The FEBS Journal

FEBS PRESS science publishing by scientists

REVIEW ARTICLE

Approaching transglutaminase from *Streptomyces* bacteria over three decades

Hans-Lothar Fuchsbauer

Department of Chemical Engineering and Biotechnology, University of Applied Sciences of Darmstadt, Germany

Keywords
catalysis; occurrence; production; properties; purification; *Streptomyces*; *Streptoverticillium*; structure; substrates; transglutaminase

Correspondence
H.-L. Fuchsbauer, Department of Chemical Engineering and Biotechnology, University of Applied Sciences of Darmstadt, Stephanstraße 7, Darmstadt 64295, Germany
Tel: +49 6151 1638181
E-mail: hans-lothar.fuchsbauer@h-da.de

(Received 25 February 2021, revised 23 April 2021, accepted 7 June 2021)

doi:10.1111/febs.16060

Transglutaminases are protein cross-linking and protein-modifying enzymes that have attracted considerable interest due to their causal involvement in various diseases and versatility in industrial applications. In particular, microbial transglutaminases (MTG) from *Streptomyces* bacteria have managed in recent years to evolve from simple food additives to specialized enzymes for the site-directed modification of therapeutic proteins. The review summarizes relevant studies from the beginning dealing with the occurrence, production, structure, catalysis, and substrate molecules of MTG enzymes. It also addresses biotechnological procedures with MTG from *S. mobaraensis* (*Sm*MTG) as the most prominent representative in focus. Reassessment of the available data revealed unexpected insights into catalysis of *Sm*MTG and other transglutaminases, suggesting selection of glutamine donor proteins by subsites at the front vestibule and the existence of distinct lysine pockets. Flexibility of the *Sm*MTG-accessible glutamine donor substrate regions seems to be more important than the glutamine environment. Nevertheless, residues in close vicinity to glutamines also determine interaction with the *Sm*MTG subsites. The apparent lack of subsites for lysine donor proteins suggests self-assembly of the substrate proteins prior to enzymatic cross-linking. The study of natural substrate proteins, especially their mutual interaction, is proposed to further illuminate catalysis of *Sm*MTG. To this end, structure and function of the characterized substrate proteins from *S. mobaraensis* are discussed in conclusion.

Preliminary remarks

The objective of the present review is microbial transglutaminase (MTG), which has attracted the author for more than thirty years. During this time, MTG enzymes from *Streptomyces mobaraensis* and other *Streptomyces* bacteria have been evolved from simple additives that improve the manufacture and properties of foods to exceptional tools that enable the site-specific attachment of new functionalities to therapeutic proteins. Especially, methods dealing with the MTG-mediated production of antibody–drug conjugates have promoted recent progress in this regard. Therefore, MTG structure as a determinant of

Abbreviations
B., *Bacillus*; Chp, chaplin; DAIP, Dispase autolysis-inducing protein; G-CSF, granulocyte colony-stimulating factor; GH, growth hormone; *g*TG, guinea pig liver transglutaminase; IFN, interferon; IP, inhibitory peptide; mAb, monoclonal antibody; Mb, myoglobin; MTG, microbial transglutaminase; Myc, transcription factor myc proto-oncogene protein; Rdl, rodlin; RNase, ribonuclease; *S.*, *Streptomyces*; Sec, general secretion pathway; Sml, *Streptomyces mobaraensis* β-lactamase; SPI, *Streptomyces* papain inhibitor; SSI, *Streptomyces* subtilisin inhibitor; SSTI, *Streptomyces* subtilisin and TAMP inhibitor; *Sv.*, *Streptoverticillium*; TAMP, transglutaminase-activating metalloprotease; TAP, tri-/tetrapeptidyl aminopeptidase; Tat, twin arginine translocation pathway; TGase, transglutaminase; αLA, α-lactalbumin.

catalytic specificity and function will be the focus of this review. Unfortunately, MTG function in the life of *Streptomyces* bacteria has not received the same attention as its use in biotechnological applications. Our knowledge is like a white map that needs to be filled with lines, and it is to be hoped that the biological role of MTG will soon be studied with more enthusiasm than in the past to fully understand catalysis.

Multifunctional transglutaminases—a family of protein-modifying enzymes

Transglutaminases (R-glutaminyl peptide: amine γ-glutamyltransferase, TGase, EC2.3.2.13) are multifunctional enzymes being present in all kingdoms. They not only cross-link proteins between glutamine and lysine residues (Fig. 1A) but also catalyze various other modification reactions. The linkage of primary amines to *endo*-glutamines yields protein conjugates with new molecular entities (B). The irreversible hydrolysis of glutamines, whether substituted (removal of amines) or not (removal of ammonia), lowers the isoelectric point of the substrate proteins by increasing the negative charge (C). *Endo*-glutamyl ester formation (D), protein disulfide isomerase, and nuclease activities are also attributed to TGases.

Pioneering work was done in the middle of the last century by Heinrich Waelsch, John Folk, Laszlo Lorand, and coworkers [1–3], after the discovery of a fibrin-clotting component in blood, the blood coagulation factor XIII (FXIII), and a tissue transglutaminase (TG_C, TG2) in guinea pig liver (*g*TG) [4–6]. The determination of N^{ε}-γ-glutamyllysine isopeptide bonds [2] and other analytical methods (Appendix S1) were then used to detect other TGases in tissue and body fluids. At a time when biotechnology had just begun, *g*TG became the most studied TGase as abundant amounts in guinea pig liver facilitated the purification procedure. Quaternary structure and the need of zymogen activation by thrombin limited the use of FXIII as model enzyme [3]. From this point of view, the crystallization of FXIII and the determination of the first transglutaminase structure by Teller and colleagues are outstanding in history of the transglutaminase enzymes [7]. TGases from animals and plants, their physiological role, and causative involvement in many diseases have been extensively studied since then but cannot be objective of this overview. Readers are referred to recent reviews on these topics [8–11].

Mammalian transglutaminases such as tissue TG2 and the thrombin-processed FXIII subunit A have a multi-domain structure consisting of N-terminal sandwich, active core, and two C-terminal barrel domains (Fig. S1). They contain many unbridged cysteines and form as inactive enzymes a compact ('close') molecular structure that is stabilized by GDP and GTP in TG2 [12]. Activation by Ca^{2+} causes conformational movements of the barrel domains toward an 'open' (active) enzyme structure that is accessible to target proteins [13]. Guinea pig liver *g*TG was also used in early attempts to modify food proteins [14], but elaborate production procedures, low oxidation stability, susceptibility for heavy metals, and enzyme activation by Ca^{2+}-restricted broad applications. In this light, the discovery of a small, Ca^{2+}-independent transglutaminase in culture supernatants of the *Streptoverticillium* strain S-8112 by two teams of the Japanese companies Ajinomoto and Amano Pharmaceutical represents a major advance in industrial protein modification procedures [15]. Molecular cloning and crystal structure analysis then revealed sequence and structure of microbial transglutaminase (MTG) [16,17], the culture of *S. mobaraensis* DSM 40847, formerly *Streptoverticillium* (*Sv.*) *mobaraense*, [18] an efficient production strain

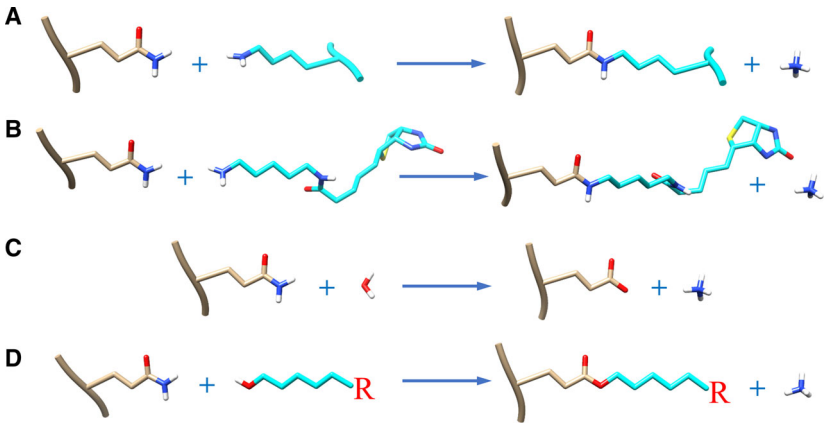

Fig. 1. Protein modification reactions mediated by transglutaminases. (A) Protein cross-linking between *endo*-glutamines and *endo*-lysines; (B) Protein conjugation with primary amines. The biotin moiety can be replaced by any payload; (C) Hydrolysis of protein glutamines; (D) Protein conjugation with ω-hydroxyceramide. The extension of the fatty acid hydrocarbon chain and the ceramide moiety are symbolized by R.

(Fig. S2A,B). While many MTG-modified foods quickly entered the global markets, the production of biomaterials and protein conjugates has attracted more attention in recent years. Using the keyword 'microbial transglutaminase', the PubMed database (https://www.ncbi.nlm.nih.gov/pubmed/) revealed a considerable but not complete number of publications (596 reports in total, effective date 04.01.2021), continuously increasing from the first (36, still influenced by a restrictive sales policy) over the second (133) to the last decade (427).

Various genome sequences of MTG-producing *Streptomyces* bacteria and many MTG sequences are now available. Several studies focused on optimizing the catalytic fold and deciphering MTG recognition sites for substrate proteins and peptides. Accordingly, discussion of the essential amino acids, affecting structural integrity and catalytic function of MTG, is a central task of this review. Of similar importance, the structural requirements of MTG for the glutamine and lysine donor proteins reflect the state-of-the-art of catalysis.

Occurrence of microbial transglutaminases in nature

A recent bioinformatic survey revealed more than 8000 protein sequences annotated as transglutaminases (TGases) in the PFAM database, also proposing classification of the putative microbial enzymes into five different groups [19]. BLASTN search using the NCBI genome database yielded six complete and 85 draft *Streptomyces* genomes exhibiting a single gene for *Sm*MTG-like transglutaminases (effective date 18.11.2019). Inspection of the gene loci suggested MTG acquisition by horizontal gene transfer (HGT) as conserved gene clusters are missing, and *tgase* gene environment varies from species to species. The lack of MTG genes in the genome of the majority of streptomycetes, especially in the genomes of the type strains *Streptomyces coelicolor* A3(2) and *Streptomyces griseus* subsp. *griseus* IFO 13350, may be further indicative for HGT.

In the past, the intracellular transglutaminase from *Bacillus subtilis* (*Bs*MTG), which cross-links GerQ of the endospore protein coat, was considered an alternative candidate for catalyzing protein modification reactions [20,21]. The small *Bs*MTG enzyme differs from *Sm*MTG in molecular size and structure without sharing structural homology (Fig. S2C). However, versatility of *Bs*MTG proved to be limited as the manufacturing procedure is challenging, possibly due to the presence of two cysteines [22]. Transglutaminases have been further characterized in *Streptococcus suis* [23], *Pseudomonas aeruginosa* [24], and *Kutzneria albida* [25]. Only the enzyme from the latter bacterium shows some structural similarity with *Sm*MTG, especially in the near of the catalytic amino acids (Fig. S2D). Moreover, the *Bordetella* dermonecrotic toxin (1464 aa in size) [26] and the *Escherichia coli* necrotizing factor 1 (1014 aa) [27] include C-terminal TGase domains (284 aa and 282 aa, respectively) that activate small rho GTPases of the host by glutamine γ-carboxamide hydrolysis or polyamine incorporation. The eucaryotic slime mold *Physarum polycephalum* [28] and oomycetes such as *Phytophora sojae* [29] may be additional sources for MTG production, but the enzymes require, like animal and plant TGases, Ca^{2+} activation. In summary, properties of the characterized transglutaminases limit their general use for industrial applications with exception of MTG from *Streptomyces* spp. and, possibly, from *K. albida*.

Production of transglutaminase by *Streptomyces* bacteria

Secretion of transglutaminase during culture of *Streptomyces* bacteria

The *Streptoverticillium* strain S-8112 from soil achieved the highest MTG activity of about 2.5 $U \cdot mL^{-1}$ after reaching the stationary growth phase [15] (for growth and genetic stability see Appendix S1). Specific antibodies revealed the export of MTG as a zymogen when *S. mobaraensis* DSM 40847 was allowed to grow for 24 h [30]. In shaking flasks, activation starts after 35–45 h and is usually completed two days later. More typically, according to the results of numerous experiments under similar conditions, the complete processing of the zymogen is a rare event, and culture of more than 70–90 h causes proteolytic degradation of *Sm*MTG. Prolonged culture further promotes the formation of inactive MTG by auto-catalytic glutamine deamidation, which gradually lowers the isoelectric point of *Sm*MTG from pI 8.9 [15] or pI 8.0 [30] to pI 6.4 (R. Pasternack, E. Dechène & H.-L. Fuchsbauer, unpublished). The recent purification of two *Sm*MTG variants differing in charge may be consistent with this observation [31].

Processing of the MTG zymogen, indicated by a decrease in protein size and an increase in transamidation activity, was observed during submerged culture of various streptomycetes. The transglutaminase-activating metalloprotease (TAMP), which was isolated from surface colonies of *S. mobaraensis*, is usually absent, and only traces were immunologically detected in the death phase of *S. mobaraensis* [32]. However, low azocasein

degradation activities (up to 0.08 $AU \cdot h^{-1}$) and MTG activation by a serine and a metalloprotease have been reported in *Streptomyces hygroscopicus* and *S. mobaraensis* [33]. The observation is consistent with the fact that a cascade of proteases, including trypsin- and chymotrypsin-like serine proteases, is involved in aerial hyphae formation of *Streptomyces* bacteria [34–36]. Unfortunately, protein accumulation mediated by cetyltrimethyl ammonium bromide precipitation was not sufficient to show a protease band [33]. In addition, autoactivation was observed in *S. hygroscopicus*, thus showing cysteine protease activity of *Sh*MTG [37]. Autoprocessed *Sh*MTG was active but could not activate the MTG zymogen.

Transglutaminase processing and regulation

The metalloprotease TAMP cleaves the *Sm*MTG zymogen at the amido-side of Phe(−4), thus leaving the tetrapeptide FRAP at the mature N terminus [32]. The tetrapeptide is removed by an AP-specific tri/tetrapeptidyl aminopeptidase (TAP) that is secreted by *S. mobaraensis* in an early stage of culture [38,39] (Table S1). Flexibility of the FRAP loop and the specific recognition sequence also enable chymotrypsin and trypsin to remove the propeptide by cleaving the Phe–Arg and Arg–Ala peptide bonds [38]. TAP, however, only cuts the remaining tripeptide from RAP-MTG [38]. Processing of AP-MTG fails for unknown reasons, although AP*pNA* is a useful substrate of TAP. A Leu/Phe aminopeptidase, secreted by *S. mobaraensis*, is possibly involved in the final processing of trypsin-activated *Sm*MTG [S14] (Table S1).

Apart from MTG and TAP, the secretome of *S. mobaraensis* contains the TAMP inhibitory protein SSTI [32], a variant of the *Streptomyces* subtilisin inhibitors (SSI) [40], and a substrate of MTG [41] (Table S1). The subtilisin-binding site of SSTI between α-helix 2 and β-sheet 3 differs from the first characterized SSI of *Streptomyces albogriseolus* S-3253 [40] in the replacement of methionine for lysine. Hence, Lys-70 enables SSTI to act not only as a subtilisin and trypsin inhibitor but also as a lysine donor protein of *Sm*MTG [42]. The conserved SLYAP motif at the N terminus of SSTI and the adjacent glutamine pair (QQSLY) were recently determined as inhibition and glutamine donor sites of TAMP and *Sm*MTG [43]. The extension peptide and the N-terminal stretch further include repetitive AP motifs that may enable TAP to truncate SSTI like TAMP-activated *Sm*MTG. In fact, SSTI labeling mediated by *Sm*MTG suggests loss of the glutamine-binding site during culture of *S. mobaraensis* [41].

Other soluble *Sm*MTG substrate proteins are produced during submerged culture of *S. mobaraensis* (Appendix S1, Table S1) [44–46]. They have in common that they are secreted substantially before *Sm*MTG activation occurs [45]. The activating protease TAMP is likely retained in the cell wall by a C-terminal convertase domain [S14] that lacks in other metalloproteases of the thermolysin family.

Purification and properties of microbial transglutaminases

The purification of transglutaminase from culture supernatants of *Streptomyces* is as simple as the recovery of His-tagged enzymes. Ethanol precipitation (30–50 vol%) followed by ion exchange chromatography is usually sufficient to obtain the highly purified enzyme (see also Appendix S1).

Microbial transglutaminase from *Streptomyces* is much smaller than mammalian and plant transglutaminases due to its single-domain structure of about 38 kDa (Tables S1 and S2) [47]. Although the isoelectric point depends on glutamine auto-deamidation during culture, the reported data of pI 8.9 [15] and pI 8.0 [30] shows that MTG is usually a positively charged protein at pH 6.0. The pH optimum ranges from pH 5.0 to pH 8.0 (Table S2). Acidification below pH 4.0 causes *Sn*MTG and MTG enzymes from other *Streptomyces* spp. to precipitate (Appendix S1, Refs S26, S27). Activity is usually lost at temperatures above 50 °C, even though thermal denaturation may occur at 37–45 °C. *Sh*MTG, characterized by DSC, revealed a transition temperature of 60.2 °C [48]. The propeptide stabilizes *Sm*MTG to resist heating at 60 °C for at least 1 h, which causes complete inactivation of the mature enzyme in parallel [30]. The high sensitivity to disulfide reagents (Appendix S1, Refs S15,S19), electrophilic inhibitors [15,S15,S16,S19], and heavy metals [15,S15] such as cystamine (2,2′-dithiobisethanamine), 5,5′-dithiobis(2-nitrobenzoic acid), *N*-ethyl maleimide, iodoacetic acid, $ZnCl_2$, $PbAc_2$, $CuCl_2$, and *para*-chloromercuribenzoate further proves that the only MTG cysteine is essential for catalysis. Transglutaminases, among them MTG, have also been shown to be inhibited by a natural melanin macromolecule from *Streptomyces lavendulae*, most likely by nucleophilic addition of the catalytic cysteine thiolate to the *o*-quinone moieties of the inhibitory molecule [49].

The need to refold inclusion bodies led to the discovery of a stable *Sm*MTG intermediate at pH 4.0. Investigation of the intermediate and other methods that support refolding of *Sm*MTG is described in the Appendix S1.

Transglutaminase catalysis

The catalysis mechanism of transglutaminases (TGase) has been extensively studied by the admirable work of Folk and colleagues using guinea pig liver *g*TG as model enzyme. The results are summarized in excellent reviews [2,50] and will be briefly revived here (Fig. S3).

In general, the active site of all TGase enzymes includes a catalytic triad consisting of cysteine, histidine, and aspartate to catalyze two consecutive *ping-pong* reactions, similar to cysteine proteases. When a glutamine donor substrate, a protein in general, has entered the catalytic core of TGase (Fig. S3A), the cysteine thiolate attacks the γ-carboxamide sidechain to form the tetrahedral oxyanion, most likely from the *re*-side, if the leaving amido group is directed to histidine or aspartate (Fig. S3B). A proton relay system, which is not understood, then enables the elimination of ammonia and ammonium ions to form the thioester intermediate (Fig. S3C). Study of the binary transition state structure is not possible because the thioester hydrolyzes in the absence of primary amines (Fig. 1C). For this reason, it remains unclear where the energy for thioester bond formation comes from. The high-energy complex then reacts with amine nucleophiles, which can be a lysine donor protein, a polyamine, or another primary amine with any payload (Fig. 1). Noticeably, cross-linking occurs via a transition state consisting of three proteins. After a second oxyanion intermediate is formed (Fig. S3D), the catalytic cysteine thiolate is displaced, and transglutaminase dissociates from the protein conjugate or the cross-linked proteins (Fig. S3E).

Structure and function of transglutaminases from *Streptomyces* bacteria

Sequence homology

The primary structure of the known *Streptomyces* transglutaminases is highly conserved (Fig. S4), sharing 72.1–78.6% [UniProt ID A0A0D7CI14 (*S. natalensis*) *vs* A0A445N4D8 (*S. netropsis*)] identity with the type enzyme (prepro-MTG) from *S. mobaraensis* DSM 40847 (P81453). An unexpected BLAST*P* result was the discovery that MTG (A5PHK1) from the recently sequenced *Saccharomonospora viridis* is identical to *Sm*MTG [51]. The thermophilic bacterium, which was originally isolated from composted manure heaps, is Gram-negative, although it belongs to the Gram-positive actinomycetes [52]. Furthermore, MTG sequences from *S. mobaraensis* are listed under at least ten UniProt ID codes, which are identical (M3C567, Q2VI01, Q6RET8; 99–100% identity), slightly different (Q6RET9, Q5UCB5, Q8KNY5, Q84AM9; 88.5–92.5%), or significantly different (Q6TGB5; 76.8%) to the reference protein P81453. The Q6E0Y3 variant shares, apart from a single point mutation (P225S), the same sequence with mature P81453 but is altered by an exchange of two terminal stretches downstream from Gly286. One of the inserted peptides, CHACLTRASSATG, contains a CxxC motif that is typical for peptidoglycan-binding protein disulfide isomerase, thioredoxin, and ferredoxin proteins. It should also be noted that the catalytic Asp-255 is replaced by glycine in the *S. mobaraensis* variant Q6RET8.

Significance of the signal peptide for the production of transglutaminase

From the very beginning, the discovery of microbial transglutaminase triggered a race to develop industrial production procedures and to apply for patents, which seems to continue to this day. The data revealed significance of the signal and the propeptides for secretion and folding of microbial transglutaminases. Appendix S1 summarizes some highlights of the recombinant production procedures (Table S3) and the protein export in *Streptomyces* bacteria.

The production of *Sm*MTG with Sec-dependent pelB accumulates inclusion bodies in *E. coli* [S39,S60], while *Sh*MTG may be translocated with both pelB and the native leader sequence [S43]. Although the signal peptides preceding *Streptomyces* MTG are highly conserved (Fig. S4), they probably favor either the Sec or the Tat secretion pathway (Table 1, Appendix S1). Among the enzymes used in recombinant production procedures (Table S3), the signal peptide sequence motifs R-R-(A/R/S)-L-(V/L) of *Sm*MTG [16,30], *Sc*MTG [S16,S18], *Sp*MTG [S21], and *Sh*MTG [S43] are shared with those of Tat-exported *Streptomyces* proteins [S56,S61–S64]. In contrast, a glutamine interrupts the twin arginine pair of *Sf*MTG [S20] and *Sn*MTG [S27], thus suggesting more a Sec-like protein transport as was shown for phospholipase D from *Streptomyces cinnamoneus* [S58] and the *Streptomyces* subtilisin inhibitor from *Streptomyces lividans* [S59].

Efficient production procedures in *E. coli* were further reported for the transglutaminase-activating metalloprotease (TAMP) [53], the MTG-processing tri/tetrapeptidyl aminopeptidase (TAP) [39], the TAMP inhibitor (SSTI) [43], the papain inhibitory protein (SPI) [54], the Dispase autolysis-inducing protein (DAIP) [55], and the β-lactamase Sml-1 [46] from *S. mobaraensis*. While TAP, SSTI, SPI, DAIP, and

Table 1. Signal peptides encoding Sec- or Tat-mediated secretion of various proteins. Basic amino acids, cysteines, and the Sec/Tat export and signal peptidase-binding sequences are printed bold. Predicted pathways are given in square brackets.

Source	Enzyme	Signal peptide sequence	Pathway	Ref.
S. mob.	Transglutaminase	M**R**I**RRRALV**FATMSAVL**C**TAGFMPSAGE**AAA**	[Tat]	[16,30]
	TAMP	M**R**PTPQ**RRAVA**TGALVAVTAMLAVGVQTTS**ANA**	[Tat]	[53]
	TAP	M**RK**AL**R**SLLAASMLIGAIGAGSATAEAASI**TAP**	[Sec]	[39]
	SSTI, TAMP inhib.	M**R**YITGGIALGSALILGSLVG**AGA**T**ASA**TP**APAPA**	[Sec]	[43]
	SPI (expansin)	M**R**EF**RR**V**RR**V**R**FAA**C**ALVAAATGITLAAGP**ASA**	[Sec]	[54]
	DAIP	M**KR**MGWAVTAAVTTIVLAQSSLA**AQA**	[Sec]	[55]
	β-Lactamase Sml-1	MP**K**LSL**R**I**R**G**K**ALGFTAALAVGAAALTGP**ATA**	[Sec]	[46]
S. cin.	Transglutaminase	MH**KRRRLL**AFATVGAVI**C**TAGFTPSVSQA**ASS**	[Tat]	[S16,S18]
S. hyg.	Transglutaminase	MY**KRRSLL**AFATVSAAIFTAGVMPSV**SHA**	[Tat]	[S43]
S. pla.	Transglutaminase	MY**KRRSLL**AFATVGALI**C**TAGVMPSV**SHA**	[Tat]	[S21]
S. fra.	Transglutaminase	MY**KR**Q**R**FLAFATVGAVI**C**TAGSTPSL**SQA**	[Sec]	[S20]
S. net.	Transglutaminase	MN**KR**Q**R**FLAFATVGAVI**C**TAGFTPSLSQAA	[Sec]	[S27]
S. chr.	Phospholipase D	LASFSP**RRRTVVK**AAAATAVLAGPLAAALP**ARA**	Tat	[S61]
S. cin.	Phospholipase D	LL**R**H**R**L**RR**LH**R**LT**R**SAAVSAVVLAALPAAP**AFA**	Sec	[S58]
S. liv.	Xylanase C	PLNGM**SRRGFL**GGAGTLALATASGLLLPGT**AHA**	Tat	[S56]
S. coe.	Agarase (dagA)	MVN**RRDLIK**WSAVALGAGAGLAGPAPA**AHA**	Tat	[S62]
S. coe.	bc_1 [Fe-S] subunit	**K**ESVI**GRRKLIR**NTMLGALTLVPLSGVV	Tat	[S63]
S. liv.	FKBP	M**RRRSLL**IAVPTGLVTLAA**C**G	Tat	[S64]
S. liv.	Subtilisin inhibitor	M**R**NTA**R**WAATLGLTATAV**C**GPLAGASLASPAT**APA**	Sec	[S57]
C. glu.	TorA, MTG export	MNNNDLFQA**SRRRFL**AQLGGLTVAGMLGPSLLTPRR**ATA**	Tat	[S41]
C. amm.	CsbA, MTG export	M**KR**M**K**SLAAALTVAGAMLAAPV**ATA**	Sec	[S40]

Sml-1 were fused to Sec-dependent pelB, TAMP export required the native signal peptide to yield the functional enzyme. The TAMP signal peptide includes like *Sm*MTG a Tat recognition sequence (RRAVA) suggesting that both enzymes prefer intracellular protein folding and membrane passage through the TatABC channel (Table 1). However, unlike many proteins from *S. coelicolor* and *S. scabiei* [S62,S65–S67], Tat preference of *Sm*MTG and TAMP is not yet proven, and only the published production procedures suggest the usefulness of native MTG signal peptides. For instance, *Sc*MTG [S44] and *Sh*MTG [S46] are secreted with high yields if the full-length genes were used to transform *S. lividans*. Moreover, overexpression of TatABC transport proteins significantly increased *Sm*MTG secretion in *Corynebacterium glutamicum* (Table S3) when the zymogen was fused with Tat-dependent TorA rather than Sec-dependent CspA signal peptide [S41].

Almost all MTG proteins further contain a single, highly conserved cysteine in the sequence of the signal peptide (Fig. S4). These cysteines can be modified by lipids to anchor proteins to the cell membrane. However, MTG enzymes lack the essential lipobox for lipoprotein diacylglycerol transferase binding [56] and are secreted without a doubt. We hypothesize that, besides the propeptide, a disulfide bridge between the catalytic and signal peptide cysteines keeps folded MTG inactive until Tat translocation is complete.

Significance of the propeptide and activation of transglutaminase

As emphasized, propeptides play an important role in protein folding and inhibition of the *Streptomyces* transglutaminases. While the inactivation peptide of *Sm*MTG consists of 45 amino acids, 3–11 additional residues are inserted between Ala-65 and Ala-66 (protein numbering with signal peptide) in most of the other variants (Fig. S4). The N termini of the pro- and mature domains are characterized by acidic residues that most likely prevent proteolytic degradation by aminopeptidases such as TAP. Removal of the first propeptide residues (^{30}ASGGDG^{35}) reduced the secretion of *Sh*MTG by *E. coli*, and further truncation by ^{36}EREGSYAETH^{45} yielded insoluble protein aggregates [S43]. The removed decapeptide includes the conserved SYAET stretch, which covers as a small α-helix the active site of *Sm*MTG from the rear vestibule (Fig. S2A) [57]. In addition, Tyr-42 (*Sh*MTG Tyr-41) and Tyr-46 (only available in *Sm*MTG) were shown to be essential for *Sm*MTG inhibition and folding [57]. The active site of *Sm*MTG is further covered by a long α-helix that forms a L-shaped structure with the small

α-helix. By this means, the conserved sequence motif ^{56}INALN^{60} fills the space to the catalytic amino acids by Ile-56, Asn-57, and Asn-60. The structure of *Sm*MTG between Pro-64 and Pro-76 (including the extension peptides of other MTG variants) remains elusive due to the high flexibility of this region. The characterized binding site for activating proteases, ^{73}FRAP^{76}, is largely shared by other transglutaminases, even though some modifications occur by amino acid deletion or substitution (Fig. S4).

The *in vitro* activation, which is essential for many recombinant production procedures, was already demonstrated with the discovery of the *Sm*MTG zymogen [30]. Bacillolysin from *Paenibacillus polymyxa* (trade names Dispase, Gentlyase), a metalloprotease of the thermolysin family M4 with P1′-specificity, cleaves the peptide bond between Ser(−5) and Phe(−4) without obvious degradation of the active enzyme (Table S4). The same peptide bond is hydrolyzed by the highly specific transglutaminase-activating metalloprotease (TAMP) from *S. mobaraensis* [32]. However, it should be noted that both proteases, Dispase and TAMP, truncate the oligohistidine tag at the C terminus of recombinantly produced *Sm*MTG [53,S69]. Phe-(−4) and Arg-(−3) of the protease-binding motif further enables MTG activation by chymotrypsin, trypsin, and other serine proteases, but degradation and loss in MTG activity were observed in many cases (Table S4). It would be of some interest to investigate whether the serine protease, discovered in *S. hygroscopicus* [33], activates the zymogen as specifically as TAMP and is inhibited by SSTI. However, this project requires the production of a series of serine proteases from *S. mobaraensis* or other MTG-producing *Streptomyces* variants in *E. coli* as purification from culture supernatants seems to be difficult due to low protease concentrations [33].

The protease-binding site was further modified to produce structurally homogenous MTG with the mature N-terminal peptide [58,S38,S49] (Table S4). In addition, specific activation was targeted by the introduction of a rhinovirus C3 cleavage site [S70], linker peptides [S71], intein [S53], the split production of propeptide and mature domain [S51], and automaturation [S54]. Whether specificity and efficiency of TAMP and Dispase were achieved by these approaches cannot be estimated due to the lack of comparative experiments.

Tertiary structure of microbial transglutaminase

The MTG structures yet determined accommodate four (mature *Sm*MTG), one (*Sm*MTG zymogen), or two molecules (inhibited *Sm*MTG) per asymmetric crystal unit [17,57,59]. N-terminal extension by methionine or deletion of Asp-1 disturbs the crystal package of the mature enzyme [17]. *Sm*MTG adopts a disk-like structure shaped by 8 anti-parallel β-strands and 12 α-helices, which surround the hydrophobic protein core of 7 β-sheets (Fig. 2A). A cleft of ~ 16 Å depth includes the catalytic triad consisting of Cys-64, Asp-255, and His-274 (mature enzyme numbering). The orientation of Cys-64 and His-274 at the cleft bottom indicates that glutamine and lysine donor proteins enter the enzyme from opposite directions (Fig. S3). The front view, the *endo*-glutamine-binding site, shows on the left side the α-helix cluster $\alpha^1/\alpha^2/\alpha^3$ with Cys-64 at the top of the α^2/α^3 turn and β-strands ($\beta^1/\beta^6/\beta^7$) behind, from which the N terminus (D1-P12) and the extended β^6/β^7 loop (N276-M288) rise (Fig. 2A). These structural elements form the left cleft wall of substantial rigidity as low B factors suggest [17]. In contrast, the right cleft wall, established by the loops between β^4/β^5 (D233-N253) and α^{12}/β^8 (E300-D306), seems to be more flexible showing the highest B factor at the β^4/β^5 loop tip. Moreover, the amino acids Tyr-62, Val-65, Trp-69, Tyr-75, Ile-240, and Phe-254 determine a hydrophobic region in close vicinity to Cys-64, called the front vestibule [17], which was predicted to interact with acyl donor proteins (Fig. 2B).

*Sm*MTG in complex with the inhibitory peptide DIPIGS*KMTG (IP), which contains the irreversibly binding glutamine mimic Ser-O-CO-CH_2Cl (S*), provided deeper insights into the binding mode of glutamine donor proteins [59]. Three amino acids of IP occupy the *Sm*MTG cleft bottom, that is, IP Ser*-6 covalently bound to Cys-64 (Ser-O-CO-CH_2-S-Cys), IP Gly-5 in backbone hydrogen bond (carboxy oxygen) with the *Sm*MTG Tyr-278 amide, and IP Ile-4 (Fig. 2C). A second hydrogen bond is formed between the IP Ser*-6 carboxy oxygen and the *Sm*MTG Cys-64 amide. These observations are consistent with an earlier docking model, which predicted H-bonds between *Sm*MTG Cys-64 or Val-65 amide and the tetrahedral *cbz*QG oxyanion intermediate as well as between the *Sm*MTG Tyr-278 amide and the *cbz* urethane oxygen [60]. In both cases, the N-terminal peptide moiety, that is, IP ^{1}DIPIG^5 and the *cbz* protecting group, faces the front vestibule.

The IP Ile-4 side chain further fills the hydrophobic pocket shaped by *Sm*MTG Tyr-62, Val-65, and Phe-254 of the predicted accommodation area for glutamine donor substrates [17] (Fig. 2C). Alanine substitution of Tyr-62 only revealed a small effect on *Sm*MTG-mediated *cbz*QG hydroxamate formation, while activity of the V65A and F254A exchange

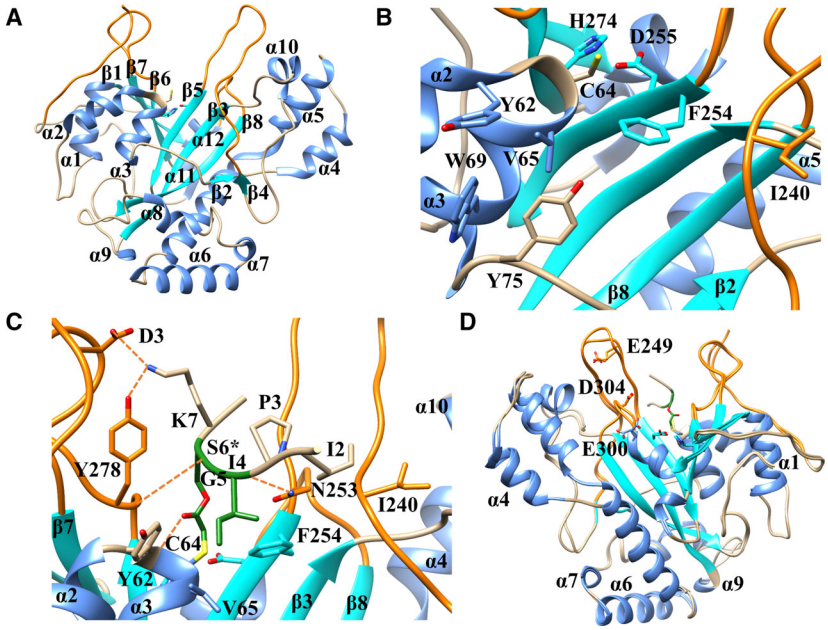

Fig. 2. Functional amino acids of microbial transglutaminase from *Streptomyces mobaraensis*. α-Helices, β-strands, and the cleft wall loops are colored cornflower blue, turquoise, and orange. Catalytic and predicted functional amino acids are shown in a stick representation. The secondary structure numbering follows PDB entry 1IU4. (A) Front view of *Sm*MTG; (B) predicted front vestibule consisting of Tyr-62, Val-65, Trp-69, Tyr-75, Ile-240, and Phe-254 [17]; (C) *Sm*MTG in complex with DIPIGS*KMTG (6GMG) [59]. The amino acids of the inhibitor, which occupy the cleft bottom of the enzyme, are colored forest green. Hydrogen bonds are depicted by orange dashed lines; (D) rear view and predicted rear vestibule of *Sm*MTG [17,59]. The structures of *Sm*MTG (1IU4) and inhibited *Sm*MTG (6GMG) are superimposed.

variants was reduced almost to zero [60]. The IP termini, which rise to the cleft entrance, also interact with the enzyme. Hydrogen bonds are established by the ε-amino group of IP Lys-7 with both the *Sm*MTG Asp-3 δ-carboxylate and the Tyr-278 hydroxyl group as well as the IP Pro-3 backbone carboxy oxygen with the *Sm*MTG Asn-253 δ-carboxamide. Moreover, orientation of IP Ile-2 to *Sm*MTG Ile-240 suggests hydrophobic interactions between the residues. In total, seven *Sm*MTG amino acids (Asp-3, Tyr-62, Cys-64, Val-65, Ile-240, Asn-253, Tyr-278) contribute to the IP binding. The side chain of Asn-253 at the entrance to Cys-64 could also act as a guide for the entering γ-carboxamide group of substrate glutamines as alanine substitution yielded complete *Sm*MTG inactivation, but this hypothesis has to be verified [60].

In *Sm*MTG, Asp-255 may be more involved in proton transfer than His-274. The distance of Cys-64 to Asp-255 is smaller, and substitution of Asp-255 with alanine reduced activity to background levels [17,60]. Apart from Cys-64 and Asp-255, the exhaustive replacement of *Sm*MTG amino acids by alanine further showed that Gly-63, Asn-253, Phe-254, Tyr-256, and His-277 are essential for maintaining the catalytic activity [60]. In contrast, the H274A exchange variant still exhibits residual activity of 40% and 90% toward small (*cbz*QG) and large (ovalbumin) substrate molecules, respectively, thus demonstrating the minor role of His-274 for *Sm*MTG catalysis. His-274 is thought to be involved in the orientation of the lysine donor substrate at the back of *Sm*MTG, which is also referred to the rear vestibule [17]. This binding site is lined by negatively charged amino acids (Glu-249, Glu-300, Asp-304), which are integrated into the more flexible β^4/β^5 and α^{12}/β^8 loops (Fig. 2D).

Remarkably, the covalent binding of the inhibitory peptide barely influenced the positions of the catalytic amino acids. The rigid cleft wall also remained unchanged by the interaction with IP Lys-7 (Fig. 2D, right side). In contrast, IP binding moved the flexible β^4/β^5 loop between Pro-244 and Phe-251 to the front vestibule (Fig. 2D, left side).

Functional amino acids influencing transglutaminase activity and thermostability

The influence of the N-terminal peptide (D1-G25) on enzyme activity was recognized at a time when the tertiary structure of *Sm*MTG was revealed [17,61,62]. Incorporation of ^{15}N-labeled NH_4Cl into food proteins

showed increased $^{14}N/^{15}N$ exchange rates when *Sm*MTG was truncated by Asp-1 [62]. Asp-1 seems to be flexible (*cf*. Fig. 2D) and to have some influence on the catalytic cysteine as NMR difference spectra of Met-*Sm*MTG and Δ^{Asp}-*Sm*MTG suggested [63]. Further removal of up to five amino acids from the N terminus proved that elimination of Asp-1, Ser-2, and Asp-3 successively increases enzyme activity without changing substrate specificity, but further shortening by Asp-4 and Arg-5 yielded inactive *Sm*MTG variants [63]. Neither small (*cbz*QG) nor large glutamine donor substrates (ovalbumin) were modified by the Δ^{1-4}- and Δ^{1-5}-*Sm*MTG variants, so the authors suspected the strong impairment of the protein structure. Noticeably, the replacement of Ser-2 by arginine (S2R), aspartate (S2D), and tyrosine (S2Y) revealed great differences in small and macromolecule binding by *Sm*MTG. While all *Sm*MTG variants mediated the formation of *cbz*QG hydroxamate, S2D and S2Y failed to incorporate ^{15}N-labeled NH_4Cl into ovalbumin [63].

In the following decades, great efforts were made to increase thermostability and activity of *Sm*MTG by bioengineering. The first random mutagenesis approach selected heat-resistant and heat-sensitive *Sm*MTG variants from a library of 5500 *E. coli* clones [64]. The most effective *Sm*MTG variants were point-mutated in the N-terminal stretch upstream of Gly-25 (Fig. S5A), increasing the half-life of FRAP-*Sm*MTG from 1.7 min (wild-type) to 4.6 min (S2P) at 60 °C. Heat-sensitive positions were also found in the terminal peptide (P9L, A10T, G25S). The combination of several hot spots then yielded the S23V/Y24N/K294L and S2P/S23Y/Y24N/H289Y/K294L *Sm*MTG exchange variants, which showed significantly higher heat-resistance compared with the wild-type enzyme at 60 °C [65,S72].

The already mentioned alanine screen further revealed the great influence of various amino acids on *Sm*MTG activity against small (*cbz*QG) and large (ovalbumin) substrate molecules [60] (Fig. S5B). While alanine substitution of Gly-63, Cys-64 (catalytic), Asn-253, Phe-254, Asp-255 (catalytic), Tyr-256, and His-277 resulted in inactive variants [residual activity of < 5% (*cbz*QG) and < 2% (ovalbumin)], reduced activity with ovalbumin was observed for Arg-26 (18%), Val-65 (10%), Tyr-75 (5%), Val-252 (6%), His-274 (9%), Asn-276 (2%), and Phe-305 (19%) mutations. A more rational approach including a library of 24 000 clones then uncovered Y75F, S199A, and M16T as the most effective *Sm*MTG exchange variants with 1.5–1.7-fold *cbz*QG transamidation activity [66]. However, benefit of the S199A variant was lost when mature *Sm*MTG and not FRAP-*Sm*MTG was mutated. Recently, the combination of disclosed *Sm*MTG hot spots confirmed earlier results but could not contribute to the further improvement of thermostability and specific activity [S73]. Moreover, a huge library of *Sm*MTG zymogen variants was generated by error-prone PCR [58]. Screening of $3x10^8$ clones by yeast surface display showed that MTG variants exhibiting significantly enhanced enzyme activity could not be produced by this comprehensive approach.

Microbial transglutaminase was further mutated for the site-specific modification of human growth hormone (*h*GH) and monoclonal antibodies (mAb) [67,68]. The Y62A exchange variant of *Sl*MTG showed enhanced PEGylation activity and improved *h*GH Gln-141 selectivity [67]. The complete transamidation of glycosylated mAb heavy chain Gln-295 was achieved using the *Sm*MTG exchange variants D4E/G250S and A212S/G250S [68]. However, enzyme amounts of 100 U/ml and incubation for 40 h are rather unusual reaction conditions favoring glutamine deamidation.

Illuminating structure of *Sm*MTG by reassessment of the available data

Great efforts were made to increase *Sm*MTG activity and selectivity by mutation, but the results remained poor. *Sm*MTG seems to be an enzyme that has matured in evolution. In our opinion, the available crystal structures [17,59] and the alanine screen results [60] shed sufficient light on essential *Sm*MTG amino acids involved in catalysis (Fig. 3), which will be briefly summarized here to conclude the section.

*Sm*MTG contains like other transglutaminases the amino acids Cys-64, Asp-255, and His-274 at the bottom of the catalytic cleft [17] (Fig. 3A). The inhibitory peptide (IP), DIPIGS*KMTG [59], is linked to the front vestibule of *Sm*MTG (B–D, note the edited warhead) and provides evidence for two subsites, which may select glutamine donor proteins and peptides (G, H). Essential *Sm*MTG amino acids, uncovered by the individual alanine replacement [60], are ^{63}GCV^{65} (C), ^{252}VNFDY^{256} (D), and ^{274}HGNH^{277} (E,F). Asn-253 may be involved in glutamine orientation to support the nucleophilic addition of Cys-64 as mentioned.

The Cys-64 thiolate most likely attacks the glutamine γ-carboxamide carbon from the *re*-side (*cf*. Fig. S3) as the γ-carboxy oxygen forms a H-bond with the Cys-64 amido hydrogen [59]. Asn-276/His-277 and Val-252 further separate glutamine and lysine-binding sites by forming a tight bottleneck (E). Another barrier consisting of conserved Arg-26/Tyr-302/Phe-305 probably completes the lysine pocket (F). It restricts the access of lysine donor proteins from the rear vestibule and suggests amino group insertion from the cleft entrance.

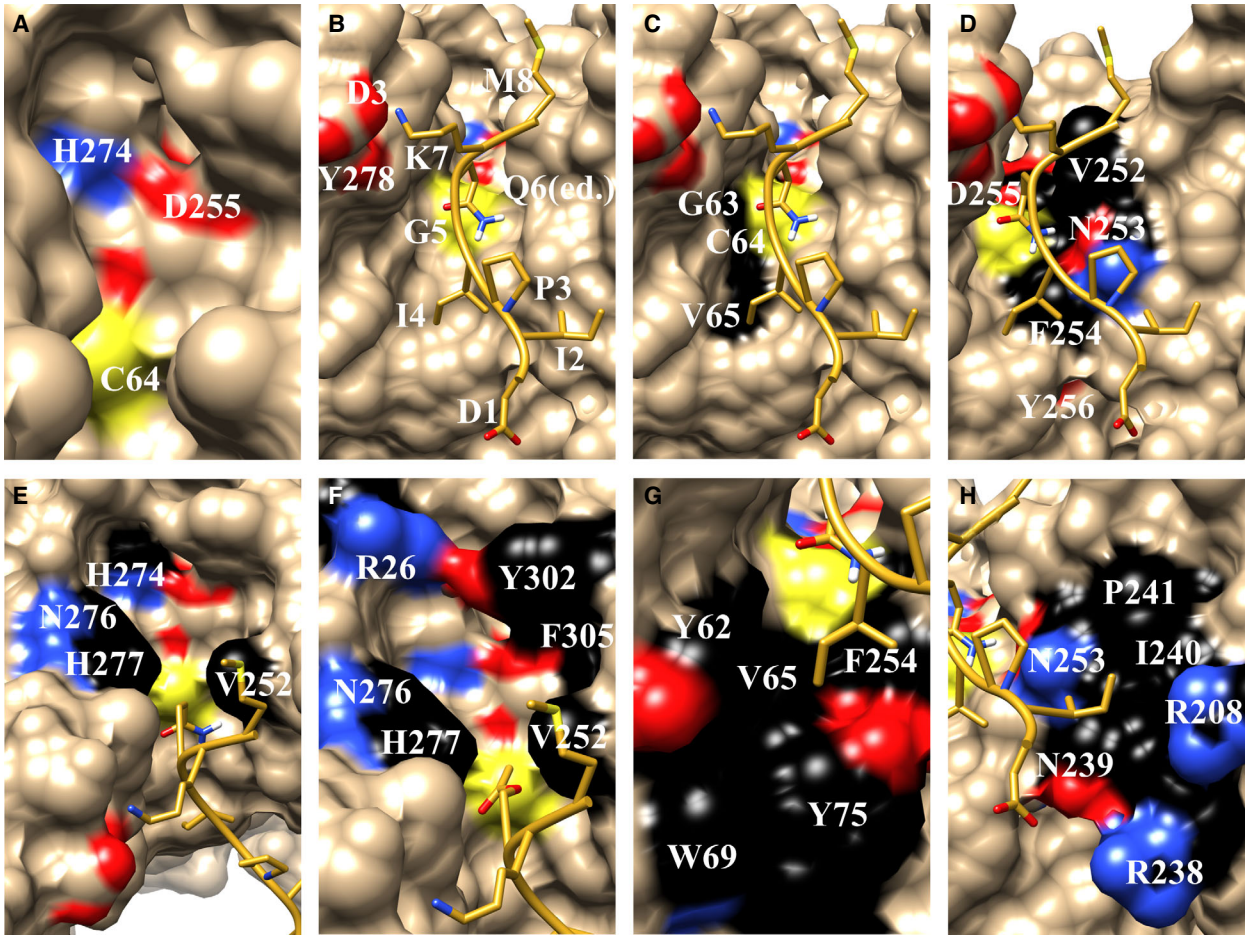

Fig. 3. Amino acids presumably involved in *Sm*MTG catalysis. (A) Catalytic triad of *Sm*MTG (PDB 1IU4); (B) inhibited *Sm*MTG (6GMG) with modified warhead (exchange of Ser*-6 for Gln); (C, D) *Sm*MTG amino acids involved in inhibitor binding; (E) amino acids separating glutamine and lysine-binding sites; (F) amino acids shaping the putative lysine pocket; (G) amino acids forming the *Sm*MTG subsite G2 (nomenclature proposal see text); (H) amino acids forming the *Sm*MTG subsite G4.

Subsites providing for substrate selectivity seem to be only available at the front vestibule (G,H). Tyr-62, Val-65, Trp-69, Tyr-75, and Phe-254 form a hydrophobic hole that accommodates IP Ile-4 [59] (G). A second *Sm*MTG subsite shaped by ^{238}RNIP^{241}, Asn-253, and Arg-208 interacts with the IP Ile-2 (H). The positively charged arginine residues also suggest interaction of the subsite with acidic amino acids.

Following the protease nomenclature of Schechter and Berger [69], we propose to refer to the glutamine accommodation subsites as G1, G2, G3, *etc*., corresponding to G2 and G4 of inhibited *Sm*MTG. Binding of glutamine donor proteins to the rear vestibule, which would require G1′, G2′, G3′ *etc*., nomenclature, can be excluded. Amine subsites (if any) are named A1, A2, A3, *etc*., correspondingly. In comparison, orientation of the substrate proteins in two directions is possible, indicating designations such as …T3′-T2′-T1′-Q-T1-T2-T3… and …T3′-T2′-T1′-K-T1-T2-T3… This means that Ile-2 and Ile-4 are the T4′ and T2′ positions of the inhibitory peptide docking to the subsites G2 and G4 of *Sm*MTG.

Interestingly, the *K. albida* transglutaminase (Fig. S2D) shares nearly the same structure with mature *Sm*MTG despite low sequence identity of 24.9% [25]. Essential stretches such as ^{45}GCI^{47} (*Sm*MTG: ^{63}GCV^{65}), ^{172}VNFDY^{176} (^{252}VNFDY^{256}), and ^{188}HANH^{191} (^{274}HGNH^{277}) are strictly conserved. The amino acids that complete the lysine pocket are even the same: Arg-208 (Arg-26) and Tyr-14 (Tyr-302). Moreover, it seems likely that glutamine selectivity is provided by two subsites around Ile-47 (*Sm*MTG-G2: Val-65) and Tyr-165 (*Sm*MTG-G4: Ile-240).

A comparison of *Sm*MTG (Fig. 3, Fig. S6A) with mammalian and fish transglutaminase shows glutamine and lysine-binding sites similarly separated by tight bridges of tryptophan residues (Fig. S6B–D). The resulting eyelets are sealed by the catalytic cysteine in the 'closed' structures of human TG2 and FXIII (Fig. S1, Fig. S6B,C). The fact that the access of acyl donor and acceptor proteins are strictly compartmentalized by conserved amino acids is best illustrated by fish TGase. In addition to the Trp-236/Trp-329 barrier, an overlying salt bridge between Glu-237 and Arg-324 is likely to provide for the individual threading of glutamine and lysine sidechains from the cleft entrance (Fig. S6D).

Since the catalytic core of transglutaminases is similar, self-assembly of the substrate proteins is hypothesized that precedes cross-linking and enables the simultaneous uptake of the glutamine and lysine residues by the enzymes, especially by *Sm*MTG (Fig. S7).

Structural requirements for transglutaminase substrates

The study of transglutaminases over several decades has shown that there is no consensus sequence at either the glutamine or lysine-binding sites of substrate proteins. While glutamines determine selectivity of the enzymes, the requirements for lysine residues and primary amines are comparably low. Significant differences in glutamine recognition by *Sm*MTG caused us to consider peptides, pharmaceutical proteins, and natural proteins separately.

Requirements for primary amines

Transglutaminases only incorporate primary amines into glutamine donor proteins and peptides, showing high affinity for histamine and glycine ethylester as examples [70,S1]. Polar functions restrict conjugation with *endo*-glutamines, especially if they are contiguous to the amino group. Several studies demonstrated the importance of aliphatic, lysine-like spacer groups, the stereochemical preference for L-lysine, and the low significance of the lysine sequence environment. Hydrophobic residues at the N-side of lysine appear to improve affinity of *g*TG and FXIIIa for the heptapeptide *ac*GGLKGGG compared with *ac*GGGK GGG [71,S74]. Large aromatic substituents attached to alkylamine hydrocarbon chains of about 7.2 Å length further support the incorporation of amines [72]. A good example may be monodansylcadaverine, which is used for protein labeling and activity measurement [S2,S3].

Similar investigations with microbial transglutaminases showed the same influence of spacer, polar groups, and branching on amine preference [73,S26,S75]. Small amino acids (Gly, Ala) at the N- or C-sides were preferred to bulky or charged residues to ligate lysine dipeptides to citraconylated α_{S1}-casein [74]. Furthermore, study of the trypsin inhibitor from *Streptomyces longisporus*, STI2, whose inhibitory and cross-linking site, Lys-70, is in a flexible loop, shows that the exchange of Ile-67 (I67S) and Asn-69 (N69R) had the most favorable effect on *Sm*MTG activity [42] (Table S5). It should be noted, however, that these mutations could impair formation of the loop-stiffening Cys-68/Cys-98 bridge. STI2-like effects of adjacent amino acids on lysine donor sites were not observed in other proteins (*cf.* Table S6). Interestingly, the STI2 stretch ^{67}ICNKL^{71} shares nearly the same sequence motif with the C-terminal peptide ^{119}LCEKL^{123} of de-calcinated α-lactalbumin (apo-α-LA), which has been identified as an acyl acceptor site for *Sm*MTG [75]. Also, in this case, the removal of Ca^{2+} could have an impact on the Cys-6/Cys-120 bridge stability. The same authors [75] further showed that *Sm*MTG has access to Lys-96 and Lys-98 of heme-less apo-myoglobin (apo-Mb). The Mb ^{96}KHK^{98} motif was frequently used as lysine donor (tag) sequence in MTG-mediated conjugation reactions (*cf.* Table 3 and Table S6).

Glutamine peptide sequences for microbial transglutaminase

Motoki and colleagues already reported in their first study that MTG from *Streptoverticillium* S-8112 recognizes the dipeptide *cbz*QG as glutamine donor substrate like guinea pig liver *g*TG [15]. Q/N exchange or peptide shortening completely abolished binding to *Sm*MTG (Table 2). The carboxy extension by glutamine, glycine, or an ester function improved *Sm*MTG affinity, but glycine insertion at the glutamine N-side had the opposite effect. Significance of the N-side was further demonstrated by G_3QG_3-based heptapeptides [76]. The reference molecule, G_3QG_3, was not a substrate of *Sm*MTG, but substitution of glycine by hydrophobic and polar groups in the positions −1 and −2 achieved 30–55% of the *cbz*QG transamidation rate (Table 2). Similarly, *N*-acetylated tripeptides obtained by combinatorial synthesis could not exceed *cbz*QG efficiency [77].

The requirement for glutamine-binding sites was further studied by *Sm*MTG-mediated labeling of threadfin bream myosin peptides [78] and huge peptide libraries obtained by phage [79] and mRNA [80] display (Table 2). Reference to *cbz*QG cannot be

Table 2. Glutamine donor sequences of MTG substrate peptides. MDC, monodansylcadaverine; MBC, monobiotin cadaverine.

Sequence position													
−5	−4	−3	−2	−1	Gln	+1	+2	+3	+4	+5	Acyl acceptor	Efficiency (%)	Ref.
				Cbz	Q	G					H_2NOH	100	[15]
				Cbz	Q	G	OEt					122	
				Cbz	Q	Q	G					288	
			Cbz	G	Q	G	G					126	
		Cbz	G	G	Q	G						27	
				Cbz	Q	G					MDC	100	[76]
				Cbz	Q	F						86	
		G	G	G	Q	G	G	G				0	
		G	L	G	Q	G	G	G				55	
		G	G	L	Q	G	G	G				35	
				Cbz	Q	G					MDC	100	[77]
			Ac	Y	Q	W						87	
			Ac	Y	Q	R						87	
			Ac	P	Q	R						111	
E	G	S	L	E	Q	E	K				MBC		[78]
			E	L	Q	A	R						
	E	A	E	F	Q	K							
			N	V	Q	G	Q	L	K				
		V	A	E	Q	E	L	V	D	ASER			
E	V	D	A	E	Q	R							
		Y	E	L	Q	R	P	Y	H	SELP	MBC		[79]
		W	A	L	Q	R	P	Y	T	LTES			
		W	A	L	Q	R	P	H	Y	SYPD			
		Y	A	L	Q	R	P	Y	H	SELP			
			R	L	Q	Q	P				Lys_6-beads		[80]
	Cbz	F	G	L	Q	S	P	Y	NH_2		*Ac*KAYANH_2	> 90[a]	[81]
	Cbz	F	G	L	Q	G	P	Y	NH_2			> 90[a]	
	Cbz	F	G	L	Q	R	P	Y	NH_2			0[a]	

[a]Product formation (%) after incubation for 40 min.

made as comparative measurements were not performed. Interestingly, the occurrence of glutamate and hydrophobic amino acids upstream from the *Sm*MTG-accessible myosin peptide glutamines corresponded with those of the most reactive heptapeptides G(L/V)GQG_3 and G_2(L/E/V)QG_3 [76,78]. Furthermore, the YELQRPYHSELP peptide from phage display has a similar upstream sequence, but glutamate substitution by alanine improved affinity of *Sm*MTG [79]. The common sequence motif, ALQRP, suggests the interaction of Ala-Leu with the hydrophobic subsite G2 (at Val-65) of *Sm*MTG, while Arg(+1) may form hydrogen bonds with Asp-3 and Tyr-278 (*cf.* Fig. 2C). Therefore, ALQRP may have the same orientation as the *Sm*MTG inhibitory peptide [59], which faces the front vestibule with the amino terminus. In contrast, the most effective peptide of the mRNA library, RLQQP, is more likely to occupy *Sm*MTG in opposite direction [80]. Recently, a combinatorial synthesis based on *cbz*FGX^1QX^2P$Y NH_2$ [81] showed that GX^1QRP-containing peptides, among them *cbz*FGLQRP$Y NH_2$ (Table 2), cannot be linked to the lysine tetrapeptide *ac*KAYANH_2 by MTG, which contradicts the results of Hitomi *et al.* [79]. The efficiency of *cbz*FGLQGP$Y NH_2$ and *cbz*FGLQSP$Y NH_2$ could be explained by the assumption that MTG bind these peptides in the N-C and C-N direction equally. Other reasons are plausible because the used MTG was not specified.

Nevertheless, usefulness of peptide sequences such as ALQRP, ALQSG, and ALQAY has been proven in the *Sm*MTG-mediated synthesis of cyclic peptides [82]. Furthermore, besides *cbz*QG [S76–S79], LLQG [S80–S85], and EAQQIVM [S86,S87], proven tags for mammalian transglutaminases [S88], WALQRPH was fused to proteins without accessible glutamines [S89,S90], thus allowing *Sm*MTG-mediated labeling, immobilization, and conjugation reactions (Table S6). Moreover, strep tag I, modified by Tyr-Lys (YKAWRHPQFGG) and other amino acids, was shown to be a useful acyl donor and

acceptor tag in *Sm*MTG-mediated protein cross-linking reactions [S91,S92].

Glutamine donor sequences of globular proteins

The high specificity of microbial transglutaminase has caused many researchers to investigate the occurrence of glutamine donor sites in pharmaceutical proteins and other proteins of interest. After many years of research with not always consistent results (Table 3), it is sobering to reconfirm the lack of glutamine consensus sequences for *Sm*MTG. The screening for glutamine donor proteins is like looking for a needle in a haystack, and every prediction seems to be arbitrary. However, the studies succeeded in discovering some general findings.

Glutamines, accessible for MTG, are thought to occur in flexible, disordered regions, which was originally shown for Gln-74 of human interleukin 2 (*h*IL-2) (Table 3, Fig. 4A) and predicted by Kashiwagi *et al.* [17,83]. The concept was then extended by the assumption that local unfolding of substrate proteins may be a general prerequisite for the interaction with proteases and transglutaminases [75]. As with glutamine peptides, the glutamine environment of proteins further determines the affinity of MTG. However, in protein substrates studied with *Sm*MTG, charged and bulky residues occupy both sides of accessible glutamines (Table 3). While N-C or C-N orientation remained largely undetermined, the *Sm*MTG preference for peptide sequence motifs (Table 2) gives some indication on the binding direction.

Like WALQRPHYSYPD [79], the donor sites of *h*IL-2 (Gln-74) [83], de-calcinated or reduced bovine α-lactalbumin (αLA, Gln-43, Gln-54) [75,84–87], hemeless apo-myoglobulin (apo-Mb, Gln-91, Gln-152) [75], or thermolysin (Gln-225, Gln-246) [75] contain uncharged, hydrophobic amino acids at the N-side of glutamine (Table 3). In contrast, amino acids of αLA (Gln-39, Gln-43, Gln-54, Gln-65), human granulocyte

Table 3. Glutamine and lysine donor sites of globular proteins for MTG binding. Glutamine and lysine residues are printed bold. *h*IL-2, human interleukin 2; α-LA, bovine alpha-lactalbumin; *h*Mb, horse myoglobin; *h*G-CSF, human granulocyte colony-stimulating factor; *h*GH, human growth hormone; mAb, monoclonal antibody; HC, heavy chain; *h*IFN, human interferon.

Protein	Gln no.	Accessible glutamines	Glutamine sequence	Lys no.	Accessible lysines	Lysine sequence	Ref.
hIL-2	6	Q74	VLNLA**Q**SKNFH	10			[83]
α-LA	6	Q39	SGYDT**Q**AIVQN	12	K5	EQLT**K**CEVFR	
		Q43	TQAIV**Q**NNDST		K13	VFREL**K**DLKGY	
		Q54	EYGLF**Q**INNKI		K16	ELKDL**K**GYGGV	
		Q65	WCKDD**Q**NPHSS		K58	FQINN**K**IWCKD	
		Q117	SEKLD**Q**WLCEK		K108	YWLAH**K**ALCSE	
					K114	ALCSE**K**LDQWL	
					K122	QWLCE**K**L	
Apo (Ca^{2+}-free)		Q54			3 K		[84]
Apo, 50 °C		Q39,43,54,65,117			K5,13,16,108,114		[85]
Apo, 30 °C		Q39,43			K13,16,108,114		[75]
Apo		Q39 > Q43,54,65			K16,58,114,122		[86]
Holo, DTT		Q54			K5		[87]
Holo, GSH		Q39,43			K13,16		
*h*Mb	6	Q91	LKPLA**Q**SHATK	19	K96	QSHAT**K**HKIPI	[75]
		Q152	KELGF**Q**G		K98	HATKH**K**IPIKY	
Thermolysin (205–316)	7	Q225	RYTGT**Q**DNGGV	7	K219	PDHYS**K**RYTGT	[75]
		Q246	AYLIS**Q**GGTHY		K210	MSDPA**K**YGDPD	
*h*G-CSF	17	Q134	ALQPT**Q**GAMPA	4			[88]
*h*GH	13	Q40	YIPKE**Q**KYSFL	9	K145	KQTYA**K**FDANS	[89]
		Q141	GQAFK**Q**TYAKF				
Calcitonin	2	Q20	ELHKL**Q**TYPRT	2			[90]
mAb, HC (deglycos.)	18	Q3	VKVSC**Q**ASGYR	35			[91]
		Q295	KPREE**Q**YNSTY				[92]
					K340	TISKA**K**GQPRE	[93]
					K447	SLSPG**K**-aa* (*aa not D,E,P)	[94]
*h*IFN α-2b	12	Q101	EACVI**Q**GVGVT	11	K164	ESLRS**K**E	[95]
Avidin	4	(Q126)	TRLRT**Q**KE	9	K58	ENTIN**K**RTQPT	[96]
					K127	RLRTQ**K**E	

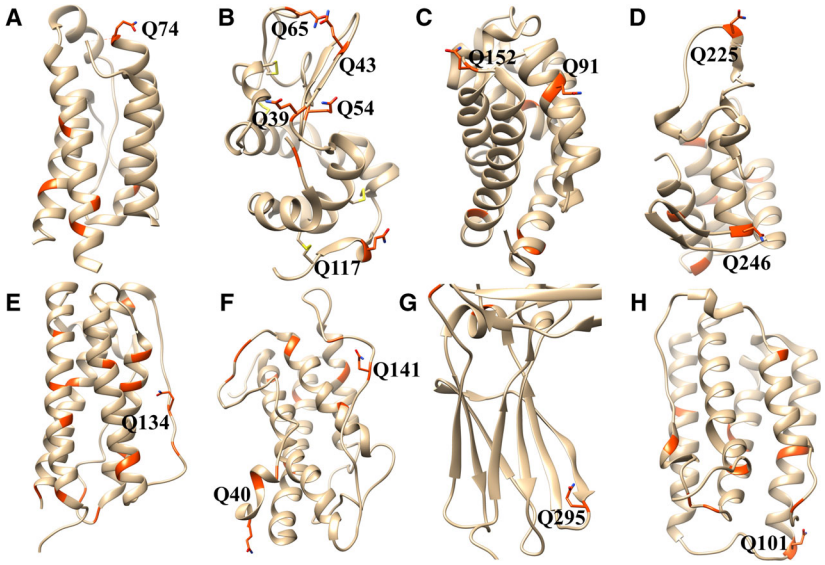

Fig. 4. Glutamine donor sites of globular proteins for *Sm*MTG binding. The glutamine residues are colored orange red, the determined binding sites are labeled and depicted in a stick representation. (A) Human interleukin 2 (PDB 1M49); (B) de-calcinated bovine α-lactalbumin (1F6R); (C) horse myoglobulin (1AZI). The porphyrin moiety is removed; (D) C-terminal domain of thermolysin (2TLX); (E) human granulocyte colony-stimulating factor (5GW9); (F) human growth hormone (1HGU); (G) part of human antibody heavy chain (1HZH); (H) human interferon α-2b (2HYM).

colony-stimulating factor (*h*G-CSF, Gln-134) [88], human growth hormone (*h*GH, Gln-141) [89], calcitonin (Gln-20) [90], the heavy chain of human monoclonal antibodies (*h*mAb-HC, Gln-295, Gln-3) [91–94], and human interferon α-2b (*h*IFNα, Gln-101) [95] are more likely to interact at the glutamine C-side with the subsites of *Sm*MTG. A few glutamines (Gln-117 of apo-αLA, Gln-40 of *h*GH, Gln-126 of avidin) [85,89,96] are surrounded by charged amino acids or bridged cysteines that usually interrupt *Sm*MTG binding (Table 3). In fact, these residues are considered as minor glutamine donor sites of *Sm*MTG, but a dearth of kinetic measurements under the same conditions makes the comparison difficult. Almost all characterized glutamines are in large, disordered loops (Gln-65 of apo-αLA, Q152 of *h*Mb, Q134 of G-CSF, Gln-141 of *h*GH) or at the border between loop and secondary structure (Q74 of *h*IL-2, Q40 of *h*GH, Q295 of *h*mAb, Q101 of *h*IFNα) (Fig. 4). The modification of helical Gln-91 of horse apo-Mb by *Sm*MTG could be explained with the removal of the porphyrin molecule. In any case, rigid structures appear to restrict *Sm*MTG access to the glutamines of globular proteins.

Although the acyl donor sites of several globular proteins were determined (Table 3), only a single tag sequence, helical ^{88}PLAQSH^{93} of *h*Mb was fused to anti-hen egg-white lysozyme antibody, alkaline phosphatase, or *E. coli* biotin ligase [S93–S96] (Table S6). The myc tag (^{410}EQKLISEEDL^{419}) of the human transcription factor myc proto-oncogene protein (*h*Myc) and the S-tag (^{32}AKFERQHMDS^{41}), a substrate-binding sequence of bovine RNaseI, were not shown to be MTG binding sites with the whole proteins [S93, S97–S99]. The helical glutamine positions and charged residues in close vicinity are likely to disturb the interaction of *Sm*MTG with *h*Myc and *b*RNaseI. Nevertheless, the discovery of unnatural *Sm*MTG glutamine donor substrates may be a big step forward in the site-specific modification of therapeutic proteins. Many reports deal with the enzymatic antibody–drug conjugation and protein PEGylation (for recent reviews see Refs [97,98]). However, due to the diversity of the identified glutamine and lysine sequences and lack of comparative measurements, the studied proteins do not provide deeper insights into the catalytic mechanism of the MTG enzymes.

Natural substrate proteins of transglutaminase from *Streptomyces mobaraensis*

The biological function of *Sm*MTG is largely unknown. Investigation of a knockout mutant from *S. hygroscopicus* revealed the requirement for transglutaminase in the development of aerial hyphae [99]. Expression of the substrate proteins and time-dependent changes in the protein level remain to be examined. The characterized substrate proteins from *Sv. mobaraense* are secreted under submerged culture conditions [41,44–46], although other *Sm*MTG substrate proteins such as the long chaplin *Sm*Chp1 and rodlin-1 (Rdl1) are likely retained in the cell wall (Table 4).

Glutamine donor sites were determined for three out of six characterized substrate proteins (Table 4). Although sequence homology is missing, hydrophobic and uncharged amino acids predominate the N-side of

the *Sm*MTG-accessible glutamines. In contrast to artificial peptides and globular proteins (Table S6 and Table 3), bulky aromatic residues are absent. Noticeably, the characterized glutamines of natural substrates are in protein regions of high flexibility (Figs. 5–7). The glutamine pair of SSTI occupies an unresolved extension (pro)-peptide close to the TAMP (*Sm*MTG-activating metalloprotease) inhibition site [43,53]. In DAIP, Gln-39 is found in a large loop ranging from Tyr-25 to His-45 [55]. Only Gln-6 of SPI_p seems to be frozen by strong attachments of the N-terminal peptide to the protein body [54]. Conformational changes, shown for the inhibitory peptide in complex with *Sm*MTG [59], may considerably enhance flexibility of the terminal SPI_p peptide, but also raise the question of why a protein should change the structure, just so the enzyme can find a binding site?

Under *in vitro* conditions, *Sm*MTG usually cross-links disordered proteins such as casein but cannot readily form the ternary encounter complex with native and natural substrate proteins, as indicated by intramolecular cross-linking or glutamine deamidation reactions [41,44,46]. To obtain cross-linked protein aggregates, substrate assembly should precede *Sm*MTG binding for several reasons: (a) A shaken culture of *S. mobaraensis* disrupts protein assembly and spore envelope formation that cell wall proteins are released; (b) detergents, which are secreted by streptomycetes [100], favor the formation of substrate protein aggregates in the presence of *Sm*MTG [41,44]; (c) the SPI_p protein is probably equipped with an N-terminal hinge region, which could facilitate access of *Sm*MTG to Gln-6 and Lys-7 [59]; (d) the catalytic center of *Sm*MTG suggests pairwise uptake of the glutamine and lysine residues as already discussed (Fig. 3); and (e) self-assembly and cross-linking by *Bs*MTG were shown for GerQ, an endospore coat protein from *B. subtilis* [21]. Therefore, if protein assembly is a prerequisite of protein cross-linking as hypothesized, the natural substrate proteins from *Sv. mobaraense* should provide more information on the interaction with *Sm*MTG through their determined function.

Is self-assembly of natural *Sm*MTG substrate proteins determined by the structure?

SSTI is a homo-dimeric multifunctional protein, whose subunits are glued together by a series of β-strands (Fig. 5A). The adjacent binding sites for *Sm*MTG (Gln-(−1)/Gln-(−2)) and TAMP (Leu-2/Tyr-3) are located at the interface of both subunits but could not be resolved due to high flexibility. The Lys-70 residue, opposite to the *Sm*MTG and TAMP-binding sites, is the serine protease inhibition and *Sm*MTG acyl acceptor site in homologous STI2 [42]. The overlay of SSTI with SSI from *S. albogriseolus* in complex with subtilisin BPN' proves that Lys-70 cannot be the glutamine acceptor site for *Sm*MTG without loss of the inhibitory activity (Fig. 5B). The opposite arrangement of both inhibition sites further implicates opposite binding of two subtilisin molecules. Assuming that Lys-70 is located at the aerial hyphae surface to parry hostile proteases, one subunit faces the cell wall. For these reasons, SSTI is likely to dissociate prior to cross-linking, which has already been observed in the inhibition of TAMP [43]. The most obvious acyl acceptor site is then Lys-34, which is flanked by two prolines and close to the tandem glutamine acyl donor site (Fig. 5C). The alternative residue, Lys-89, appears to be stiffened by strand β4 and may be more involved in substrate protein assembly.

Table 4. Glutamine and predicted lysine donor sites of natural *Sm*MTG substrate proteins. Accessible glutamines and the predicted lysines are printed bold.

Protein	Gln no.	Accessible glutamines	Acyl donor sequence	Lys no.	Predicted lysines	Acyl acceptor sequence	Ref.
SSTI	6[a]	Q(−2)	PAPAA**Q**QSLYA	8[a]	K34	LNCMP**K**PSGTH	[43]
		Q(−1)	APAAQ**Q**SLYAP		(K89)	EGQRV**K**YEHTF	
DAIP	5	Q39	TTGTL**Q**SVSYT	10	K14	APSCT**K**VTGDG	[55]
		(Q298)	YGTYF**Q**AYGTD		K160	WKPLG**K**LAFGP	
		(Q345)	GLEEV**Q**IHH				
SPI_p	3	Q6	DIPIG**Q**KMTGK	6	K7	IPIGQ**K**MTGKT	[54]
					K107	NGITW**K**FVR	
β-Lactamase	15			19			[46]
Chaplin-1	7[b]	Q227	AHRPA**Q**QAAVE	6			[105]
		Q228	HRPAQ**Q**AAVEE				
Rodlin-1	8			4			[105]

[a]Per subunit.; [b]Mature protein without glutamines of the signal peptide (1 Q) and the C-terminal peptide (2 Q), which are removed by signal peptidases and sortases, respectively.

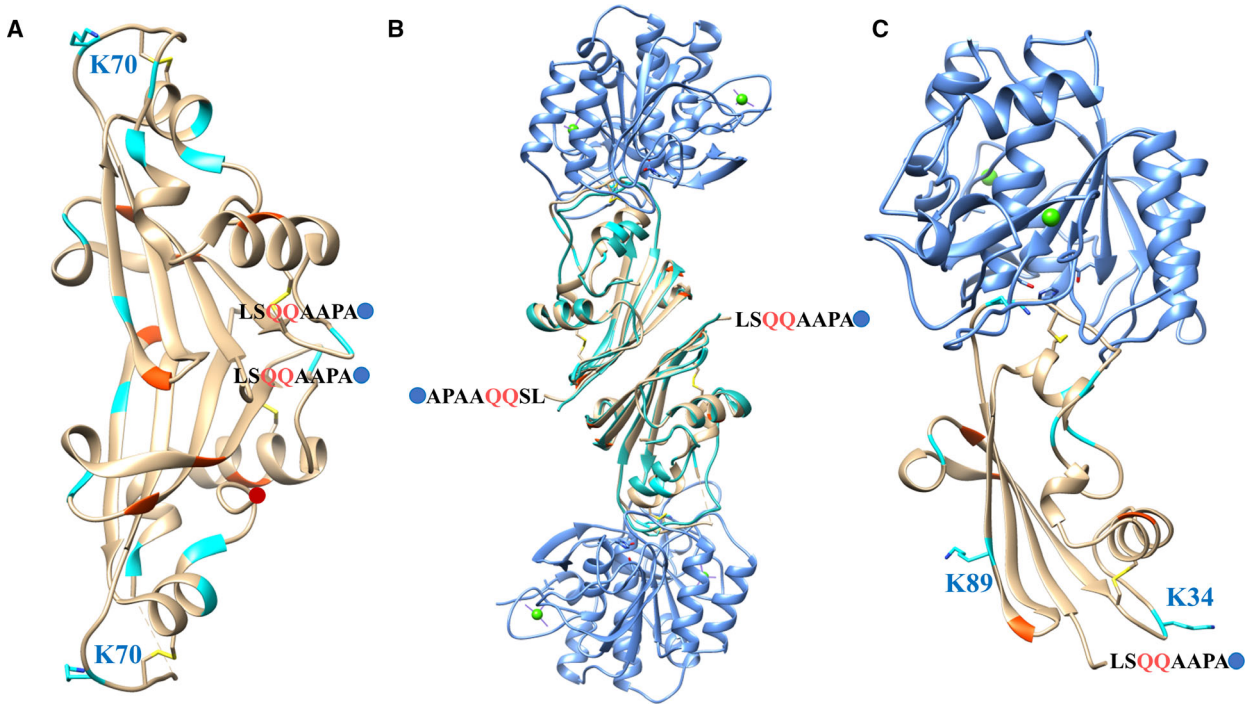

Fig. 5. Structure of SSTI from *Streptomyces mobaraensis*. Glutamine and lysine residues are shown in orange red and turquoise. The N-terminal peptides are depicted by blue and red filled circles. (A) SSTI (PDB 6I0I) showing the subtilisin/trypsin binding site and the sequence of the unresolved extension peptide for *Sm*MTG and TAMP binding; (B) overlay of SSTI from *S. mobaraensis* (tan) with SSI (SIC2) from *Streptomyces albogriseolus* (light sea green) in complex with subtilisin BPN′ (cornflower blue); (C) SSTI subunit in complex with subtilisin BPN′ showing putative acyl acceptor sites for *Sm*MTG. *Sa*SSI was removed for clarity.

It should also be noted that both SSTI cysteine bridges are located near the subtilisin and the predicted MTG lysine-binding sites.

The Dispase autolysis-inducing protein (DAIP) is a seven-bladed β-propeller with high affinity for neutral metalloproteases [44,55] (Fig. 6A/B). Protein twisting results in inhibition or autolysis of bound bacillolysin, thermolysin, aureolysin, and pseudolysin [101,102]. Three DAIP glutamines were biotinylated by *Sm*MTG with sufficient velocity, but Gln-39 was the most reactive residue by far [55]. Recently, the crystal structure of DAIP in complex with the C-terminal fragment of thermolysin was determined [101] (Fig. 6C). The overlay with thermolysin provided evidence that Gln-298 and Gln-345 are covered by the target proteases and inaccessible for *Sm*MTG (Fig. 6D). Moreover, Phe-297, adjacent to Gln-298, was shown to be involved in twisting of the bound protease substrates [102]. If DAIP also participates in the defense of proteases at the hyphal surface of *S. mobaraensis*, the upper (protease-binding) side of DAIP should be oriented to the air as proposed for SSTI. Based on this model, two DAIP lysines, Lys-14 and Lys-160, are predicted to be the *Sm*MTG acyl acceptor sites (Table 4, Fig. 6D). Remarkably, the only cysteine bridge of DAIP is next to Lys-14 as in the case of SSTI Lys-34 and Lys-70.

The biological function of the *Streptomyces* papain inhibitor (SPI) is unclear. The SPI protein (SPI_p) was originally purified from culture supernatants of *S. mobaraensis* with coeluting chymostatin derivatives as the active compounds (SPI_{ac}) [45,103]. The recombinant production demonstrated inactivity of SPI_p [54], but evidence, showing the interaction of the SPI_{ac} compounds with SPI_p, could not be obtained so far [103]. Nevertheless, the double-ψ-β-barrel protein SPI_p, structurally related to the N-terminal D1 domain of expansins (Fig. 7A), was used as model substrate to shed light on *Sm*MTG catalysis [54]. Exhaustive substitution of charged amino acids suggested the absence of strong SPI_p interactions with *Sm*MTG. Similar to SSTI, the sole acyl donor site, Gln-6, is opposite to an inverse thioredoxin motif, ^{74}KCPSC^{78}, which can be speculated to be the potential (if any) binding site for the SPI_{ac} aldehydes. In this hypothetical case, the chymostatin analogues are bound either by nucleophilic

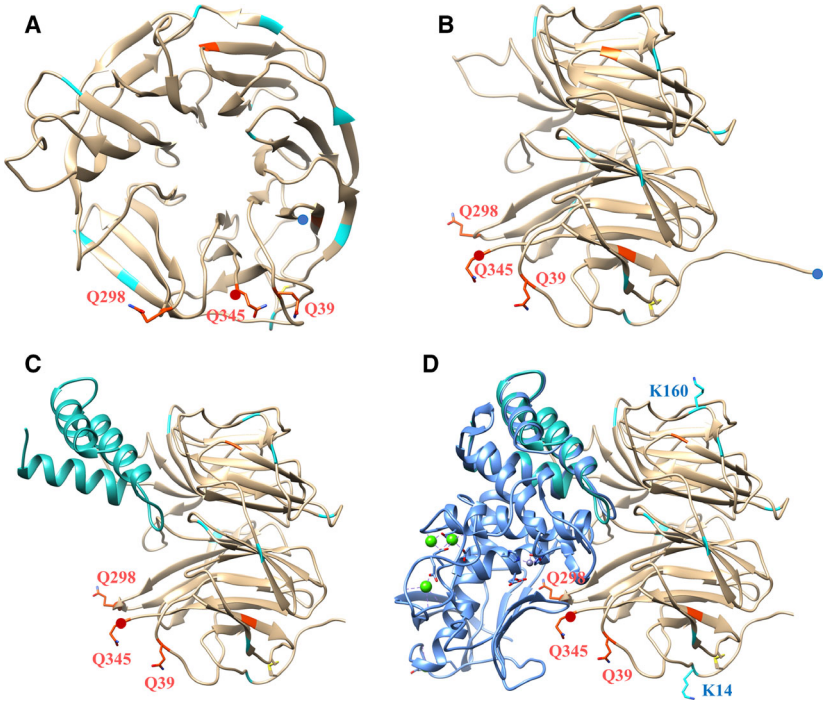

Fig. 6. Structure of DAIP from *Streptomyces mobaraensis*. Glutamine and lysine residues are shown in orange red and turquoise. (A) Top view (upper side) of the seven-bladed β-propeller (PDB 5FZP) showing the reactive glutamines in a stick representation. The N-terminal peptides are depicted by blue and red filled circles; (B) side view of DAIP rotated by ~ 90°; (C) DAIP in complex with the C-terminal thermolysin domain (CTD, sea-blue, 6FHP); (D) overlay of the DAIP CTD complex with thermolysin (cornflower blue, 1TLX). The proposed lysine donor sites are indicated.

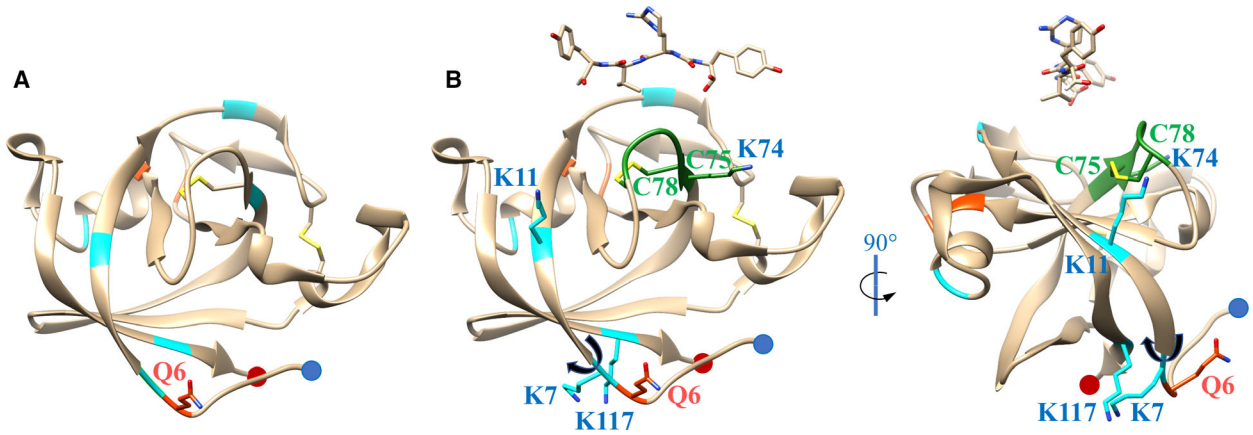

Fig. 7. Structure of SPI from *Streptomyces mobaraensis*. The SPI_p termini are indicated by blue and red filled circles, and the glutamine and lysine residues are colored orange red and turquoise. (A) View on the sole glutamine donor site (Gln-6) of inactive SPI_p (PDB 5NTB); (B) proposed functional lysines of SPI_p. The putative hinge between Lys-7 and Met-8 is indicated by a black arrow. An active SPI_{ac} variant, H-Tyr-Val-Cam-NHCONH-Tyr-OH (Cam, L-capreomycidine), is shown above the green colored SPI_p ^{74}KCSPC^{78} motif.

addition to the reduced cysteines Cys-75/Cys-78 or to the adjacent lysines Lys-11/Lys-74 (Fig. 7B). Further assuming that SPI_{ac} is incorporated into the hyphal chaplin-rodlin envelope with ^{74}KCPSC^{78} oriented to the air, the most probable acyl acceptor sites of *Sm*MTG are Lys-7 and Lys-117 (Table 4).

Other natural substrate proteins (Sml-1, Chp1, Rdl1) of *Sm*MTG that have been characterized so far remain to be studied in more detail. The β-lactamase Sml-1 has an N-terminal helix interspersed with glutamines and glutamine pairs [46]. It could be an attractive model protein to substantiate the hypothesis of Fontana and colleagues [75] that local unfolding (triggered by protein assembly) supports cross-linking by transglutaminases. The aerial hyphae proteins, the chaplins and rodlins, may be the missing links of

natural protein assembly, but study of the intrinsically disordered proteins is challenging. The glutamine pair (Gln-227/Gln-228) of Chp1 near the sortase-binding motif ^{238}LAXTG^{242} is most likely the acyl donor site for *Sm*MTG binding [104,105,S13]. However, glutamines of the fibril-forming zipper peptide (Gln-5) [106] and both chaplin domains (Gln-21, Gln-68) are also slowly biotinylated by *Sm*MTG in solution. The attempt to determine the *Sm*MTG-accessible glutamines of compressed Chp1 on PVDF membranes yielded unclear results, even though stable Langmuir–Blodgett films were reproducibly established [105]. We assume that small chaplins, lipids, teichoic acids, or other hyphal compounds are missing to stabilize the *Sm*Chp1 structure during film transfer.

In conclusion, it should be mentioned that glutamine tags derived from the natural substrate proteins SPI_p (Gln-6) and DAIP (Gln-39) were successfully fused with the monoclonal antibody trastuzumab to produce antibody–drug conjugates [S100]. The most efficient tag sequence for *Sm*MTG binding was the N-terminal peptide of SPI_p (Table S6).

Concluding remarks

Although transglutaminase from *S. mobaraensis*, *Sm*MTG, has emerged as a powerful tool for the site-specific modification of functional proteins in recent years, the enzyme is not fully understood. We had to realize that the mature enzyme needs, like proteases, a chaperone-like folding aid, usually the propeptide, to adopt the active structure. High yields of the *Sm*MTG zymogen have been obtained with several hosts, but protein export and proteolytic activation still await optimization. Especially, the appropriate signal peptide, the translocation proteins involved, and the natural activation proteases could contribute to more efficient *Sm*MTG production procedures.

Mature *Sm*MTG appears to have an already optimized structure. The activity was not decisively improved by mutagenesis, only the thermoresistance to some extent. Undoubtedly, the catalysis-determining residues revealed by exhaustive alanine substitution belong to the most conserved amino acids in *Streptomyces* transglutaminases. They clearly define distinct glutamine and lysine uptake pockets that seem to occur equally in mammalian, fish, and *K. albida* transglutaminases. Additional subsites at the front vestibule provide selectivity to modify only the physiologically relevant glutamines of substrate proteins and peptides.

The binding of *Sm*MTG to substrate proteins remains enigmatic due to the lack of conserved glutamine and lysine recognition motifs. The exchange of charged amino acids for alanine refuted strong interactions between the glutamine donor substrate SPI_p and the enzyme. Flexibility and local unfolding seem rather to be the prerequisites for *Sm*MTG to attach to glutamines of native proteins. Small and uncharged amino acids around substrate glutamines may further promote the attachment of *Sm*MTG, even though bulky and negatively charged amino acids are present in the characterized glutamine sequences of artificial peptides and non-natural substrate proteins. Potentially, the arginine-containing *Sm*MTG subsite may attract anionic residues but this cannot be predicted as N-C or C-N orientation was not revealed. Only the inhibitory peptide derived from SPI_p was shown to direct the terminal amino group to the front vestibule of *Sm*MTG. Orientation of both termini to the cleft entrance of *Sm*MTG without coverage of the lysine pocket leads to the assumption that self-assembly of the substrate proteins precedes cross-linking. Major conformational changes during protein association, triggered by specific hinge regions or secondary structure unfolding, could further support *Sm*MTG-mediated cross-linking. The natural substrate proteins SPI_p and Sml-1 might be good candidates to verify the hypothesis.

To gain new insights in catalysis, the biological role of *Sm*MTG has to be studied, at least in our opinion. *Sm*MTG is probably involved in the formation of aerial hyphae and spores since gene disruption yielded bald colonies in *S. hygroscopicus*. The cell wall of aerial mycelium is covered by hydrophobic proteins, the chaplins (coelicolor hydrophobic aerial proteins) and rodlins (rod-forming proteins), that assemble to amyloid-like rodlet structures. *S. mobaraensis* exhibits genes for 3 rodlins and 6 chaplins. *Sm*Chp1 and *Sm*Rdl1 were shown to be glutamine donor substrates of *Sm*MTG, even though as soluble, disordered proteins. Other natural substrate proteins against serine (SSTI, SPI), cysteine (SPI), and metalloproteases (SSTI, DAIP) are predicted to act at the surface of aerial hyphae. A future task could be the investigation of such natural *Sm*MTG substrate proteins to understand cross-linking of self-assembled aggregates. Analysis of the aerial hyphae may help to discover other promising protein candidates that build the outer protein envelope similar to the ~ 70 endospore coat proteins of *B. subtilis*.

Acknowledgements

The author gratefully acknowledges the close and friendly collaboration with wonderful, highly motivated young scientists over more than 32 years, who,

like him, were infected by *Streptomyces mobaraensis* and transglutaminase. The molecular graphics and analyses were performed using the royalty-free UCSF Chimera program developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

Conflict of interest

The authors declare no conflict of interest.

References

1 Kornguth SE & Waelsch H (1963) Protein modification catalyzed by transglutaminase. *Nature* **198**, 188–189. https://doi.org/10.1038/198188a0

2 Folk JE & Finlayson JS (1977) The ε-(γ-glutamyl) lysine crosslink and the catalytic role of transglutaminases. In Advances in Protein Chemistry (Anfinsen CB, Edsall JT & Richards FM, eds), Vol. **31**, pp. 1–133. Academic Press, New York, NY. https://doi.org/10.1016/s0065-3233(08)60217-x

3 Lorand L, Losowsky MS & Miloszewski KJM (1980) Human factor XIII: fibrin-stabilizing factor. *Prog Haemost Thromb* **5**, 245–290.

4 Laki K & Lorand L (1948) On the solubility of fibrin clots. *Science* **108**, 280. https://doi.org/10.1126/science.108.2802.280

5 Borosook H, Deasy CL, Haagen-Smit AJ, Keighley G & Lowy PH (1950) The uptake *in vitro* of C14-labeled glycine, l-leucine, and l-lysine by different components of guinea pig liver homogenate. *J Biol Chem* **184**, 529–543.

6 Siekevitz P (1952) Uptake of radioactive alanine *in vitro* into the proteins of rat liver fractions. *J Biol Chem* **195**, 549–565.

7 Yee VC, Pedersen LC, Le Trong I, Bishop PD, Stenkamp RE & Teller DC (1994) Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII. *Proc Natl Acad Sci USA* **91**, 7296–7300. https://doi.org/10.1073/pnas.91.15.7296

8 D'Eletto M, Rossin F, Fedorova O, Farrace MG & Piacentini M (2018) Transglutaminase type 2 in the regulation of proteostasis. *Biol Chem* **400**, 125–140. https://doi.org/10.1515/hsz-2018-0217

9 Quinn BR, Yunes-Medina L & Johnson GVW (2018) Transglutaminase 2: friend or foe? The discordant role in neurons and astrocytes. *J Neurosci Res* **96**, 1150–1158. https://doi.org/10.1002/jnr.24239

10 Mitchell JL & Mutch NJ (2019) Let's cross-link: diverse functions of the promiscuous cellular transglutaminase factor XIII-A. *J Thromb Haemost* **17**, 19–30. https://doi.org/10.1111/jth.14348

11 Del Duca S, Aloisi I, Parrotta L & Cai G (2019) Cytosceleton, transglutaminase, and gametophytic self-incompatibility in the *Malinae* (*Rosaceae*). *Int J Mol Sci* **20**, 209. https://doi.org/10.3390/ijms20010209

12 Cassadio R, Polverini E, Mariani P, Spinozzi F, Carsughi F, Fontana A, de Laureto PP, Matteucci G & Bergamini CM (1999) The structural basis for the regulation of tissue transglutaminase by calcium ions. *Eur J Biochem* **262**, 672–679. https://doi.org/10.1046/j.1432-1327.1999.00437.x

13 Pinkas DM, Strop P, Brunger AT & Khosla C (2007) Transglutaminase 2 undergoes a large conformational change upon activation. *PLoS Biol* **5**, e327. https://doi.org/10.1371/journal.pbio.0050327

14 Ikura K, Sasaki R & Motoki M (1992) Use of transglutaminase in quality-improvement and processing of food proteins. *Comment Agric Food Chem* **2**, 389–407.

15 Ando H, Adachi M, Umeda K, Matsuura A, Nonaka M, Uchio R, Tanaka H & Motoki M (1989) Purification and characteristics of a novel transglutaminase derived from microorganisms. *Agric Biol Chem* **53**, 2613–2617. https://doi.org/10.1080/00021369.1989.10869735

16 Washizu K, Ando K, Koikeda S, Hirose S, Matsuura A, Takagi H, Motoki M & Takeuchi K (1994) Molecular cloning of the gene for microbial transglutaminase from *Streptoverticillium* and its expression in *Streptomyces lividans*. *Biosci Biotech Biochem* **58**, 82–87. https://doi.org/10.1271/bbb.58.82

17 Kashiwagi T, Yokoyama K, Ishikawa K, Ono K, Ejima D, Matsui H & Suzuki E (2002) Crystal structure of microbial transglutaminase from *Streptoverticillium mobaraense*. *J Biol Chem* **277**, 44252–44260. https://doi.org/10.1074/jbc.m203933200

18 Gerber U, Jucknischke U, Putzien S & Fuchsbauer HL (1994) A rapid and simple method for the purification of transglutaminase from *Streptoverticillium mobaraense*. *Biochem J* **299**, 825–829. https://doi.org/10.1042/bj2990825

19 Giordano D & Facchiano A (2019) Classification of microbial transglutaminases by evaluation of evolution trees, sequence motifs, secondary structure topology and conservation of catalytic residues. *Biochem Biophys Res Commun* **509**, 506–513. https://doi.org/10.1016/j.bbrc.2018.12.121

20 Monroe A & Setlow P (2006) Localization of the transglutaminase cross-linking sites in the *Bacillus subtilis* spore coat protein GerQ. *J Bacteriol* **188**, 7609–7616. https://doi.org/10.1128/JB.01116-06

21 Fernandes CG, Martins D, Hernandez G, Sousa AL, Carolina C, Tranfield EM, Cordeira TN, Serrano M, Moran CP & Henriques AO (2019) Temporal and spatial regulation of protein cross-linking by the pre-assembled substrates of a *Bacillus subtilis* spore coat

transglutaminase. *PLoS Genet* **15**, e1007912. https://doi.org/10.1371/journal.pgen.1007912

22 Kobayashi K, Hashiguchi K, Yokozeki K & Yamanaka S (1998) Molecular cloning of the transglutaminase gene from *Bacillus subtilis* and its expression in *Escherichia coli*. *Biosci Biotechnol Biochem* **62**, 1109–1114. https://doi.org/10.1271/bbb.62.1109

23 Yu J, Pian Y, Ge J, Guo J, Zheng Y, Jiang H, Hao H, Yuan Y, Jiang Y & Yang M (2015) Functional and structural characterization of the antiphagocytic properties of a novel transglutaminase from *Streptococcus suis*. *J Biol Chem* **290**, 19081–19092. https://doi.org/10.1074/jbc.m115.643338

24 Uruburu M, Mastrangelo E, Bolognesi M, Ferrara S, Bertoni G & Milani M (2019) Structural and functional characterization of TgpA, a critical protein for the viability of *Pseudomonas aeruginosa*. *J Struct Biol* **205**, 18–25. https://doi.org/10.1016/j.jsb.2018.12.004

25 Steffen W, Ko FC, Patel J, Lyamichev V, Albert TJ, Benz J, Rudolph MG, Bergmann F, Streidl T, Kratzsch P *et al.* (2017) Discovery of a microbial transglutaminase enabling highly site-specific labeling of proteins. *J Biol Chem* **292**, 15622–15635. https://doi.org/10.1074/jbc.m117.797811

26 Fukui A & Horiguchi Y (2004) *Bordetella* dermonecrotic toxin exerting toxicity through activation of the small GTPase rho. *J Biochem* **136**, 415–419. https://doi.org/10.1093/jb/mvh155

27 Buetow L, Flatau G, Chiu K, Bouquet P & Ghosh P (2001) Structure of the rho-activating domain of *Escherichia coli* cytotoxic necrotizing factor 1. *Nature Struct Biol* **8**, 584–588. https://doi.org/10.1038/89610

28 Klein JD, Guzman E & Kuehn GD (1992) Purification and partial characterization of transglutaminase from *Physarum polycephalum*. *J Bacteriol* **174**, 2599–2605. https://doi.org/10.1128/jb.174.8.2599-2605.1992

29 Reiss K, Kirchner E, Gijzen M, Zocher G, Löffelhardt B, Nürnberger T, Stehle T & Brunner F (2011) Structural and phylogenetic analyses of the GP42 transglutaminase from *Phytophora sojae* reveal an evolutionary relationship between oomycetes and marine *Vibrio* bacteria. *J Biol Chem* **286**, 42585–42593. https://doi.org/10.1074/jbc.m111.290544

30 Pasternack R, Dorsch S, Otterbach JT, Robenek IR, Wolf S & Fuchsbauer HL (1998) Bacterial pro-transglutaminase from *Streptoverticillium mobaraense*. Purification, characterization, and sequence of the zymogen. *Eur J Biochem* **257**, 570–576. https://doi.org/10.1046/j.1432-1327.1998.2570570.x

31 Jin M, Chen Z, Wang Z, Huang J, Chang Z & Gao H (2018) Separation of two microbial transglutaminase isomers from *Streptomyces mobaraensis* using pH-mediated cation exchange chromatography and their characterization. *J Chromatogr B Analyt Technol Biomed Life Sci* **1097–1098**, 111–118. https://doi.org/10.1016/j.jchromb.2018.09.003

32 Zotzel J, Keller P & Fuchsbauer HL (2003) Transglutaminase from *Streptomyces mobaraensis* is activated by an endogenous metalloprotease. *Eur J Biochem* **270**, 3214–3222. https://doi.org/10.1046/j.1432-1033.2003.03703.x

33 Zhang D, Wang M, Wu J, Cui L, Du G & Chen J (2008) Two different proteases from *Streptomyces hygroscopicus* are involved in transglutaminase activation. *J Agric Food Chem* **56**, 10261–10264. https://doi.org/10.1021/jf8008519

34 Kato J, Suzuki A, Yamazaki H, Ohnishi Y & Horinouchi S (2002) Control by A-factor of a metalloendopeptidase gene involved in aerial mycelium formation in *Streptomyces griseus*. *J Bacteriol* **184**, 6016–6025. https://doi.org/10.1128/jb.184.21.6016-6025.2002

35 Kato J, Chi W, Ohnishi Y, Hong S & Horinouchi S (2005) Transcriptional control by A-factor of two trypsin genes in *Streptomyces griseus*. *J Bacteriol* **187**, 286–295. https://doi.org/10.1128/jb.187.1.286-295.2005

36 Tomono A, Tsai Y, Ohnishi Y & Horinouchi S (2005) Three chymotrypsin genes are members of the AdpA regulon in the A-factor regulatory cascade in *Streptomyces griseus*. *J Bacteriol* **187**, 6341–6353. https://doi.org/10.1128/jb.187.18.6341-6353.2005

37 Cui W, Yang X, Fang Y, Zhou S, Liu S, Du G, Du K, Chen J, Tao G & Zhou Z (2013) Discovery of two transglutaminases derived from same zymogen for the *Streptomyces hygroscopicus* and analysis of their formation processes. *J Sci Food Agric* **93**, 1711–1717. https://doi.org/10.1002/jsfa.5956

38 Zotzel J, Pasternack R, Pelzer C, Ziegert D, Mainusch M & Fuchsbauer HL (2003) Activated transglutaminase from *Streptomyces mobaraensis* is processed by a tripeptidyl aminopeptidase in the final step. *Eur J Biochem* **270**, 4149–4155. https://doi.org/10.1046/j.1432-1033.2003.03809.x

39 Umezawa Y, Yokoyama K, Kikuchi Y, Date M, Ito K, Yoshimoto T & Matsui H (2004) Novel prolyl tri/tetra-peptidyl aminopeptidase from *Streptomyces mobaraensis*: substrate specificity and enzyme gene cloning. *J Biochem* **136**, 293–300. https://doi.org/10.1093/jb/mvh129

40 Ikenaka T, Odani S, Sakai M, Nabeshima Y, Sato S & Murao S (1974) Amino acid sequence of an alkaline protease inhibitor (*Streptomyces* subtilisin inhibitor) from *Streptomyces albogriseolus* S-3253. *J Biochem* **76**, 1191–1209. https://doi.org/10.1093/oxfordjournals.jbchem.a130672

41 Schmidt S, Adolf F & Fuchsbauer HL (2008) The transglutaminase activating metalloprotease inhibitor from *Streptomyces mobaraensis* is a glutamine and

lysine donor substrate of the intrinsic transglutaminase. *FEBS Lett* **582**, 3132–3138. https://doi.org/10.1016/j.febslet.2008.07.049

42 Taguchi S, Nishihama K, Igi K, Ito K, Taira H, Motoki M & Momose H (2000) Substrate specificity analysis of microbial transglutaminase using proteinaceous protease inhibitors as natural model substrates. *J Biochem* **128**, 415–425. https://doi.org/10.1093/oxfordjournals.jbchem.a022769

43 Jüttner NE, Schmelz S, Anderl A, Colin F, Classen M, Pfeifer F, Scrima A & Fuchsbauer HL (2020) The N-terminal peptide of the transglutaminase-activating metalloprotease inhibitor from *Streptomyces mobaraensis* accommodates both inhibition and glutamine cross-linking sites. *FEBS J* **287**, 708–720. https://doi.org/10.1111/febs.15044

44 Sarafeddinov A, Schmidt S, Adolf F, Mainusch M, Bender A & Fuchsbauer HL (2009) A novel transglutaminase substrate from *Streptomyces mobaraensis* triggers autolysis of neutral metalloproteases. *Biosci Biotechnol Biochem* **73**, 993–999. https://doi.org/10.1271/bbb.80769

45 Sarafeddinov A, Arif A, Peters A & Fuchsbauer HL (2011) A novel transglutaminase substrate from *Streptomyces mobaraensis* inhibiting papain-like cysteine proteases. *J Microbiol Biotechnol* **21**, 617–626.

46 Zindel S, Ehret V, Ehret M, Hentschel M, Witt S, Krämer A, Fiebig D, Jüttner N, Fröls S, Pfeifer F *et al.* (2016) Involvement of a novel class C beta-lactamase in the transglutaminase mediated cross-linking cascade of *Streptomyces mobaraensis* 40847. *PLoS One* **11**, e0149145. https://doi.org/10.1371/journal.pone.0149145

47 Kanaji T, Ozaki H, Takao T, Kawajiri H, Ide H, Motoki M & Shimonishi Y (1993) Primary structure of microbial transglutaminase from *Streptoverticillium* sp. strain s-8112. *J Biol Chem* **268**, 11565–11572.

48 Cui L, Du G, Zhang D & Chen J (2008) Thermal stability and conformational changes of transglutaminase from a newly isolated *Streptomyces hygroscopicus*. *Biores Technol* **99**, 3794–3800. https://doi.org/10.1016/j.biortech.2007.07.017

49 Ikura K, Minami K, Otomo C, Hashimoto H, Natsuka S, Oda K & Nakanishi K (2000) High molecular weight transglutaminase inhibitor produced by a microorganism (*Streptomyces lavendulae* Y-200). *Biosci Biotechnol Biochem* **64**, 116–124. https://doi.org/10.1271/bbb.64.116

50 Folk JE (1983) Mechanism and basis for specificity of transglutaminase-catalyzed ε-(γ-glutamyl)lysine bond formation. *Adv Enzymol Relat Areas Mol Biol* **54**, 1–56. https://doi.org/10.1002/9780470122990.ch1

51 Pati A, Sikorski J, Nolan M, Lapidus A, Copeland A, Glavina Del Rio T, Lucas S, Chen F, Tice H, Pitluck S *et al.* (2009) Complete genome sequence of *Saccharomonospora viridis* type strain (P101^T). *Stand Genomic Sci* **1**, 141–149. https://doi.org/10.4056/sigs.20263

52 Schuurmans DM, Olson BH & San Clemente CL (1956) Production and isolation of thermoviridin, an antibiotic produced by *Thermoactinomyces viridis* n.sp. *Appl Microbiol* **4**, 61–66.

53 Juettner NE, Classen M, Colin F, Hoffmann SB, Meyners C, Pfeifer F & Fuchsbauer HL (2018) Features of the transglutaminase-activating metalloprotease from *Streptomyces mobaraensis* DSM 40847 produced in *Escherichia coli*. *J Biotechnol* **281**, 115–122. https://doi.org/10.1016/j.jbiotec.2018.07.004

54 Juettner NE, Schmelz S, Bogen JP, Happel D, Fessner WD, Pfeifer F, Fuchsbauer HL & Scrima A (2018) Illuminating structure and acyl donor sites of a physiological transglutaminase substrate from *Streptomyces mobaraensis*. *Prot Sci* **27**, 910–922. https://doi.org/10.1002/pro.3388

55 Fiebig D, Schmelz S, Zindel S, Ehret V, Beck J, Ebenig A, Ehret M, Fröls S, Pfeifer F, Kolmar H *et al.* (2016) Structure of the Dispase autolysis-inducing protein from *Streptomyces mobaraensis* and glutamine cross-linking sites for transglutaminase. *J Biol Chem* **291**, 20417–20426. https://doi.org/10.1074/jbc.m116.731109

56 Thompson BJ, Widdick DA, Hicks MG, Chandra G, Sutcliffe IC, Palmer T & Hutchings MI (2010) Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium *Streptomyces coelicolor*. *Mol Microbiol* **77**, 943–957. https://doi.org/10.1111/j.1365-2958.2010.07261.x

57 Yang M, Chang C, Wang JM, Wu TK, Wang Y, Chang C & Li TT (2011) Crystal structure and inhibition studies of transglutaminase from *Streptomyces mobaraense*. *J Biol Chem* **286**, 7301–7307. https://doi.org/10.1074/jbc.m110.203315

58 Deweid L, Neureiter L, Englert S, Schneider H, Deweid J, Yanakieva D, Sturm J, Bitsch S, Christmann A, Avrutina O *et al.* (2018) Directed evolution of a bond-forming enzyme: ultrahigh-throughput screening of microbial transglutaminase using yeast surface display. *Chemistry* **24**, 15195–15200. https://doi.org/10.1002/chem.201803485

59 Juettner NE, Schmelz S, Kraemer A, Knapp S, Becker B, Kolmar H, Scrima A & Fuchsbauer HL (2018) Structure of a glutamine donor mimicking inhibitory peptide shaped by the catalytic cleft of microbial transglutaminase. *FEBS J* **285**, 4684–4694. https://doi.org/10.1111/febs.14678

60 Tagami U, Shimba N, Nakamura M, Yokoyama K, Suzuki E & Hirokawa T (2009) Substrate specificity of microbial transglutaminase as revealed by three-dimensional docking simulation and mutagenesis.

Protein Eng Des Sel **22**, 747–752. https://doi.org/10.1093/protein/gzp061
61 Shimba N, Yokoyama K & Suzuki E (2002) NMR-based screening method for transglutaminases: rapid analysis of their substrate specificities and reaction rates. *J Agric Food Chem* **50**, 1330–1334. https://doi.org/10.1021/jf010995k
62 Shimba N, Yamada N, Yokoyama K & Suzuki E (2002) Enzymatic labelling of arbitrary proteins. *Anal Biochem* **301**, 123–127. https://doi.org/10.1006/abio.2001.5485
63 Shimba N, Shinohara M, Yokoyama K, Kashiwagi T, Ishikawa K, Ejima D & Suzuki E (2002) Enhancement of transglutaminase activity by NMR identification of its flexible residues affecting the active site. *FEBS Lett* **517**, 175–179. https://doi.org/10.1016/S0014-5793(02)02616-9
64 Marx CC, Hertel TC & Pietzsch M (2008) Random mutagenesis of a recombinant microbial transglutaminase for the generation of thermostable and heat-sensitive variants. *J Biotechnol* **136**, 156–162. https://doi.org/10.1016/j.jbiotec.2008.06.005
65 Böhme B, Moritz B, Wendler J, Hertel TC, Ihling C, Brandt W & Pietzsch M (2020) Enzymatic activity and thermoresistance of improved microbial transglutaminase variants. *Amino Acids* **52**, 313–326. https://doi.org/10.1007/s00726-019-02764-9
66 Yokoyama K, Utsumi H, Nakamura T, Ogaya D, Shimba N, Suzuki E & Taguchi S (2010) Screening for improved activity of a transglutaminase from *Streptomyces mobaraensis* created by a novel rational mutagenesis and random mutagenesis. *Appl Microbiol Biotechnol* **87**, 2087–2096. https://doi.org/10.1007/s00253-010-2656-6
67 Zhao X, Shaw AC, Wang J, Chang C, Deng J & Su J (2010) A novel high-throughput screening method for microbial transglutaminases with high specificity toward Gln141 of human growth hormone. *J Biomol Screen* **15**, 206–212. https://doi.org/10.1177/1087057109356206
68 Dickgiesser S, Rieker M, Mueller-Pompalla D, Schröter C, Tonillo J, Warszawski S, Raab-Westphal S, Kühn S, Knehans T, Könning D *et al.* (2020) Site-specific conjugation of native antibodies using engineered microbial transglutaminase. *Bioconjug Chem* **31**, 1070–1076. https://doi.org/10.1021/acs.bioconjchem.0c00061
69 Schechter I & Berger A (1967) On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157–162. https://doi.org/10.1016/s0006-291x(67)80055-x
70 Pincus JH & Waelsch H (1968) The specificity of transglutaminase. II. Structural requirements of the amine substrate. *Arch Biochem Biophys* **126**, 44–52. https://doi.org/10.1016/0003-9861(68)90557-2
71 Schrode J & Folk JE (1979) Stereochemical aspects of amine substrate attachment to acyl intermediates of transglutaminases. Human blood plasma enzyme (activated coagulation factor XIII) and guinea pig liver enzyme. *J Biol Chem* **254**, 653–661.
72 Lorand L, Parameswaran KN, Stenberg P, Tong YS, Velasco PT, Jônsson NÅ, Mikiver L & Moses P (1979) Specificity of guinea pig liver transglutaminase for amine substrates. *Biochemistry* **18**, 1756–1765. https://doi.org/10.1021/bi00576a019
73 Ohtsuka T, Sawa A, Kawabata R, Nio N & Motoki M (2000) Substrate specificities of microbial transglutaminase for primary amines. *J Agric Food Chem* **48**, 6230–6233. https://doi.org/10.1021/jf000302k
74 Nonaka M, Matsuura Y & Motoki M (1996) Incorporation of lysine and lysine dipeptides into α_{S1}-casein by Ca^{2+}-independent microbial transglutaminase. *Biosci Biotechnol Biochem* **60**, 131–133. https://doi.org/10.1271/bbb.60.131
75 Spolaore B, Raboni S, Molina AR, Satwekar A, Damiano N & Fontana A (2012) Local unfolding is required for the site-specific protein modification by transglutaminase. *Biochemistry* **51**, 8679–8689. https://doi.org/10.1021/bi301005z
76 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. https://doi.org/10.1271/bbb.64.2608
77 Malesevic M, Migge A, Hertel TC & Pietsch M (2015) A fluorescence-based array screen for transglutaminase substrates. *ChemBioChem* **16**, 1669–1674. https://doi.org/10.1002/cbic.201402709
78 Hemung B, Li-Chan ECY & Yongsawatdigul J (2008) Reactivity of fish and microbial transglutaminases on glutaminyl sites of peptides derived from threadfin bream myosin. *J Agric Food Chem* **56**, 7510–7516. https://doi.org/10.1021/jf800856g
79 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**, 379–383. https://doi.org/10.1016/j.abb.2008.06.014
80 Lee J, Song C, Kim D, Park I, Lee S, Lee Y & Kim B (2013) Glutamine (Q)-peptide screening for transglutaminase reaction using mRNA display. *Biotechnol Bioeng* **110**, 353–362. https://doi.org/10.1002/bit.24622
81 Corporale A, Selis F, Sandomenico A, Jotti GS, Tonon G & Ruvo M (2015) The LQSP tetrapeptide is a new highly efficient substrate of microbial transglutaminase for the site-specific derivatization of peptides and proteins. *Biotechnol J* **10**, 154–161. https://doi.org/10.1002/biot.201400466

82 Touati J, Angelini A, Hinner MJ & Heinis C (2011) Enzymatic cyclisation of peptides with a transglutaminase. *ChemBioChem* **12**, 38–42. https://doi.org/10.1002/cbic.201000451
83 Sato H, Hayashi E, Yamada N, Yatagai M & Takahara Y (2001) Further studies on the site-specific protein modification by microbial transglutaminase. *Bioconjug Chem* **12**, 701–710. https://doi.org/10.1021/bc000132h
84 Matsumura Y, Chanyongvorakul Y, Kumazawa Y, Ohtsuka T & Mori T (1996) Enhanced susceptibility to transglutaminase reaction of α-lactalbumin in the molten globule state. *Biochim Biophys Acta* **1292**, 69–76. https://doi.org/10.1016/0167-4838(95)00197-2
85 Nieuwenhuizen WF, Dekker HL, De Koning LJ, Gröneveld T, De Koster CG & De Jong GAH (2003) Modification of glutamine and lysine residues in holo and apo α-lactalbumin with microbial transglutaminase. *J Agric Food Chem* **51**, 7132–7139. https://doi.org/10.1021/jf0300644
86 Lee D, Matsumoto S, Matsumura Y & Mori T (2002) Identification of the ε-(γ-glutamyl)lysine cross-linking sites in α-lactalbumin polymerized by mammalian and microbial transglutaminases. *J Agric Food Chem* **50**, 7412–7419. https://doi.org/10.1021/jf020529a
87 Rachel NM, Toulouse JL & Pelletier JN (2017) Transglutaminase-catalyzed bioconjugation using one-pot metal-free bioorthogonal chemistry. *Bioconjug Chem* **28**, 2518–2523. https://doi.org/10.1021/acs.bioconjchem.7b00509
88 Maullu C, Raimondo D, Caboi F, Giorgetti A, Sergi M, Valentini M, Tonon G & Tramontano A (2009) Site-directed enzymatic PEGylation of the human granulocyte colony-stimulating factor. *FEBS J* **276**, 6741–6750. https://doi.org/10.1111/j.1742-4658.2009.07387.x
89 Mero A, Spolaore B, Veronese FM & Fontana A (2009) Transglutaminase-mediated PEGylation of proteins: direct identification of the sites of protein modification by mass spectrometry using a novel monodisperse PEG. *Bioconjug Chem* **20**, 384–389. https://doi.org/10.1021/bc800427n
90 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 Control Release* **154**, 27–34. https://doi.org/10.1016/j.jconrel.2011.04.024
91 Marculescu C, Lakshminarayanan A, Gault J, Knight JC, Folkes LK, Spink T, Robinson CV, Vallis K, Davis BG & Cornelissen B (2019) Probing the limits of Q-tag bioconjugation of antibodies. *Chem Commun* **55**, 11342–11345. https://doi.org/10.1039/c9cc02303h
92 Jeger S, Zimmermann K, Blanc A, Gruenberg J, Honer M, Hunziker P, Struthers H & Schibli R (2010) Site-specific and stoichiometric modification of antibodies by bacterial transglutaminase. *Angew Chem Int Ed Engl* **49**, 9995–9997. https://doi.org/10.1002/anie.201004243
93 Spycher PR, Ahmann CA, Wehrmüller JE, Hurwitz DR, Kreis O, Messmer D, Ritler A, Küchler A, Blanc A, Béhé M *et al.* (2017) Dual, site-specific modification of antibodies using solid-phase immobilized microbial transglutaminase. *ChemBioChem* **18**, 1923–1927. https://doi.org/10.1002/cbic.201700188
94 Spidel JL, Vaessen B, Albone EF, Cheng X, Verdi A & Kline JB (2017) Site-specific conjugation to native and engineered lysines in human immunoglobulins by microbial transglutaminase. *Bioconjug Chem* **28**, 2471–2484. https://doi.org/10.1021/acs.bioconjchem.7b00439
95 Spolaore B, Raboni S, Satwekar AA, Grigoletto A, Mero A, Montagner IM, Rosato A, Pasut G & Fontana A (2016) Site-specific transglutaminase-mediated conjugation of interferon α-2b at glutamine or lysine residues. *Bioconjug Chem* **27**, 2695–2706. https://doi.org/10.1021/acs.bioconjchem.6b00468
96 Spolaore B, Damiano N, Raboni S & Fontana A (2014) Site-specific derivatization of avidin using microbial transglutaminase. *Bioconjug Chem* **25**, 470–480. https://doi.org/10.1021/bc400378h
97 Schneider H, Deweid L, Avrutina O & Kolmar H (2020) Recent progress in transglutaminase-mediated assembly of antibody-drug conjugates. *Anal Biochem* **595**, 113615. https://doi.org/10.1016/j.ab.2020.113615
98 Doti N, Caporale A, Monti A, Sandomenico A, Selis F & Ruvo M (2020) A recent update on the use of microbial transglutaminase for the generation of biotherapeutics. *World J Microbiol Biotechnol* **36**, 53. https://doi.org/10.1007/s11274-020-02829-y
99 Chen K, Zhang D, Liu S, Wang NS, Wang M, Du G & Chen J (2013) Improvement of transglutaminase production by extending differentiation phase of *Streptomyces hygroscopicus*: mechanism and application. *Appl Microbiol Biotechnol* **97**, 7711–7719. https://doi.org/10.1007/s00253-012-4614-y
100 Gräfe U, Reinhardt G, Krebs D, Roth M & Bormann EJ (1982) Biochemical characteristics of none-streptomycin-producing mutants of *Streptomyces griseus*. Lipids and fatty acid composition of vegetative mycelia. *Z Allg Mikrobiol* **22**, 97–106. https://doi.org/10.1002/jobm.19820220204
101 Fiebig D, Storka J, Roeder M, Meyners C, Schmelz S, Blanckenfeldt W, Scrima A, Kolmar H & Fuchsbauer HL (2018) Destructive twisting of neutral metalloproteases: the catalysis mechanism of the

Dispase autolysis-inducing protein from *Streptomyces mobaraensis* DM 40847. *FEBS J* **285**, 4246–4264. https://doi.org/10.1111/febs.14647

102 Fiebig D, Anderl A, Al Djaizani S, Kolmar H & Fuchsbauer HL (2020) Dissecting capture and twisting of aureolysin and pseudolysin: functional amino acids of the Dispase autolysis-inducing protein. *Biochem J* **477**, 2595–2606. https://doi.org/10.1042/bcj20200407

103 Jüttner NE, Bogen JP, Bauer TA, Knapp S, Pfeifer F, Hüttenhain SH, Meusinger R, Krämer A & Fuchsbauer HL (2020) Decoding the papain inhibitor from *Streptomyces mobaraensis* as being hydroxylated chymostatin derivatives: purification, structure analysis and putative biosynthetic pathway. *J Nat Prod* **83**, 2983–2995. https://doi.org/10.1021/acs.jnatprod.0c00201

104 Anderl A, Kolmar H & Fuchsbauer HL (1919) The metal-binding properties of the long chaplin from *Streptomyces mobaraensis*: a bioinformatic and biochemical approach. *J Inorg Biochem* **202**, 110878. https://doi.org/10.1016/j.jinorgbio.2019.110878

105 Anderl A (2020) Characterization of cell wall proteins from transglutaminase-producing *Streptomyces mobaraensis* (translation). *Doctoral thesis*, Technische Universität Darmstadt, Germany. http://d-nb.info/1212584082

106 Goldschmidt L, Teng PK, Riek R & Eisenberg D (2010) Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci USA* **107**, 3487–3492. https://doi.org/10.1073/pnas.0915166107

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. *Streptomyces* transglutaminase.
Fig. S1. Structure of human transglutaminases.
Fig. S2. Structure of microbial transglutaminases.
Fig. S3. Mechanistic model for *Sm*MTG catalysis.
Fig. S4. Primary structure of *Streptomyces* transglutaminases.
Fig. S5. Amino acids influencing thermostability and activity of *Sm*MTG (PDB 1IU4).
Fig. S6. Top view on glutamine and lysine-binding sites of transglutaminases.
Fig. S7. Model of glutamine and lysine orientation in the catalytic cleft of *Sm*MTG.
Table S1. Characterized proteins and peptides from culture supernatants of *Streptomyces mobaraensis*.
Table S2. Properties of microbial transglutaminases from *Streptomyces* bacteria.
Table S3. Selected procedures for the recombinant production of microbial transglutaminase.
Table S4. Activation of microbial transglutaminases by proteases.
Table S5. Effect of amino acids upstream from a protein lysine donor site of *Sm*MTG.
Table S6. Glutamine and lysine tags for protein and polymer conjugation reactions mediated by microbial transglutaminase.