Comment on binding affinity determines substrate specificity and enables discovery of substrates for N-Myristoyltransferases

Thierry Meinnel

Comment on "Binding Affinity Determines Substrate Specificity and Enables Discovery of Substrates for N-Myristoyltransferases"

Journal:	*ACS Catalysis*
Manuscript ID	cs-2022-01060h.R2
Manuscript Type:	Correspondence/Rebuttal
Date Submitted by the Author:	n/a
Complete List of Authors:	Meinnel, Thierry; Institute for Integrative Cell Biology, Genome Biology

Correspondence

Comment on "Binding Affinity Determines Substrate Specificity and Enables Discovery of Substrates for N-Myristoyltransferases"

Thierry Meinnel[iD,1]

[1] Université Paris Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette cedex, France

Corresponding Author

thierry.meinnel@i2bc.paris-saclay.fr

[iD] ORCID 0001-5642-8637 (T. Meinnel)

Keywords: catalytic mechanism, enzyme kinetics, conformational change, selectivity, myristoylation

ABSTRACT

Establishing the protein posttranslational modification (PTM) landscape at the proteome scale relies on the target specificity of the relevant enzyme catalysts. Su et al. (*ACS Catal*, 2021:14877) proposed that "the k_{cat}/K_m value is not the best parameter to determine the *in vivo* substrate specificity of an enzyme. Instead, the binding affinities of substrates are more important for determining the substrate specificity of enzymes in a physiological setting". The authors extended their conclusions to any "substrate pairs for enzymes that catalyze PTM". This study provides a springboard for the discussion of the relative merits of different approaches used to identify protein modification targets. My point of view is that the specificity constant k_{cat}/K_m remains a highly relevant parameter for defining specificity, while knowledge of the catalytic mechanism - including limiting and synergistic steps - is crucial for reliable data interpretation. Enzyme catalysis and specificity cannot be reduced solely to the formation of an encountered complex that makes the reaction between two partners more likely. I highlight how reactants promote conformational changes that significantly contribute to the final specificity and whose impact can only be assessed using kinetic approaches. There is also a need to integrate data with *in vivo* availability of each competing substrate, and protein data resources must be regularly updated to validate any PTM discovery. These conclusions apply to the substrate specificity of any catalyst.

The protein modification field continues to expand, with several hundred modifications and their landscapes now described [1-3]. In this journal, Su *et al.* recently suggested [4] that approaches that measure target binding affinities to their cognate catalysts are superior for protein modification target discovery than those based on k_{cat}/K_m, the so-called specificity constant or catalytic efficiency [5-6]. In the face of these new developments, is measuring k_{cat}/K_m really obsolete for protein modification target discovery? My point of view is that there is still insufficient proof that cognate catalyst target affinity approaches are superior for this purpose. This is partly due to the catalyst that was chosen to demonstrate the advantages of the approach and because affinity-based technologies still require advances in sensitivity to meet the challenges of characterizing modifications at the proteome-wide scale. Regardless, relevant information on the enzyme's catalytic mechanism and kinetics in individual biological contexts is still required to properly mine binding data and integrate them into the biological context.

Identification of new NMT targets based on binding affinity strategies

In their recent report [4], Su *et al.* proposed that binding affinities represent a valuable strategy for the discovery of protein modification catalyst substrates. In support of their proposal, they used N-myristoyltransferase (NMT) as their case study. NMT is a major catalyst involved in adding a C14 fatty acid to (mainly) proteins featuring an N-terminal glycine originating usually from N-terminal methionine removal [7-8] but also from proteolytic cleavages arising from caspases in the context of apoptosis [9-11]. C14 (myristate) is usually a minor fatty acid compared with C16 (palmitate) or C18 (stearate) [11]. Over the last thirty years, NMT has therefore been referred to as a myristoyltransferase, as it shows high preference for the low-abundance saturated fatty acid C14 (see Fig. 5A in [12]). NMT is a useful case study for protein modifications, as it allows easy *in vitro* screening with short peptide arrays, mainly due to its cotranslational mechanism that uses unfolded polypeptide substrates that are easy to synthesize and assess [13]. The myristoylated proteome is a medium-abundance proteome, representing up to 2% of protein isoforms in eukaryotes [14]. Recent studies have indicated that NMT can also catalyze lysine myristoylation, but only one target, ARF6, has been described to date [15-16].

There are therefore novel NMT substrates to discover, as also suggested by *in silico* analyses [3, 11], so new efforts and novel approaches are required for this discovery effort.

Su *et al.* reported that acetyl-CoA, not traditionally thought of as CoA donor for NMT, behaves as an NMT substrate *in vitro*. The authors compared the k_{cat}/K_m values of myristoyl-CoA and acetyl-CoA with model peptides derived from the small G-protein ARF6. The data indicated that NMT has similar catalytic efficiencies for both the non-cognate (acetyl-CoA) and cognate (myristoyl-CoA) acyl donors (Table 1 in [4]). From competition experiments, the authors reported that myristoyl-CoA is favored over acetyl-CoA and that acetyl-CoA has a binding constant three orders of magnitude greater (10 μM) than myristoyl-CoA (15 nM). They explained these differences in catalytic efficiencies and binding affinities by the fact that the k_{cat}/K_m in NMT's sequential ordered Bi-Bi mechanism [17] measures the k_1 forward kinetic constant of the substrate, here the CoA donor, whereas the binding constant reflects k_{-1}/k_1 (**Figure 1A**). The authors proposed that a much faster backward constant (k_{-1}) of acetyl-CoA compared with myristoyl-CoA explains why the cognate CoA donor is favored in competition experiments *in vitro* and likely *in vivo*. The authors also noticed that the overexpressed protein purified from *Escherichia coli* tightly retains myristoyl-CoA, which is indicative of the acyl-CoA binding selectivity of NMT.

Su *et al.* next noticed that, within NMT1 and NMT2 – the two NMT isoforms in humans - interactome, 28 of 52 interactors harbored Met-Gly starts, 19 of which have already been described as myristoylated. This remarkable enrichment was used to show how the strong binding of Met-Gly proteins to NMT is a primary selection mechanism. The authors further indicated that, of the other nine Met-Gly-starting and one Met-Lys-starting interactors yet to be identified as NMT substrates, three were novel (LRATD1/LRATD2/ERICH5) while two others (CADM4/PHEAT2) reacted with Alk12 (a clickable analog or myristate) but were unlikely to be substrates.

Is binding efficiency a valuable approach for the discovery of new substrates?

A first conclusion of the Su *et al.* report is that binding efficiency is a valuable tool for new substrate discovery. Proof-of-concept came from the unexpected finding that acetyl-CoA is as efficient a substrate (i.e., similar k_{cat}/K_m values for the CoA donor) *in vitro* as myristoyl-CoA, the cognate acyl donor. Nevertheless, in competition experiments performed at a saturating concentration of 50 µM of each CoA donor, myristoyl-CoA was preferred over acetyl-CoA. Su *et al.* next showed that the binding affinity of acetyl-CoA was three orders of magnitude lower than that of myristoyl-CoA (10 µM *vs* 15 nM). The authors concluded that binding affinity appears a more reliable parameter than catalytic efficiency to assess NMT specificity towards the acyl-CoA donor. Nevertheless, the authors did not consider that the free cytosolic myristoyl-CoA concentration lies in the low nanomolar range in the cellular context of eukaryotes [21], whereas the acetyl-CoA concentration is of the order of 10 µM [22]. As a result, in the eukaryotic cytosol - where NMT is active either free or in a ribosome-bound state - both CoA donors are present at concentrations in the order of magnitude of their respective binding affinities. If binding affinity alone drives selectivity, then more or less half of the myristoylation sites should also be acetylated. A few proteins display both modifications *in cellulo* in support of this observation [10]. Nevertheless, the acetylation was considered to arise from competition with an N-acetyltransferase such as NatA rather than from NMT acetylation itself [10]. Though acetylation by NMT is certainly novel, the general model is actually that myristoylation is the only N-terminal modification attached to the protein targets. Therefore, mechanisms other than affinity must contribute to NMT's preference for myristoyl-CoA. With currently available data, two non-exclusive hypotheses are likely.

This first hypothesis involves the acyl-CoA-binding protein ACBD6 as a major actor in NMT's CoA donor selectivity. First, this acyl carrier is known to interact with both mammalian NMT1s [23]. Next, ACBD6 is very selective for C14 chains over longer chains [24] and protects NMT from non-cognate acyl-CoA addition [25]. Consequently, NMT's selectivity *in cellulo* for CoA donors may not be related to the binding affinity but rather to the action of ACBD6. Indeed, ACBD6, through its interaction with NMT, filters out non-cognate CoA donors. Therefore, the channeling effect of the

myristoyl-CoA:ACBD6:NMT complex leads to local increases in myristoyl-CoA concentration, further reinforcing the selection of the C14 acyl chain over any other CoA donor.

A second hypothesis relies on the observation that the *in vitro* binding selectivity of acyltransferases for their acyl-CoA donor is often poor [11, 26-28]. This is in contrast with the tight selectivity exhibited by all known acyltransferases observed *in cellulo*. For instance, CoA derivatives of C12:0-C17:0 fatty acids bind to NMT *in vitro* with similar affinity in the 10 nM range (Table 1 in Ref. [29]). If selectivity is through binding affinity alone - not taking into account that C14 is a minor fatty acid *in vivo* compared with, for example, C16 - then NMT would be defined as specific for any long chain acyl substrate, possibly via the only CoA binding interaction network. This is not the case, as *in vitro* catalytic efficiency assessments have systematically confirmed the *in vivo* observation of a strict C14 preference for NMT acylation. For example, C16 is added 40-fold less efficiently *in vitro* than C14, despite binding similarly [30] showing cooperativity between the fatty acyl-CoA and peptide-binding sites [31]. This also holds true for NMT purified from *E. coli* for crystallographic purposes (NB, most of these constructs are devoid of the complete B'A' loop starting at residue 115 of human NMT, strongly accelerating product release; see below and Ref. [15]). Indeed, the intracellular concentration of acetyl-CoA in *E. coli* is always above 30 µM, much greater than the *in vitro* binding constant [32]. *E. coli* does not display myristoyl-CoA-dependent protein myristoylation [33]. Furthermore, bacteria do not display the required fatty acyl synthase release machinery and couple acyl transfer to the donor directly through acyl carrier protein (ACP) [11]. With respect to myristoyl-CoA availability, *E. coli* cells actually rely on the yeast extract of the growth medium to import myristate, while C14 is a poorly represented fatty acid in yeast [11]. In addition, there is poor solubility of fatty acids in solution in the absence of albumin. Intracellular concentrations of free myristoyl-CoA are therefore expected to be extremely low in bacteria, likely in the sub-nanomolar range, below the binding constant. Besides, myristate needs to be added to the growth medium to allow incorporation into protein targets when expressed heterogeneously with NMT [33]. Finally, NMT tightly retains only myristoyl-CoA when overexpressed

in *E. coli* [4]. This cannot be explained by the relative concentrations of acetyl- and myristoyl-CoA available in *E. coli* and the associated binding constants as determined *in vitro*. Instead - beyond binding itself and before peptide binding - there must be tight CoA donor selection. Therefore, NMT manages to filter out non-cognate CoA donors even in the absence of ACBD6, which is absent in prokaryotes. How can we explain the difference between *in vitro* and *in vivo* observations when considering CoA donor selectivity? Molecular crowding, pH, or salt concentrations may strongly influence enzyme behaviors *in vitro* and *in vivo* [34]. This is especially true when dealing with flexible orientations of mobile loops and acid-base behaviors of crucial chemical groups involved in salt bridges and protonation/deprotonation cycles, such as those seen in the NMT catalytic mechanism (see below and [15]).

Modeling acetyl-CoA and myristoyl-CoA specificity

The data indicate that neither the binding constant of acetyl-CoA nor the kinetic parameters of CoA derivatives are faithful surrogates of CoA selectivity by NMT. Another measurement is clearly required *in vitro* to better reflect and assess this unique property. In a Bi-Bi ordered model such as that of NMT, the k_{cat}/K_m of the first substrate - here the CoA donor - equals the k_1 forward kinetic constant [20]. Still, it is striking that NMT's k_1 constant is controlled by diffusion, i.e., of the order of 10^8 $M^{-1}.s^{-1}$ [18] (**Figure 1A**), whereas the determined k_{cat}/K_m values for the CoA derivatives are only of the order of 10^4 $M^{-1}.s^{-1}$ [4], four orders of magnitude lower, indicating a poor correlation between k_{cat}/K_m and k_1. Instead, the k_{cat}/K_m of the second substrate - here the protein/peptide - is $k_3k_5/(k_5+ k_{-3})$ [20]. As NMT $k_5<<k_{-3}$ (**Figure 1A**), then k_3k_5/k_{-3} or $k_5/K_{D(peptide)}$, so we examined whether this k_{cat}/K_m better accounts for the observations than that of the CoA, as it fully reflects both the rate-limiting step (k_5) and the overall binding of the selected substrates ($K_{D(peptide)}$).

Both kinetic and structural data are useful when interpreting this issue. Upon binding myristoyl-CoA, the disordered Ab loop changes conformation (see [15, 35-36] and SupFig 1 in Ref. [15]) with two alternative conformations and only the open one leading to peptide binding. Next, the Ab loop closure

ensures tight retention of the substrates to NMT. There is therefore synergistic coupling between the conformational change due to myristoyl-CoA binding and peptide binding. As a result, binding of the second ligand (the peptide) only in the presence of the first ligand (the CoA derivative) reveals this step. This is pure catalysis that cannot easily be assessed unless rapid kinetic analyses such as stopped-flow are used, which is a very low-throughput approach not suited to screening. This explains why this step - k_3 is 100-fold lower than k_1 (**Figure 1A**) - is limiting for the formation of the ternary encounter complex, the reacting complex.

The NMT::peptide::acyl-CoA complex with a closed Ab loop features a tense myristoyl-CoA with interbend compaction that promotes a distorted thioester plane of the reactive group of the CoA donor, making it prone to reaction [15]. It is most likely that the ordering of the Ab loop induced by the CoA donor is only properly stabilized in the presence of the cognate myristoyl-CoA derivative. As acetyl-CoA does not make contacts in the myristate-recognition hydrophobic channel, it is also expected that this also favors premature dissociation of the non-cognate CoA substrate featuring increased backward kinetic constant k_{-3}. With this knowledge, the Bi-Bi model can be simplified as shown in **Figure 1B,C**. In this simplified model, CoA is considered as a cofactor that accounts for the aforementioned tight selectivity of NMT *in vivo* for C14 acyl CoA derivatives, and it is easier to understand why NMT's overall selectivity is mostly driven beyond formation of the NMT::myristoyl -CoA complex. The catalytic efficiencies of 2.5 x 10^4 and 1.1 x 10^3 $M^{-1}s^{-1}$ for the peptide, as deduced from the available data (Table 1 in Ref. [4]) with myristoyl-CoA and acetyl CoA, respectively, reflect the kinetic data of the model in **Figure 1B,C**. The $k_{cat}/K_{m(\text{peptide})}$ is thus reduced by 23-fold for acetyl-CoA with respect to myristoyl-CoA. This value is in keeping with the competition experiment performed at saturation concentrations of the peptide and the two CoA donors showing an at least 12-fold myristoylation over acetylation (see Figure 2 in Ref.[4]). Therefore, the conclusion of the report would be more applicable to NMT selection if using another k_{cat}/K_m parameter that properly fits the

overall selection mechanism of NMT. This is another limitation of using the specificity constant k_{cat}/K_m when dealing with a two-substrate reaction such as seen in protein modification.

Added value of the binding affinity to identify new NMT targets

While the Su *et al.* study was obtained with CoA donor preselection, the authors next applied their hypothesis of the binding affinity representing a driver of substrate selectivity to the second site of selectivity in NMT, the peptide-binding site. This site lies in a distinct domain of the NMT tertiary structure, and the peptide only binds once the CoA donor has already bound NMT according to the sequential ordered Bi-Bi model of NMT [17, 19] (**Figure 1A**). The rate-limiting step of the reaction (k_5=0.1 s^{-1}) is much slower than the chemical transformation (k_4=16 s^{-1}) likely in the course of product dissociation [18], i.e., myristoyl-peptide release in the case of human NMT [15, 19]. The complex of NMT with the myristoylated polypeptide product devoid of CoA is the last complex depicted with HsNMT1, and its dissociation is most likely to be the limiting step of NMT catalysis caused by a second conformational change in an extreme N-terminal NMT loop (B'A') [15]. The binding constant (k_{-3}/k_3) of the polypeptide to the NMT-MyrCoA complex is in the micromolar range [30], two orders of magnitude higher than the myristoyl-CoA binding constant. The K_d of the peptide therefore lies in the range of the K_m value, unlike the situation seen with the cognate CoA donor. This also indicates that the K_m of the peptide is a direct readout of the binding constant. This is in keeping with the limiting step lying beyond formation of the ternary complex.

In this context, it is quite challenging to transfer the observation made with the CoA donor to the polypeptide acceptor. The enrichment of myristoylated proteins appears, therefore, more likely to arise due to the tight binding of the reaction product - if diffusion-limited, as in the case of the binding step [18], then in the nanomolar range - to the enzyme. Such tight binding results from the entropic effect caused by the synergy between the affinity of both the peptide and the myristate moieties into a single molecule product. I conclude that the study does not demonstrate that the binding affinity underpins

the discovery of NMT substrates. Only the strong binding affinity of the product to the enzyme explains the data, and this is specific to NMT due to the unique behavior of the B'A' N-terminal loop.

Finally, the authors mined the NMT interactome of putative NMT substrates and claimed to identify new substrates. The authors missed that all these "new" substrates had already been identified in previous studies [10, 37-38]. I acknowledge that none of the missing three entries were in the 121 experimentally validated myristoylated human proteins available in UniProt (**Table S1**). This further strengthens the urgent need to improve data annotation in protein data resources. Finally, the most complete dataset (834 IDs from humans and *A. thaliana*), which includes both k_{cat}/K_m and *in vivo* data, is not yet in UniProt [10]. To contribute to this collective effort, I have compiled all the available data and evidence on the human NMT-myristoylated proteome (433 protein entries in **Table S1**), including from post-translational myristoylation (43 entries). **Table 1** summarizes the main highlights and shows that the human myristoylated proteome is now well advanced (2.1 % of the human proteome), and 85% of the myristoylated proteome is now supported by strong experimental evidence.

The NMT interactors introduce another dimension and reinforce the conclusions based only on k_{cat}/K_m for three entries (ARL11, FBX17, MEAK7; **Table S1**). In its own, however, this interaction network is not directly exploitable (**Table 1**) and requires a second validation round, similar to that performed by Su *et al*. [4], perhaps also by exploiting available predictive software. For instance, in Ref. [10], six of the seven non-myristoylated MG-starting peptides of the NMT interaction set did not display a significant probability score so were disregarded. The last one (RFTN2) had a significant positive score and, while the peptide was assessed, it had such a low k_{cat}/K_m value that it was not retained in the myristoylated proteome (SupDataset 1 in Ref.[10]). It would have been interesting to establish whether the full-length RFNT2 protein is also negative by *in vivo* metabolic labeling and therefore compare the relative sensitivity of the binding and specificity constant approaches.

Is binding affinity more important than $k_{ca}t/K_m$ values when determining substrate specificity *in vivo*?

The Su *et al.* study stated that "the acylCoA specificity of human NMT demonstrates that the k_{cat}/K_m value is not the best parameter for determining the *in vivo* substrate specificity of an enzyme". The myristoylated proteome is now progressively emerging to reveal a near-complete range of substrates and is likely one of the few modification sub-proteomes with such completeness. The most efficient strategies that have achieved this aim involved click chemistry approaches with reactive alkyl myristate precursors [9, 37, 39], proteomic strategies with subcellular enrichment of the membrane fractions [40], or a combination of approaches involving structural, kinetic, proteomic, and bioinformatics analyses [11]. In these studies [10], the k_{cat}/K_m values of the myristoylation reaction measured *in vitro* with hundreds of model peptides derived from each proteome have regularly been used to define the cognate subproteome. In such experiments, the measured parameter is the k_{cat}/K_m of the peptide, and the myristoyl-CoA donor concentration is fixed and almost saturating. The capacities of different peptides to be myristoylated have been compared, and a large range of catalytic efficiency values have been obtained. Early approaches were performed with arrays of model substrates and few peptides derived from actual proteins (Table 2 in Ref.[13]). Recent analyses use only peptides derived from proteomes. With the human dataset, for instance, positive peptides and negatives were obtained from 568 peptides preselected by a first filter step using a dedicated bioinformatics tool based on machine learning (SupTable 1 in Ref.[10]). Based on the analysis of the proteins extracted from the cellular proteome and validation by mass spectrometry or other approaches including radiolabeling, a 100% fit was observed, validating the approach. These data show that k_{cat}/K_m measurements *in vitro* were and still are a relevant strategy for discovering or validating NMT substrates with significant sampling and discovery rate (**Table 1**).

Conclusions

In their study, Su *et al.* report two novel observations significant to the field. First, they noticed that NMT may also use acetyl-CoA as an acyl donor to make N-acetylated proteins *in vitro*. This feature is one of the missing links when considering the ancient origin of NMT in early primitive eukaryotes [14]. NMTs are part of the so-called general control non-repressible 5 (GCN5)-related N-acetyltransferase (GNAT) family, a family of transferases that usually use acetyl-CoA as the donor. NMTs are formed from the fusion of two GNAT domains, with the N-terminal domain showing acyl-CoA binding affinity and transfer. This fusion likely occurred early in eukaryogenesis from a prokaryotic acetyl-CoA transferase in the GNAT family. Nevertheless, the authors revealed that the higher binding affinity of myristoyl-CoA is strongly preferred *in vitro* and that this likely explains why it is incorporated *in cellulo* instead of acetyl-CoA. During primitive eukaryogenesis, NMT has therefore acquired features to discriminate between CoA donors including selectivity synergy between the two substrates.

The second novel observation is that the human NMT interactome is highly enriched with NMT substrates. Twenty-two of 52 interactors represent 42% of interactors with either or both human NMTs from the BioPlex human interactome based on affinity purification of C-terminally tagged baits [41-42], which is amazing enrichment, as the myristoylated proteome represents ~2.1% of the human proteome [10]. If one cross-references BioGRID and STRING data of interactors with either human NMT1 or NMT2, the number of interactors is significantly higher (161), with 31 (19%) known NMT substrates (**Table 1**). If one focuses on both interactors of NMT1 and NMT2, 19 out of 34 (56%) are NMT substrates (**Table 1**).

While Su *et al.* report two very interesting observations related to NMT, these are unrelated to each other and functionally rely on molecular mechanisms distinct from binding affinity. It should be emphasized that two limiting conformational rearrangements of two N-terminal loops (Ab and B'A') likely contribute to the unique behavior of NMT upon binding or release of reactants and products, and previous reports highlight the role of substrate-induced conformational changes in protein modification catalysts [43]. First, with respect to CoA derivative binding, complex selectivity through synergy with

the peptide occurs at the level of the formation of the ternary and not the binary complex, and the acyl-CoA binder ACBD6 contributes to counter-select non-cognate CoA derivatives *in vivo* (**Figure 1**). Second, with respect to peptide binding, product dissociation is the rate-limiting step, as it displays strong, nanomolar affinity; in both cases, this step likely explains the enrichment described. The affinity of the ternary complex (NMT:MyrCoA:Peptide) is in the 10 μM range, at least three orders of magnitude greater than that of the product. With its unusually strong affinity for the reaction product, NMT is a special case.

ACKNOWLEDGMENT

The work of the team investigating myristoylation has been supported over recent years by the French National Research Agency (ANR) PalMyProt (ANR-2010-BLAN-1611-01), DynaMYT (ANR-20-CE44-0013), and Fondation ARC (ARCPJA32020060002137) grants to TM.

SUPPORTING INFORMATION

Supporting Information is available for this paper. The Supporting Information is available free of charge on the ACS Publications website.

Table S1, the human myristoylated proteome, includes full annotation and 93 references dealing with the complete protein set (433 entries) of the human myristoylated proteome.

COMPETING INTERESTS

The author declares no competing interests.

ABBREVIATIONS

ACBD, acyl-CoA binding domain containing

Acetyl-CoA, AcCoA

GNAT, (GCN5)-related N-acetyltransferase

NMT, N-myristoyltransferase

Myristoyl-coA, MyrCoA

SCX, Strong Cation eXchange

REFERENCES

1. Khoury, G. A.; Baliban, R. C.; Floudas, C. A., Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep* **2011,** *1*.

2. Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; Ge, Y.; Gunawardena, J.; Hendrickson, R. C.; Hergenrother, P. J.; Huber, C. G.; Ivanov, A. R.; Jensen, O. N.; Jewett, M. C.; Kelleher, N. L.; Kiessling, L. L.; Krogan, N. J.; Larsen, M. R.; Loo, J. A.; Ogorzalek Loo, R. R.; Lundberg, E.; MacCoss, M. J.; Mallick, P.; Mootha, V. K.; Mrksich, M.; Muir, T. W.; Patrie, S. M.; Pesavento, J. J.; Pitteri, S. J.; Rodriguez, H.; Saghatelian, A.; Sandoval, W.; Schlüter, H.; Sechi, S.; Slavoff, S. A.; Smith, L. M.; Snyder, M. P.; Thomas, P. M.; Uhlén, M.; Van Eyk, J. E.; Vidal, M.; Walt, D. R.; White, F. M.; Williams, E. R.; Wohlschlager, T.; Wysocki, V. H.; Yates, N. A.; Young, N. L.; Zhang, B., How many human proteoforms are there? *Nat Chem Biol* **2018,** *14*, 206.

3. Devabhaktuni, A.; Lin, S.; Zhang, L.; Swaminathan, K.; Gonzalez, C. G.; Olsson, N.; Pearlman, S. M.; Rawson, K.; Elias, J. E., TagGraph reveals vast protein modification landscapes from large tandem mass spectrometry datasets. *Nat Biotechnol* **2019,** *37* (4), 469-479.

4. Su, D.; Kosciuk, T.; Yang, M.; Price, I. R.; Lin, H., Binding affinity determines substrate specificity and enables discovery of substrates for N-myristoyltransferases. *ACS Catal* **2021**, 14877-14883.

5. Koshland, D. E., The application and usefulness of the ratio k(cat)/K(M). *Bioorg Chem* **2002,** *30* (3), 211-3.

6. Eisenthal, R.; Danson, M. J.; Hough, D. W., Catalytic efficiency and kcat/KM: a useful comparator? *Trends Biotechnol* **2007,** *25* (6), 247-9.

7. Giglione, C.; Boularot, A.; Meinnel, T., Protein N-terminal methionine excision. *Cell Mol Life Sci* **2004,** *61* (12), 1455-1474.

8. Giglione, C.; Fieulaine, S.; Meinnel, T., N-terminal protein modifications: bringing back into play the ribosome. *Biochimie* **2015,** *114*, 134-146.

9. Thinon, E.; Serwa, R. A.; Broncel, M.; Brannigan, J. A.; Brassat, U.; Wright, M. H.; Heal, W. P.; Wilkinson, A. J.; Mann, D. J.; Tate, E. W., Global profiling of co- and post-translationally N-myristoylated proteomes in human cells. *Nat Commun* **2014,** *5*, 4919.

10. Castrec, B.; Dian, C.; Ciccone, S.; Ebert, C. L.; Bienvenut, W. V.; Le Caer, J.-P.; Steyaert, J.-M.; Giglione, C.; Meinnel, T., Structural and genomic decoding of human and plant myristoylomes reveals a definitive recognition pattern. *Nat Chem Biol* **2018,** *14* (7), 671-679.

11. Giglione, C.; Meinnel, T., Mapping the myristoylome through a complete understanding of protein myristoylation biochemistry. *Prog Lipid Res* **2022,** *85*, 101139.

12. Kishore, N. S.; Wood, D. C.; Mehta, P. P.; Wade, A. C.; Lu, T.; Gokel, G. W.; Gordon, J. I., Comparison of the acyl chain specificities of human myristoyl-CoA synthetase and human myristoyl-CoA:protein N-myristoyltransferase. *J Biol Chem* **1993,** *268* (7), 4889-902.

13. Towler, D. A.; Gordon, J. I.; Adams, S. P.; Glaser, L., The biology and enzymology of eukaryotic protein acylation. *Annu Rev Biochem* **1988,** *57*, 69-99.

14. Meinnel, T.; Dian, C.; Giglione, C., Myristoylation, an ancient protein modification mirroring eukaryogenesis and evolution. *Trends Biochem Sci* **2020,** *45* (7), 619-632.

15. Dian, C.; Pérez-Dorado, I.; Rivière, F.; Asensio, T.; Legrand, P.; Ritzefeld, M.; Shen, M.; Cota, E.; Meinnel, T.; Tate, E. W.; Giglione, C., High-resolution snapshots of human N-myristoyltransferase in action illuminate a mechanism promoting N-terminal Lys and Gly myristoylation. *Nat Commun* **2020,** *11* (1), 1132.

16. Kosciuk, T.; Price, I. R.; Zhang, X.; Zhu, C.; Johnson, K. N.; Zhang, S.; Halaby, S. L.; Komaniecki, G. P.; Yang, M.; DeHart, C. J.; Thomas, P. M.; Kelleher, N. L.; Fromme, J. C.; Lin, H., NMT1 and NMT2 are lysine myristoyltransferases regulating the ARF6 GTPase cycle. *Nat Commun* **2020,** *11* (1), 1067.

17. Rudnick, D. A.; McWherter, C. A.; Rocque, W. J.; Lennon, P. J.; Getman, D. P.; Gordon, J. I., Kinetic and structural evidence for a sequential ordered Bi Bi mechanism of catalysis by Saccharomyces cerevisiae myristoyl-CoA:protein N-myristoyltransferase. *J Biol Chem* **1991,** *266* (15), 9732-9.

18. Farazi, T. A.; Manchester, J. K.; Gordon, J. I., Transient-state kinetic analysis of Saccharomyces cerevisiae myristoylCoA:protein N-myristoyltransferase reveals that a step after chemical transformation is rate limiting. *Biochemistry* **2000,** *39* (51), 15807-16.

19. Rocque, W. J.; McWherter, C. A.; Wood, D. C.; Gordon, J. I., A comparative analysis of the kinetic mechanism and peptide substrate specificity of human and Saccharomyces cerevisiae myristoyl-CoA:protein N-myristoyltransferase. *J Biol Chem* **1993,** *268* (14), 9964-71.

20. Segel, I. H., *Enzyme kinetics: Behavior and analysis of rapid equilibrium and steady-state enzyme systems*. John Wiley & Sons, Inc.: New York, 1993.

21. Knudsen, J.; Neergaard, T. B.; Gaigg, B.; Jensen, M. V.; Hansen, J. K., Role of acyl-CoA binding protein in acyl-CoA metabolism and acyl-CoA-mediated cell signaling. *J Nutr* **2000,** *130* (2S Suppl), 294s-298s.

22. Lee, J. V.; Carrer, A.; Shah, S.; Snyder, N. W.; Wei, S.; Venneti, S.; Worth, A. J.; Yuan, Z. F.; Lim, H. W.; Liu, S.; Jackson, E.; Aiello, N. M.; Haas, N. B.; Rebbeck, T. R.; Judkins, A.; Won, K. J.; Chodosh, L. A.; Garcia, B. A.; Stanger, B. Z.; Feldman, M. D.; Blair, I. A.; Wellen, K. E., Akt-dependent metabolic reprogramming regulates tumor cell histone acetylation. *Cell Metab* **2014,** *20* (2), 306-319.

23. Soupene, E.; Schatz, U. A.; Rudnik-Schöneborn, S.; Kuypers, F. A., Requirement of the acyl-CoA carrier ACBD6 in myristoylation of proteins: Activation by ligand binding and protein interaction. *PLoS One* **2020,** *15* (2), e0229718.

24. Soupene, E.; Kuypers, F. A., ACBD6 protein controls acyl chain availability and specificity of the N-myristoylation modification of proteins. *J Lipid Res* **2019,** *60* (3), 624-635.

25. Soupene, E.; Kao, J.; Cheng, D. H.; Wang, D.; Greninger, A. L.; Knudsen, G. M.; DeRisi, J. L.; Kuypers, F. A., Association of NMT2 with the acyl-CoA carrier ACBD6 protects the N-myristoyltransferase reaction from palmitoyl-CoA. *J Lipid Res* **2016,** *57* (2), 288-98.

26. Jennings, B. C.; Linder, M. E., DHHC protein S-acyltransferases use similar ping-pong kinetic mechanisms but display different acyl-CoA specificities. *J Biol Chem* **2012,** *287* (10), 7236-45.

27. Greaves, J.; Munro, K. R.; Davidson, S. C.; Riviere, M.; Wojno, J.; Smith, T. K.; Tomkinson, N. C.; Chamberlain, L. H., Molecular basis of fatty acid selectivity in the zDHHC family of S-acyltransferases revealed by click chemistry. *Proc Natl Acad Sci U S A* **2017,** *114* (8), E1365-e1374.

28. Stix, R.; Lee, C. J.; Faraldo-Gómez, J. D.; Banerjee, A., Structure and mechanism of DHHC protein acyltransferases. *J Mol Biol* **2020,** *432* (18), 4983-4998.

29. Bhatnagar, R. S.; Schall, O. F.; Jackson-Machelski, E.; Sikorski, J. A.; Devadas, B.; Gokel, G. W.; Gordon, J. I., Titration calorimetric analysis of AcylCoA recognition by myristoylCoA:protein N-myristoyltransferase. *Biochemistry* **1997,** *36* (22), 6700-8.

30. Bhatnagar, R. S.; Jackson-Machelski, E.; McWherter, C. A.; Gordon, J. I., Isothermal titration calorimetric studies of Saccharomyces cerevisiae myristoyl-CoA:protein N-myristoyltransferase. Determinants of binding energy and catalytic discrimination among acyl-CoA and peptide ligands. *J Biol Chem* **1994,** *269* (15), 11045-53.

31. Rudnick, D. A.; McWherter, C. A.; Adams, S. P.; Ropson, I. J.; Duronio, R. J.; Gordon, J. I., Structural and functional studies of Saccharomyces cerevisiae myristoyl-CoA:protein N-myristoyltransferase produced in Escherichia coli. Evidence for an acyl-enzyme intermediate. *J Biol Chem* **1990,** *265* (22), 13370-8.

32. Takamura, Y.; Nomura, G., Changes in the intracellular concentration of acetyl-CoA and malonyl-CoA in relation to the carbon and energy metabolism of Escherichia coli K12. *J Gen Microbiol* **1988,** *134* (8), 2249-53.

33. Ha, V. L.; Thomas, G. M.; Stauffer, S.; Randazzo, P. A., Preparation of myristoylated Arf1 and Arf6. *Methods Enzymol* **2005,** *404*, 164-74.

34. Schnell, S.; Turner, T. E., Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog Biophys Mol Biol* **2004,** *85* (2-3), 235-60.

35. Weston, S. A.; Camble, R.; Colls, J.; Rosenbrock, G.; Taylor, I.; Egerton, M.; Tucker, A. D.; Tunnicliffe, A.; Mistry, A.; Mancia, F.; de la Fortelle, E.; Irwin, J.; Bricogne, G.; Pauptit, R. A., Crystal structure of the anti-fungal target N-myristoyl transferase. *Nat Struct Biol* **1998,** *5* (3), 213-21.

36. Wu, J.; Tao, Y.; Zhang, M.; Howard, M. H.; Gutteridge, S.; Ding, J., Crystal structures of Saccharomyces cerevisiae N-myristoyltransferase with bound myristoyl-CoA and inhibitors reveal the functional roles of the N-terminal region. *J Biol Chem* **2007,** *282* (30), 22185-94.

37. Broncel, M.; Serwa, R. A.; Ciepla, P.; Krause, E.; Dallman, M. J.; Magee, A. I.; Tate, E. W., Multifunctional reagents for quantitative proteome-wide analysis of protein modification in human cells and dynamic profiling of protein lipidation during vertebrate development. *Angew Chem Int Ed Engl* **2015,** *54* (20), 5948-51.

38. Goya Grocin, A.; Serwa, R. A.; Sanfrutos, J. M.; Ritzefeld, M.; Tate, E. W., Whole proteome profiling of N-myristoyltransferase activity and inhibition using sortase A. *Mol Cell Proteomics* **2019,** *18* (1), 115-126.

39. Kallemeijn, W. W.; Lanyon-Hogg, T.; Panyain, N.; Goya Grocin, A.; Ciepla, P.; Morales-Sanfrutos, J.; Tate, E. W., Proteome-wide analysis of protein lipidation using chemical probes: in-gel fluorescence visualization, identification and quantification of N-myristoylation, N- and S-acylation, O-cholesterylation, S-farnesylation and S-geranylgeranylation. *Nat Protoc* **2021,** *16* (11), 5083-5122.

40. Majeran, W.; Le Caer, J. P.; Ponnala, L.; Meinnel, T.; Giglione, C., Targeted profiling of A. thaliana sub-proteomes illuminates new co- and post-translationally N-terminal myristoylated proteins. *Plant Cell* **2018,** *30* (3), 543-562.

41. Huttlin, E. L.; Bruckner, R. J.; Paulo, J. A.; Cannon, J. R.; Ting, L.; Baltier, K.; Colby, G.; Gebreab, F.; Gygi, M. P.; Parzen, H.; Szpyt, J.; Tam, S.; Zarraga, G.; Pontano-Vaites, L.; Swarup, S.; White, A. E.; Schweppe, D. K.; Rad, R.; Erickson, B. K.; Obar, R. A.; Guruharsha, K. G.; Li, K.; Artavanis-Tsakonas, S.; Gygi, S. P.; Harper, J. W., Architecture of the human interactome defines protein communities and disease networks. *Nature* **2017,** *545* (7655), 505-509.

42. Schweppe, D. K.; Huttlin, E. L.; Harper, J. W.; Gygi, S. P., BioPlex display: an interactive suite for large-scale AP-MS protein-protein interaction data. *J Proteome Res* **2018,** *17* (1), 722-726.

43. Fieulaine, S.; Boularot, A.; Artaud, I.; Desmadril, M.; Dardel, F.; Meinnel, T.; Giglione, C., Trapping conformational states along ligand-binding dynamics of peptide deformylase: the impact of induced fit on enzyme catalysis. *PLoS Biol* **2011,** *9* (5), e1001066.

44. Takamitsu, E.; Fukunaga, K.; Iio, Y.; Moriya, K.; Utsumi, T., Cell-free identification of novel N-myristoylated proteins from complementary DNA resources using bioorthogonal myristic acid analogues. *Anal Biochem* **2014,** *464*, 83-93.

45. Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B. J.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; Dolma, S.; Coulombe-Huntington, J.; Chatr-Aryamontri, A.; Dolinski, K.; Tyers, M., The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci* **2021,** *30* (1), 187-200.

46. Huttlin, E. L.; Bruckner, R. J.; Navarrete-Perea, J.; Cannon, J. R.; Baltier, K.; Gebreab, F.; Gygi, M. P.; Thornock, A.; Zarraga, G.; Tam, S.; Szpyt, J.; Gassaway, B. M.; Panov, A.; Parzen, H.; Fu, S.; Golbazi, A.; Maenpaa, E.; Stricker, K.; Guha Thakurta, S.; Zhang, T.; Rad, R.; Pan, J.; Nusinow, D. P.; Paulo, J. A.; Schweppe, D. K.; Vaites, L. P.; Harper, J. W.; Gygi, S. P., Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell* **2021,** *184* (11), 3022-3040.e28.

47. Huttlin, E. L.; Ting, L.; Bruckner, R. J.; Gebreab, F.; Gygi, M. P.; Szpyt, J.; Tam, S.; Zarraga, G.; Colby, G.; Baltier, K.; Dong, R.; Guarani, V.; Vaites, L. P.; Ordureau, A.; Rad, R.; Erickson, B. K.; Wühr, M.; Chick, J.; Zhai, B.; Kolippakkam, D.; Mintseris, J.; Obar, R. A.; Harris, T.; Artavanis-Tsakonas, S.; Sowa, M. E.; De Camilli, P.; Paulo, J. A.; Harper, J. W.; Gygi, S. P., The BioPlex network: a systematic exploration of the human interactome. *Cell* **2015,** *162* (2), 425-440.

TABLE

Table 1: Impact of various approaches for the validation of the human myristoylated proteome and its overall coverage.

Data	Uni Prot [a]	Lite rature survey [b]	k_{cat}/K_m [c]	Cross validation [d]	NMT1/2 interactors [e]	SVM predicted [f]	Total [g]	
							Entries (N-ter)	Proteo forms
Number	121	236	568	170	161	64	433 (2.1%)	644
Validated	119		308		3 (31)	4		

[a] experimentally verified N-myristoylated proteins currently available in UniProtKB (2022-15-01); each entry was manually validated; four entries (DDX46, KCNJ2, RL15, RS8) wrongly extracted from Refs. [9, 44] were discarded, while two entries (HCK, ZEP1) each with two isoforms with distinct N-termini were considered as separate.

[b] manual survey and collection over the years.

[c] data from Ref. [10]. k_{cat}/K_m refers to a large peptide assay formed from human and *Arabidopsis* sequences challenged by the corresponding NMTs. 568 corresponds to the number of human sequences out of the 2048 on the array. The 568 sequences were selected out of 2075 proteins starting with Met-Gly, the only pattern potentially leading to N-myristoylation (see SupTable 1 in Ref. [10]). 308 corresponds to the number of sequences with associated k_{cat}/K_m values high enough to support myristoylation in cellulo.

[d] data with both literature and k_{cat}/K_m cross-validation.

[e] data are compiled from human NMT1 and NMT2 interactomes retrieved at the BioGRID (https://thebiogrid.org/; Ref. [45]), which integrates BioPlex 1.0/2.0/3.0 datasets [41, 46-47] data and other datasets, and STRING (https://string-db.org/) resources; the validated data correspond to the validation made in Ref. [4]; the value in brackets corresponds to the validation arising from the literature and k_{cat}/K_m reports (see data in **Table S1**).

[f] data from Ref. [10]. SVM refers to a support vector machine classifier, which was used to select the most likely sequences. The machine learning tool identified another 64 putative targets in humans which were not probed in the k_{cat}/K_m approach (see footnote [c]). Depending on whether they were high, medium, or low confidence, these predictions are indicated as HC, MC, or LM, respectively, in **Table S1**. "4" refers to the number of pieces of experimental evidence that have since demonstrated the relevance of the *in silico* prediction.

[g] myristoylated proteome validated through various approaches (including all SVM predicted); data from the various approaches are compiled. The proteoforms of a given entry refer to proteins with the

same N-terminus (see corresponding column in **Table S1**). The number corresponds to the sum of all of them. The neXtprot release (v2.41.0; 2021-11-19) reports 20,380 protein entries and 42,365 isoforms.

FIGURE LEGENDS

Figure 1 The kinetic mechanism of NMT.

Data are from Ref.[18] with yeast NMT. Human NMT also displays an ordered Bi-Bi mechanism [19]. CoA is released first, followed by the myristoylpeptide in human NMT [15, 19].

Figure 2 Simplified models and kinetic parameters of NMT with MyrCoA and Ac-CoA.

Panel A: Simplified model and relevant kinetic parameters for k_{cat}/K_m[protein] measurement (see Chapter IX in Ref. [20]).

Panel B: Proposed model and kinetic constants for protein acetylation *in vitro* by NMT. Data are from Ref. [4] for $K_m=k_{-3}/k_3$ and k_5).

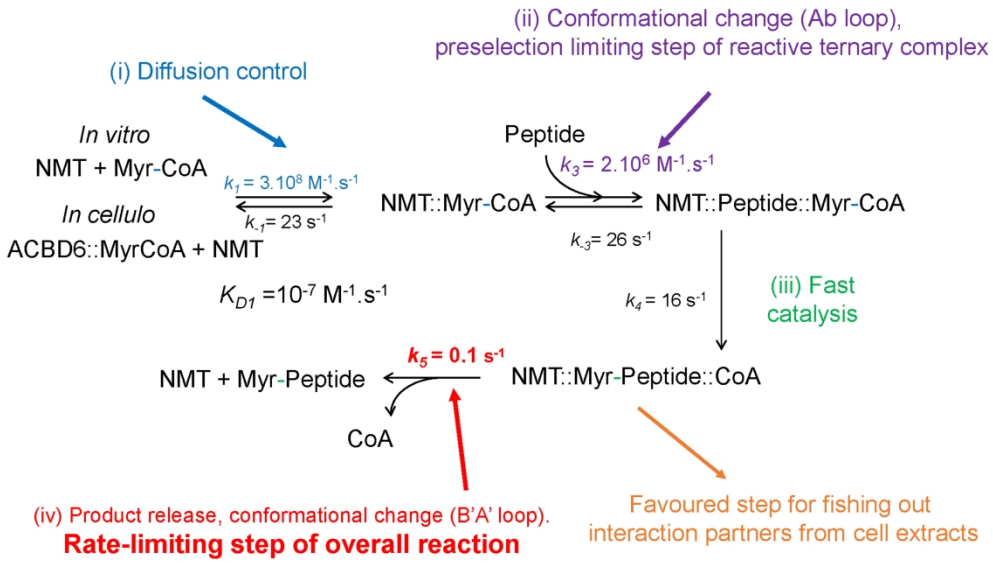


Figure 1

180x107mm (600 x 600 DPI)

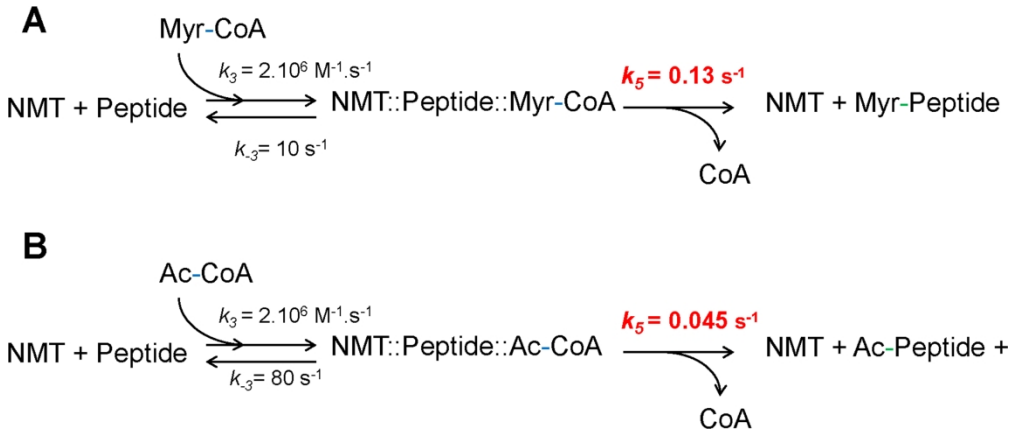
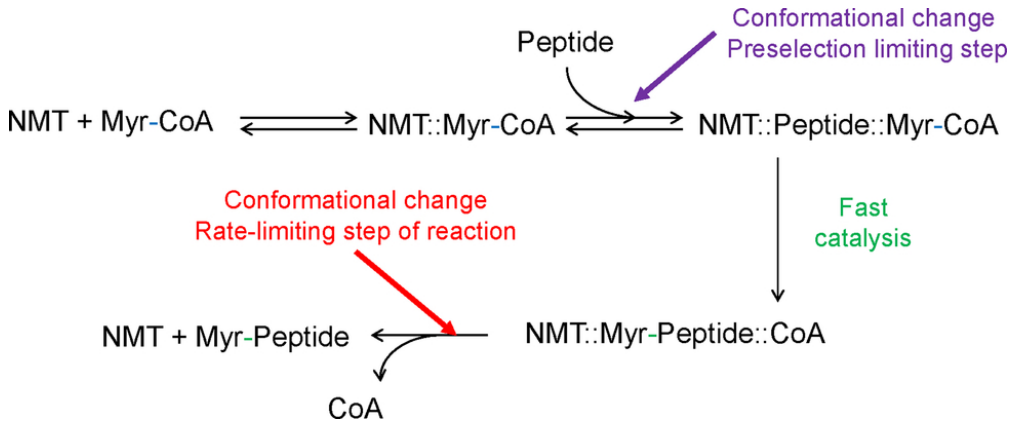


Figure 2

127x53mm (600 x 600 DPI)



TOC

82x33mm (300 x 300 DPI)