Université de Montréal

New insights into the substrate specificities of microbial transglutaminase: a biocatalytic perspective

par
Maria Gundersen

Département de Biochimie, Faculte de Médecine

Mémoire présenté à la Faculté
James G. Omichinski, Joelle N. Pelletier et James D. Wuest
en vue de l'obtention du grade de maîtrise en Biochimie option libre

Décembre, 2013

Identification of Jury:

Professor James G. Omichinski (President)

Professor Joelle Pelletier (2nd member)

Professor James D. Wuest (3rd Member)

Résumé

La transglutaminase microbienne (Microbial transglutaminase : MTG) est fortement exploitée dans l'industrie textile et alimentaire afin de modifier l'apparence et la texture de divers produits. Elle catalyse la formation de liaisons iso-peptidiques entre des protéines par l'entremise d'une réaction de transfert d'acyle entre le groupement γ-carboxamide d'une glutamine provenant d'un substrat donneur d'acyle, et le groupement ε-amino d'une lysine provenant d'un substrat accepteur d'acyle. La MTG est tolérante à un large éventail de conditions réactionnelles, ce qui rend propice le développement de cette enzyme en tant que biocatalyseur. Ayant pour but le développement de la MTG en tant qu'alternative plus soutenable à la synthèse d'amides, nous avons étudié la réactivité d'une gamme de substrats donneurs et accepteurs non-naturels.

Des composés chimiquement diversifiés, de faible masse moléculaire, ont été testés en tant que substrats accepteurs alternatifs. Il fut démontré que la MTG accepte une large gamme de composés à cet effet. Nous avons démontré, pour la première fois, que des acides aminés non-ramifiés et courts, tels la glycine, peuvent servir de substrat accepteur. Les α-acides aminés estérifiés Thr, Ser, Cys et Trp, mais pas Ile, sont également réactifs. En étendant la recherche à des composés non-naturels, il fut observé qu'un cycle aromatique est bénéfique pour la réactivité, bien que les substituants réduisent l'activité. Fait notable, des amines de faible masse moléculaire, portant les groupements de forte densité électronique azidure ou alcyne, sont très réactives. La MTG catalyse donc efficacement la modification de peptides qui pourront ensuite être modifiés ou marqués par la chimie 'click'. Ainsi, la MTG accepte une variété de substrats accepteurs naturels et non-naturels, élargissant la portée de modification des peptides contenant la glutamine.

Afin de sonder le potentiel biocatalytique de la MTG par rapport aux substrats donneurs, des analogues plus petits du peptide modèle Z-Gln-Gly furent testés; aucun n'a réagi. Nous avons toutefois démontré, pour la première fois, la faible réactivité d'esters en tant que substrats donneurs de la MTG. L'éventuelle amélioration de cette réactivité permettrait de faire de la MTG un biocatalyseur plus général pour la synthèse d'amides.

Mots clés:

Lien amide, biocatalyse, biotransformation, transglutaminase, arrimage moléculaire, criblage de substrats, ingénierie de substrats.

Abstract:

Microbial transglutaminase (MTG) is used extensively in the food and textile industry to alter the appearance and texture of products. MTG catalyses the formation of isopeptide linkages between proteins by an acyl transfer reaction between the γ-carboxamide group of a glutamine 'acyl-donor' substrate, and the ε-amino group of a lysine 'acyl-acceptor' substrate. MTG is tolerant to a broad range of reaction conditions and is therefore suitable for further development as a biocatalyst. Toward developing MTG as a "green" alternative for amide synthesis, we have investigated a range of non-native donor and acceptor substrates to probe the scope of MTG reactivity.

Small, chemically varied compounds were tested as alternative acyl-acceptor substrates. We observed a broad acceptor specificity. We show, for the first time, that very short-chain alkyl-based amino acids such as glycine can serve as acceptor substrates. The esterified α-amino acids Thr, Ser, Cys and Trp – but not Ile – also show reactivity. Extending the search to non-natural compounds, an aromatic ring was observed to be beneficial for reactivity, although ring substituents reduced reactivity. Overall, bonding of the amine to a less hindered carbon increases reactivity. Importantly, very small amines carrying either the electron-rich azide or the alkyne groups required for click chemistry were highly reactive as acceptor substrates, providing a ready route to minimally modified, 'clickable' peptides. These results demonstrate that MTG is tolerant to a variety of chemically varied natural and non-natural acceptor substrates, which broadens the scope for modification of glutamine-containing peptides.

To further probe the biocatalytic potential of MTG in terms of the donor substrate, smaller analogues of the model substrate Z-Gln-Gly were tested. We did not find product formation with substrates smaller than the model substrate. We observed, for the first time, trace esterase

activity with MTG. Future improvement of this activity would render MTG a more attractive, general biocatalyst for amide bond formation.

Key words:

Amide bond, biocatalysis, biotransformations, microbial transglutaminase, docking, substrate screening, substrate engineering.

Table of contents

List of Tables:

List of Figures:

List of acronyms, abbreviations list:

Transglutaminase (TGase)

Microbial transglutaminase (MTG)

Polyethylene glycol (PEG)

Active pharmaceutical ingredients (APIs)

Carboxybenzyl-glutaminyl-glycine (Z-Gln-Gly)

Deionized water (MilliQ)

This thesis is dedicated to André.

-Takk for dine smil, all din hjelp, mange gode stunder, og millioner av sekunder.

Acknowledgements:

I would first and foremost like to thank my supervisor Joelle N. Pelletier for generously sharing her time for every aspect of my masters. The discussions throughout my masters have been invaluable, both for the development of the project and my personal development as a scientist. Furthermore she has offered me many exceptional academic opportunities which I will bring with me for my future career. Her supervision exceeded my highest expectations. Thank you Joelle.

Klavdja Annabel Fignolé was a great asset for my masters. She aided in part of the experimental work as part of her exchange for her masters. Her exemplary work is found throughout this thesis, both in experimental work and as a discussion partner which lead to new experiments. Thank you!

André Deslauriers has aided in every aspect of this thesis; experimental work, as a discussion partner, revisions and translations and practical help. But most of all I would like to thank him for his continuous support and encouragement, merci André.

Fabrice Galaud and Robert Cordeau have been invaluable in technical assistance and as discussion partners throughout my thesis and especially for synthesis of chemicals.

I am grateful for all the time and effort Natalie M. Rachel has offered me. For her tremendous help in revising Chapter 2, she has also aided in experimental work for several of the chapters,

and as a discussion partner. I specifically thank her for aiding in the finalizing of parts of this thesis; the completion would not have been possible without her.

Finally I would like to thank Flora Rosenow and Marten Meyer-Mews for their generous aid in correcting parts of this thesis and moral support and encouragement throughout my masters.

Chapter 1
Introduction

In this thesis I will explore the reactivity of the enzyme microbial transglutaminase. We do this in order to probe if this enzyme can serve as a more general catalyst for the synthesis of non-natural amide bonds. The ultimate goal of this work is to aid in developing a new biocatalyst and, as such, a tool for greener chemistry in a variety of applications.

1.1 Green Chemistry

"Sustainable Development: the ability to make development sustainable to ensure that it meets the needs of the present without compromising the ability of future generations to meet their own needs." [1]

Over the last decades, it has become evident that human activity is significantly altering the climate of the planet, which in turn may have detrimental effects on life on earth [2]. The term 'planetary boundaries' has been defined as the safe operating space for humanity, where nine boundaries have been identified, including global warming, chemical pollution and global fresh water use. It follows that transgression of any of these boundaries may be catastrophic, and humanity as a whole may jeopardize its very own existence [2]. Maintaining a safe operating level by remaining within these boundaries is therefore essential. The chemical industry directly impacts upon several of these boundaries; thus the transition towards chemical sustainability may be vital for preventing detrimental effects on life. Over 80,000 chemicals are currently on the

global market; however, the toxicity of the majority of those chemicals has not been assessed, and even less is known about their interaction with each other and the environment [2]. As such, it is evident that the synthesis of safer chemicals and a broader knowledge of their impact is paramount for green chemistry.

Green chemistry has grown, over the last 20 years, to become an integral part of chemistry and the chemical industry. It is defined as "the design of chemical products and processes that reduce or eliminate the use and generation of hazardous substances" [3]. Green chemistry is a philosophy which guides decision makers from all areas of chemistry to make sustainable choices, in aspects ranging from chemical design to waste disposal. Green chemistry should therefore not be seen as a quest for one single optimum method, but rather as a process towards a more sustainable industry. Decision makers can obtain guidance from the twelve principles of green chemistry (Table 1.1) [3,4]. Two key points of green chemistry are the design of safe chemicals and non-hazardous syntheses of those chemicals.

In the context of this thesis, the focus lies on safe chemical synthesis. Many methods are available for assessing material efficiency of a chemical process, including calculation of the *E-factor* (product mass/sum of the mass of reactants and auxiliaries), *reaction mass efficiency* (product mass/reactant mass), and *atom economy* (product molar mass/reactant molar mass). The variety of methods illustrates the complexity of defining how green a method is. These simple methods of assessing and comparing various methods provide initial estimations, but in order to gain the understanding needed for design of truly benign new chemicals and chemical processes, a more holistic view of the 12 principles of green chemistry may be required [4]. Here we will investigate one option towards the greener synthesis of chemicals, biocatalysis, in accordance with our initial goal of developing MTG as a green alternative to amide synthesis

Table 1.1: The 12 principles of green chemistry [3]

Principle
1. Prevention
2. Atom economy
3. Less hazardous chemical syntheses
3. Designing safer chemicals
4. Safer solvents and auxiliaries
5. Design for energy efficiency
6. Use of renewable feedstocks
7. Reduce derivatives
8. Catalysis
9. Design for degradation
10. Real-time analysis for pollution prevention
12. Inherently safer chemistry for accident prevention

1.2 Biocatalysis

Although in its broadest sense, one could define biocatalysis as pertaining to all reactions which occur in nature, here we define biocatalysis as a reaction proceeding with the aid of a biocatalyst, whether in its pure isolated form, or as whole cell catalyst, which is applied by human intervention to obtain a specific desired product. Biocatalysis is an emerging field within green chemistry, where a biocatalyst - typically an enzyme - is used to alter or create new molecules. It encompasses many of the 12 principles of green chemistry by its very nature, by such qualities as requiring mild reaction conditions, using biodegradable and non-toxic catalysts and producing fewer side reactions.

According to our definition, biocatalysis dates back to prehistoric times, starting with the early fermentation of alcoholic beverages around 7000 BC in China [5]. Thus, it dates back further than the discipline of chemistry itself, which is estimated to have appeared around 2000 years ago. But, as chemistry developed over the past two millenia, industrial biocatalysis lagged behind. The complexity related to the application of labile biomolecules and the poor

understanding of how cells function chemically have hindered its development. Indeed, this is demonstrated through brewing, where methods have remained almost unchanged over the millennia. For example, the yeast species used for fermentation, *Saccharomyces cerevisiae*, has been found across all continents. This indicates that this species has traditionally been used uniformly for brewing [6]. From these early beginnings of biocatalysis, it would take 9000 years before Rosenthaler used a biocatalyst for the synthesis of the R stereoisomer of mandelonitrile from a plant extract [7,8], marking the transition into the era of modern biocatalysis in 1908. Below, the main approaches to improving biocatalysis are presented: condition screening (including substrate screening), the directed evolution of enzymes, computational approaches and synthetic biology.

1.2.1 Condition screening

A natural starting point for novel biocatalytic processes is substrate screening, where non-natural starting materials are tested, often in combination with non-native reaction conditions. Numerous methods for substrate screening can be found in the literature. Examples include expansion of the substrate scope of monooxygenases to evaluate the biocatalytic potential of the enzyme [9]. When the number of substrates tested becomes too great, high-throughput screening is applied. This technique is widely applied in the pharmaceutical industry in search of novel inhibitors, where time reduction and cost per sample is crucid to success. These methods are easily transferred to substrate screening, as illustrated by the screening for new substrates of hydrolases by Bornscheuer and colleagues [10]. The co-product of the reaction was coupled with a second reaction, producing a readily-detectable fluorescent substrate [10]. Other novel strategies for high-throughput screening include techniques such as automation and small/nano scale

experiments [11]. In this thesis, we performed substrate screening (Chapters 2 and 3), albeit not at a high-throughput level.

1.2.2 Directed evolution of enzymes

Another area of biocatalysis which is highly dependent on screening methods is directed evolution. The genetic revolution of the past decades has enabled scientists to unravel the complexity of the cellular machinery. From these findings, our understanding of the structure and function of enzymes has rapidly expanded. The pioneering work of Pim Stemmer and Frances Arnold (Caltech) in molecular biology has enabled biocatalysis to take the role it has today, where enzymes are broadly applied in novel areas [8,12]. Their work opened up a new area of molecular biology, namely 'directed evolution', where enzymes are rapidly 'evolved' towards novel traits, or optimized for a particular reaction or reaction conditions (Figure 1.1). Novel traits include altered or optimized substrate specificity, increased stability under new reaction conditions such as elevated temperatures or tolerance towards organic solvents [8,12]. This is done through introducing changes in the DNA sequence, which then leads to variations in the amino acid sequence of proteins and as such may change the properties of the final protein product. In this thesis we have not explored this option, but it remains an important aspect of the future development of the enzyme as a general amide catalyst, especially to alter the substrate specificity of the enzyme.

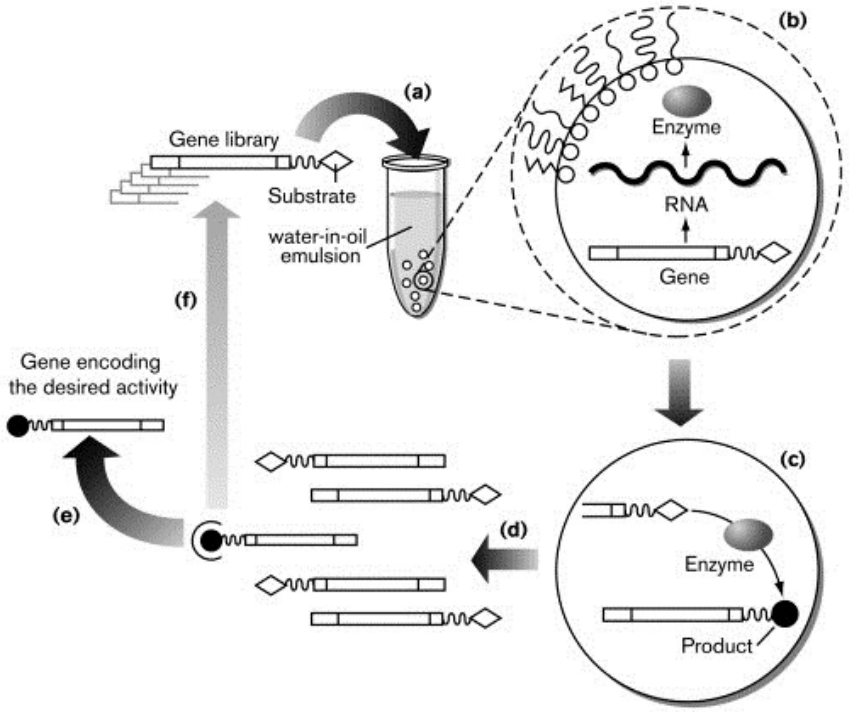


Figure 1.1: Overview of directed evolution [13]

1.2.3. Computational approaches

An emerging area that is currently changing the field of biocatalysis pertains to *in silico* methods, which serve as a support to directed evolution. These computational tools combine knowledge of enzyme structure and function with chemical and physical knowledge to predict, through extensive calculations, the behavior and interactions of proteins. Computational algorithms exist to predict the outcome of mutations, by replacing one amino acid residue for another, and predicting the outcome [14]. *In silico* methods offer potential broad scope, as billions of mutants (in fact, astronomical numbers) are rapidly be screened at a low cost [15]. Due to the limitations in computing power, modern *in silico* methods turn to classical physics (rather than quantum) to predict how an enzyme will change as a result of mutations. However, we are currently seeing

the rise of the next generation of computational methods, which applies quantum mechanics to active-site residues [16]. This may yield the precision needed for more accurate simulations.

In this body of work we apply an *in silico* method, where we use molecular docking to explore new substrates (Chapter 4). Molecular docking calculates the binding energies between a ligand and a target protein [17]. We docked small molecules as potential substrates on our target enzyme surface, to investigate binding, where the target remains static and the ligand is flexible. Through molecular docking one may rapidly identify new substrates for enzymes at low cost, to broaden the biocatalytic potential of an enzyme without altering the enzyme itself. Nonetheless, docking algorithms provide only approximations, as important effects such as solvation are not explicitly considered.

A more computationally-intensive application of computer-aided methods is *de novo* design of biocatalysts, where residues are built into and around a designed active site to change substrate specificity or even give rise to new reactivity. This method predicts the physical and chemical space needed to fit a particular substrate and then calculates which changes in an enzyme would facilitate such an environment [14]. A recent example of *de novo* design from David Baker and colleagues (U. Washington) is the engineering of a Diels-Alderase, a novel enzymatic reaction [18]. To date, the resulting new enzymes offer very low reactivity, comparable only to poor chemical catalysts. Nonetheless, increasing computational power will contribute to the rapid expansion of these highly promising tools [18]. With such tools at hand, we could experience the next great advances in the field of biocatalysis [14,15]. Thus, advancements could be made both on expanding the range of substrates and the stability of enzymes.

While biocatalysis has grown in academia with increasing knowledge and new tools, important advances have also been made in biocatalytic process chemistry, in particular in areas

such as enzyme immobilization on solid surfaces, and co-factor regeneration [19]. Some large scale biocatalytic operations outside of the pharmaceutical industry include the synthesis of aspartame [19–21] and nicotinamide [20,22]. Without the successful implementation of industrial strategies for enzymes, the development of application-based science is in vain. Nonetheless, more development of innovative techniques is required to fully exploit the potential of biocatalysis [23] and to effectively transfer biocatalysis from academia to industrial applications [8].

1.2.4 Synthetic biology

Synthetic biology is a field that successfully combines the depth of biochemical knowledge with the application-directed orientation of engineering, and is likely to become an integral part of the future of biocatalysis. In synthetic biology, cells are seen as programmable units that can be engineered to produce complex molecules by altering cellular pathways [24,25]. Examples include the synthesis of complex biofuels (biodiesel), through highly engineered *E. coli* cells [26]. Because synthetic biology depends on highly complex intracellular mechanisms and signaling pathways, scientists have only recently obtained enough biochemical knowledge to begin to explore the full potential of this field. Synthetic biology will expand significantly in the coming years, as a result of its recently demonstrated utility in biocatalytic applications such as synthesis of active pharmaceutical ingredients. Complex synthesis routes using whole cells are expected to greatly reduce costs and increase the green potential of the chemical industry [24,25].

To apply our enzyme in a synthetic biology synthesis route remains distant, as its cellular over expression would likely lead to undesired protein-protein cross linking. Significant enzyme engineering would be required to prevent it from reacting with cellular components, and it has not been a focus of this thesis to further enable the enzyme for synthetic biology.

1.3 Amide bond formation

A suitable target for biocatalysis is the synthesis of amide bonds. Amides are widespread in nature and are vital to maintaining life, as they are found in the peptide backbone of proteins [27]. The success of amides in nature can be attributed to their inherent stability, which also makes them an attractive functional group in all areas of synthetic chemistry. It is, in fact, the most abundant functional group in the pharmaceutical industry [27,28]. An example of an amide-bearing pharmaceutical compound is Atorvastatin (commercialized as Lipitor), a cholesterol-lowering drug that was the leading blockbuster drug from 2003 to the expiry of its patent in 2011 [29]. Although amide bonds are extensively used in organic synthesis, many challenges remain in making these bonds. Amides are most often synthesized through the condensation of an amine with a carboxylic acid. This reaction does not proceed readily, and therefore depends on activation of the carboxylic acid to render it more susceptible to nucleophilic attack by the amine. These methods naturally depend on stoichiometric quantities of activating (or coupling) reagents. Therefore, this reaction suffers from poor atom efficiency and is often costly, but it nonetheless remains the most common synthetic reaction in the pharmaceutical industry [27–29]. In 2007 'Amide formation avoiding poor atom economy reagents' was voted the top challenge of organic chemistry by the ACS Green Chemistry Institute. Still, the call for catalytic methods for amide synthesis remains an important challenge [28].

Protein biosynthesis involves formation of the peptide backbone by linking the α-amino group of one amino acid to the activated α-carboxylate of another; activation occurs by esterifying the carboxylate to a tRNA. Although this method is elegant and highly functional, it is a complex, highly specialized system that does not transfer well outside the environment of a cell. A biocatalytic alternative to the formation of amide bonds is proposed here. There are several examples in nature of enzymes that catalyze amide bond formation, such as nitrile

hydrolases[22], peptidases, proteases[30] and transglutaminases [19–21,31]. In this thesis we focus on microbial transglutaminase for the formation of amide bonds.

1.4 Transglutaminases

Transglutaminases (TGase) (EC 2.3.2.13) are enzymes that catalyze the formation of amide bonds through iso-peptide linkages between proteins in nature, thus crosslinking proteins to form insoluble aggregates. Best known is perhaps the human factor XIII, which coagulates blood through the crosslinking of fibrin. The transglutaminase reaction relies on an acyl transfer reaction between an acyl-donor substrate, the γ-carboxyamide group of glutamine, and an acyl-acceptor substrate, the ε-amino group of lysine, yielding an isopeptide bond (Figure 1.2) that is highly resistant towards enzymatic proteolysis and mechanical stress [32].

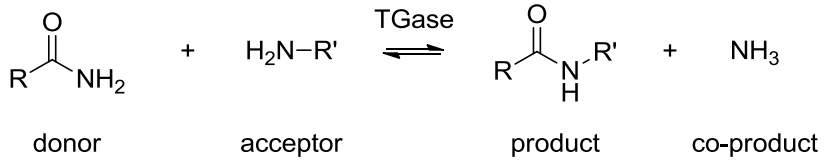


Figure 1.2: Transglutaminase reaction

Eukaryotic transglutaminases are calcium-dependent enzymes, requiring calcium concentrations above physiological conditions to be activated. Thus, their activation occurs only under conditions disrupting cellular homeostasis, such as skin regeneration during wound healing, and blood clotting. They are also involved in a number of human diseases, including neurodegeneration, and neoplastic, autoimmune and skin diseases. These enzymes are therefore important therapeutic targets. Eight mammalian transglutaminases have been identified, and six

of them thoroughly characterized. They show high structural homology, where all consist of four domains: a β-sandwich, the core domain with the active site and regulatory sites, and two *C*-terminal β-barrels, with a total mass ranging from 77 kDa to 90 kDa. The active site is also conserved, with a cysteine-based catalytic triad consisting of cysteine, histidine, and aspartic acid or aspargine. Moreover, the enzyme belongs to the papain-like superfamily [32,33].

1.5 Microbial transglutaminase

This thesis will examine the microbial transglutaminase (MTG) from *Streptomyces mobaraensis*, which structurally bears little resemblance to the mammalian variants and is likely a product of convergent evolution. This is made evident by the size of MTG, which is 38 kDa, or about half the size of its mammalian homologue [32,34,35]. Its crystal structure [36] reveals that it is a monomeric, single-domain enzyme, having no homology with any of the 4 domains found in the mammalian TGase [33]. Both the mammalian and the microbial enzymes must be activated post-translationally. This is an important feature because undesired protein crosslinking activity may be detrimental to the cell; indeed, transglutaminase activity is known to be involved in mammalian apoptosis mechanisms [35,37,38]. MTG is a calcium-independent enzyme, which contains an *N*-terminal pro-sequence that folds over the active site and renders the pro-enzyme inactive. It is cleaved off by a protease to activate the enzyme.

MTG has a compact globular structure composed of 11 α-helices and 8 β-strands, with the active site located in a central cleft (Figure 1.3). The surface of the enzyme is mostly covered with positively-charged residues, except the active site cleft which contains five acidic residues: Asp3, Asp4, Glu249, Asp255, and Glu300, illustrated in red in Figure 1.4. Furthermore, we find a number of aromatic residues in the active cleft: Trp59, Tyr62, Trp69, Tyr75, Tyr278, Tyr291,

and Tyr302, illustrated in green in Figure 1.4 [36]. Several of these residues have been found to be important in substrate binding, particularly larger protein substrates [39]. This knowledge guided many of the experiments in the following chapters (Chapters 2-5).

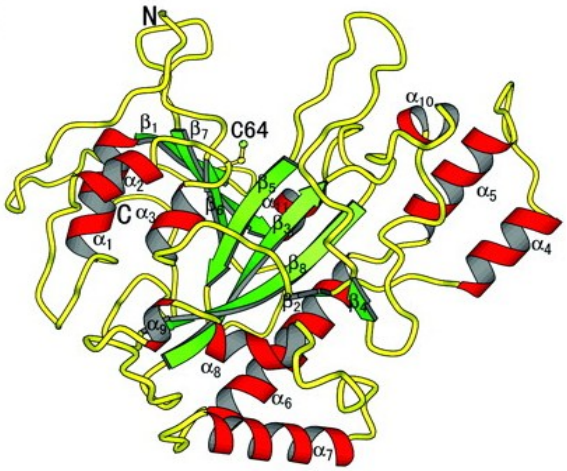


Figure 1.3: Tertiary structure of MTG [36], Coordinates from PDB 1IU4

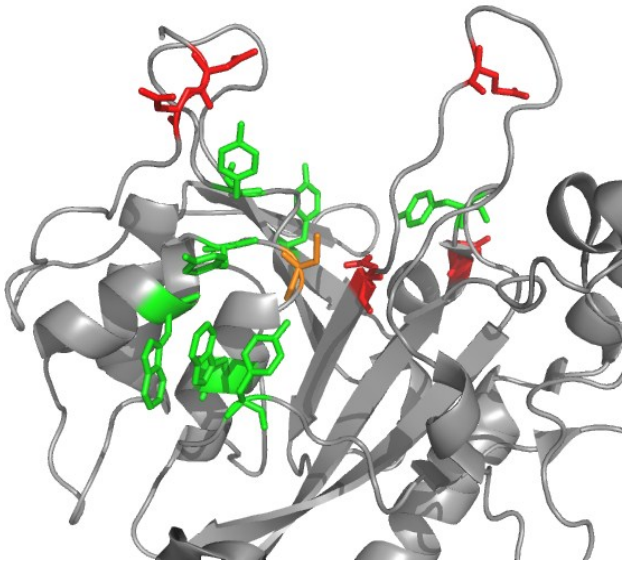


Figure 1.4: Active site of MTG. The reactive cysteine is shown in yellow, acidic residues in red, and aromatics in green. The structure is illustrated using Pymol with the coordinates from PDB 1IU4

The reaction mechanism is cysteine-based, where Cys64 attacks the amide carbonyl of the acyl-donor substrate, forming a tetrahedral enzyme-substrate intermediate. A more complete reaction mechanism involving an inverted catalytic triad has been proposed (Figure 1.5). Following formation of the tetrahedral intermediate, Asp255 donates a proton to the oxyanion hole, and allows for a collapse of the tetrahedral intermediate to form the enzyme-bound thioamide, releasing the starting amide nitrogen in the form of ammonia. An amine acyl-acceptor substrate – habitually a protein-bound lysine – aided by deprotonation by Asp255, then attacks the thioamide carbonyl. The crosslinked product is released and the free enzyme is regenerated [35,36].

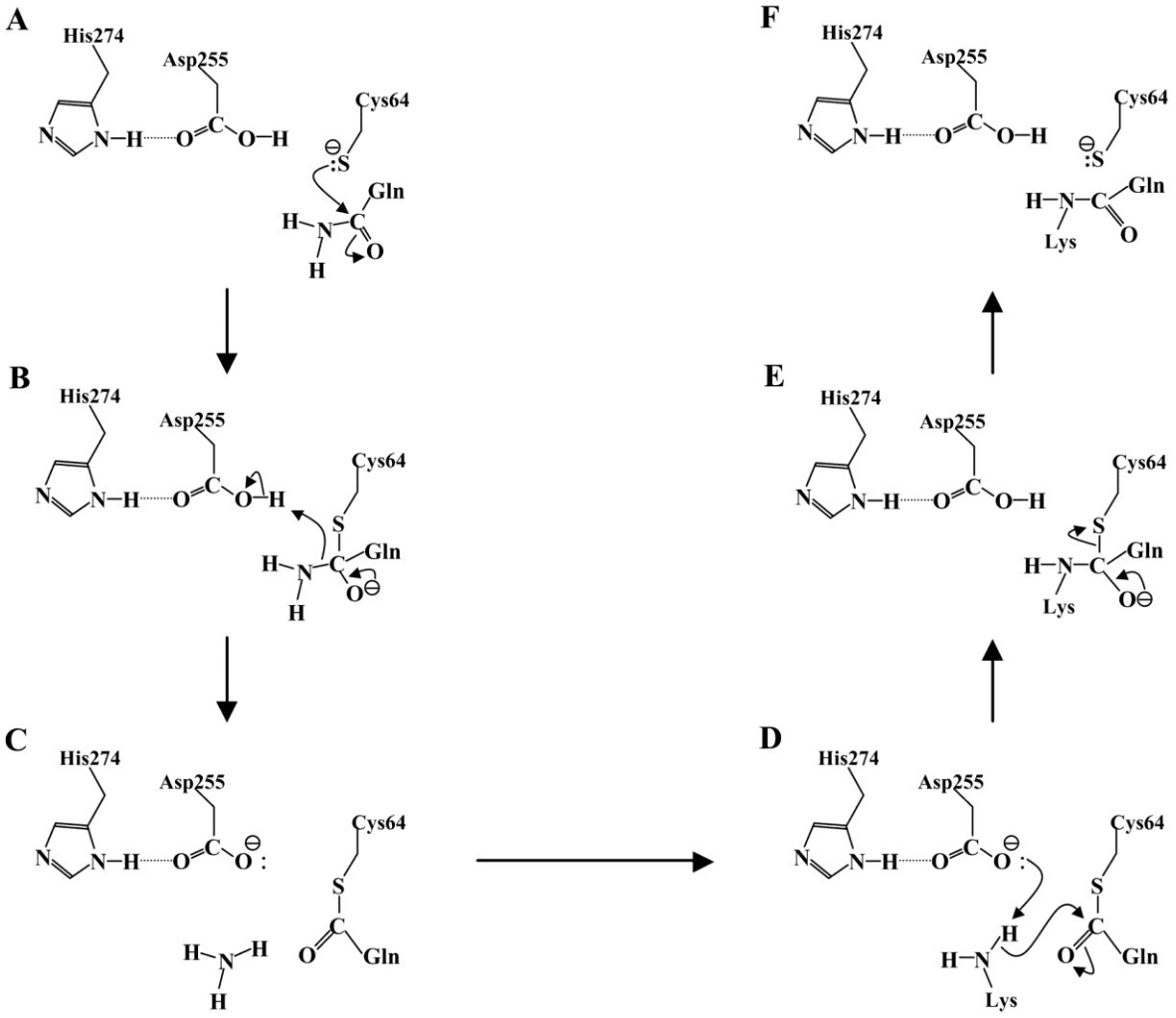


Figure 1.5: Reaction mechanism of MTG [36]

MTG has been in commercial use since the late 1980's in the textile and food industry. The crosslinking of beef myosin, casein, and crude actomyosin whey-, soy- and wheat-proteins results in a texturization of the product. This property is in turn used alter the texture and appearance of products [40]. For example, MTG is used in the seafood industry, where the texture of products may impede handling and further usage in processed foods. Because of its widespread applicability, it is currently produced industrially on a large scale by companies such

as Ajinomoto (Tokyo, Japan) and Zhejiang Yiwan Biolabs (Zhejiang, China). The success of MTG in the food and textile industry can be attributed to its inherent stability, as illustrated by a number of favorable properties, described below. It remains active over long periods of time: we observed up to 60% of activity remaining after 3 days of incubation at 37 °C (Chapter 2). In comparison, its mammalian homologue loses all activity within a few hours under similar conditions (R. A. Chica, J. W. Keillor, and J. N. Pelletier, unpublished). MTG is also active under broad pH and temperature ranges, between pH 4 to 10, and 10 °C to 70 °C, where the pH optimum was found to be between pH 6-7. Further, the highest activity was identified at 55 °C, but the incubation of MTG at temperatures over 40 °C for 10 minutes resulted in a decreased activity [41,42]. Finally MTG is stable and active under high concentrations of certain organic solvents, such as DMSO and ethanol, which enables MTG to be used with non-polar substrates, thereby greatly increasing its biocatalytic scope [43].

MTG has garnered interest over the last decade for its application to novel, non-native processes, perhaps because of the many favorable properties mentioned above, and due to the stability of the isopeptide bond formed, both *in vivo* and under harsh conditions. Those applications often rely on crosslinking of compounds other than two native proteins. It is known that MTG is promiscuous towards its acyl-acceptor substrate [32,35,44], but less so towards its acyl-donor substrate [32,45,46]. To date, the acyl-acceptor and acyl-donor substrate-binding sites are not known, other than the fact that they must be proximal to the long cleft at the bottom of which the reactive cysteine is found. Several research groups have investigated peptidic substrates, searching for specific, high-affinity sequences to serve as a labeling tag. Grafting such a peptide to a target protein would enable MTG-catalyzed, site-specific protein labeling, for *in situ* labeling of proteins at this specific sequence. Three attempts to identify a specific acyl-donor

sequence are discussed here. Ohtsuka and colleagues conducted a systematic study using synthetic acyl-donor (Gln-containing) peptides to test reactivity of MTG as well as transglutaminases from guinea pig liver and from the fish, *red sea bream* (*Pagrus major*) [46]. The heptamer peptide library GXXQXXG (X=G, A, S, L, V, F, Y, R, N, E, L) was tested with all three TGases. The results indicate that MTG is sensitive to the types of residues which precede the glutamine. The authors further found that MTG did not react with glutamine analogues if the side-chain length was reduced or increased [46]. Further, two studies using phage display investigated acyl-donor substrate specificity. Neither identified a clear pattern in the residues surrounding the Gln in the donor sequence and as such giving rise to a specific acyl-donor identity. Finally, Lee and colleagues recently reported two pentapeptides, RLQQP and RTQPA, toward which MTG exhibits a high reactivity, with 93% and 53% conversion of 1 mM equimolar donor and acceptor after 40 min at 37 °C [47]. The K_M of the model substrate Z-Gln-Gly for MTG was found to be 27 mM [48]; the high reactivity of the two pentapeptides at 1 mM suggests that the binding of these peptides is stronger than with Z-Gln-Gly [45,47].

In quest of novel applications of MTG, a number of new areas have been explored, some of which are listed here. These include site-specific PEGylation of proteins, where a terminally aminated polyethylene glycol (PEG) was conjugated at a single acyl-donor site on a target protein. This was done by engineering the target protein to contain only one glutamine suitable for recognition by MTG [43]. Two studies involved the modification of antibodies through labeling either a protein-bound glutamine or lysine on the antibody with either a fluorescent label, a metal chelator or radionuclides [49,50]. Similar experiments involved making DNA-protein conjugates, where DNA linked to the donor compound Z-Gln-Gly was reacted with a lysine on the protein substrate alkaline phosphatase. The enzyme remained active through the linkage [51]. MTG has

also been used for formation of hydrogels for tissue engineering and drug delivery. In those reactions, peptides 2 to 11 amino acids long were linked together by reaction of an internal glutamine with the *N*-teminus of another molecule of the same peptide, resulting in a polymerization reaction [40,52]. Finally MTG has been used in areas which fall closer to the typical green chemistry category, through applications such as polymer synthesis in degradable bioplastics. Bioplastics incorporate the cross-linking capabilities of MTG together with degradable biomass, to make an environmentally benign product [53].

In this thesis, the substrate specificity of MTG was explored in greater depth, by means of substrate screening and molecular docking, in order to expand its synthetic utility for the formation of amide bonds between diverse compounds. Specifically, Chapter 2 describes acceptor substrate screening of small non-natural amines. In Chapter 3, we further discuss the impact of the results from Chapter 2 in the context of biocatalysis and green chemistry, and compare these trends with molecular docking results to predict the reactivity of novel amine substrates. In Chapter 4 we briefly discuss the significance of screening the donor substrate specificity, and present some preliminary data indicating trends found in small amide donor substrates. Finally Chapter 5 wraps-up the thesis by concluding on major findings in the previous chapters and highlights directions for future work.

Chapter 2 :
Microbial transglutaminase displays broad acyl-acceptor substrate specificity

This chapter presents the first extensive study of amine acceptor substrates of MTG with varied functional groups. We performed these experiments to aid in defining the substrate scope both with respect to natural substrates such as amino acids, and with respect to non-natural substrates. Therefore this chapter supports our goal of probing the potential of MTG as a broader biocatalyst for the formation of amide bonds.

Author contributions are: Joelle N. Pelletier supplied financial and intellectual support, and contributed to writing the manuscript. Jeffrey W. Keillor assisted in the initial development of the project and writing the manuscript. All other contributions are mentioned in the acknowledgments. The majority of the work both in terms of experimental data collection, analysis and interpretation of data and writing of the manuscript was done by myself.

A revised version of the manuscript has been submitted to Applied Microbiology and Biotechnology.

Microbial transglutaminase displays broad acyl-acceptor substrate specificity

Maria T. Gundersen[1,2,3], Jeffrey W. Keillor[3,4] and Joelle N. Pelletier[1,2,3,5]*

From the [1]Département de biochimie, Université de Montréal, Montréal, Québec H3C 3J7, [2]PROTEO, the Québec Network for Protein Function, Structure and Engineering, [3]CGCC, the Center in Green Chemistry and Catalysis, [4]Department of Chemistry, University of Ottawa, Ottawa, ON K1N 6N5 and[5]Département de chimie, Université de Montréal, Montréal, Québec H3C 3J7

*Corresponding author: Joelle N. Pelletier

2.1 Abstract

The great importance of amide bonds in industrial synthesis has encouraged the search for efficient catalysts of amide bond formation. Microbial transglutaminase (MTG) is heavily utilized in crosslinking proteins in the food and textile industries, where the side-chain of a glutamine reacts with the side-chain of a lysine, forming a secondary amide bond. Long alkylamines carrying diverse chemical entities can substitute for lysine as acyl-acceptor substrates, to link molecules of interest onto peptides or proteins. Here, we explore short and chemically varied acyl-acceptor substrates, to better understand the nature of non-natural substrates that are tolerated by MTG, with the aim of diversifying biocatalytic applications of MTG. We show, for the first time, that very short-chain alkyl-based amino acids such as glycine can serve as acceptor substrates. The esterified α-amino acids Thr, Ser, Cys and Trp – but not Ile – also showed reactivity. Extending the search to non-natural compounds, a ring near the amine group – particularly if aromatic – was beneficial for reactivity, although ring substituents reduced reactivity. Overall, amines attached to a less hindered carbon increased reactivity. Importantly, very small amines carrying either the electron-rich azide or the alkyne groups required for click chemistry were highly reactive as acyl-acceptor substrates, providing a robust route to minimally modified, 'clickable' peptides. These results demonstrate that MTG is tolerant to a variety of chemically varied natural and non-natural acyl-acceptor substrates, which broadens the scope for modification of Gln-containing peptides and proteins.

Keywords: Amide bond formation, microbial transformations, biocatalysis, peptide modification

2.2 Introduction

Transglutaminases (EC 2.3.2.13) catalyze the formation of peptide linkages by promoting an acyl transfer reaction between an acyl-donor substrate, the γ-carboxyamide group of glutamine, and an acyl-acceptor substrate, the ε-amino group of lysine, thus forming a new amide bond. In nature, these enzymes crosslink proteins to form insoluble protein aggregates [54]. Microbial transglutaminase (MTG) from *Streptomyces mobaraensis* is a calcium independent enzyme that has been used commercially since the late 1980's in the textile and the food industries, to alter the texture and appearance of products *via* protein crosslinking [40]. More recently, MTG has been utilized in novel non-native processes such as site-specific protein PEGylation [43], antibody modification [49,50], formation of DNA-protein conjugates [51], and the formation of hydrogels for tissue engineering and drug delivery [40,52]. It is also applied to 'green' applications, such as the biocatalytic synthesis of degradable bioplastics [53].

Biocatalysis is increasingly applied in areas spanning the preparation of bulk chemicals to high-value synthesis of chiral active pharmaceutical ingredients (APIs) [19,21]. Amide bonds are highly represented in APIs, as well as in bulk polymers and commodity chemicals. Nonetheless, current methods of amide synthesis are generally characterized by high waste and cost, highlighting the importance of making their synthesis more sustainable. Indeed, 'amide formation avoiding poor atom economy reagents' was voted the top challenge of organic chemistry in 2007 by the ACS Green Chemistry Institute [28]. To this effect, new catalytic methods for amide synthesis are called for [28,29].

A number of enzymes catalyze amide bond formation from different starting materials. Well known examples include proteases, peptidases, transglutaminases, nitrile hydrolases and lipoic acid ligases. Competing hydrolysis reactions frequently impede the efficiency of enzymatic

amide bond formation. Transglutaminases, which natively catalyze the formation of amide bonds rather than their hydrolysis (as is the case for proteases and peptidases), offer an intrinsic advantage in this respect. The mature form of MTG is produced with high quality on a large scale, and its industrial success can be attributed to its ease of handling and stability over a broad range of pH and temperature [41,42]. Furthermore, it is active in the presence of organic co-solvents [43], thus allowing the use of poorly water-soluble compounds. Its stability and ease of application enable the use of high substrate concentrations and catalyst recycling, both important considerations in the context of biocatalyzed conversions.

In spite of the extensive industrial use of MTG, little is known about its ability to use small, non-proteogenic molecules as substrates. Ohtsuka and colleagues investigated a restricted range of small, non-proteogenic acyl-acceptor substrates, confirming that MTG can react with a number of primary amines, mainly natural compounds or their analogues [44]. Here, we investigate the breadth of acyl-acceptor substrate specificity of MTG in order to expand its scope and utility as a green biocatalyst for amide synthesis, by broadening the classes of compounds investigated. In particular, we revisited alkyl-based amino acids to uncover new reactivity, expanded the range of known reactive α-amino acids and identified synthetically attractive amine substrates of non-biological origin. Our results increase the understanding of the specific characteristics of non-proteogenic acceptor substrates, and broaden the scope of MTG as a 'green' catalyst for amide synthesis.

2.3 Materials and methods

2.3.1 Materials

The plasmid pDJ1-3 was kindly provided by Professor M. Pietzsch (Martin-Luther-Universität, Halle-Wittenberg, Germany). pDJ1-3 encodes the pro-enzyme of MTG from *Streptomyces mobaraensis* inserted between the *Nde*I and *Xho*I restriction sites of the vector pET20b [55]. Deionized water (18 Ω) was used for all experiments. HPLC solvents were of analytical grade, and products used for the expression and purification of MTG were of biological grade. Azides were synthesized in the laboratory of J.W.K., according to known literature procedures for the preparation of ethylamine azide [56] and propylamine azide [57]. Other chemicals used were purchased from the suppliers listed below. 4-Methoxybenzamide (97% purity) and glycine ethyl ester hydrochloride were purchased from Acros Organics (Waltham, USA). Benzylamine, *m*-anisidine, N^{α}-acetyl-L-lysine methyl ester hydrochloride, □-alanine, 6-aminocaproic acid, 5-aminovaleric acid, aniline, *O*-benzylhydroxylamine hydrochloride, L-threonine methyl ester hydrochloride, L-tryptophan methyl ester hydrochloride, sarcosine, ammonium carbamate, *N*-ethylmethylamine, cyclohexane methylamine, cyclohexylamine, glycine hydrochloride, D-serine methyl ester hydrochloride, L-serine methyl ester hydrochloride, trypsin from bovine pancreas (10,000 BAEE U/mg), cadaverine, γ-aminobutyric acid, hydroxylamine and *p*-xylenediamine were purchased from Sigma-Aldrich (St Louis, USA). Aminoacetonitrile was purchased from Bachem (Bubendorf, Switzerland). Carboxybenzyl-glutaminyl-glycine (Z-Gln-Gly) and benzylazide were from Alfa Aesar (Ward Hill, USA). Glutathione (reduced) and thiamine were from Bioshop (Burlington, Canada). L-Cysteine ethyl ester hydrochloride and L-isoleucine methyl ester hydrochloride were purchased from Chem-Impex (Wood Dale, USA).

Propargylamine was purchased from Fisher (Waltham, USA). Formic acid (98% purity) was from Fluka Analytical (St Louis, USA). 3-Chloro-4-fluorobenzylamine was purchased from Lancaster (Ward Hill, USA).

2.3.2 Expression and purification of MTG

The plasmid pDJ1-3 encodes MTG with its *N*-terminal pro-sequence and a *C*-terminal hexa-histidine tag. It was transformed into *E. coli* BL21(DE3), using standard procedures [58] and maintained with 100 μg/mL ampicillin. MTG was expressed in autoinducing medium as follows: a5-mL culture was propagated overnight at 37°C with agitation at 240 RPM in ZYP-0.8G medium [59], and used to inoculate 500 mL of ZYP-505 medium [59]. The culture was incubated at 240 RPM, first at 37°C for 2h, then overnight at 22°C. The culture was centrifuged at 5,000 RPM (Sorvall centrifuge RC 5C Plus, SLA-3000 rotor) at 4°C for 15 min, and the cells resuspended in 40 mL of 0.2M Tris-HCl pH 6.0. The cells were disrupted by sonication over ice (three cycles of 30s pulse at 20% intensity / 1 min pause) using a Branson sonicator and further by one pass through a Constant Systems cell disruptor set at 27 kPSI and cooled to 4°C. MTG was then activated by cleavage of the pro-enzyme leader sequence through incubation in a 1:9 ratio (v/v) of trypsin (1mg/mL) to unpurified MTG for 45 min, at 30°C. The activated MTG was purified using a 5 mL His-trap Ni-NTA column (GE Healthcare) equilibrated in 50 mM phosphate buffer pH 8.0, with 300 mM NaCl, and eluted with an imidazole gradient (0-140 mM), on an Åkta FPLC (GE Healthcare). The purified, activated MTG was dialyzed against 0.2 M Tris-HCl buffer, pH 6.0. The average yield was 90 mg of activated MTG per liter of culture, with > 85% purity as estimated by resolution on 10% SDS-PAGE followed by staining with Coomassie blue. Aliquots were snap-frozen and stored at -80°C in 15% glycerol.

2.3.3 Determination of MTG activity

The activity of the purified, activated MTG was quantified using the hydroxamate assay, as previously described [60]. Briefly, MTG was incubated with 30 mM Z-Gln-Gly and 100 mM hydroxamate at 37°C for 10 min. The reaction was quenched with a solution containing 2.0 M $FeCl_3$•6 H_2O, 0.3 M trichloroacetic acid and 0.8 M HCl. The resulting iron complex was detected by its absorbance at 525 nm. One unit (U) of MTG produces 1 μmol of L-glutamic acid γ-monohydroxamate per min at 37°C.

2.3.4 Reaction of MTG with various acceptor substrates

Amide acyl-donor substrate Z-Gln-Gly (40 mM), amine acyl-acceptor substrate (100mM) and 10 mM glutathione were combined in 0.2 M Tris-HCl buffer in a final volume of 350 μL. The pH of each reaction was adjusted to the pK_a of the tested acceptor substrate amine, if within the range of pH 6-9. Otherwise, pH 6 or pH 9 were used, as most closely matched to the amine pK_a. For this purpose, the pK_a values were calculated online through the SPARC calculator [61]. The pH values of the reaction mixtures are indicated in Table A1.1, Annex 1. MTG (2 U/mL final concentration) was added to the substrate mixture, and the reaction was incubated in closed 1.5-mL Eppendorf tubes at 37°C for up to 72h. A control reaction without MTG was run for each acyl-acceptor substrate. All experiments were performed in triplicate using MTG isolated from three independent MTG purifications (nine times total).

2.3.5 Product detection

Donor substrate consumption was monitored by HPLC-MS. Formic acid was added to an equal volume (0.1 mL) of reaction mixture and allowed to stand at room temperature for 5 min to quench the reaction. The volume of the quenched sample was adjusted up to 1 mL with H_2O; 0.15 mL of this solution was then combined with 0.15 mL internal standard solution (1 g/L 4-methoxybenzamide in neat DMSO), and the volume was adjusted up to 1.5 mL with H_2O. Each sample was run over a hand-packed Ni-NTA column (0.5 mL bed volume) to remove MTG, and then filtered through a 0.2-µm PTFE filter to remove any particles. Samples (10 µL) were injected on a Synergi 4-µm polar-RP 80 Å, 50 × 2.00 mm LC column (Phenomenex), using a Waters 2545 HPLC apparatus, and eluted with a 5-70% MeOH/H_2O gradient. Masses were detected under positive ionization with a Waters 3100 single quadrupole mass detector. The consumption of donor substrate was determined by standardization with the 4-methoxybenzamide internal standard, relative to the concentration of the donor substrate in the control reaction (no MTG) for the same acceptor. The mass corresponding to the expected product was also detected. NMR analysis of the product was performed to confirm identity of the expected product for the following acceptor substrates: 5-aminovaleric acid, propargylamine and 6-aminocaproic acid. For that purpose, the MTG-catalyzed reactions were run on a larger scale (1.4 mL). Preparative HPLC was performed with a Synergy polar-RP 80 Å, 100 × 21.20 mm AXIA packed column (Phenomenex) on a Waters 1525 HPLC with elution using a 5-90% MeOH/H_2O gradient containing 0.1% formic acid. The product was detected with a Waters 2487 dual absorbance detector. Fractions were further analyzed with direct injection mass spectrometry with a Waters 3100 single quadrupole mass detector, and those containing the mass corresponding to the expected products where pooled. Methanol was evaporated then the

samples were lyophilized to yield the isolated product in a powder form. ^{1}H- and ^{13}C-NMR spectra of the products were acquired on a Bruker Avance II 700 MHz spectrometer.

2.3.6 Click chemistry

The purified Z-γ-propargyl-Gln-Gly product served as the alkyne substrate for the azide-alkyne copper-catalyzed Huisgen cycloaddition. The reaction was performed using benzyl azide, according to the general procedure described by Himo and colleagues [62]. Briefly, a 0.25 M solution of the starting materials was made in 1 mL (1:1) H_2O/*t*-butanol containing 25 mM sodium ascorbate and 2.5 mM copper(II) sulfate pentahydrate [62]. The reaction was monitored by LC-MS over 24 h as described above.

2.4 Results

2.4.1 Establishment of the reaction conditions

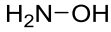
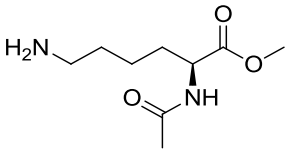
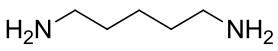
To gain greater insight into the reactivity of MTG with non-native acyl-acceptor substrates, screening was performed by reacting the protected dipeptide acyl-donor substrate Z-Gln-Gly with a variety of amines as acyl-acceptor substrates. Reaction pH was adjusted in the range of pH 6-9 to aid compound dissolution while maintaining nucleophilicity of the reactive amine; MTG has been reported to exhibit > 80% activity following a 10-min incubation between pH 5 and 10 [41]. MTG has been reported to be optimally active at 70°C, but it loses activity over 10 min incubation at that temperature [41], so reactions were run at 37°C, since MTG has been reported to maintain full activity upon 10 min incubation at 40°C [41]. Product formation was reported as consumption of Z-Gln-Gly donor substrate, as it could be monitored irrespective of the acceptor substrate tested.

We verified MTG activity using the standard assay with hydroxylamine (1) (Table 2.1) as the acceptor substrate [60], for up to 72 h, as representing a maximal acceptable time frame for most biocatalytic processes. After 72 h of incubation at pH 6 and 37°C, 60% of MTG activity remained, confirming its stability. Hydroxylamine (1) is chemically and sterically distinct from the native peptide- or protein-bound lysine substrates of MTG. Using hydroxylamine as a reference for small molecule acceptor substrate specificity, the donor substrate was consumed to > 99% after 24 h. Lysine, the native acceptor residue of MTG, was tested under the form of di-protected N^{α}-acetyl-L-lysine-methyl ester (2), to verify reaction of the ε-amine: it reacted to > 98% after 24 h. Cadaverine (3) is an analogue of lysine, as it mimics the butylamine side-chain. Substituted cadaverines have long been used as substrates of MTG [32,45,46,49,50,54,63–65]. Its reactivity under the conditions used herein was as high as that of lysine and hydroxylamine (> 99% after 24 h). Cadaverine (3) is a symmetric diamine and therefore has twice the amine concentration of simple amines. This may favor product formation by increasing the initial concentration of reactive amine. In addition, mono-acylated amine may serve as a substrate for a second acylation event. The formation of the diamide product was qualitatively investigated, where the donor concentration was increased to 60 mM and the acceptor concentration reduced to 30 mM. Mass spectrometry confirmed formation of the disubstituted product along with monosubstituted product (data not shown).

The low rate of MTG-mediated hydrolysis relative to mammalian transglutaminases [41] is one of its advantageous properties. Formation of the hydrolysis product Z-Glu-Gly was monitored during all reactions; little or no hydrolysis was observed, even after 72 h (Supporting Information, Table A1.1). Hydrolysis occurred almost exclusively under conditions of low acceptor substrate reactivity, particularly at pHs between 6 and 8 where MTG shows maximal

activity [41]. Under those conditions, the donor substrate concentration remained high, increasing its susceptibility to hydrolysis.

Table 2.1: Reactivity of reference acyl-acceptor compounds

Acceptor compound	Structure	Donor consumed in 72 h (%)
1		> 99 [a]
2		> 99 [a]
3		> 99 [a]

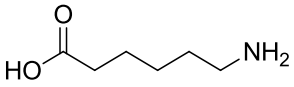
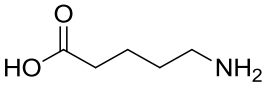
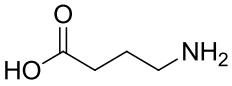
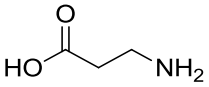
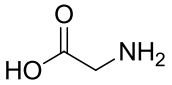
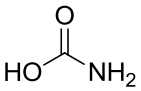
[a] > 98% reacted in 24 h

2.4.2 Amino acids with varying intervening chain length

We verified the reactivity of acceptor substrates with various substituents in close proximity to the reactive amine. We first determined the minimal distance allowed between the reactive amine and a negatively charged substituent. Because of the presence of negatively-charged amino acids in the vicinity of the reactive cysteine of MTG (Figure 2.1), a negative charge such as a carboxylate may be detrimental to reactivity. Ohtsuka and colleagues previously screened amino acids with alkyl chains of one to seven carbons separating the amino group from the carboxylic acid [44]. We partly confirmed their observations, testing alkyl spacers between zero and five carbons in length (Table 2.2). As previously noted [44], 6-aminocaproic acid (4) was the acceptor with the highest reactivity in that series (71% donor consumption after 72 h). Reactivity decreased with decreasing acyl-chain length (Table 2.2; Figure 2.2). 5-Aminovaleric acid (5), γ-aminobutyric acid (6), β-alanine (7) and glycine (8) yielded donor consumptions of 39%, 12%, 4.8% and 8.4%, respectively; this is the first report of reactivity with the short-chain amino acids

(6-8). The higher reactivity of glycine (8) relative to β-alanine (7) may be due to reacting glycine at pH 7.7 and β-alanine at pH 9; under the latter conditions enzyme activity is reduced by 50% [41]. Carbamic acid (9) yielded no detectable product or donor consumption.

Table 2.2: Reactivity of amino acids with varying chain length as acyl-acceptor substrates

Acceptor compound	Structure	Donor consumed in 72 h (%)
4		70.6 ± 6.4
5		38.5 ± 3.2
6		11.8 ± 6.7
7		4.8 ± 1.3
8		8.4 ± 2.4
9		nd [a]

[a] nd: not detected (< 2% product)

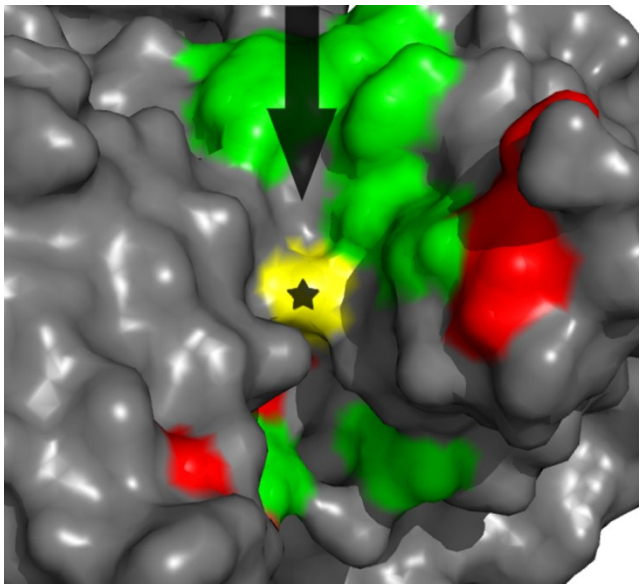


Figure 2.1: The active site of MTG (1IU4) (Kashiwagi et al., 2002) in gray surface rendering, with the reactive cysteine indicated in yellow with a black asterisk, and aromatic and acidic residues surrounding the active site indicated in green and red, respectively.

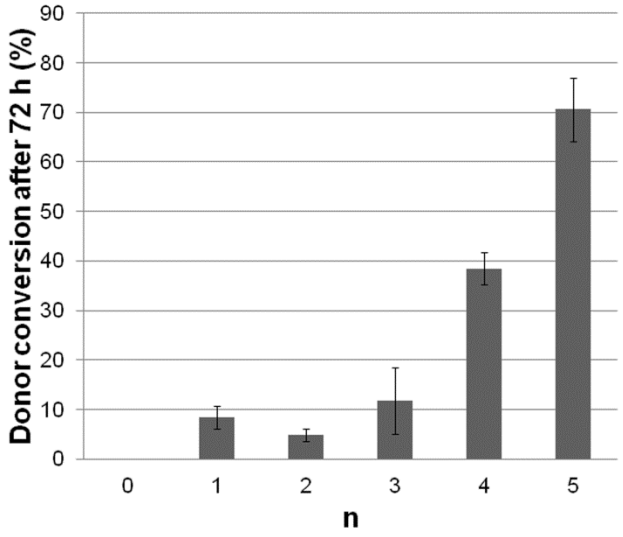


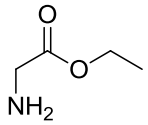
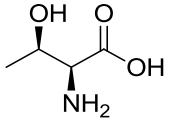
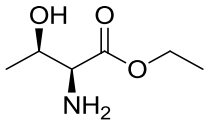
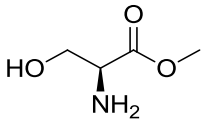
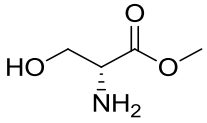
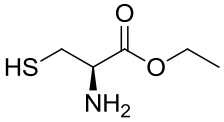
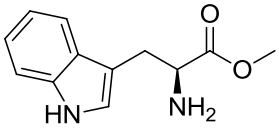
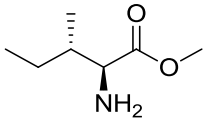
Figure 2.2: Reactivity of alkyl-based amino acids as acyl-acceptor substrates over 72 h. Reactivity for amino acids having the structure H2N-(CH2)n-CO2H with varying spacer length n = 0 to 5. Error bars indicate standard deviation from the mean from three experiments.

2.4.3 α-Amino acid acceptors

We probed the reactivity of MTG toward chemically diverse α-amino acids (Table 2.3). Because a negatively charged carboxylate near the reactive amine was unfavorable for reactivity (Table 2.2), esterified amino acids were tested, as previously demonstrated [44]. Esterification effectively increased reactivity, as shown by the faster reaction of glycine ethyl ester (10) (Table 2.3: > 99 % in 4 h) relative to its parent glycine (8) (Table 2.2: 8.4% in 72 h). The beneficial effect of esterification is also evident on comparison of unreactive L-threonine (11) with L-threonine ethyl ester (12), which yielded 4.1% product in 72 h. The lower reactivity of L-threonine ethyl ester (12) relative to glycine ethyl ester (10) suggests that the steric hindrance due to β-branching is incompatible with high reactivity, as observed with the unreactive valine ethyl ester [44].

The reactivity of the unhindered L-serine methyl ester (13), D-serine methyl ester (14) and L-cysteine ethyl ester (15) over 72 h was 5.8%, 3.0% and 2.7%, respectively. These similar values are much lower than the > 99% in 4 h observed for glycine ethyl ester. The bulky tryptophan methyl ester (16) also yielded a reactivity of 4.5% over 72 h, similar to the unhindered, polar amino acids. However, the non-polar, β-branched L-isoleucine methyl ester (17) did not form any detectable product. This result indicates that the bulky bicyclic indole, three carbons away from the reactive amine, is less detrimental to acceptor reactivity than is β-branching, two carbons away from the amine.

Table 2.3: Reactivity of α-amino acids as acyl-acceptor substrates.

Acceptor compounds	Structure	Donor consumed in 72 h (%)
10		> 99 [a]
11		nd [b]
12		4.1 ± 9.5
13		5.8 ± 1.4
14		3.0 ± 1.9
15		2.7 ± 1.7
16		4.5 ± 1.1
17		nd [b]

[a] > 99% reacted in 4 h

[b] nd: not detected (< 2% product)

2.4.4 Aromatic acceptors

In addition to negatively charged residues, the active site of MTG also contains a number of nearby aromatic residues (Figure 2.1) [36]. Hypothesizing that they may contribute to substrate binding, we tested simple aromatic amines. Benzylamine (18) and aniline (19) showed reactivity of 94% and 19% respective substrate consumption over 72 h (Table 2.4). In contrast, the non-aromatic cyclohexylmethylamine (20) and cyclohexylamine (21) were approximately 50% as reactive as their respective aromatic analogues, and were of reactivity comparable to other compounds containing an amino group attached to a secondary carbon (Tables 2.2 and 2.4).

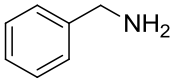
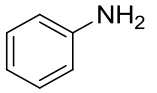
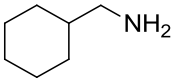
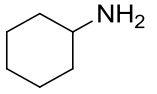
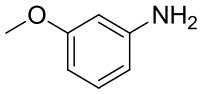
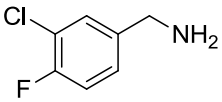
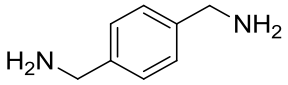
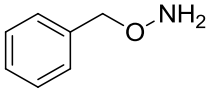
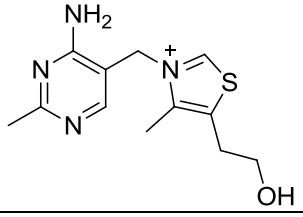
Conjugated substituents on the aromatic ring yielded a reduction in reactivity. *m*-Anisidine (22) reacted slowly, with 11% conversion, a reduction of approximately 40% relative to its unsubstituted analogue aniline. Similarly, the dihalogenated 3-chloro-4-fluorobenzylamine (23) showed 23% reactivity, a 75% reduction relative to its unsubstituted benzylamine analogue. Despite having a doubled amine concentration like cadaverine (3) (Table 2.1), the diamine *p*-xylenediamine (24) yielded a reactivity of 63%, which is a 30% reduction compared to the monosubstituted benzylamine (18). As for the reaction of cadaverine, both amines of *p*-xylenediamine (24) were reactive, yielding both mono and diamide products (data not shown).

We attempted to combine the beneficial effect of the benzyl ring with that of the highly reactive hydroxylamine (1) by testing the reactivity of *O*-benzylhydroxylamine (25). The longer linker between the benzene and the amine combined with the reactivity of the hydroxylamine were expected to increase the reactivity of MTG relative to benzylamine (18) and aniline (19). Surprisingly, *O*-benzylhydroxylamine (25) yielded only 2.6% conversion. This may be due to the reduced polarity of the hydroxylamine group in *O*-benzylhydroxylamine, though it appears more

likely that the bulk of *O*-benzylhydroxylamine relative to hydroxylamine (1) – the smallest compound tested – may be the cause of reduced reactivity.

Finally, to explore the extent to which aromaticity compensates for steric hindrance, we tested the reactivity of thiamine (26), also known as vitamin B_1. Its primary amine belongs to an aminopyrimidine ring, which is further attached to a thiazole ring. Thiamine was unreactive; its greater bulk or higher polarity than benzylamine (18) may hinder binding to the active site.

Table 2.4: Reactivity of aromatic amino acids as acyl-acceptor substrates.

Acceptor compound	Structure	Donor consumed in 72 h (%)
18		93.6 ± 1.3
19		18.9 ± 4.4
20		46.3 ± 1.6
21		3.1 ± 1.6
22		11.3 ± 1.4
23		23.1 ± 5.7
24		62.7 ± 2.4
25		2.6 ± 5.6
26		nd [a]

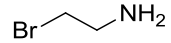
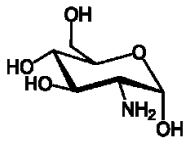
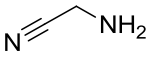
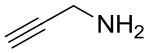
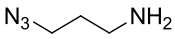
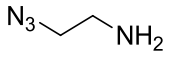
[a] nd: not detected (< 2% product)

2.4.5 Electron-rich amines

Halogenated moieties provide a starting point for synthetic modification. Furthermore, halogens are increasingly found in pharmaceutical compounds to increase lipophilicity and thus improve drug uptake [66]. Above, we showed that 3-chloro-4-fluorobenzylamine (23) was reactive as an acyl-acceptor substrate (Table 2.4), illustrating the utility of MTG for installing modifiable moieties on the donor peptide. The smaller 2-bromoethylamine (27) did not yield detectable product (Table 2.5), potentially due to the electron density of bromine, or its bulk, near the amine. While Ohtsuka and colleagues [44] previously reported the MTG-catalyzed addition of pentylamine or hexylamine-linked sugar moieties to proteins, glucosamine (28) did not yield any product. This may result from the amine being adjacent to a secondary carbon and/or to the hydrophilicity in the immediate vicinity of the amine. Nitriles are a further functional group of high value in synthesis of APIs, as their high permeability and bioavailability, combined with low metabolite rates in the cell, makes them a favorable substituent [67,68]. In light of the above results, we reasoned that its lack of bulky substituents or negative charge should make aminoacetonitrile (29) compatible with reactivity. Indeed, it reacted to completion within 24 h (Table 2.5), which makes it one of the most reactive compounds identified, as previously reported for tissue transglutaminase [69]. Having demonstrated the reactivity of MTG toward aminoacetonitrile, we probed its reactivity toward the similarly small, π electron-rich propargylamine (30), which yielded 90% reaction after 72 h (Table 2.5). This is consistent with the high reactivity observed for propargylamine with tissue transglutaminase [70]. The resulting reaction product, Z-*N*□-propargyl-Gln-Gly, can subsequently serve as the alkyne in a click reaction with a suitable azido-substituted compound [70]. To increase the flexibility imparted by MTG in preparing modified peptides for click reactions, we tested its reactivity toward two simple amino-azides: propylamine azide (31) and ethylamine azide (32), which reacted to an

extent of 99% and 98%, respectively, over 72 h. These results demonstrate that MTG tolerates the presence of both alkynes and azides in close proximity to the reactive amine of the acceptor substrate.

Table 2.5: Reactivity of halogenated amines and labeling agents as acyl-acceptor substrates.

Acceptor compound	Structure	Donor consumed in 72 h (%)
27		nd[a]
28		nd[a]
29		> 99 [b]
30		90.1 ± 0.9
31		98.5 ± 0.5
32		97.7 ± 0.5

[a] nd: not detected (< 2% product)

[b] > 99% reacted in 24 h

2.4.6 Acceptors other than primary amines

We investigated limits of MTG's acyl-acceptor substrate specificity by testing compounds other than primary amines as potential acyl-acceptor substrates (Annex 1, Table A1.2). Neither of the secondary amines *N*-ethylmethylamine (33) or sarcosine (34) yielded product. Non-primary amines are poorer nucleophiles than primary amines; their increased steric hindrance and the carboxylate of sarcosine (34) may be a further deterrent to reactivity. We also probed the reactivity of an unhindered alcohol and a thiol as alternative nucleophiles: neither butanol (35) nor butanethiol (36) yielded product. These results indicate that weaker

nucleophiles than amines cannot serve as acceptor substrates for MTG and that an unhindered primary amine is required as the reactive species.

2.5 Discussion

The reaction conditions investigated indicate that MTG is a good biocatalyst candidate for running lengthy reactions (72 h). To gain insight into the tolerance of MTG to chemically and sterically varied acceptor substrates, we tested its reactivity using acceptor substrates with substituents proximal to the reactive amine, using the donor substrate Z-Gln-Gly. Donor substrate binding is the first event in this ordered reaction, and the bound donor substrate is therefore an integral part of the active site with respect to acceptor binding.

2.5.1 Very short chain amino acids can serve as acceptor substrates

In investigating the minimal distance allowed between the reactive amine and a negatively charged substituent (compounds 4-9), we detected conversion not only for the long chain amino acids 4 and 5 but, contrary to the report of Ohtsuka [44], also for the amino acids 6-8 having only one to three intervening carbons. This may be due to the higher enzyme concentrations and increased reaction time used here. Thus, even a single intervening carbon between the carboxylate and amino groups is sufficient to allow some reactivity with MTG as an acceptor substrate. While the reactivity with amino acids 4-8 was modest, its optimization would be great utility for general synthetic biocatalytic applications.

Decreasing reactivity with decreasing spacer length (Figure 2.2) is consistent with the reactivity of MTG with other negatively-charged acceptor substrates such as amino alkyl sulfonates and phosphates [44,71,72], and likely results from electrostatic repulsion with negatively charged residues in the active site (Figure 2.1) [36]. Repulsion may be augmented by the bound

donor substrate Z-Gln-Gly, which harbours a negatively-charged carboxylate. However, in the absence of structural confirmation of the mode of substrate binding to MTG, this remains speculative.

2.5.2 Side-chain volume reduces reactivity of α-amino acid acceptor substrates

α-Amino acids and their derivatives are heavily used in the synthesis of high value APIs because they provide a readily accessible pool of chemically varied chiral starting materials. The comparison of glycine and L-threonine with their respective esters illustrates that esterification effectively increased reactivity, confirming previous observations [44]. The reaction rate with glycine ethyl ester (10) (Table 2.3: > 99% in 4 h) even surpassed that of N^{α}-acetyl-L-lysine methyl ester (2) (Table 2.1: 80% in 4 h), which served as a mimic of the natural, protein-bound lysine substrate, consistent with previous reports [44]. This observation is consistent with use of an *N*-terminal Gly or a poly-Gly tag as an acceptor substrate in protein-protein conjugations [71], and raises the question as to whether *N*-terminal Gly may also be a native acceptor substrate of MTG. While MTG natively catalyses protein side-chain cross-linking, a potential, high-value application is the biocatalysis of peptide bonds between the α-amino and α-carboxyl groups of natural or non-natural amino acids, as we have demonstrated using mutants of tissue transglutaminase [73]. Despite esterification, the reactivity of the α-amino group of a set of chemically diverse α-amino acids was at least 18-fold lower than that of glycine ethyl ester. Thus, amines linked to a primary carbon are more reactive than amines linked to secondary carbons. Considering that alanine ethyl ester showed only a 3-fold decrease in reactivity [44], this illustrates the negative impact of increasing side-chain volume. Nonetheless, the reactivity of the bulky tryptophan methyl ester (16) was comparable to that of the serine (13, 14) and cysteine (15) ethyl esters. Indeed, a bulky substituent farther removed from the reactive amine was better

tolerated than β-branching two carbons away from the amine, as illustrated by the unreactive L-isoleucine methyl ester (17). Nonetheless, the reactivity of the β-branched L-threonine ethyl ester (12) (4.1% in 72 h) indicates that β-branching can be tolerated to some extent. Overall, its robustness and reactivity toward a number of amino acids make MTG a good candidate for further development as a peptide bond catalyst, although the stringent acyl-donor substrate specificity of MTG currently precludes reaction ofan amino acid α-carboxyl group. We note that no significant stereoselectivity was observed upon reaction of the L or D isomers of serine ethyl ester (13 and 14); MTG may require more voluminous substituents for stereoselectivity to be observed.

2.5.3 Unsubstituted aromatic amines are more reactive acyl-acceptor substrates

In addition to a number of negatively charged residues in and around the active site, the MTG active site also holds a number of aromatic residues, several of which have been shown to be important for donor substrate binding (Figure 2.1)(Tagami et al., 2009; [36]. We hypothesized that they may also aid binding of acceptor substrates, *via* hydrophobic interactions and/or π-stacking. The higher reactivity of benzylamine (18) than aniline (19) (Table 2.4) is in agreement with the trends we observed above, where an amino group attached to a primary carbon yielded higher reactivity than an amino group attached to a more highly substituted carbon – a secondary carbon in the case of the α-amino acids investigated, and an aromatic ring carbon in the case of aniline (19). It is also consistent with the higher nucleophilicity of benzylamine (18) relative to aniline (19). The higher reactivity of aniline relative to α-amino acids, and of benzylamine and aniline relative to their non-aromatic analogues cyclohexylmethylamine (20) and cyclohexylamine (21), appears to be related to aromaticity rather than to the cyclic structure.

Increased reactivity may thus result from the formation of favorable π-π interactions between the acceptor substrate and MTG and/or the carboxybenzyl group of the MTG-bound donor substrate, from the planar geometry of the benzyl ring, as opposed to the puckered cyclohexane ring, or to a combination of those two factors. Ring substituents on substrates 22 to 24 reduced reactivity, again indicating the detrimental effect of increased steric bulk. Despite reduced reactivity of *p*-xylenediamine (24) relative to benzylamine (18), we demonstrated the feasibility of using MTG for homodisubstitution of the acyl-acceptor substrates *p*-xylenediamine (24) and cadaverine (3), where the resulting diamine has either a rigid (24) or a flexible (3) linker.

2.5.4 MTG is reactive toward small, π electron-rich amines

We have observed reactivity with acyl-acceptor compounds that provide avenues for further synthesis. High reactivity was confirmed for the small, unbranched, π electron-rich amines aminoacetonitrile (29), propargylamine (30) as well as for propylamine azide (31) and ethylamine azide (32) (Table 2.5), opening the door to click chemistry. The Cu(I)-catalyzed azide/alkyne [3+2]-cycloaddition reactions are established techniques for specific labeling of proteins *in vivo*. These selective reactions rapidly proceed with high accuracy under mild conditions [74] and have been applied to specific labeling of proteins with fluorescent labels and sugars, as we demonstrated with tissue transglutaminase [70]. Performing click chemistry with proteins generally requires non-canonical amino acids containing either an azido or alkyne moiety [75], making the method costly and poorly accessible. The specific addition of the azido or alkyne moiety to the target protein with a transglutaminase can constitute a significant advantage [70].

The reactivity of MTG with small amino-azides is of particular interest because many commercially-available labels are alkyne-substituted. We note that, following 4 h incubation of

MTG with 0.1 M of either azide (31) or (32), no further conversion was observed (data not shown), suggesting rapid initial reaction followed by inactivation of the enzyme. Indeed, azides can serve as inhibitors of cysteine proteases (Le et al., 2006). To demonstrate the utility of an MTG-mediated alkyne-modified peptide in a copper catalyzed cycloaddition reaction, we reacted the purified Z-*N*□-propargyl-Gln-Gly with benzylazide under previously reported conditions [62]. Product formation was confirmed by MS; no degradation of the starting materials was observed. MTG thus demonstrates flexibility towards substitution of peptides for further modification by click chemistry. Furthermore, we demonstrate that MTG tolerates the π electron density of the nitrile, alkyne and azido groups in proximity to the reactive amine.

In conclusion, this work significantly increases our knowledge of the acyl-acceptor specificity of MTG. Our results expand the range of acyl-acceptor substrates known to be accepted – or not accepted – by MTG, for formation of a secondary amide bond with a peptide-bound glutamine. We have shown, for the first time, that even very short-chain alkyl-based amino acids such as glycine can serve as acceptor substrates, with reactivity increasing with the chain length separating the amine and the carboxyl groups. We observed reactivity with the α-amino group of chemically diverse α-amino acids. Additional steric hindrance in the immediate vicinity of the amine was detrimental to reactivity. Nonetheless, the bulky, aromatic Trp – but not Ile – showed reactivity; this suggests that steric bulk farther removed from the reactive amine is better tolerated. Extending the search to non-natural compounds, an aromatic ring was beneficial for reactivity of the acceptor substrate. Ring substitution reduced reactivity, apparently as a result of steric hindrance. Finally, only primary amines were reactive, and bonding of the amine to a less hindered carbon increased reactivity.

While specific molecules were shown to be reactive, they likely indicate a trend where other, similar compounds will exhibit comparable reactivity. Importantly, very small amines carrying either the electron-rich nitrile, azide or the alkyne groups required for click chemistry were highly reactive as acceptor substrates, facilitating the synthesis of minimally modified, 'clickable' peptides. These results demonstrate that MTG is tolerant to a variety of chemically varied natural and non-natural acceptor substrates. We expect that the reactivity of the acyl-acceptor substrates observed here with the Z-Gln-Gly dipeptide as the acyl-donor substrate is predictive of their relative reactivity toward reactive, protein-bound glutamines, as previously demonstrated with mammalian transglutaminase [70] and MTG (Lee et al., 2013). These results broaden the scope for modification of Gln-containing peptides and the use of MTG as a biocatalyst.

2.6 Acknowledgments

This work was supported by funding from the Natural Science and Engineering Research Council of Canada and CGCC, the Center in Green Chemistry and Catalysis. We thank Fabrice Galaud and Pierre Lavallée, in the Combinatorial Chemistry Laboratory at the Université de Montréal, for their generous technical help, Alexandra Furtos in the Regional Mass Spectrometry Center at Université de Montréal for her assistance, Maria Stoïca and Dr. Christophe Pardin for the synthesis of azide compounds and acknowledge Klavdja Annabel Fignolé for assisting in experimental work and Natalie Rachel for commenting and revising the manuscript.

2.7 References (As listed in publication)

Abe H, Goto M, Kamiya N (2010) Enzymatic single-step preparation of multifunctional proteins. Chem. Commun. 46:7160–2.

Best MD (2009) Click chemistry and bioorthogonal reactions: unprecedented selectivity in the labeling of biological molecules. Biochemistry 48:6571–84.

Clouthier CM, Pelletier JN (2012) Expanding the organic toolbox: a guide to integrating biocatalysis in synthesis. Chem. Soc. Rev. 41:1585–605.

Fichert T, Yazdanian M, Proudfoot JR (2003) A structure–permeability study of small drug-like molecules. Bioorg. Med. Chem. Lett. 13:719–722.

Fleming FF, Yao L, Ravikumar PC, Funk L, Shook BC (2010) Nitrile-containing pharmaceuticals: efficacious roles of the nitrile pharmacophore. J. Med. Chem. 53:7902–17.

Folk JE, Cole PW (1966) Transglutaminase: Mechanistic features of the active site as determined by kinetic and inhibitor studies. Biochim. Biophys. Acta, Enzymol. Biol. Oxid. 122:244–264.

Gnaccarini C, Ben-Tahar W, Mulani A, Roy I, Lubell WD, Pelletier JN, Keillor JW (2012) Site-specific protein propargylation using tissue transglutaminase. Org. Biomol. Chem. 10:5258–65.

Griffin M, Casadio R, Bergamini CM (2002) Transglutaminases: Nature's biological glues. Biochem. J 368:377–396.

Hernandes MZ, Cavalcanti SMT, Moreira DRM, De Azevedo Junior WF, Leite ACL (2010) Halogen atoms in the modern medicinal chemistry: hints for the drug design. Current Drug Targets 11:303–14.

Himo F, Lovell T, Hilgraf R, Rostovtsev V V, Noodleman L, Sharpless KB, Fokin V V (2005) Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates. J. Am. Chem. Soc. 127:210–6.

Hu B-H, Messersmith PB (2003) Rational design of transglutaminase substrate peptides for rapid enzymatic formation of hydrogels. J. Am. Chem. Soc.125:14298–9.

Inverarity IA, Hulme AN (2007) Marked small molecule libraries: a truncated approach to molecular probe design. Org. Biomol. Chem. 5:636–43.

Jeger S, Zimmermann K, Blanc A, Grunberg J, Honer M, Hunziker P, Struthers H, Schibli R (2010) Site Specific and Stoichiometric Modification of Antibodies by Bacterial Transglutaminase. Angew. Chem. Int. Ed. 49:1–4.

Karickhoff SW, Carreira LA, Hilal SH (2011) Sparc On-Line Calculator.

Kashiwagi T, Yokoyama K-I, Ishikawa K, Ono K, Ejima D, Matsui H, Suzuki E (2002) Crystal structure of microbial transglutaminase from Streptoverticillium mobaraense. J. Biol. Chem. 277:44252–60.

Keillor JW, Chica RA, Chabot N, Vinci V, Pardin C, Fortin E, Gillet SMFG, Nakano Y, Kaartinen MT, Pelletier JN, Lubell WD (2008) The bioorganic chemistry of transglutaminase — from mechanism to inhibition and engineering. Can. J. Chem. 86:271–276.

Kitaoka M, Tsuruda Y, Tanaka Y, Goto M, Mitsumori M, Hayashi K, Hiraishi Y, Miyawaki K, Noji S, Kamiya N (2011) Transglutaminase-mediated synthesis of a DNA-(enzyme)n probe for highly sensitive DNA detection. Chemistry 17:5387–92.

Kolb HC, Finn MG, Sharpless KB (2001) Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew. Chem. Int. Ed. 40:2004–2021.

Le GT, Abbenante G, Madala PK, Hoang HN, Fairlie DP (2006) Organic azide inhibitors of cysteine proteases. J. Am. Chem. Soc.128:12396–7.

Leblanc A, Gravel C, Labelle J, Keillor JW (2001) Kinetic studies of guinea pig liver transglutaminase reveal a general-base-catalyzed deacylation mechanism. Biochemistry 40:8335–8342.

Lee J-H, Song C, Kim D-H, Park I-H, Lee S-G, Lee Y-S, Kim B-G (2013) Glutamine (Q)-peptide screening for transglutaminase reaction using mRNA display. Biotechnol. Bioeng. 110:353–62.

Marx CK, Hertel TC, Pietzsch M (2007) Soluble expression of a pro-transglutaminase from Streptomyces mobaraensis in Escherichia coli. Enzyme Microb. Technol. 40:1543–1550.

Mero A, Schiavon M, Veronese FM, Pasut G (2011) A new method to increase selectivity of transglutaminase mediated PEGylation of salmon calcitonin and human growth hormone. J. Controlled Release 154:27–34.

Mindt TL, Jungi V, Wyss S, Friedli A, Pla G, Novak-Hofer I, Grünberg J, Schibli R (2008) Modification of different IgG1 antibodies via glutamine and lysine using bacterial and human tissue transglutaminase. Bioconjugate Chem. 19:271–8.

Ohtsuka T, Ota M, Nio N, Motoki M (2000) Comparison of substrate specificities of transglutaminases using synthetic peptides as acyl donors. Biosci. Biotechnol., Biochem. 64:2608–2613.

Ohtsuka T, Sawa A, Kawabata R, Nio N, Motoki M (2000) Substrate specificities of microbial transglutaminase for primary amines. J. Agric. Food. Chem. 48:6230–3.

Park KD, Morieux P, Salomé C, Cotten SW, Reamtong O, Eyers C, Gaskell SJ, Stables JP, Liu R, Kohn H (2009) Lacosamide isothiocyanate-based agents: novel agents to target and identify lacosamide receptors. J. Med. Chem. 52:6897–911.

Pasternack R, Dorsch S, Otterbach JT, Robenek IR, Wolf S, Fuchsbauer HL (1998) Bacterial pro-transglutaminase from Streptoverticillium mobaraense--purification, characterisation and sequence of the zymogen. Eur. J. Biochem. / FEBS 257:570–6.

Pasternack R, Laurent HP, Rüth T, Kaiser A, Schön N, Fuchsbauer HL (1997) A fluorescent substrate of transglutaminase for detection and characterization of glutamine acceptor compounds. Anal. Biochem. 249:54–60.

Pattabiraman VR, Bode JW (2011) Rethinking amide bond synthesis. Nature 480:471–479.

Porta R, Di Pierro P, Sorrentino A, Mariniello L (2011) Promising Perspectives for Transglutaminase In “Bioplastics” Production. J. Biotechnol. Biomaterial 01:1–4.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.

Schmid A, Dordick JS, Hauer B, Kiener A, Wubbolts M, Witholt B (2001) Industrial biocatalysis today and tomorrow. Nature 409:258–68.

Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. Protein Expression Purif. 41:207–234.

Sugimura Y, Yokoyama K, Nio N, Maki M, Hitomi K (2008) Identification of preferred substrate sequences of microbial transglutaminase from Streptomyces mobaraensis using a phage-displayed peptide library. Arch. Biochem. Biophys. 477:379–83.

Tagami U, Shimba N, Nakamura M, Yokoyama K-I, Suzuki E-I, Hirokawa T (2009) Substrate specificity of microbial transglutaminase as revealed by three-dimensional docking simulation and mutagenesis. Protein Eng. Des. Sel. 22:747–52.

Tanaka T, Kamiya N, Nagamune T (2004) Peptidyl linkers for protein heterodimerization catalyzed by microbial transglutaminase. Bioconjugate Chem. 15:491–497.

Tanaka T, Kamiya N, Nagamune T (2005) N-terminal glycine-specific protein conjugation catalyzed by microbial transglutaminase. FEBS lett. 579:2092–6.

Tominaga J, Kemori Y, Tanaka Y, Maruyama T, Kamiya N, Goto M (2007) An enzymatic method for site-specific labeling of recombinant proteins with oligonucleotides. Chem. Commun. 401–3.

Umezawa Y, Ohtsuka T, Yokoyama K (2002) Comparison of Enzymatic Properties of Microbial Transglutaminase from Streptomyces sp. Food Sci. Technol. Res. 8:113–118.

Valeur E, Bradley M (2009) Amide bond formation: beyond the myth of coupling reagents. Chemical Society Reviews 38:606–31.

Yokoyama K, Nio N, Kikuchi Y (2004) Properties and applications of microbial transglutaminase. Appl. Microbiol. .Biotechnol. 64:447–454.

Zhu Y, Tramper J (2008) Novel applications for microbial transglutaminase beyond food processing. Trends in Biotechnol. 26:559–65.

Chapter 3
Predictability of novel acceptor substrate reactivity

3.1. Introduction

This chapter presents a more detailed analysis of certain points presented in Chapter 2, and puts them in perspective with the overall goal of this thesis: using MTG as a general amide bond catalyst for the advancement of green chemistry. The exploration of substrate specificity is important because it provides an insight into the applications where the enzyme may be used, on novel acceptor substrates. Only one paper has been previously published on the acceptor substrate specificity of MTG. Because that work was of a narrower scope, the present work may therefore aid in steering the future applications of the enzyme.

3.2. Discussion

Chapter 2 had two main focuses: investigating both the reactivity of MTG toward application-driven chemicals and the use of a variety of functional groups to understand the chemical nature of MTG's acceptor specificity. To put this work into perspective and highlight the importance of these findings as well as the available literature, we have outlined four rules that govern acceptor substrate reactivity (Table 3.1). Only one other study[44] has previously investigated small non-natural amines similar to our study, although a number of articles have been published using larger substrates.

Throughout Chapter 2 we saw that unhindered primary amines have the highest reactivity as acceptor substrates for MTG. In literature, lysine has been proposed to be the natural substrate of the enzyme. Where the side-chain amine is reacted, this is an amine bound to a long unhindered alkyl chain. We also observed high reactivity with long, unhindered alkyl amines such as cadaverine and α-amino-protected lysine (Table 2.1), or the small unhindered substrate hydroxylamine (Table 2.1). In a synthetic context, this observation has led to the frequent use of the pentylamine cadaverine as a linker to provide a free unhindered amine of high reactivity that can be bound to a larger group for use as a tag[32,45,46,49,50,54,63–65]. In our study we also observed that additional bulk reduced reactivity. Examples of this are found in Table 2.3 where a lower reactivity is observed with β-branched amino acids compared to linear side-chains. The same trend is found when the bulk is further from the amine, for example with the aromatic compounds in Table 2.4, where any substitution on the aryl ring leads to a reduced reactivity. We therefore have strong evidence confirming that unhindered primary amines are a requirement for the enzyme in order to yield high reactivity.

Our study continued by investigating the chemical nature of the active site. We found that the active site contained both several negatively charged and aromatic residues (Figure 3.1). We began by investigating the impact of the negatively charged residues by reacting acceptor substrates containing a negative charge in close proximity to the amine (Table 2.2). This had previously been studied by Ohtsuka and colleagues[44] which found, as we did, that a negative charge close to the amine was detrimental for reactivity. Further, both Ohtsuka and we found that decreasing the spacer length between the amine and the charge reduced reactivity significantly (Table 2.2). We believe this originates from electrostatic repulsion between the charged residues in the active site and/or the donor compound, with the charged residue on the acceptor

compound. Because the donor compound used in both studies is the negatively charged compound Z-Gln-Gly, we cannot isolate this effect to the enzyme itself. As such, this rule of reactivity may be highly dependent on the donor substrate used.

Aromatic residues are also dominant in the active site, as seen in Figure 3.1. In Chapter 2 we investigated the reactivity of aromatic aryl amines, and compared their reactivity to similar saturated cyclohexylamines (Table 2.4). We found the aromatic acceptor compounds to have the highest reactivity of the two. We speculate that this may be due to a favourable aromatic π-π stacking, between the acceptor and the active site and/or the aromatic group on the donor substrate. This trend therefore, like the previous trend, may be dependent on the nature of the donor substrate. We also note that the planar structure of the aryl ring, compared to the puckered nature of the cyclohexanes may also relieve some steric strain, in accordance with rule 1, and thus also influence the elevated reactivity.

Finally an aspect that was only briefly aluded to in Chapter 2 is our fastest reacting substrate, glycine ethyl ester, which also provided the highest reaction yield in the Othsuka study[44]. This is a particularly interesting substrate, because it gives us an insight into the preferred chemical nature of the acceptor substrate for this enzyme. From previous work on transglutaminases, we know that the long aliphatic side chain of lysine is a suitable substrate for the enzyme. However the high reactivity of glycine ethyl ester provides an insight that reactivity may increase with substrates that allow for hydrogen bonding. The use of a *N*-terminal polyglycine as an acceptor substrate has previously been reported. Tanaka and colleagues reacted 1-, 3- and 5-glycine *N*-terminal tag on a green fluorescent protein with another protein donor substrate[71]. They found that a 3- or 5-polyglycine tag was more reactive than a single glycine tag[71]. This shows that aliphatic unhindered amines with the possibility of hydrogen bonding may serve as good

acceptor substrates for MTG. Ohtsuka's and our study indicate that the hydrogen bonding may indeed be increasing reactivity. This could either be due to an increased interaction with the active site or a better solubility in the reaction media.

This may indicate that the natural acceptor target for MTG may be *N*-terminal polyglycine, rather than lysine. It is not known what the role of MTG is in the cell. If MTG should have a more specialized role in the cell, it would depend on a specific recognition pattern. Because MTG has been found to be more promiscuous towards acceptor substrates[76] it is has often been speculated that any specificity may come from the donor peptide[39,45,47]. However the result above may indicate that the specificity may in part also originate from the acceptor substrate.

Table 3.1: Trends in reactivity of acceptor substrates.

	Trends for increased reactivity
1	Unhindered amine
2	No negative charge in close proximity
3	Presence of aromatic group
4	Hydrogen bonding

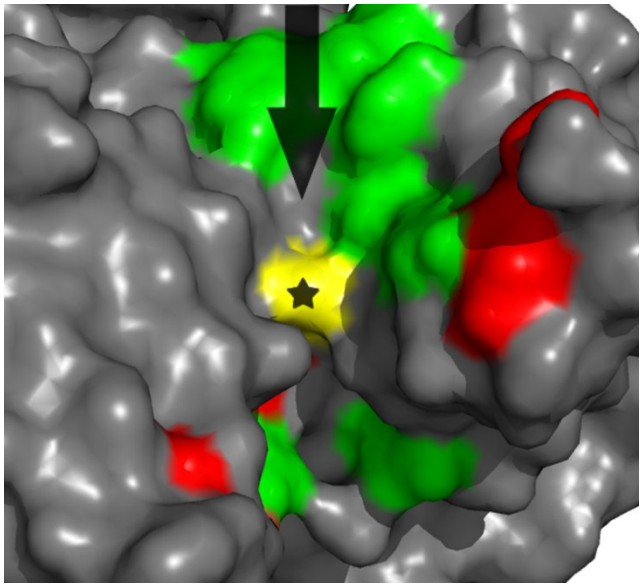


Figure 3.1: The active site of MTG (1IU4) (Kashiwagi et al., 2002) in gray surface rendering, with the reactive cysteine indicated in yellow with a black asterisk, and aromatic and acidic residues surrounding the active site indicated in green and red, respectively. The active site cleft is indicated by the arrow.

In this thesis our aim was to use MTG as a more general amide bond catalyst, therefore the results from Chapter 2 take on their full significance when they can be used in a predictive manner. The rules of reactivity in Table 3.1 give a good indication of the predicted reactivity of novel acceptor substrates. In order to evaluate the value of these rules we attempted to verify them with another common method for determining substrate binding: molecular docking. It must be noted that the methods are not directly comparable as our rules define reactivity, and the docking study defines binding.

Molecular docking has been used to dock the donor substrate Z-Gln-Gly was done with the software Molecular Operating Environment (MOE)[39]. This study showed that the donor substrate is held in place by a numer of aromatic and hydrophobic inteeraction in addition to hydrogen bonds. To the best of our knowledge, no docking study of MTG with acceptor substrates has been published. We therefore undertook a molecular docking study of acceptor substrates to further build on our experimental dataset. While the reaction mechanism requires that the acyl-

donor substrate (which includes a Gln) binds to the active site before the acyl-acceptor substrate, we verified whether the apoenzyme could still serve for identifying acceptor substrates. Currently no crystal structure of the enzyme with a donor substrate bound exists, therfore the apoenzyme remains the only available template with which to conduct docking. We compared the reactivity after 72h of 12 randomly chosen acceptors from Chapter 2, with the calculated scores from Molegro Virtual Docker (MolDock Score) in Table 3.2. The MolDock score does not have a physical significance, and as such only represents an arbitrary range where a lower score indicates a more favorable binding. Figure 3.2 illustrates the correlation between the docking scores and the reactivity; values are taken from Table 3.2. The data does not indicate any correlation between the docking scores and the experimental reactivity, under the conditions used here. We thus conclude that docking onto the apoenzyme with the parameters used here is not useful to predict the reactivity of potential acceptor substrates.

Table 3.2: Molegro Virtual Docker scores and reactivity of selected acceptor compounds.

Compound	reactivity (% after 72 h)	MolDock Score
Cadaverine	100	-57
Glucosamine	0	-51
Sarcosine	0	-45
L-Threonine	0	-53
Aminoacetonitrile	100	-34
γ-Aminobutyric acid	8.4	-38
6-Aminocaproic acid	71	-66
Benzylamine	94	-54
m-Anisidine	11	-57
Propargylamine	90	-36
γ-Aminobutyric acid	12	-55
N^{α}-Acyl-L-lysine-methylester	100	-81

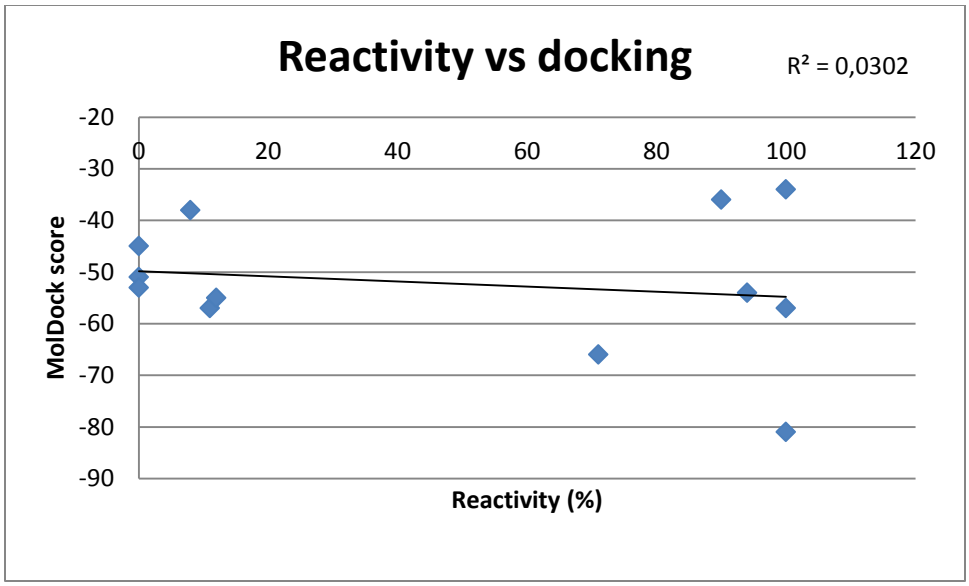


Figure 3.2: Correlation of Molegro Virtual Docker scores vs. reactivity of acceptors, as reported in Table 3.2.

3.4. Conclusion

To use MTG as an amide bond catalyst for synthetic substrates, we must understand its substrate specificity, and more importantly differenciate a reactive substrate from a non-reactive substrate. Here a number of beneficial properties are listed in Table 3.1 above. They may serve as a guide for future experiments with the native MTG, when it comes to choosing a good acceptor substrate. We compared selected experimental results from Chapter 2 with the molecular docking of the software Molegro Virtual Docker. We did not find a correlation between the experimental reactivity and the docking scores, and as such our general rules derived from experimental observations to guide reactivity maybe be more appropriate.

Chapter 4
Reactivity of small acyl-donor substrates

4.1 Introduction

In this body of work we aim to investigate the potential of MTG as a broader amide catalyst for non-natural substrates. To fully explore this, it is essential to explore both partners of the reaction, the acceptor and the donor. We have previously (Chapters 2 and 3) studied acceptor substrate specificity; here we begin to explore the donor substrate reactivity.

We wish to probe MTG's reactivity beyond glutamine-containing peptides or proteins, towards a more general amide-containing compound or even a compound where the electrophilic amide is replaced with an alternative electrophile. Several groups have investigated peptide-based substrates larger than Z-Gln-Gly. In addition to using MTG to label or modify glutamine-containing proteins, discussed in Chapter 1, the donor substrate has been explored in terms of longer peptides [45–47]. Several groups have also used functionalized Z-Gln-Gly analogues, such as ZQG linked with DNA [51,65] or biotinylated ZQG [64]. To the best of our knowledge, no donor substrate smaller than Z-Gln-Gly has been reported. Here we react smaller donor substrates to broaden the potential of MTG as a biocatalyst.

4.2 Materials and methods

4.2.1 Materials

MTG was prepared as described in Chapter 2. *N*-Z-L-Glu-(*γ*-*p*-nitrophenyl ester)-Gly was kindly synthesized in the laboratory of Jeffrey W. Keillor (University of Ottawa, Ottawa, Canada). Other chemicals used were purchased from the suppliers listed below. 4-Methoxybenzamide (97%) and glycine ethyl ester hydrochloride were purchased from Acros Organics (Waltham, USA). Glutamine, Z-glutamine, N^{α}-acetyl-L-lysine methyl ester hydrochloride, trypsin from bovine pancreas (10,000 BAEE U/mg) and hydroxylamine were purchased from Sigma-Aldrich (St Louis, USA). Carboxybenzyl-glutaminyl-glycine (Z-Gln-Gly) was from Alfa Aesar (Ward Hill, USA). Glutathione (reduced) was from Bioshop (Burlington, Canada). Formic acid 98% purity was from Fluka Analytical (St Louis, USA).

Synthesis of most of the donor compounds was necessary as they were not commercially available; synthetic methods and results are described in appendix 3.

4.2.2 Reaction conditions

All enzyme assays were as described in Chapter 2, with the exception of acceptor concentrations. Briefly, 72 mM N^{α}-acetyl-L-lysine-methyl ester or 94 mM glycine ethyl ester was reacted with 40 mM donor substrate in a 350 µL Eppendorf tube with 2 U/mL MTG, at 37 °C for 72 h. The reaction was monitored with HPLC-MS with the addition of an internal MS standard after the reaction was terminated.

4.3 Results & discussion

4.3.1 Reducing the size of Z-Gln-Gly

In accordance with experiments in previous chapters, the use of smaller donor substrates would broaden the potential applications of MTG as a biocatalyst. We began this work by expanding on previous findings from Chapter 2, where we observed that the presence of a negative charge was detrimental for reactivity of acceptors, and further that reactivity was recovered when the charge was neutralized by esterification. Therefore, we used free acids and esterified donor substrates in the following experiments. Six smaller donor substrates (Figure 4.1) were tested with two acceptor substrates N^{α}-acetyl-L-lysine-methyl ester and glycine ethyl ester. These two acceptor substrates were chosen because of their high reactivity, determined in Chapter 2. No product formation corresponding to the amide serving as the acyl donor was observed with any of the new donors: Gln (1), Gln(OMe) (2), Gln(OEt) (3), Z-Gln (4), Z-Gln(OMe) (5) or Z-Gln(OEt) (6). We conclude that under our conditions, the varied, smaller peptidic substrate analogues of Z-Gln-Gly did not provide detectable reactivity with the γ-carboxamide of the glutamine.

To verify if esterification was beneficial for donor reactivity, we methylated the original donor substrate Z-Gln-Gly used in previous chapters. Although the esterification was incomplete, we nonetheless reacted the mixed substrate Z-Gln-Gly and Z-Gln-Gly(OMe). The data (annex 3) suggest that reactivity is not significantly improved by esterification.

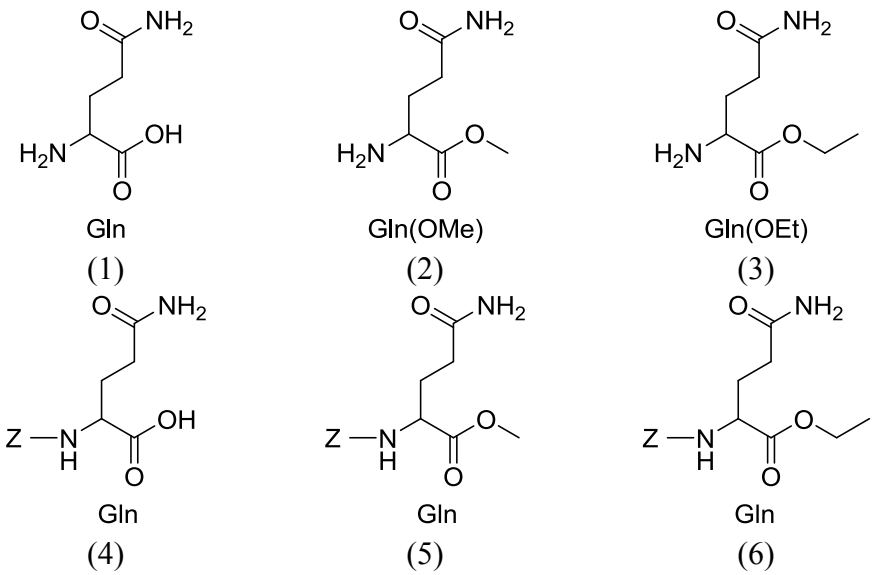


Figure 4.1: Structures of smaller Z-Gln-Gly analogue substrates.

4.3.2 Ester reactivity

To better explore the catalytic potential of MTG as an amide catalyst, it is clear that the enzyme must not only be reactive towards a broad substrate range, but it must be able to catalyze the bond formation from low cost, readily available functional groups. Esters serve as suitable substrates because, among other possibilities, this could facilitate the development of MTG as a peptide catalyst, and may prove advantageous relative to current atom-inefficient peptide synthesis methods. Specifically, the *C*-terminus of esterified amino acids would serve as acyl-donor substrates for reaction with the *N*-terminus of free amino acids (amines), resulting in the selective synthesis of peptides. A proposed method for MTG catalyzed peptide synthesis is outlined in Chapter 5.

During the reaction mentioned above, where we use donor substrates which contain both amide and ester functional groups, we observed on several occasions a small peak in the MS spectra which corresponds to the ester serving as the reactive acyl-donor substrate instead of the amide. Specifically this was observed with the compounds Z-Gln(OMe) (5) and Z-Gln-

Gly(OMe). No such peaks were found in the blank samples without the enzyme present. This has previously been observed with a mammalian transglutaminase[77]. This is interesting because as mentioned above, esters could serve as readily available, low cost substrates. To further explore this path we reacted MTG with the colometric activated ester substrate *N*-Z-L-Glu(*γ*-*p*-nitrophenylester)Gly (data shown in appendix 4). We observed that this compound functions as a donor substrate for MTG, albeit a poor one, and the competing hydrolysis reaction is prominent. Nonetheless, these observations indicate that MTG may be suitable as an amide bond catalyst using an ester donor substrate, although further development is required.

4.4 Conclusion

We explored smaller analogues of the donor substrate, but did not find any of them to be reactive under our conditions. Several observations were made which indicate that MTG may use an ester instead of an amide as an acyl-donor substrate, which opens a possibility that MTG might be used in peptide synthesis. We conclude that MTG has potential as a biocatalyst for specific applications, but much work remains before we fully understand the interaction between MTG and acyl-donor substrates.

Chapter 5
Conclusions & Future Work

5.1 Conclusions

The objective of this thesis has been to investigate the potential of using MTG as an amide bond catalyst. This is an area of organic chemistry where there is room for novel benign methods. To meet our goal we tested the reactivity of small acyl-acceptor and acyl-donor compounds. Over the last decade MTG has been used for many new applications (Chapter 1) therefore understanding the reactivity scope of the enzyme thus does not only enable the enzyme for amide catalysis for non-natural substrates, as is our objective, but furthermore may permit the use of MTG for other novel applications. Our work demonstrated for example that compounds suitable for protein labeling with click chemistry were highly reactive as acceptor substrates. Although these results are important for application driven use of MTG, this section will summarizing general trends of acceptor substrate reactivity, in accordance with our objective, defined above.

We began by exploring acyl-acceptor reactivity by testing compounds carrying a range of substituents in close proximity to the reactive amine, as MTG was known to be promiscuous towards acyl-acceptor substrates[34]. This is the first extensive study of this kind; only one other study[44] has been carried out on acceptor specificity. Our study supports the results found in literature and adds on to the knowledge of MTG by defining general trends that may be used to predict the reactivity of novel acceptor substrates (Chapters 2 and 3). We have defined 4 rules (Table 3.1, Chapter 3) that guide reactivity briefly described below. Firstly we observed that an

unhindered primary amine is required for reactivity, therefore amines adjacent to primary carbons were found to have the highest reactivity. Further we found a negative charge to be detrimental to acyl-acceptor reactivity, when less than two carbons separated the negatively charged functional group (carboxylic acid), and the reacting amine. Reactivity was also reduced with decreasing spacer length between the amine and acid, found both in our study (Chapter 2) and literature[44]. Furthermore we recognized that reactivity increased with the presence of an aromatic group compared to similar non-aromatic acceptors. We compared reactivity between benzyl rings and cyclohexanes; the former was consistently found to be more reactive (Table 2.4, Chapter 2). Finally a comparison of our results with previous literature[44,71] has led to the conclusion that sterically unhindered aliphatic compounds capable of hydrogen bonding have a higher reactivity. This is demonstrated by the high reactivity of glycine ethyl ester[44] and protein bound polyglycine[71] substrates. This may originate from the cellular peptidic target substrates of MTG.

Although these results give an insight into the reactivity of MTG they cannot be taken as definite rules of reactivity. The first event of the acyl transfer reaction is the binding of the donor substrate; the latter is therefore bound in the active site when the acceptor binds, and is strongly influencing the active site. Therefore a study with other donor substrates comparing acceptor reactivity must be carried out to define the general rules of MTG's reactivity.

Chapter 4 explores briefly donor substrate reactivity towards smaller donor substrates, with our objective of defining substrate specificity in mind. We did not find smaller donor substrates than Z-Gln-Gly to be reactive under our conditions, confirming that MTG is less promiscuous towards the donor substrate than towards the acceptor substrate. However we observed preliminary results towards an ester serving as the acyl donor in place of the amide. This is an

important finding because it enables the use of a more readily available substrate as a starting material, and therefore increases the potential use of MTG as an amide catalyst. More experimental work is required to confirm this observation.

The body of this work has probed the use of MTG as an amide bond catalyst, using non-natural substrates. We have defined a broad range of substrates that may be used as acyl acceptor substrates, and defined a set of rules which predicts the reactivity of novel acceptor substrates. Requirements for donor substrate reactivity remain stringent, and enzyme engineering may be required to enable broader substrate specificity. In conclusion our results support the hypothesis that MTG may be a good candidate for a more general non-natural amide bond catalyst, however much work remains both in terms of enzyme engineering, and defining the donor substrate reactivity.

5.2 Future work

Throughout the course of this work, many additional ideas have been developed to address remaining challenges in understanding the biocatalytic potential of microbial transglutaminase (MTG). Below, some promising ideas are presented, which may enable a greater comprehension of the reactivity of MTG toward small-molecule substrates.

5.2.1 Determination of the general structural attributes of acyl-donor substrates

It has been of interest for some time to identify a highly specific, efficient acyl-donor substrate for MTG, for use in selective labeling of proteins. If MTG were to serve as an agent to label specific proteins with small compounds, we should be able to add to the target protein highly reactive Gln-containing peptide that would serve as the acyl-donor, to which an amine-

containing compound could be attached. Currently, the environment of a Gln allowing efficient reactivity with MTG remains enigmatic, as previously discussed (Chapters 1 and 4). In contrast with mammalian tissue transglutaminases, where Gln-containing peptides that react with high selectivity have been identified [45–47]. No highly reactive acyl-donor substrate of MTG has been identified, despite several investigations using synthetic peptides and phage display [45,46]. With the exception of a more reactive donor sequence determined by Lee and colleagues[47], yet in all these studies there is no consistent pattern for the amino acids surrounding the Gln, and the reactivity of the peptides was not fully characterized. As a result, the prediction of highly reactive acyl-donor peptide sequences for MTG is currently not possible.

Because there is no pattern that emerges from the numerous peptide sequences identified as MTG substrates, we suggest that MTG does not recognize a primary structure as a preferred substrate, but rather a 3-dimensional structure, which depends on the secondary or tertiary structure of the Gln-containing target protein. We propose a method to investigate this, where cell lysates are reacted with an acyl-acceptor substrate containing a selective tag. For example, biotinylated cadaverine could help with the identification of donor substrates in a pull-down assay. The reacted proteins would be extracted from the lysate, and identified by MS-MS to allow identification of the reactive glutamine. Alternatively the method could be used to identify acceptor compounds by protein bound lysine or the N-terminal of the protein, by using Z-Gln-Gly-Gly-Biotin in place of the biotinylated cadaverine. This would be repeated several times to narrow the search to those proteins that are repeatedly identified. The sequences identified would be cross-referenced with crystal structures available in the Protein Data Bank. Finally, specific patterns will be searched in the proteins, both in terms of the structural elements in the area of the reactive Gln, and in terms of the chemical properties of neighboring amino acids. A

recognition software such as VASCo (freely available from the University of Graz) could be used.

5.2.2 Ester-based acyl-donor substrates

In Chapter 4, section 4.3, we observed trace activity of MTG toward acyl-donor substrates where an ester group replaced the amide. Specifically, traces of a product corresponding to the ester serving as a substrate were detected by HPLC-MS with the two substrates Z-Gln-Gly(OMe) and Z-Gln(OMe) Reactivity of the substrate and Z-L-Glu(γ-p-nitrophenylester)Gly was monitored by spectrophotometry, and our results indicate that the activated ester functions as a donor substrate for MTG. These observations are of particular interest because esters are more readily available substrates than amides. Among other possibilities, this could facilitate the development of MTG as a peptide catalyst, and may prove to be more robust than current, atom-inefficient peptide synthesis methods (Chapter 1). Briefly, we can envision a method where:

1. An amino acid is immobilized on a resin *via* its *N*-terminus.

2. The amino acid is carboxy-methylated by conventional methods, to render it reactive as an acyl-donor substrate of MTG.

3. Immobilized MTG (for example on magnetic beads, previously demonstrated with MTG [78]) is added along with an un-protected amino acid; its free amine will serve as an acyl-acceptor substrate of MTG, while the free carboxylate remains unreacted.

4. The esterified *C*-terminus of the immobilized amino acid and the N-terminus of the free amino acid form an amide bond, catalyzed by MTG.

5. Any unreacted amino acid is removed by washing and MTG is separated from the resin-bound nascent peptide in the reaction vessel. Steps 2-5 are repeated to form an elongated peptide.

This method is attractive, as MTG is economical and has a high stability, which would render it active for a number of reactions.

In order to investigate the feasibility of this method, it is necessary to clearly establish the reactivity of esters as acyl-donor substrates with MTG. If ester activity is confirmed, this would serve as a starting point to engineer MTG for higher ester reactivity and/or broader substrate specificity towards ester-based acyl-donor substrates. It will also be necessary to engineer broader tolerance of MTG for accepting free amino acids as acyl-acceptor substrates, as Chapter 2 and previous reports[44,46] demonstrated that not all amino acids are recognized as acyl-acceptors by native MTG. This could be done by a broad screening guided by the survival assay outlined in the next section.

5.2.3 Selection assay for MTG activity based on cell survival

Finally, a method for a selection assay for MTG activity based on cell survival is proposed. The development of a selection assay would ease MTG engineering, by enabling screening of millions of MTG variants, rather than a few hundred through individual screening. We propose the MTG-catalyzed modification of an antibiotic. The antibiotic would be chemically modified with a primary amine linker in a functionally non-essential part, to serve as an acyl-acceptor substrate for MTG (Figure 5.1); this linker should not prevent its action as an antibiotic. Mutants of MTG would be added, along with a bulky acyl-donor substrate (Figure 5.1). If the acyl-donor substrate is successfully linked to the acyl-acceptor antibiotic, the antibiotic will be inactivated, by hindering the access of the antibiotic to the antibiotic target, (Figure 5.2), which would further lead to cell survival. One concern which should be taken in to consideration, is the slow reaction rate of MTG. The target should thus be a 'slow' target such as antibiotic chemotherapy targets aimed at cell replication. Furthermore for the blocker to function,

the binding site of the antibiotic on the target should be sterically hindered, such as a deep binding pocket. Finally MTG must be expressed in its mature, active form, as has been demonstrated previously [79,80]. Through modification, this assay could also be applied to other cross-linking enzymes and thus serve as a general assay.

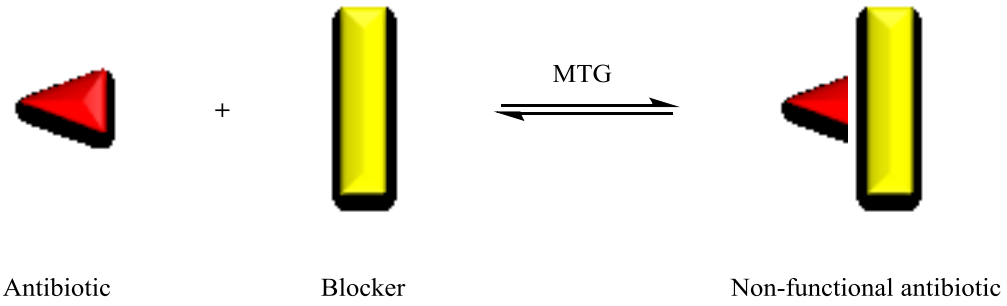


Figure 5.1: Concept of the MTG survival assay. The antibiotic is modified with a primary amine linker to serve as an acyl-acceptor substrate; the blocker is a peptide-based, Gln-containing acyl-donor substrate. Reaction between the antibiotic and the blocker by an active MTG renders the antibiotic non-functional and the cell survives.

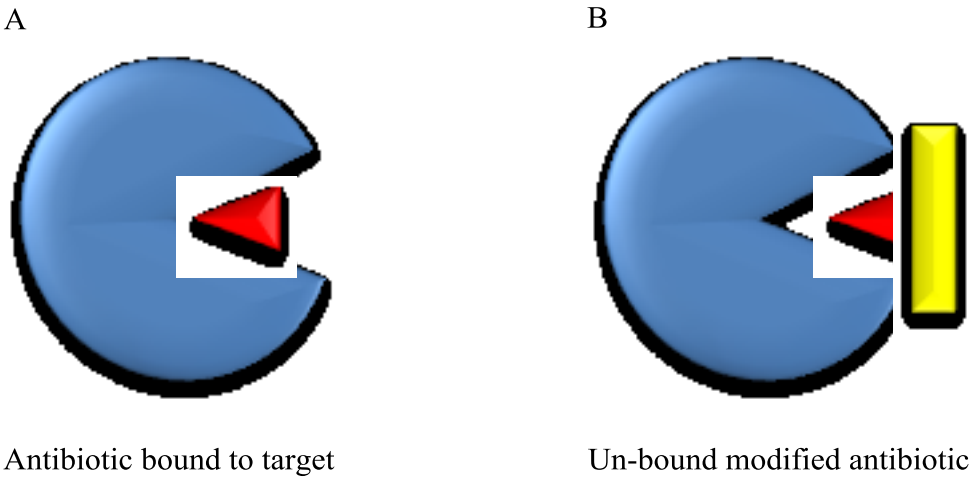


Figure 5.2: Antibiotic interactions with target. A) Un-modified antibiotic binds to target which leads to cell death. B) Modified antibiotic cannot bind to target, cells survive.

Chapter 6
References

1. World Commission *Our common future*. (Oxford Univ. Press, Oxford, 1987).
2. Rockström, J. *et al.* Planetary Boundaries : Exploring the Safe Operating Space for Humanity. *Ecol. Soc.* 14, 2-32 (2009).
3. Anastas, P. T. & Warner, J. C. *Green Chemistry: Theory and Practice*. (Oxford Univ. Press, Oxford, 1998).
4. Mulvihill, M. J., Beach, E. S., Zimmerman, J. B. & Anastas, P. T. Green Chemistry and Green Engineering: A Framework for Sustainable Technology Development. *Annu. Rev. Env. Resour.* **36**, 271–293 (2011).
5. Meussdoerffer, F. G. *A Comprehensive History of Beer Brewing*. (Roy. Soc. Chem. Cornwall, 2009).
6. Legras, J.-L., Merdinoglu, D., Cornuet, J.-M. & Karst, F. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol. Ecol.* **16**, 2091–102 (2007).
7. Rosenthaler, L. Durch Enzyme bewirkte asymmetrische Synthesen. *Biochem. Z* **14**, 238–253 (1908).
8. Bornscheuer, U. T. *et al.* Engineering the third wave of biocatalysis. *Nature* **485**, 185–94 (2012).
9. Riebel, A. *et al.* Expanding the set of rhodococcal Baeyer-Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Appl. Microbiol. Biot.* **95**, 1479–89 (2012).
10. Konarzycka-Bessler, M. & Bornscheuer, U. T. A high-throughput-screening method for determining the synthetic activity of hydrolases. *Angew.Chem. Int.Edit.* **42**, 1418–20 (2003).
11. Mayr, L. M. & Bojanic, D. Novel trends in high-throughput screening. *Curr. Opin. Pharmacol.* **9**, 580–8 (2009).
12. Nestl, B. M., Nebel, B. .A & Hauer, B. Recent progress in industrial biocatalysis. *Curr. Opin. Chembiol.* **15**, 187–93 (2011).
13. Arnold, F. H. & Volkov, A. A. Directed evolution of biocatalysts. *Curr. Opin. Chembiol.* **3**, 54–59 (1999).
14. Schueler-Furman, O., Wang, C., Bradley, P., Misura, K. & Baker, D. Progress in modeling of protein structures and interactions. *Science* **310**, 638–42 (2005).
15. Lutz, S. Beyond directed evolution-semi-rational protein engineering and design. *Curr. Opin. Chembiol.* **21**, 734–743 (2010).
16. Murphy, R., Phillip, D. & Friesner, R. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large• scale modeling of chemistry in protein environments. *J. Comput. Chem.* **21**, 1442–1457 (2000).

17. Taylor, R. D., Jewsbury, P. J. & Essex, J. W. A review of protein-small molecule docking methods. *J. Comput. Aid. Mol. Des.* **16**, 151–66 (2002).
18. Siegel, J. B. *et al.* Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–13 (2010).
19. Schmid, A. *et al.* Industrial biocatalysis today and tomorrow. *Nature* **409**, 258–68 (2001).
20. Schoemaker, H. E., Mink, D. & Wubbolts, M. G. Dispelling the myths-biocatalysis in industrial synthesis. *Science* **299**, 1694–7 (2003).
21. Clouthier, C. M. & Pelletier, J. N. Expanding the organic toolbox: a guide to integrating biocatalysis in synthesis. *Chem. Soc. Rev.* **41**, 1585–605 (2012).
22. Kobayashi, M. & Shimizu, S. Nitrile hydrolases. *Curr. Opin. Chembiol.* 95–102 (2000).
23. Woodley, J. M. Reaction and process engineering. *Enzyme Catalysis in Organic Synthesis, Third Edition* 217–247 (Wiley-VCH, Weinheim, 2012).
24. Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nat. Rev. Genet.* **11**, 367–79 (2010).
25. Purnick, P. E. M. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Bio.* **10**, 410–22 (2009).
26. Steen, E. J. *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559–62 (2010).
27. Allen, C. L. & Williams, J. M. J. Metal-catalysed approaches to amide bond formation. *Chem. Soc. Rev.* **40**, 3405–15 (2011).
28. Pattabiraman, V. R. & Bode, J. W. Rethinking amide bond synthesis. *Nature* **480**, 471–479 (2011).
29. Valeur, E. & Bradley, M. Amide bond formation: beyond the myth of coupling reagents. *Chem. Soc. Rev.* **38**, 606–31 (2009).
30. Heck, T. *et al.* Kinetic Analysis of L-Carnosine Formation by β-Aminopeptidases. *Adv. Synth. Catal.* **352**, 407–415 (2010).
31. Straathof, A. J. J., Panke, S. & Schmid, A. The production of fine chemicals by biotransformations. *Curr. Opin. Biotech.* **13**, 548–56 (2002).
32. Griffin, M., Casadio, R. & Bergamini, C. M. Transglutaminases : Nature ' s biological glues. *Biochem. J.* **368**, 377–396 (2002).
33. Yee, V. C. *et al.* Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII. *P.Natl Acad. Sci. USA.* **91**, 7296–300 (1994).
34. Touati, J., Angelini, A., Hinner, M. J. & Heinis, C. Enzymatic cyclisation of peptides with a transglutaminase. *Chembiochem.* **12**, 38–42 (2011).
35. Shleikin, A. G. & Danilov, N. P. Evolutionary-biological peculiarities of transglutaminase. Structure, physiological functions, application. *J. Evol. Biochem. Phys.* **47**, 1–14 (2011).
36. Kashiwagi, T. *et al.* Crystal structure of microbial transglutaminase from *Streptoverticillium mobaraense*. *J. Biol. Chem.* **277**, 44252–60 (2002).
37. Oliverio, S. *et al.* Tissue transglutaminase-dependent posttranslational modification of the retinoblastoma gene product in promonocytic cells undergoing apoptosis. *Mol. Cell Biol.* **17**, 6040–8 (1997).

38. Fabbi, M. *et al.* Tissue transglutaminase is a caspase substrate during apoptosis. Cleavage causes loss of transamidating function and is a biochemical marker of caspase 3 activation. *Cell Death Differ.* **6**, 992–1001 (1999).
39. Tagami, U. *et al.* Substrate specificity of microbial transglutaminase as revealed by three-dimensional docking simulation and mutagenesis. *Protein Eng. Des. Sel.* **22**, 747–52 (2009).
40. Zhu, Y. & Tramper, J. Novel applications for microbial transglutaminase beyond food processing. *Trends Biotechnol.* **26**, 559–65 (2008).
41. Umezawa, Y., Ohtsuka, T. & Yokoyama, K. Comparison of Enzymatic Properties of Microbial Transglutaminase from *Streptomyces* sp. *Food Sci. Technol. Res.* **8**, 113–118 (2002).
42. Yokoyama, K., Nio, N. & Kikuchi, Y. Properties and applications of microbial transglutaminase. *Appl. Microbiol. Biot.* **64**, 447–454 (2004).
43. Mero, A., Schiavon, M., Veronese, F. M. & Pasut, G. A new method to increase selectivity of transglutaminase mediated PEGylation of salmon calcitonin and human growth hormone. *J. Control. Release.* **154**, 27–34 (2011).
44. Ohtsuka, T., Sawa, A., Kawabata, R., Nio, N. & Motoki, M. Substrate specificities of microbial transglutaminase for primary amines. *J. Agric. Food. Chem.* **48**, 6230–3 (2000).
45. Sugimura, Y., Yokoyama, K., Nio, N., Maki, M. & Hitomi, K. Identification of preferred substrate sequences of microbial transglutaminase from *Streptomyces mobaraensis* using a phage-displayed peptide library. *Arch. Biochem. Biophys.* **477**, 379–83 (2008).
46. Ohtsuka, T., Ota, M., Nio, N. & Motoki, M. Comparison of substrate specificities of transglutaminases using synthetic peptides as acyl donors. *Biosci. Biotechnol., Biochem.* **64**, 2608–2613 (2000).
47. Lee, J.-H. *et al.* Glutamine (Q)-peptide screening for transglutaminase reaction using mRNA display. *Biotechnol. Bioeng.* **110**, 353–62 (2013).
48. Ohtsuka, T., Umezawa, Y. & Nio, N. Comparison of deamidation activity of transglutaminases. *J. Food Sci.* **66**, 25–29 (2001).
49. Mindt, T. L. *et al.* Modification of different IgG1 antibodies via glutamine and lysine using bacterial and human tissue transglutaminase. *Bioconjugate Chem.* **19**, 271–8 (2008).
50. Jeger, S. *et al.* Site Specific and Stoichiometric Modification of Antibodies by Bacterial Transglutaminase. *Angew.Chem. Int.Edit.* **49**, 1–4 (2010).
51. Kitaoka, M. *et al.* Transglutaminase-mediated synthesis of a DNA-(enzyme)n probe for highly sensitive DNA detection. *Chemistry* **17**, 5387–92 (2011).
52. Hu, B.-H. & Messersmith, P. B. Rational design of transglutaminase substrate peptides for rapid enzymatic formation of hydrogels. *J. Am. Chem. Soc.* **125**, 14298–9 (2003).
53. Porta, R., Di Pierro, P., Sorrentino, A. & Mariniello, L. Promising Perspectives for Transglutaminase In "Bioplastics" Production. *J. Biotechnol. Biomaterial* **01**, 1–4 (2011).
54. Pasternack, R. *et al.* Bacterial pro-transglutaminase from *Streptoverticillium mobaraense*-purification, characterisation and sequence of the zymogen. *Eur. J. Biochem. / FEBS* **257**, 570–6 (1998).
55. Marx, C. K., Hertel, T. C. & Pietzsch, M. Soluble expression of a pro-transglutaminase from *Streptomyces mobaraensis* in Escherichia coli. *Enzyme Microb. Tech.* **40**, 1543–1550 (2007).

56. Inverarity, I. a & Hulme, A. N. Marked small molecule libraries: a truncated approach to molecular probe design. *Org. Biomol. Chem.* **5**, 636–43 (2007).
57. Park, K. D. *et al.* Lacosamide isothiocyanate-based agents: novel agents to target and identify lacosamide receptors. *J. Med. Chem.* **52**, 6897–911 (2009).
58. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning : a laboratory manual*. (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1989).
59. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expres. Purif.* **41**, 207–234 (2005).
60. Folk, J. E. & Cole, P. W. Transglutaminase: Mechanistic features of the active site as determined by kinetic and inhibitor studies. *Biochim. Biophys. Acta, Enzymol. Biol. Oxid.* **122**, 244–264 (1966).
61. Karickhoff, S. W., Carreira, L. A. & Hilal, S. H. Sparc On-Line Calculator. (2011). http://archemcalc.com/sparc/
62. Himo, F. *et al.* Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates. *J. Am. Chem. Soc.* **127**, 210–6 (2005).
63. Pasternack, R. *et al.* A fluorescent substrate of transglutaminase for detection and characterization of glutamine acceptor compounds. *Anal. Biochem.* **249**, 54–60 (1997).
64. Abe, H., Goto, M. & Kamiya, N. Enzymatic single-step preparation of multifunctional proteins. *Chem. Commun.* **46**, 7160–2 (2010).
65. Tominaga, J. *et al.* An enzymatic method for site-specific labeling of recombinant proteins with oligonucleotides. *Chem. Commun.* 401–3 (2007).doi:10.1039/b613793h
66. Hernandes, M. Z., Cavalcanti, S. M. T., Moreira, D. R. M., De Azevedo Junior, W. F. & Leite, A. C. L. Halogen atoms in the modern medicinal chemistry: hints for the drug design. *Curr. Drug Targets* **11**, 303–14 (2010).
67. Fichert, T., Yazdanian, M. & Proudfoot, J. R. A structure–Permeability study of small drug-like molecules. *Bioorg. Med. Chem. Lett.* **13**, 719–722 (2003).
68. Fleming, F. F., Yao, L., Ravikumar, P. C., Funk, L. & Shook, B. C. Nitrile-containing pharmaceuticals: efficacious roles of the nitrile pharmacophore. *J. Med. Chem.* **53**, 7902–17 (2010).
69. Leblanc, A., Gravel, C., Labelle, J. & Keillor, J. W. Kinetic studies of guinea pig liver transglutaminase reveal a general-base-catalyzed deacylation mechanism. *Biochem.* **40**, 8335–8342 (2001).
70. Gnaccarini, C. *et al.* Site-specific protein propargylation using tissue transglutaminase. *Org. Biomol. Chem.* **10**, 5258–65 (2012).
71. Tanaka, T., Kamiya, N. & Nagamune, T. N-terminal glycine-specific protein conjugation catalyzed by microbial transglutaminase. *FEBS lett.* **579**, 2092–6 (2005).
72. Tanaka, T., Kamiya, N. & Nagamune, T. Peptidyl linkers for protein heterodimerization catalyzed by microbial transglutaminase. *Bioconjugate Chem.* **15**, 491–497 (2004).
73. Keillor, J. W. *et al.* The bioorganic chemistry of transglutaminase — from mechanism to inhibition and engineering. *Canadian Journal of Chemistry* **86**, 271–276 (2008).
74. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew.Chem. Int.Edit.* **40**, 2004–2021 (2001).
75. Best, M. D. Click chemistry and bioorthogonal reactions: unprecedented selectivity in the labeling of biological molecules. *Biochem.* **48**, 6571–84 (2009).

76. Touati, J., Angelini, A., Hinner, M. J. & Heinis, C. Enzymatic cyclisation of peptides with a transglutaminase. *Chembiochem* **12**, 38–42 (2011).
77. De Macédo, P., Marrano, C. & Keillor, J. W. A direct continuous spectrophotometric assay for transglutaminase activity. *Anal. Biochem.* **285**, 16–20 (2000).
78. Moriyama, K., Sung, K., Goto, M. & Kamiya, N. Immobilization of alkaline phosphatase on magnetic particles by site-specific and covalent cross-linking catalyzed by microbial transglutaminase. *J. Biosci. Bioeng.* **111**, 650–3 (2011).
79. Liu, S. *et al.* The order of expression is a key factor in the production of active transglutaminase in Escherichia coli by co-expression with its pro-peptide. *Microb. Cell Fact.* **10**, 112 (2011).
80. Liu, S. *et al.* The pro-region of *Streptomyces hygroscopicus* transglutaminase affects its secretion by Escherichia coli. *FEMS microbiology letters* **324**, 98–105 (2011).
81. Thomas, J. P. US patent 2008/0253997. (2008).

Annex 1: Supporting information Chapter 2

Microbial transglutaminase displays broad acyl-acceptor substrate specificity

Maria T. Gundersen[1,2,3], Jeffrey W. Keillor[4] and Joelle N. Pelletier[1,2,3,5]*

From the [1]Département de biochimie, Université de Montréal, Montréal, Québec H3C 3J7, [2]PROTEO, the Québec Network for Protein Function, Structure and Engineering, [3]CGCC, the Center in Green Chemistry and Catalysis, [4]Department of Chemistry, University of Ottawa, Ottawa, ON K1N 6N5 and [5]Département de chimie, Université de Montréal, Montréal, Québec H3C 3J7

*Corresponding author: Joelle N. Pelletier

Table A1.1: Reaction pH and formation of the hydrolysis product Z-Glu-Gly.

	Acyl-acceptor substrate	Donor consumed in 72 h (%)	Reaction pH (calculated pK_a) [a]	Hydrolysis product (%)
1	Hydroxylamine	> 99	6 (5.97)	0
2	N^{α}-acyl-l-lysine-methyl ester	> 99	9 (10.05)	0
3	Cadaverine	> 99	9 (10.36)	0
4	6-Aminocaproic acid	70.6 ± 6.4	9 (10.31)	0
5	5-Aminovaleric acid	38.5 ± 3.2	9 (10.15)	0
6	γ-Aminobutyric acid	11.8 ± 6.7	9 (10.43)	0
7	β-Alanine	4.8 ± 1.3	9 (9.34)	0
8	Glycine	8.4 ± 2.4	7.7 (7.7)	4.6 ± 0.8
9	Carbamic acid	nd [b]	6 (nc) [c]	– [d]
10	Glycine ethyl ester	> 99	7.7 (7.71)	0
11	L-Threonine	nd	7 (7.06)	–
12	L-Threonine ethyl ester	4.1 ± 9.5	7 (6.9)	24.6 ± 4.2
13	L-Serine methyl ester	5.8 ± 1.4	7 (6.9)	20.3 ± 3.1
14	D-Serine methyl ester	3.0 ± 1.9	7 (6.9)	24 ± 1.7
15	L-Cysteine ethyl ester	2.7 ± 1.7	6.5 (6.71)	19.6 ± 1.2
16	L-Tryptophan methyl ester	4.5 ± 1.1	8 (8)	41.3 ± 2.1
17	L-Isoleucine methyl ester	nd	7.5 (7.7)	–
18	Benzylamine	93.6 ± 1.3	9 (9.34)	0
19	Aniline	18.9 ± 4.4	6 (4.67)	17.2 ± 0.8
20	Cyclohexylmethylamine	46.3 ± 1.6	9 (10.41)	0
21	Cyclohexylamine	3.1 ± 1.6	9 (10.39)	0
22	*m*-Anisidine	11.3 ± 1.4	6 (4.35)	9.9 ± 1.4
23	3-Chloro-4-fluorobenzylamine	23.1 ± 5.7	9 (8.79)	0
24	*p*-Xylenediamine	62.7 ± 2.4	9 (9.75)	0
25	*O*-Benzylhydroxylamine	2.6 ± 5.6	6 (4.25)	0
26	Thiamine	nd	7 (nc)	–
27	2-Bromoethylamine	nd	8.5 (8.5)	–
28	Glucosamine	nd	9 (12.27)	–
29	Aminoacetonitrile	> 99	6 (4.59)	0
30	Propargylamine	90.1 ± 0.9	9 (9.47)	0
31	Propylamine azide	98.5 ± 0.5	9 (nc)	0
32	Ethylamine azide	97.7 ± 0.5	9 (nc)	0
33	N-Ethylmethylamine	nd	9 (10.85)	–
34	Sarcosine	nd	8 (8.06)	–
35	Butanol	nd	7 (nc)	–
36	Butanethiol	nd	7 (nc)	–

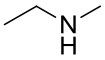
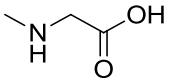
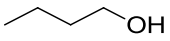
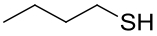
[a] pK_a values were calculated online at: http://archemcalc.com/sparc/ (Hilal, Said, S. W. Karickhoff and L. A. Carreira (1995) Quant. Struc. Act. Rel., 14, 348).

[b] nd: not detected.

[c] nc: not calculated.

[d] – : not measured.

Table A1.2: Reactivity of acyl-acceptors other than primary amines.

Acceptor compound	Structure	Donor consumed in 72h (%)
33		nd [a]
34		nd
35		nd
36		nd

[a] nd: not detected (< 2 % product)

Annex 2: Supporting information Chapter 3

A.2 Docking

All docking experiments were done with the trial version of MolDock (Aarhus, Denmark), The crystal structure of MTG (PDB file 1UI4) was used for all docking simulations. The ligand 3D structures were drawn using ChemDraw 3D Pro 8.0. Docking simulations were run on a HP ProBook 4515s, running Molegro Virtual Docker 4.0.031,the Moldock Score [Grid] scoring function was used. Grid resolution was 0.30 Å with a radius of 30 around the active site of MTG. The search algorithm used was Moldock Optimizer, with default settings. With 10 runs and 100 iterations, where the top 10 poses for the donor docking was manually inspected.

Annex 3: Supporting information Chapter 4

A3.1 Materials and Methods

A3.1.1 Materials

The enzyme was prepared as described in Chapter 2. *N*-Z-L-Glu(*γ*-*p*-nitrophenylester)Gly was kindly synthesized in the laboratory of Jeffrey W. Keillor (University of Ottawa, Ottawa, Canada)[69]. Other chemicals used were purchased from the suppliers listed below, and were not modified unless specifically noted. 4-Methoxybenzamide (97%), glycine ethyl ester•HCl and triisopropylsilane were purchased from Acros Organics (Waltham, USA). Ethyl iodide, Glutamine, Methyl iodide, N,N-Dimethylformamide, piperidine, Z-Glutamine, N^{α}-acetyl-L-lysine methyl ester•HCl, trypsin from bovine pancreas (10,000 BAEE U/mg) and hydroxylamine were purchased from Sigma-Aldrich (St Louis, USA). Carboxybenzyl-glutaminyl-glycine (Z-Gln-Gly) was from Alfa Aesar (Ward Hill, USA). Glutathione (reduced) was from Bioshop (Burlington, Canada). Sodium bicarbonate was purchased from Fisher (Waltham, USA). Formic acid 98% purity was from Fluka Analytical (St Louis, USA).

Synthesis of most of the donor compounds was necessary as they were not commercially available; synthetic methods are described below.

A3.1.2 Synthetic method: Esterification

A3.1.2.1 General synthesis conditions

The method was adapted from US patent 2008/0253997, section 0068 [81]. 1 eq substrate (1), was dissolved in dimethylformamide with $NaHCO_3$ 5.8 eq, and stirred at room temperature for

30 min. Methyl/Ethyl iodide 1.86 eq was added, followed by continuous stirring overnight. The reaction mixture was diluted to 10 mL with water and stirred for 40 min. The solid was collected by filtration over a Buchner funnel with filter paper, washed well with cold water and extracted with ethyl acetate:water (80 mL : 30 mL). The aqueous phase was extracted with 25 mL ethyl acetate. The combined organic extracts were washed with 10 mL water, dried over sodium sulfate, evaporated and dried at room temperature under vacuum for 72h, to yield the esterified product.

A3.1.2.2 Product detection

For all products synthesized, masses were detected under positive ionization of direct injection with a Waters 3100 single quadrupole mass detector. The mass corresponding to the expected product was detected and monitored in correspondence with starting material consumption.

A3.2 Results

Table A3.1: Rate of product formation with N-Z-L-Glu(γ-p-nitrophenylester)Gly substrate.

Donor	MTG	Acceptor	Rate of product formation (μM/min)
✓	×	×	0.66
✓	✓	×	1.06
✓	×	✓	0.22
✓	✓	✓	0.82