine, Aha, Hpg or Cpa, appear to be the paradigm of a new class of MetAP inhibitors that are an integral part of the protein N terminus. Met analogues carrying bioorthogonal groups such as Aha or Hpg, block MetAP and are therefore retained for chemical modification of the protein N terminus. In addition, since NME plays an important role in cancer development,[35,36] structurally and chemically distinct Met analogues that block NRE may well represent an interesting approach in anti-cancer drug design. Therefore, tuning NRE with Met analogues may find useful applications in the future.

Acknowledgements

We thank Prof. Christian Klein for kindly providing E. coli *MetAP. This work was supported by the BioFuture Program of the Federal Ministry of Education and Research of Germany and Munich Center for Integrated Protein Sciences (CIPSM).*

Keywords: amino acids · DNA · methionine aminopeptidase · methionine excision · noncanonical amino acids · post-translational modifications

[1] M. Kozak, *Gene* **2005**, *361*, 13–37.
[2] E. Schmitt, J. Guillon, T. Meinnel, Y. Mechulam, F. Dardel, S. Blanquet, *Biochimie* **1996**, *78*, 543–554.
[3] T. Meinnel, Y. Mechulam, S. Blanquet, *Biochimie* **1993**, *75*, 1061–1075.
[4] M. Kozak, *Microbiol. Mol. Biol. Rev.* **1983**, *47*, 1–45.
[5] S. Ochoa, *Naturwissenschaften* **1976**, *63*, 347–355.
[6] M. A. Algire, J. R. Lorsch, *Curr. Opin. Chem. Biol.* **2006**, *10*, 480–486.
[7] P. Londei, *FEMS Microbiol. Rev.* **2005**, *29*, 185–200.
[8] B. S. Laursen, H. P. Sorensen, K. K. Mortensen, H. U. Sperling-Petersen, *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 101–123.
[9] L. L. Spremulli, A. Coursey, T. Navratil, S. E. Hunter, M. Kivie, *Prog. Nucleic Acid Res. Mol. Biol.* **2004**, *77*, 211–261.
[10] J. Marín-Navarro, A. Manuell, J. Wu, S. P. Mayfield, *Photosynth. Res.* **2007**, *94*, 359–374.
[11] M. Kozak, *Gene* **1999**, *234*, 187–208.
[12] T. Meinnel, C. Giglione, *Proteomics* **2008**, *8*, 626–649.
[13] F. Sherman, J. W. Stewart, S. Tsunasawa, *Bioessays* **1985**, *3*, 27–31.
[14] J. T. Brosnan, M. E. Brosnan, *J. Nutr.* **2006**, *136*, 1636S-1640S.
[15] C. Giglione, A. Boularot, T. Meinnel, *Cell. Mol. Life Sci.* **2004**, *61*, 1455–1474.
[16] R. A. Bradshaw, W. W. Brickey, K. W. Walker, *Trends Biochem. Sci.* **1998**, *23*, 263–267.
[17] J. A. Vetro, Y.-H. Chang, *J. Cell. Biochem.* **2002**, *85*, 678–688.
[18] F. Frottin, A. Martinez, P. Peynot, S. Mitra, R. C. Holz, C. Giglione, T. Meinnel, *Mol. Cell. Proteomics* **2006**, *5*, 2336–2349.
[19] A. Varshavsky, *Trends Biochem. Sci.* **2005**, *30*, 283–286.
[20] A. Varshavsky, *Genes Cells* **1997**, *2*, 13–28.
[21] A. Varshavsky, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12142–12149.
[22] A. Varshavsky, *Cell* **1992**, *69*, 725–735.
[23] D. K. Gonda, A. Bachmair, I. Wunning, J. W. Tobias, W. S. Lane, A. Varshavsky, *J. Biol. Chem.* **1989**, *264*, 16700–16712.
[24] M. H. Richmond, *Microbiol. Mol. Biol. Rev.* **1962**, *26*, 398–420.
[25] D. G. Barker, C. J. Bruton, *J. Mol. Biol.* **1979**, *133*, 217–231.
[26] J. Trupin, H. Dickerman, M. Nirenberg, H. Weissbach, *Biochem. Biophys. Res. Commun.* **1966**, *24*, 50–55.
[27] J. M. Old, D. S. Jones, *FEBS Lett.* **1976**, *66*, 264–268.
[28] J. L. Brown, *Biochim. Biophys. Acta Nucleic Acids Protein Synth.* **1973**, *294*, 527–529.
[29] K. L. Kiick, J. C. M. van Hest, D. A. Tirrell, *Angew. Chem.* **2000**, *112*, 2232–2236; *Angew. Chem. Int. Ed.* **2000**, *39*, 2148–2152.
[30] J. C. M. van Hest, K. L. Kiick, D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, *122*, 1282–1288.
[31] N. Budisa, O. Pipitone, I. Siwanowicz, M. Rubini, P. P. Pal, T. A. Holak, M. L. Gelmi, *Chem. Biodiversity* **2004**, *1*, 1465–1475.
[32] L. Merkel, Y. Cheburkin, B. Wiltschi, N. Budisa, *ChemBioChem* **2007**, *8*, 2227–2232.
[33] A. Wang, N. W. Nairn, R. S. Johnson, D. A. Tirrell, K. Grabstein, *ChemBioChem* **2008**, *9*, 324–330.
[34] D. I. Andersson, K. Bohman, L. A. Isaksson, C. G. Kurland, *Mol. Gen. Genet.* **1982**, *187*, 467–472.
[35] R. A. Bradshaw, E. Yi, *Essays Biochem.* **2002**, *38*, 65–78.
[36] P. Selvakumar, A. Lakshmikuttyamma, J. R. Dimmock, R. K. Sharma, *Biochim. Biophys. Acta Rev. Cancer* **2006**, *1765*, 148–154.

Received: September 8, 2008
Published online on December 9, 2008

