⌘ *Author's Choice*

Discovery of a microbial transglutaminase enabling highly site-specific labeling of proteins

Wojtek Steffen[‡1], Fu Chong Ko[‡], Jigar Patel[§], Victor Lyamichev[§], Thomas J. Albert[§], Jörg Benz[¶], Markus G. Rudolph[¶], Frank Bergmann[‡], Thomas Streidl[‡], Peter Kratzsch[‡], Mara Boenitz-Dulat[‡], Tobias Oelschlaegel[‡], and Michael Schraeml[‡]

From [‡]Roche Diagnostics GmbH, CPS, Nonnenwald 2, 82377 Penzberg, Germany, [§]Roche Sequencing, NimbleGen, Madison, Wisconsin 53719, and [¶]F. Hoffmann-La Roche Ltd., pRED, Grenzacherstrasse 124, 4070 Basel, Switzerland

Edited by Joseph Jez

Microbial transglutaminases (MTGs) catalyze the formation of Gln–Lys isopeptide bonds and are widely used for the cross-linking of proteins and peptides in food and biotechnological applications (*e.g.* to improve the texture of protein-rich foods or in generating antibody-drug conjugates). Currently used MTGs have low substrate specificity, impeding their biotechnological use as enzymes that do not cross-react with nontarget substrates (*i.e.* as bio-orthogonal labeling systems). Here, we report the discovery of an MTG from *Kutzneria albida* (KalbTG), which exhibited no cross-reactivity with known MTG substrates or commonly used target proteins, such as antibodies. KalbTG was produced in *Escherichia coli* as soluble and active enzyme in the presence of its natural inhibitor ammonium to prevent potentially toxic cross-linking activity. The crystal structure of KalbTG revealed a conserved core similar to other MTGs but very short surface loops, making it the smallest MTG characterized to date. Ultra-dense peptide array technology involving a pool of 1.4 million unique peptides identified specific recognition motifs for KalbTG in these peptides. We determined that the motifs YRYRQ and RYESK are the best Gln and Lys substrates of KalbTG, respectively. By first reacting a bifunctionalized peptide with the more specific KalbTG and in a second step with the less specific MTG from *Streptomyces mobaraensis*, a successful bio-orthogonal labeling system was demonstrated. Fusing the KalbTG recognition motif to an antibody allowed for site-specific and ratio-controlled labeling using low label excess. Its site specificity, favorable kinetics, ease of use, and cost-effective production render KalbTG an attractive tool for a broad range of applications, including production of therapeutic antibody-drug conjugates.

Conventional chemical strategies for the modification of therapeutic and diagnostic proteins often lack site specificity, linkage stability, and stoichiometric control, giving rise to heterogeneous conjugates that may cause interference (*e.g.* with immunoreactivity or with the stability of a therapeutic agent).

The industrial development of therapeutic and diagnostic reagents in the coming years will see a massive increase in sophisticated applications requiring stable and truly site-specific conjugation. Research into novel enzymatic methods offering an attractive and cost-effective alternative to established chemical strategies is therefore of paramount importance (1–3).

Microbial transglutaminase (MTG),[2] first described by researchers of Ajinomoto Co., Inc. in 1989 (4) is one of the most widely used enzymes for the cross-linking of proteins and peptides in many food and biotechnological applications (5–7). MTG was first discovered in and later extracted from the organism *Streptomyces mobaraensis*, and recombinant *S. mobaraensis* MTG represents the bulk of industrially used MTGs today (6).

Transglutaminases catalyze the formation of a stable isopeptide bond between an acyl group (*e.g.* a glutamine side chain) and an alkyl amine (*e.g.* a lysine side chain). In the absence of reactive amine groups, the enzymatic reaction with water leads to deamination of glutamine side chains (5, 8, 9). In contrast to mammalian transglutaminases, bacterial enzymes do not require cofactors, such as Ca^{2+} or GTP, and function over a broad range of pH, buffers, and temperatures (7).

Most characterized MTGs are from *Streptomyces* or *Bacilli*. They share high sequence homology and have similar substrate specificities with typical molecular masses of ≥ 38 kDa. Being a cross-linking enzyme involved in spore coat formation, MTG displays broad substrate specificity for both acyl donor and alkyl amine groups. Although approaches for the high-throughput screening of improved transglutaminase substrates via phage panning or mRNA display have been reported (10, 11), no comprehensive peptide array-based approaches of peptides larger than trimers (12) have been conducted.

Because only the substrate specificities of the enzyme from *S. mobaraensis* and homologous enzymes are known, a bio-orthogonal conjugation approach (*e.g.* simultaneous labeling of a biomolecule using two or more different label substrates and two or more transglutaminase species) is currently not possible.

This work describes the biochemistry, substrate identification, and crystal structure of an active, highly compact (26-kDa) MTG from *Kutzneria albida (*KalbTG) unrelated to any other MTG. The recombinant production of KalbTG in *Escherichia coli* in the presence of its natural inhibitor NH_4^+ enabled the high-throughput screening of substrate peptides via peptide array (13). Specificity of this novel transglutaminase for the array-determined substrate sequences is demonstrated by efficient incorporation of labels into engineered target molecules and the poor or undetectable turnover of the enzyme with substrates recognized by conventional MTGs. The high activity and low molecular mass of KalbTG signifies a key advantage for mass production and enzymatic labeling purposes.

Together, these properties make KalbTG highly attractive for a broad range of applications, including the versatile, cost-effective, and site-specific conjugation of biomolecules with various label molecules (*e.g.* production of therapeutic antibody-drug conjugates or chemiluminescent antibodies for *in vitro* diagnostic purposes).

Results

Discovery of a novel microbial transglutaminase

Using the amino acid sequence of *S. mobaraensis* protein–glutamine γ-glutamyltransferase as a query, a search for homologs of this enzyme was performed. This yielded the hypothetical gene product KALB_7456 from bacteria *K. albida* DSM 43870, a spore-forming Gram-positive bacterium that was sequenced in 2014 (14). Comparing the primary structures of the *S. mobaraensis* and *K. albida* gene products showed 30% similarity with a distinct conservation of active site residues (supplemental Fig. 1*A*), indicating that the structure and function of this enzyme may be preserved. The *K. albida* gene product is significantly smaller than MTG, amounting to a calculated molecular mass of 30.1 kDa and a molecular mass of 26.4 kDa in the active form. Because MTG is produced as inactive proenzyme and processed by extracellular proteases, such as dispase, to yield the 38-kDa active form (15, 16), we assumed a similar activation mechanism for the hypothetical *K. albida* microbial transglutaminase and used the ProP 1.0 server to analyze the probability for signal and propeptide sequences in the N-terminal region of the protein (supplemental Fig. 1*B*). VAAPTPR ↓ AP was the only predicted propeptide cleavage site, which corresponds fairly well with the dispase site SAGPSFR ↓ AP in MTG but putatively has no dispase reactivity because Phe is a required residue in the dispase recognition motif (17, 18). Additionally, a signal peptide is predicted by the ProP 1.0 server with a high-probability cleavage site, GLPTLIA ↓ TT. However, this sequence bears no resemblance to the significantly longer MTG prese-quence or other known signal peptides.

Parallel construct evaluation allows economical recombinant production of KalbTG

To rapidly screen expression conditions for the hypothetical transglutaminase KalbTG, we inserted the synthetic gene into multiple expression vectors designed for the soluble cytosolic or periplasmic expression in *E. coli* using the fragment exchange system (19). Initial screening in the 5-ml scale provided clear evidence that proteins with the anticipated electrophoretic mobility of full-length KalbTG fusions were expressed and that a fusion with tandem SlyD chaperones as described previously (20) yielded the highest amount of soluble protein among all of the constructs tested (data not shown). By suppressing the toxic cross-linking activity of elevated KalbTG concentrations with the addition of ammonium salts, the production process was adaptable to a 10-liter fermenter scale. Purification by standard methods (Q-Sepharose, nickel-nitrilo-triacetic acid) yielded several hundred mg of a highly pure enzyme preparation. The purified and activated enzyme remained stable at 4 °C and over multiple freeze–thaw cycles, with the melting point determined at 48.9 °C using dynamic scanning calorimetry (DSC) (supplemental Fig. 2).

The NimbleGen peptide array delivers a holistic and reproducible approach for transglutaminase substrate discovery

We confirmed that both the SlyD-fused and the mature KalbTG possess basic microbial transglutaminase activity of at least 1.65 units/mg by assaying it with the Zediras MTG-ANiTA kit (M001, compared with 4.30 units/mg by the MTG supplied with the kit and a 0.07-unit/mg blank value with BSA). Next, we searched for specific recognition motifs by assaying KalbTG with the NimbleGen peptide array technology. The turnover of the transamidation reaction between 1.4 million unique 5-mer peptides and biotinylated amine donor *N*-(biotinyl)cadaverine used as a substitute for a Lys substrate was quantified via fluorescence measurement of CyTM5–streptavidin binding.

A control experiment with only CyTM5–streptavidin did not show any significant binding due to the fact that all potential streptavidin binders were removed from the 5-mer array design.

The experiments were performed on two arrays in parallel, and the sequences of the peptides with the highest turnovers were determined (Fig. 1, *A* and *B*). The 9 best peptides were resynthesized and tested for KalbTG activity in a stand-alone glutamate dehydrogenase (GLDH)-coupled assay (supplemental Table 1). The assay was specifically optimized for the analysis of potential high-affinity peptides and utilizes the ammonia released in the transglutaminase reaction as a substrate in the GLDH-catalyzed and NADH-dependent reductive amination of α-ketoglutarate (21). This confirmed YRYRQ and RYRQR as the best 5-mer substrates (KalbTG Q-tag), with turnover rates of 3.52 ± 0.08 pmol of NADH/s and 3.60 ± 0.12 pmol of NADH/s, respectively. Lys-containing substrate YKYRQ exhibited the highest rates in the GLDH assay (4.00 ± 0.18 pmol/s NADH) but showed a relatively low turnover on the peptide array. We suspect that this may be an artifact caused by
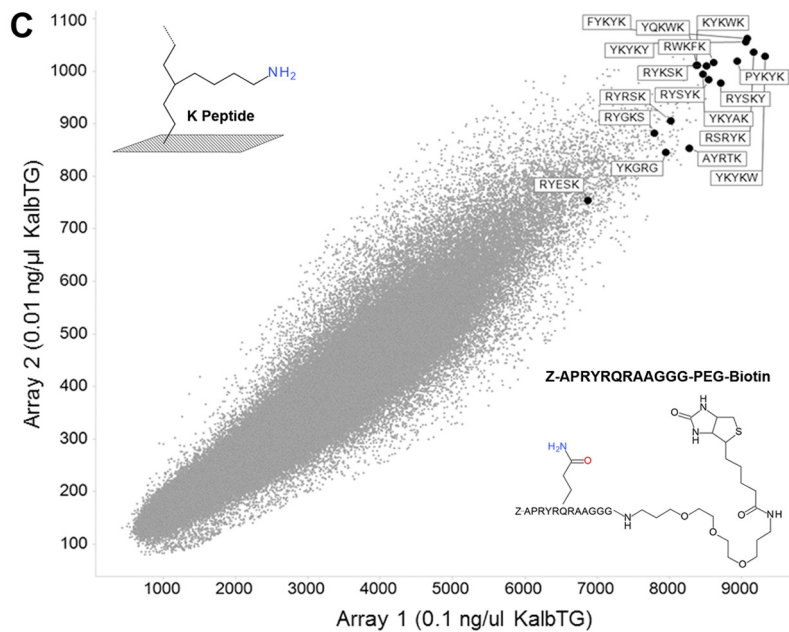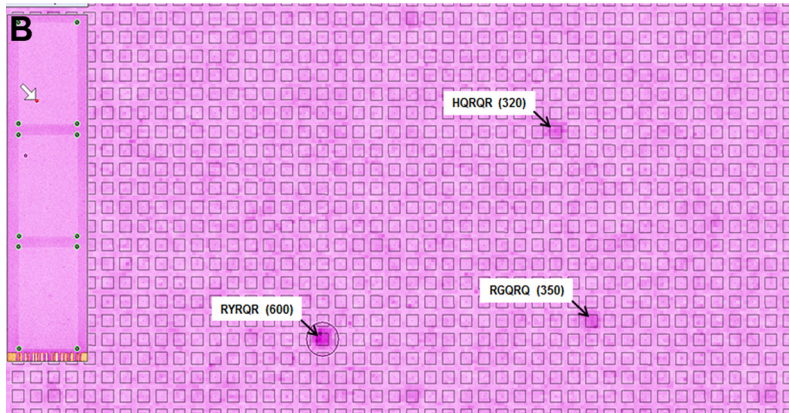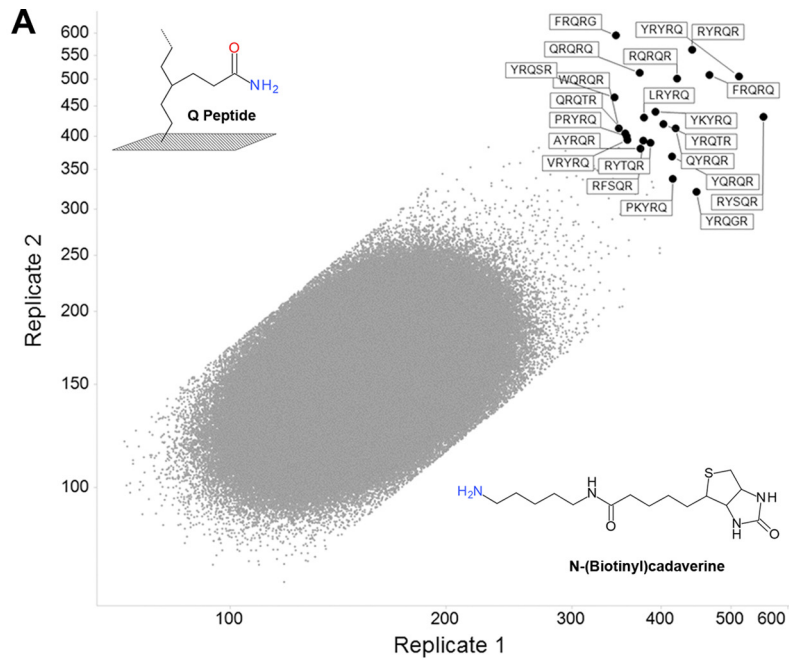
Lys cross-reactivity in solution and thus omitted this motif in further analysis. Surprisingly, no activity could be detected with the well-known MTG recognition motif MLAQGS (22) or the MTG substrate DYALQ discovered on our peptide array (Fig. 2 and supplemental Table 1).

A second round of maturation on the array yielded APRYRQRAA as the best 9-mer substrate, which was then resynthesized as biotinylated peptide to act as acyl donor for the discovery of optimized Lys recognition motifs (KalbTG K-tag) back on the 5-mer peptide array (Fig. 1C). Again, six of the best Lys peptides from the array were resynthesized and tested in the GLDH-coupled assay, now using a peptide containing the optimized Gln recognition sequence YRYRQ as acyl donor (supplemental Table 1). The highest turnover (4.47 ± 0.16 pmol NADH/s) in the GLDH assay was observed with the sequence RYESK.

KalbTG is highly specific for matured glutamine substrates and enables efficient applications for bio-orthogonal conjugation

Because the peptide array can deliver readout for a comprehensive set of 5-mer peptides at once, a single data set each suffices to evaluate how enzymes differ in substrate specificity. The top KalbTG Gln substrates (Fig. 2A) can thus be found in the midfield of the signal distribution on the array performed with MTG (Fig. 2B), and, vice versa, the best-performing MTG Gln substrates (Fig. 2C) exhibit only signal close to background level on the KalbTG array (Fig. 2D and supplemental Table 1). To confirm that the two transglutaminase enzymes have orthogonal Gln substrate preferences and to quantify the amount of cross-reactivity, we determined the kinetics of both enzymes in the presence of varying concentrations of Z-GGGYRYRQGGGG and Z-GGGDYALQGGGG substrate peptides (Fig. 2E). MTG exhibited similar K_m values in the 0.6–0.9 mM range for both substrates, whereas k_{cat} was significantly higher with the preferred DYALQ substrate (1.39 s^{-1} *versus* 0.93 s^{-1} with YRYRQ), resulting in catalytic efficiencies (k_{cat}/K_m) of 1.64×10^3 and 1.44×10^3 M^{-1} s^{-1}, respectively (Table 1). Compared with the engineered MTG enzyme from Zedira, KalbTG appears to have a lower substrate binding efficiency (K_m of 2 mM) but higher turnover (k_{cat} of 1.92 s^{-1}), leading to k_{cat}/K_m of 0.89×10^3 M^{-1} s^{-1}. KalbTG appeared to be completely unreactive toward MTG substrate Z-GGGDYALQGGGG; thus, no kinetic parameters could be determined.

Next, we applied the array and in-solution data to perform site-specific labeling on protein substrates. The molecular chaperone SlyD is an ideal scaffold for labeling approaches because epitope-containing loops can be grafted onto the

FKBP-type domain, which optimally presents them to binders or enzymes (23). We produced a chimeric protein consisting of the *Thermus thermophilus* FKBP-type domain and the KalbTG recognition sequence RYRQR. Labeling with a 10-fold excess of KalbTG K-tag-Cy3 and a substrate/enzyme ratio of 72:1 afforded ~70% yield of a labeled protein species after 15 min (Fig. 3A). This yield remained constant over a time course of 60 min. The molecular mass shift from 13 to 19 kDa was observed by SDS-PAGE, corresponding exactly to the incorporation of a single 6-kDa label molecule. An identically constructed FKBP-type domain, containing the MTG sequence DYALQ instead of RYRQR, showed no incorporation of label when incubated with KalbTG (Fig. 3A), signifying that the reaction is limited to the site of the KalbTG recognition motif and that none of the five other glutamines intrinsic to the FKBP-type domain are recognized. We furthermore assayed the pH dependence of the labeling reaction at pH 6.2, 6.8, 7.4, 8.0, 8.5, and 9 (Fig. 3B). The highest labeling efficiency after 15 min was found at pH 7.4, with activity trailing off at pH ≥8.5. These findings correspond well to the published pH preferences of MTG (15).

We used the high sequence specificity of KalbTG to conjugate a 6-kDa KalbTG K-tag-Cy3 label to the YRYRQ site of a 7-kDa substrate peptide comprising both the KalbTG and MTG Gln motifs. The reaction was run for 30 min to saturate the YRYRQ site. Analysis by SDS-PAGE confirmed that the label was integrated at a single site (Fig. 3C, lane 2). The subsequent incubation for 15 min with MTG and a 6-kDa MTG K-tag-Cy5 label in the same reaction vessel resulted in the formation of a dually labeled conjugate in high yield, with nearly all single-labeled species having visibly been converted to dually labeled (Fig. 3C, lane 3).

As an example for labeling of antibodies for use in therapeutic and immunodiagnostic applications, we inserted the KalbTG Q-tag to the heavy chain C terminus of the IgG used in the Elecsys TSH electrochemiluminescence immunoassay. IgG biotinylation via KalbTG was assessed by complex formation with fluorescein-labeled streptavidin (SA-FLUO). Successful complexation (IgG-Bi–SA-FLUO and IgG-Bi–SA-FLUO–Bi-IgG) is observed as a fluorescent double peak in the elution profile of an analytical size-exclusion chromatography (Fig. 4C). As a control, the same IgG modified C-terminally with the MTG Q-tag LLQGA published by Rinat-Pfizer Inc. (7) was biotinylated via MTG. It exhibited a similar pattern in the SA-FLUO analytics (Fig. 4A); however, compared with the KalbTG experiment, a higher excess of biotin label and higher enzyme/IgG ratio was needed to achieve this result. When incubated with KalbTG, the LLQGA-tagged IgG showed no SA-FLUO

Figure 1. KalbTG substrate discovery. *A,* Gln substrate discovery. Fluorescence signal generated by KalbTG in the presence of biotinylated amine-donor substrate. Substrate structures are shown schematically, and the TG-reactive groups are *colored*. The plot shows correlation between two replicates of 1.4 million unique 5-mer peptides. 22 peptides with highest average fluorescence signal are tagged by their respective 5-mer peptide sequence. *B,* image analysis of KalbTG Gln substrate array (*A*). *Left,* image of the whole array containing 2,721,464 features (*spots*) with 5-mer peptides synthesized in a *chessboard pattern*. The *white arrow* shows position of one peptide of the RYRQR pair. The *main panel* shows an *enlarged area* around the RYRQR feature. RYRQR (*circled*) and two other peptides with detectable KalbTG activity are marked by *black arrows* with fluorescent signal shown in *parenthesis*. The average background signal over the array is ~180. *C,* K-substrate discovery. Fluorescence signal generated on 5-mer array in the presence of KalbTG and biotinylated Gln donor substrate (Z-APRYRQRAAGGG-PEG-biotin). Substrate structures are shown schematically, and the TG-reactive groups are *colored*. Data are plotted as a correlation between average signal for unique 5-mer Lys peptides generated on two independent arrays with 0.1 ng/μl (*x axis*) or 0.01 ng/μl (*y axis*) KalbTG, respectively. Tagged by respective 5-mer sequence are 11 peptides with the highest average fluorescence signal and six single-Lys peptides tested in the GLDH-coupled assay. The data were filtered to remove noise as described under "Experimental procedures."

§ASBMB

J. Biol. Chem. (2017) 292(38) 15622–15635 **15625**
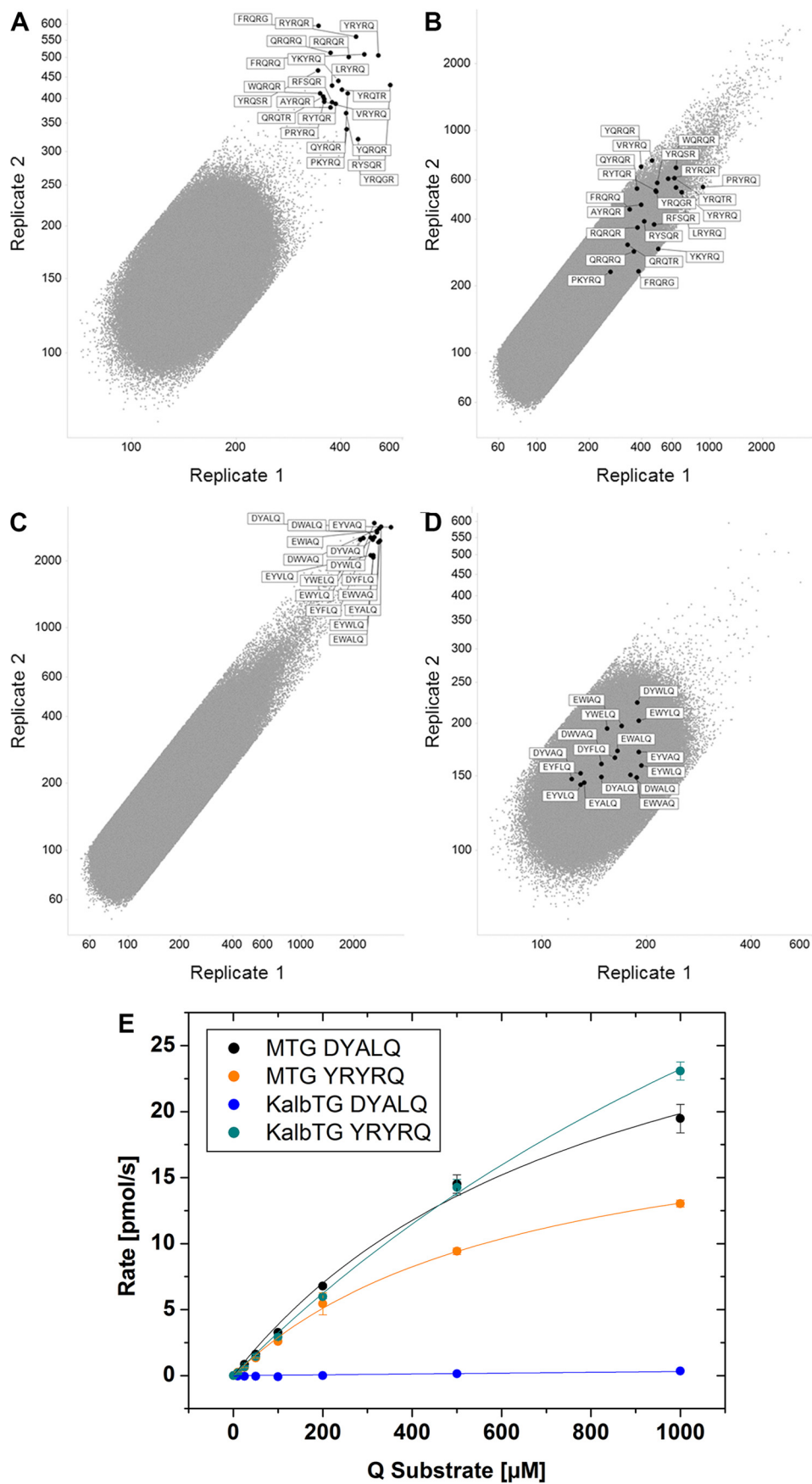
Table 1
KalbTG/MTG cross-reactivity and kinetic parameters

Parameters	MTG DYALQ	MTG YRYRQ	KalbTG DYALQ	KalbTG YRYRQ
V_{max} (pmol/s)	37	22	ND[a]	74
K_m (μM)	847	644	ND	2151
k_{cat} (s^{-1})	1.39	0.93	ND	1.92
k_{cat}/K_m (M^{-1} s^{-1})	1640	1440	ND	890

[a] ND, not detectable.

complex formation. Instead, a single absorbance peak of the unmodified IgG was observed (Fig. 4B). When further reducing the biotin label excess, it became apparent that incorporating the array-discovered KalbTG K-tag peptide in the biotin label adds another level of specificity to the reaction; IgG was almost completely covalently modified when using a 5-fold molar excess (equaling 2.5-fold excess/tagged heavy chain) of K-tag-biotin label (Fig. 4D), whereas only partial modification was observed when using a commercially available biotin-dPEG(23)-NH$_2$ label in 5-fold excess (Fig. 4E). Presumably, a significant portion of the Q-tag glutamine side chain is hydrolyzed by the transglutaminase in the absence of a suitable amine donor. The degrees of IgG biotin modification were confirmed by mass spectrometry (data not shown). Finally, the viability of biotin-modified IgG was tested in the Elecsys TSH (sandwich) assay. The chemically biotinylated capturing antibody in the R1 compartment of the original reagent pack was replaced with the IgG biotinylated via MTG or KalbTG. Both antibodies performed comparably to the original antibody of the commercial assay (Fig. 4F).

The compact crystal structure of KalbTG suggests a peptide substrate binding mode

To gain further insight into how the small KalbTG sequence can fold into such an efficient MTG and whether the substrate specificity as determined by the peptide array can be explained, we determined the crystal structure of KalbTG to a resolution of 1.9 Å. Superposition of KalbTG with its next closest homolog, the MTG from *S. mobaraensis* (24), shows a similar disc-shaped core structure (root mean square deviation of 1.5 Å) of a central β-sheet with flanking α-helices and a surface depression forming the active site cleft. However, KalbTG is more compact, having much shorter surface loops (Fig. 5A). In fact, KalbTG is the smallest MTG reported to date, even smaller by 2 kDa than the structurally unrelated *Bacillus subtilis* TG (25)(PDB entry 4P8I; 245 amino acids, 28.3 kDa), showing that the same degree of structural economy for MTGs can be reached by convergent evolution. Thermophilic proteins often have short surface loops, which reduce the entropy of the protein, and this property might also contribute to the high affinity of KalbTG.

The catalytic Cys-Asp-His triad of KalbTG is located at the bottom of the active site groove (Fig. 5A). The groove is wide enough to be covered by a kinked helical propeptide in the unprocessed enzyme (Fig. 5B), similar to what has been observed with the *S. mobaraensis* MTG zymogen (26), indicating that a similar zymogenic mechanism may be present in KalbTG. Strong difference electron density extending from Cγ of the catalytic Cys-82 (supplemental Fig. 3) showed that this residue is modified by a thiol-reactive compound, the origin of which is unknown. The electron density allows tracing of four atoms, and these were tentatively modeled as a β-mercaptoethanol disulfide adduct of Cys-82. The isosteric cysteamine or a larger compound that is disordered beyond the fourth atom cannot be excluded. A modified Cys in the active site of MTGs is also apparent when inspecting the electron density maps for Cys-110 in *S. mobaraensis* (PDB entry 3IU0), Cys-116 in *B. subtilis* (PDB entry 4PA5; modeled as cysteamine adduct in chain A), and Cys-302 in *Streptococcus suis* (PDB entry 4XZ7; adduct not modeled for one molecule in the asymmetric unit and as a string of water molecules in the second molecule). As in all other transglutaminase structures, the catalytic Cys-82 in KalbTG is located at the N terminus of an α-helix, which reduces its pK_a and increases its nucleophilicity (supplemental Fig. 4) for attack on the substrate glutamine. Of note, the catalytic Cys-82 in KalbTG is embedded 1.7 Å deeper in the active cleft than its *S. mobaraensis* MTG counterpart, raising the question of how the Gln side chain of the substrate peptide might reach it.

Attempts to model the sequence YRYRQR in α-helical conformation, guided by the structure of the propeptide (PDB entry 3IU0; Fig. 5B) were unsuccessful because the Gln side chain would not come close enough for the required attack by Cys-82. By contrast, an extended, β-strand conformation of the peptide can be placed into the active site of KalbTG with no steric repulsions and several possible hydrogen bonds between peptide and KalbTG (Fig. 5C). In this conformation, the Arg side chains would all point away from KalbTG, indicating that negative potential on KalbTG might help in electrostatic steering of positively charged peptides to the active site. This assumption is supported by the fact that both Lys and Arg are prominent in the array-discovered substrate peptides (supplemental Table 1). In the model, the second Tyr of the YRYRQR sequence packs on the His-188 side chain of KalbTG. Other active peptide sequences found in the peptide array have Phe and Gln at this position, which can entertain similar hydrophobic interactions. By contrast, the inactive peptides DYALQ and MLAQG have Ala and Leu, respectively, at this position, which would engage in fewer van der Waals interactions with His-188. A structure of a KalbTG–peptide complex will be required to

Figure 2. Array cross-reactivity of KalbTG with MTG substrates. *A*, correlation of fluorescence signal between two replicates, generated by KalbTG on a 2.8-million 5-mer array in the presence of biotinylated amine-donor substrate. 22 peptides with the highest signal intensity are highlighted within the plot. *B*, correlation plot of fluorescence signal generated by MTG on a 5-mer array in the presence of biotinylated amine-donor substrate. The fluorescence signals of the 22 peptides that exhibited the highest signals in the KalbTG array (*A*) are highlighted within the plot. *C*, plot of fluorescence signal generated by MTG on a 5-mer array in the presence of biotinylated amine-donor substrate. The 16 highest fluorescence signals are tagged by their respective 5-mer peptide sequence. *D*, plot of fluorescence signal generated by KalbTG on a 5-mer array in the presence of biotinylated amine-donor substrate. The fluorescence signals of the 16 peptides that exhibited the highest signals in the MTG array are highlighted within the plot. The average background signal over the array is ~180. *E*, plots of MTG and KalbTG activity obtained by measuring rates of NADH oxidation at 340 nm and 37 °C using various concentrations of Z-GGGDYALQGGGG (0–1 mM) and Z-GGGYRYRQGGGG (0–1 mM) in the presence of amine-donor substrate cadaverine (1 mM) in the GLDH-coupled assay. *Error bars*, S.D.
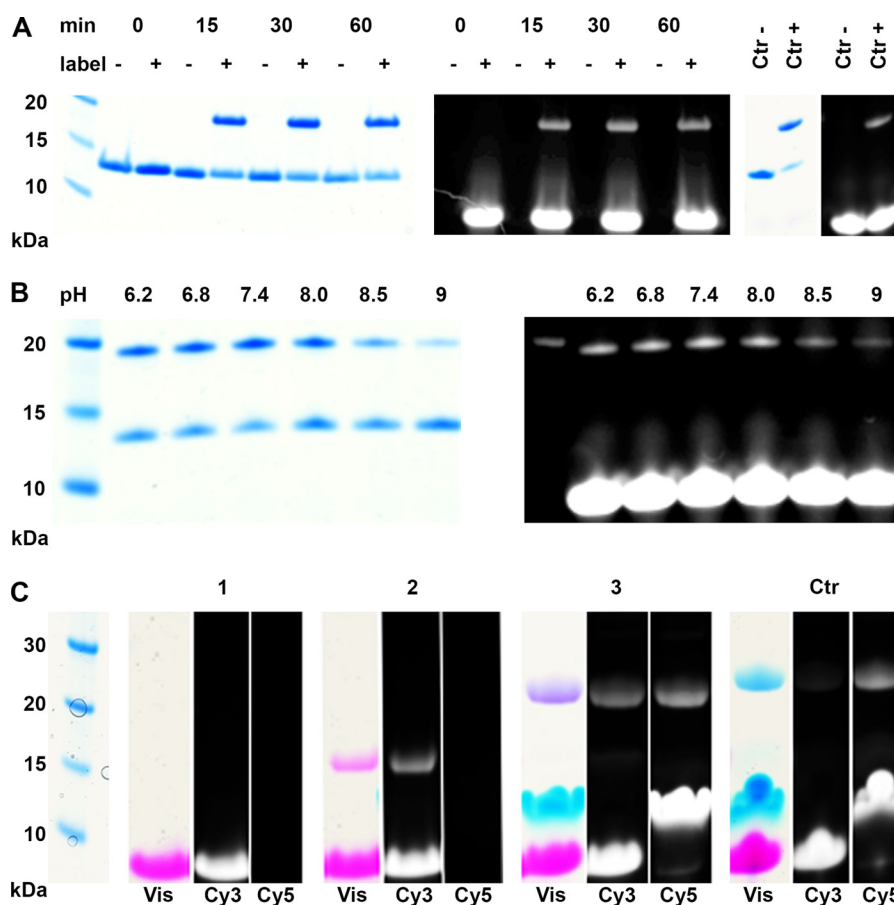
Figure 3. Site-specific and bio-orthogonal conjugation with KalbTG. *A*, time course of Cy3 labeling of a Q-tagged *T. thermophilus* SlyD moiety. Successful monolabeling is observed by the shift in electrophoretic mobility corresponding to the 5.9-kDa molecular mass of the label and by the fluorescent signal of the labeled species. A 0 – 60 min time course was performed with 10-fold label excess and controls without label addition. *Ctr−* and *Ctr+*, control reactions with SlyD containing the MTG Q-tag (DYALQ) and KalbTG or MTG, respectively. *B*, pH profile of KalbTG labeling efficacy. The highest labeling yield after 15 min of reaction time is observed at pH 7.4. *C*, dual site-specific functionalization of the construct YRYRQ-PEG27-Xa cleavage site-PEG27-PEG27-DYALQ (6.6 kDa) with two fluorophores; a mixture of peptide construct and a 10-fold excess of KalbTG K-tag-Cy3 label (*1*, 5.9 kDa) is incubated with KalbTG enzyme for 30 min to afford the monolabeled construct (*2*, 12.5 kDa). Then, with no intermediate blocking or purification steps, MTG enzyme and a 10-fold excess of MTG K-tag-Cy5 label (5.7 kDa) are added, and the mixture is incubated for 15 min to afford the dually labeled construct (*3*, 18.2 kDa). When changing the reacting order by first incubating the peptide construct and the 10-fold excess of each label with MTG for 15 min, a dually, mostly MTG K-tag-Cy5-labeled construct is afforded (*Ctr*, 18 kDa). Note that low electrophoretic mobility of the cyanine labels leads to higher apparent molecular weights compared with the protein marker.

verify the proposed binding mode. However, initial co-crystallization and soaking experiments to this extent have so far been unsuccessful.

Discussion

We describe the design, production, and structural characterization of a novel microbial transglutaminase KalbTG as well as the high-throughput screening of substrate peptides via an ultra-dense peptide array approach. The array-determined substrate sequences were further used, together with MTG and its preferred substrates, for orthogonal conjugation of biomolecules.

Establishing a viable and robust enzymatic industrial scale method for site-specific conjugation approaches, such as antibody-drug conjugates, poses high demands on the coupling enzyme; it requires a high catalytic efficiency and substrate specificity and has to be economical to produce, preferably of low molecular mass, and independent of cofactors. These requirements are met by KalbTG; it is as efficient as previously known MTGs but exhibits increased specificity and

developability. Furthermore, the enzyme requires no additives and works well in standard buffers, such as Tris, MOPS, or PBS.

Site-specific labeling with a promiscuous enzyme can be the method of choice if (*a*) the substrate is not recombinantly modified with a conjugation tag, in which case naturally occurring residues are used, and (*b*) the labeling site and label ratio can be controlled or are not of critical importance for the application at hand. Examples of this type of application are the conjugation of payloads to deglycosylated/glycosylated IgG (9), biotinylation of antibodies via glutamine residues (27), or using the Gln-295 within the heavy chain of IgGs as a substrate (28). However, using a nonspecific enzyme severely limits the range of possible applications (29) and may lead to unwanted side reactions, such as cross-linking of IgGs. Our approach combines an enzyme with naturally high substrate specificity, KalbTG, with an effective substrate screening for recombinant tags.

Biotechnological optimization and industry-scale recombinant production of MTGs are difficult, as illustrated by a plethora of recent publications (30 – 39). Genetic modification of the
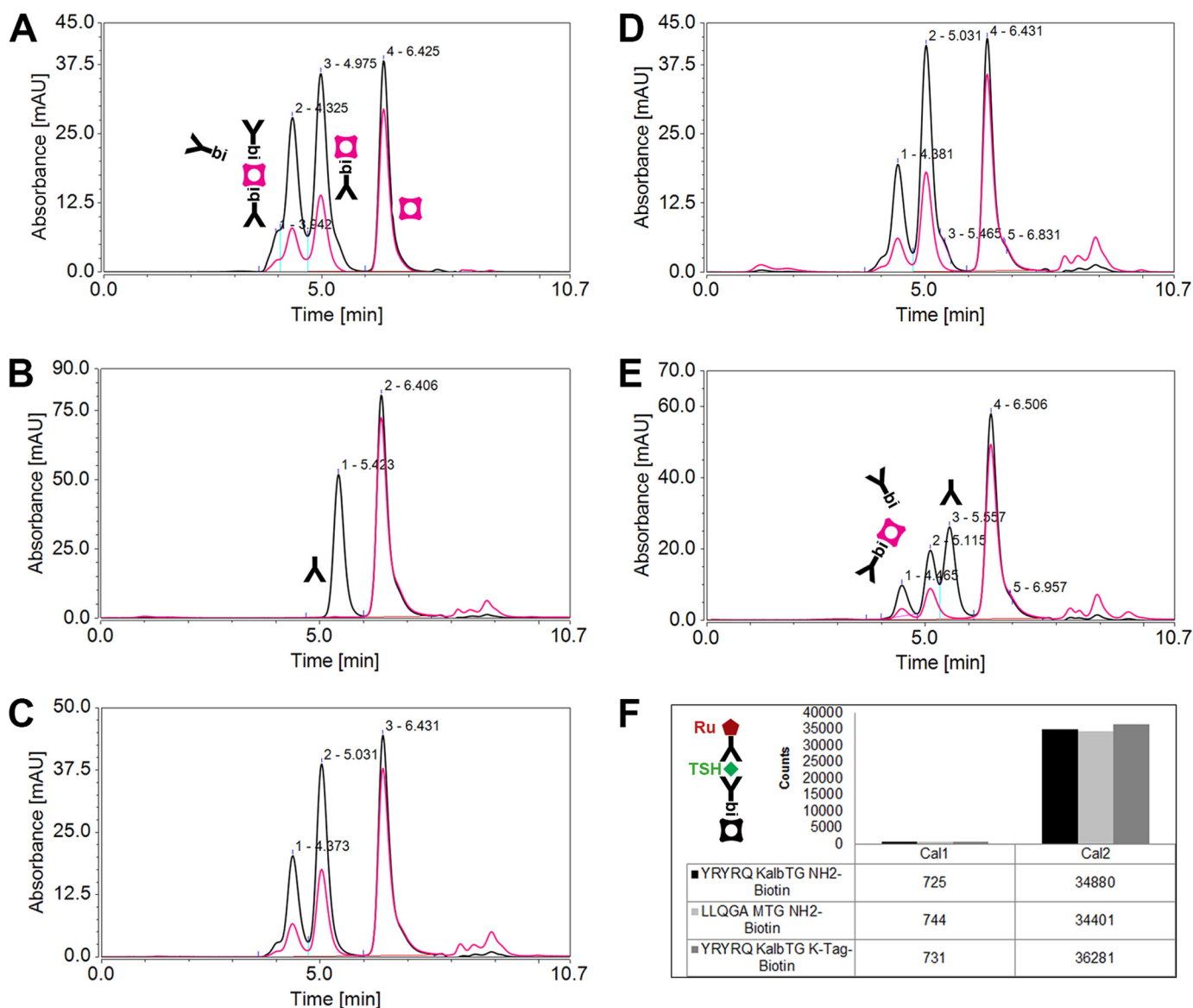
Figure 4. Enzymatic biotinylation by MTG and KalbTG and functional characterization of anti-TSH IgGs tagged C-terminally either with LLQGA or YRYRQ motifs. Shown are size-exclusion chromatography profiles of SA-FLUO–IgG-biotin complexes and unbound antibodies. *Black chromatograms*, protein absorbance (280 nm); *red chromatograms*, fluorescein absorbance (494 nm). *A*, SA-FLUO analytics of LLQGA-tagged antibody (TU1.20), labeled with MTG and a 100-fold excess of biotin-dPEG(23)-NH$_2$. *Peaks 2, 3*, and *4*, SA-FLUO molecules complexed with two, one, or no biotinylated IgGs, respectively. *Peak 1*, larger, undefined SA-FLUO–IgG complexes. *B*, SA-FLUO analytics of LLQGA-tagged antibody (TU1.20), incubated with KalbTG and a 100-fold excess of biotin-dPEG(23)-NH$_2$. *Peaks 1* and *2*, unbiotinylated IgG and unbound SA-FLUO, respectively. *C*, SA-FLUO analytics of YRYRQGGS-tagged antibody (TU1.20), labeled with KalbTG and a 25-fold excess of biotin-dPEG(23)-NH$_2$. *Peaks 1, 2*, and *3*, SA-FLUO molecules complexed with two, one, or no biotinylated IgGs, respectively. *D* and *E*, enhanced selectivity of KalbTG for K-tag amino donors. SA-FLUO analytics of YRYRQGGS-tagged antibodies (TU1.20) conjugated by KalbTG with a 5-fold excess of either K-tag-biotin label (*D*) or biotin-dPEG(23)-NH$_2$ label (*E*). *Peaks 1, 2*, and *4*, SA-FLUO molecules complexed with two, one, or no biotinylated IgGs, respectively. *Peak 3*, unbiotinylated IgG. *F*, performance of the purified IgG-biotin conjugates from *A* (LLQGA MTG NH$_2$-biotin), *C* (YRYRQ KalbTG NH$_2$-biotin), and *D* (YRYRQ KalbTG K-tag-biotin) in the TSH Elecsys immunoassay. The TSH sandwich assay principle is shown schematically.

pro-domain residues Tyr-14, Asp-20, Ile-24, and Asn-25 has been reported to increase production levels of soluble and active MTG in *E. coli*. These modifications maintain tight pro-domain interaction with the enzyme to avoid the inherent cross-linking activity of MTG that is toxic to the host cell (40) and also complicates *in vitro* handling. Overall, the MTG production processes involving the inactive pro-form have not fundamentally changed since they were originally described (16).

In our study, we found no evidence of the predicted propeptide sequence TTAQAAAVAAPTPR to bind or block the enzyme. Because the predicted sequence may not comprise the whole propeptide, a zymogen-activation approach for KalbTG would require significant effort. However, we found that the addition of NH$_4^+$, the product of the transglutaminase reaction, strongly inhibits activity of KalbTG. This simple step increased the yield and facilitated downstream processing, crucially enabling us to produce soluble transglutaminase without a specific pro-sequence and activating it by a simple dialysis step. This represents a marked improvement over any published MTG production process to date.

Peptide arrays can be highly efficient for high throughput enzyme characterization (12, 41). In this study, we synthesized

Figure 5. Structural characterization of KalbTG. *A*, KalbTG is shown as an *orange ribbon*, and *S. mobaraensis* MTG is shown in *light blue* with the 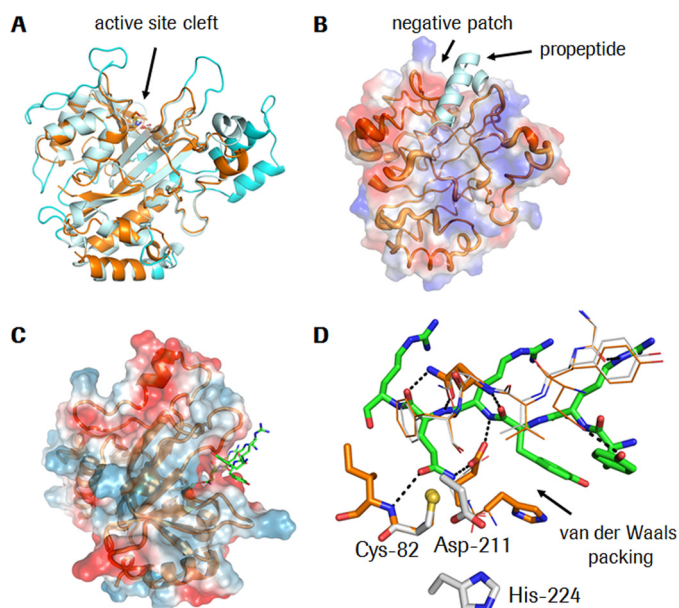longer loops (compared with KalbTG) *colored* in *cyan*, highlighting the much more compact structure of KalbTG. *B*, the surface potential of KalbTG shows a negatively charged area next to the substrate and propeptide binding site. The propeptide of *S. mobaraensis* MTG (kinked and helical) is placed at the equivalent position of KalbTG to highlight the overall similar shape of this pocket. The backbone trace of KalbTG varies in thickness as a function of *B*-values, indicating that loop regions further away from the active site are more flexible than the area around the active site, which is not involved in crystal contacts. *C*, model of the extended sequence YRYRQR (*green*) in the active site. The Gln side chain is in close vicinity to the catalytic Cys-82. The positively charged peptide is steered to the negative patch at the active site. *D*, detailed view of the peptide model (*green sticks*) from *C*. The starting KalbTG structure is shown in *gray*, and the model after energy minimization is shown in *orange*, with key residues interacting with the peptide highlighted in *thick sticks*. A number of possible hydrogen bonds are drawn as *dashed black lines*. The peptide and its surrounding residues were energy-minimized in MOLOC (www.moloc.ch)[3] (54) while keeping the remainder of KalbTG fixed. Minor movements of KalbTG are required to adapt the substrate binding site to the peptide. Figures were prepared with PyMOL.

arrays with millions of spatially addressable peptides using a light-directed, digitally controlled process and developed methods for *in situ* analysis of enzyme activity and substrate specificity for both KalbTG and MTG. Importantly, this screen allowed selection of transglutaminase substrate peptides to facilitate an orthogonal labeling with enzymes formerly known as being highly promiscuous. We consider ultra-high-density peptide arrays to be an enabling technology that will contribute to a wide field of enzyme/substrate research, including transferases, ligases, and proteases.

Whereas the compact structure of KalbTG with its comparatively shorter surface loops should have lower entropy and thus higher catalytic efficiency, this is not observed when comparing the activities of MTG and KalbTG (Table 1), indicating the presence of mutually compensating effects. However, the smaller KalbTG is an advantage for biotechnological production. Whereas in other MTGs, a helical pro-peptide is necessary to physically block the active site and avoid detrimental activity (26), elevated concentrations of NH_4^+ are sufficient to block KalbTG activity as long as it is required. This simple mechanism avoids the biotechnologically tedious step of proteolytic maturation. The negative surface potential close to the active

site might constitute a binding site for inhibitory NH_4^+. However, these are difficult to distinguish from water in electron density maps, and no NH_4^+ has been identified in the KalbTG structure based on geometric criteria (four hydrogen bond acceptors in a tetrahedral geometry). Interestingly, we found the sequence YRYRAR, inherent in a KalbTG surface loop, to closely resemble our preferred array-discovered Gln substrate peptides. However, rigidity and position of the loop prevent its interaction with the active site, therefore raising the question of what role it may play in substrate recruitment.

The KalbTG structure and the peptide model set a starting point for rational engineering of further improved or altered substrate specificity. In combination with peptide array-based high-throughput substrate screening, this will enable the creation of tailor-made enzyme–substrate pairs with extremely versatile and orthogonal uses.

Our dual-labeling experiments confirm that KalbTG and MTG constitute an orthogonal conjugation system with unparalleled ease of use, yield, and efficiency. Furthermore, as exemplified in the antibody conjugation experiments and TSH Elecsys immunoassay, KalbTG enables true site-specific labeling with minimal reaction and reagent requirements, making it highly attractive for the industrial-scale synthesis of complex protein conjugates of interest in therapeutic or diagnostic applications.

An important question that naturally arises when an artificial peptide sequence is introduced into a molecule intended for potential therapeutic use is whether it may provoke immunogenic reactions. A peptide search on Proteome database UniprotKB yielded 3,985 matches of the pentapeptide YRYRQ, of which a single one is in *Homo sapiens*. A substring search of both YRYRQ and RYESK in the Immune Epitope Database (http://www.iedb.org)[3] (53) yielded no results. As a comparison, the previously published MTG Q-tag LLQGA (2) yielded 55,145 matches in UniprotKB, 360 of which are found in human proteins. Six records of immunogenic epitopes containing the substring were found.

The introduced Q-tag peptide sequence described here can therefore be seen as a relatively rare motif in nature, putatively less prone to elicit immune responses by cross-reactive neutralizing antibodies to endogenous proteins. Furthermore, allowing specific single-amino acid substitutions in the tag sequences may strongly modulate potential immunogenicity while retaining decent transglutaminase reactivity.

The immune response (one factor being HLA/MHC reactivity) to protein therapeutics is complex and difficult to predict, leaving the "immunogenicity question" a hot topic in antibody-drug conjugate development to be closely monitored in future studies (42).

Experimental procedures

Bioinformatic methods

The web interface of NCBI Protein BLAST (43) was used to search for sequences similar to the MTG of *S. mobaraensis*. The

---

[3] Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

amino acid sequence of *S. mobaraensis* protein–glutamine γ-glutamyltransferase (Uniprot accession number P81453) was entered as a query. Manual screening of the results for *E* values $< 10^{-10}$ and polypeptide sequences shorter than that of *S. mobaraensis* MTG yielded hypothetical gene product KALB_7456 from bacterial strain *K. albida* DSM 43870 (GenBank™ accession number AHI00814.1, Uniprot accession number W5WHY8). Sequence alignment of the *S. mobaraensis* and the *K. albida* sequences with Clustal Omega version 1.2.1 (44) yielded a value of 32% in the percent identity matrix and identified conservation of the catalytically active residues of MTG (Cys-140, Asp-331, and His-350; P81453 numbering). The ProP 1.0 server from the Technical University of Denmark (45) was used to predict the propeptide and signal sequences of the hypothetical *K. albida* microbial transglutaminase. VAAPTPR ↓ AP was the only predicted propeptide cleavage site with a score (0.513) above the threshold.

Production of KalbTG

The gene sequence encoding for hypothetical *K. albida* microbial transglutaminase was codon-optimized for *E. coli* expression (Roche Sequence Analysis Web interface), chemically synthesized (GeneArt, ThermoFisher, Regensburg), and cloned via fragment exchange cloning (19) into a vector, conferring two N-terminal moieties of sensitive-to-lysis D chaperones (SlyD, Uniprot entry P0A9K9), truncated after Asp-165 (20), followed by a protease factor Xa cleavage site and including a C-terminal octa-His tag. The vector is based on the pQE-80 series by Qiagen, including isopropyl 1-thio-β-D-galactopyranoside-inducible protein expression by T5 promotor and conferring resistance to ampicillin. Plasmid preparation and transformation of chemically competent *E. coli* Bl21 Tuner cells with the expression plasmid were performed according to standard molecular biology protocols (46).

Fermentation was carried out at 35 °C for 26 h, until an A_{600} of 44 was reached. Cells were harvested and resuspended in buffer containing 50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM DTT, and 10 mM $(NH_4)_2SO_4$. Cells were disrupted by a high-pressure homogenizer at 800 bars. The resulting cellular extract was pretreated with 1–3% Polymin-G20 and then loaded onto a Q-Sepharose XL column (strong anion exchange matrix; GE Healthcare Life Sciences) at a protein concentration of ~30 mg/ml. Bound protein was washed with 20 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM DTT, 10 mM $(NH_4)_2SO_4$, and 150 mM NaCl and then eluted with a 30-column volume gradient from 150 to 500 mM NaCl. The eluate was dialyzed (10,000 molecular weight cutoff) against 20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 0.1 mM DTT, 10 mM $(NH_4)_2SO_4$, 500 mM NaCl; concentrated; and loaded onto a nickel-nitrilotriacetic acid column. Bound, His-tagged protein was washed with 20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 0.1 mM DTT, 10 mM $(NH_4)_2SO_4$, 500 mM NaCl, 25 mM imidazole and eluted with a 20-column volume gradient from 25 to 200 mM imidazole. Purified protein was dialyzed (10,000 molecular weight cutoff) against 20 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, and 10 mM $(NH_4)_2SO_4$, pH 8.0; concentrated to 1.77 mg/ml; analyzed by an SDS-PAGE and GLDH activity assay; and frozen in 10-mg aliquots at −80 °C. Prior to use, $(NH_4)_2SO_4$ was removed by dialysis with a 10-kDa molecular weight cutoff filter to yield the active enzyme. For some applications, Factor Xa proteolysis was performed to remove the 2× SlyD portion from the KalbTG construct.

Peptide array

A library of 1,360,732 unique 5-mer peptides was designed by using all combinations of 18 natural amino acids, excluding cysteine and methionine as well as any dimer or a longer repeat of the same amino acid and any peptide containing HR, RH, HK, KH, RK, KR, HP, and PQ sequences. The library was synthesized in duplicate on the same array by using maskless light-directed peptide array synthesis (Roche Nimblegen). Each 5-mer peptide was flanked from both the N and C terminus by 3-amino acid linkers synthesized by using a mixture of Gly and Ser at a 3:1 ratio. Peptide synthesis was accomplished through light-directed array synthesis in a Roche NimbleGen maskless array synthesizer using an amino-functionalized substrate as reported previously (13). Prior to use, arrays were incubated in blocking buffer (10 mM Tris-HCl, pH 7.4, 1% alkali-soluble casein (EMD Chemicals), 0.05% Tween 20) at room temperature for 1 h.

To test KalbTG specificity for Gln substrate, *N*-(biotinyl)cadaverine (Zedira) was used as a substitute for a Lys substrate to biotinylate Gln peptides on a peptide array. A KalbTG-labeling reaction was performed in 1,200 μl of 100 mM Tris-HCl, pH 8, 1 mM DTT, 50 μM *N*-(biotinyl)cadaverine, 0.2 ng/μl KalbTG in a SecureSeal™ chamber (Grace Bio-Labs) at 37 °C for 45 min. After incubation, the chamber was removed, and the array was washed in 20 mM Tris-HCl, pH 7.8, 0.2 M NaCl, 1% SDS for 1 min, followed by a 1-min wash in 20 mM Tris-HCl. Biotin linked to the array was stained with 0.3 μg/ml CyTM5–streptavidin (GE Healthcare) in blocking buffer at room temperature for 1 h. Cy5 fluorescence intensity was measured with a MS200 scanner (Roche Nimblegen) at a resolution of 2 μm and wavelength of 635 nm.

Because array synthesis is digitally programmed, a new array can be designed and synthesized in a matter of few days. This allowed us to quickly verify and mature sequence motifs selected with the 5-mer arrays. In the case of maturation of the APRYRQRAA peptide, the new array design was created by extending the YRYRQ and RYRQR 5-mer sequences with all possible combinations of 2 amino acids from both the N and C terminus (*i.e.* each 5-mer sequence was extended to 160,000 9-mers with an invariant core motif). 9-Mers with the highest KalbTG activity on array were used to design another new array that included all possible single and double substitutions of selected 9-mers with all 20 natural amino acids. This step allowed us to "mutate" the original 5-mer core sequence and evaluate restrictions imposed by the 5-mer array peptide selection.

To test KalbTG specificity for Lys substrates, chemically synthesized Z-APRYRQRAAGGG-PEG-biotin peptide was used as a Gln substrate to biotinylate Lys peptides. Array biotinylation was done as described above with 0.01 or 0.1 ng/μl KalbTG and 0.8 μM peptide in 100 mM Tris-HCl, pH 8, 1 mM DTT, 0.05% Tween 20 at 37 °C for 15 min.

New transglutaminase for site-specific coupling

MTG reactions on the peptide array were performed in 100 mM Tris-HCl, pH 8, 1 mM DTT, 50% protein-free blocking buffer (Thermo Scientific) with 25 μM N-(biotinyl)cadaverine and 0.2 ng/μl MTG at 42 °C for 1 h.

Control experiments with CyTM5–streptavidin only were performed to show that no nonspecific streptavidin binding to the array occurred. Imperfections of the array surface typically result in a high signal intensity noise affecting ~1–3% of peptide features. To remove noise, peptides for which $|(S_1 - S_2)/(S_1 + S_2)| > 0.2$ were excluded from analysis, where S_1 and S_2 represent signal intensity of replicates 1 and 2, respectively.

Because it is not technically possible to evaluate the yields and quality of each peptide on the array, a quality control process was developed where various peptides and their substitution variants that bind to streptavidin are synthesized on an array.[4] The relative binding signals of streptavidin to these peptides is evaluated from array to array and is used to assess indirect quality of synthesis and overall performance of a synthesis run.

GLDH-coupled assay

To determine whether the KalbTG peptides selected in the array assay were also preferred substrates in a solution reaction and to quantify cross-reactivity of KalbTG and MTG with various substrates, a continuous GLDH-coupled assay for MTG activity (21) was applied. For glutamine substrate evaluation, the assay was performed in a transparent 96-well microtiter plate in the presence of 500 μM α-ketoglutarate, 500 μM or 1 mM cadaverine as amine donor substituting for a Lys peptide, 2 units/ml of GLDH, 500 μM NADH, and Gln-containing substrate peptide (Z-GGG-QRWRQGGGG, Z-GGGWRYRQGGGG, Z-GGGYRYRQG-GGG, Z-GGGRYRQRGGGG, Z-GGGRYSQRGGGG, Z-GGGF-RQRQGGGG, Z-GGGRQRQRGGGG, Z-GGGFRQRGGGGG, Z-GGGQRQRQGGGG, Z-GGGYKYRQGGGG, Z-GGGQYRQ-RGGGG, Z-GGGDYALQGGGG, or Z-GGGMLAQGSGGG) concentrations ranging between 0 and 1 mM in 200 mM MOPS, 1 mM EDTA, pH 7.2 (total volume per well 200 μl).

For amine substrate evaluation, assay conditions were identical, but 100 μM each of amine substrate (Z-GGG-RYSKYGGGG, Z-GGGAYRTKGGGG, Z-GGGRYRSKGGGG, Z-GGGYKGRGGGGG, Z-GGGRYGKSGGGG, Z-GGGRYES-KGGGG, Z-GGGPGRYKGGGG, Z-GGGARSKLGGGG, or cadaverine) and 200 μM Gln donor (Z-GGGYRYRQGGGG or Z-GGGDYALQGGGG) were used.

Reactions were started by the addition of 5 μg/ml MTG (Zedira) or KalbTG, and the oxidation of NADH was continuously recorded at 340 nm for 60 min using a Biotek Synergy H4 microplate reader, temperature-controlled at 37 °C, with short shaking intervals before each measurement. After a short lag phase where the GLDH was saturated by TG-mediated release of ammonia, linear rates of absorbance *versus* time, corresponding to TG turnover, were observed and subjected to Michaelis–Menten kinetic analysis. Rates of absorbance (milli-absorbance units/min) were converted into molar rates of NADH turnover (pmol/s) using the formula (previously deter-

mined by an NADH standard curve), turnover rate = ((|absorbance rate|) $\times$ 1.111).

Labeling assays

The chaperone SlyD from *Thermus thermophilus* (Uniprot number Q5SLE7) was used as a labeling scaffold for KalbTG, by recombinant grafting of a KalbTG glutamine donor sequence (Q-tag) onto the FKBP-type domain, which yielded the polypeptide sequence MKVGQDKVVTIRYTLQVEGEVL-DQGELSYLHGHRNLIPGLEEALEGREEGEAFQAHVPAE-KAYGAGSGGGGRYRQRGGGGGSSGKDLDFQVEVVKVR-EATPEELLHGHAHHHHHHHH.

The protein was produced in *E. coli* Bl21 Tuner and purified by standard techniques (HisTrap, Superdex 200 pg).

Labeled peptides were chemically synthesized, to be composed of a Z-protecting group at the N-terminal amino group, a transglutaminase lysine donor sequence (K-tag) on the N-terminal sequence part, and a Cy3 or Cy5 fluorescent dye at the C terminus, spaced by a linker sequence.

The primary chemical structures were as follows: Z-RYESK-G*O2Oc*EUEUEUEUEUEUEUEUEUEUEUEUEUEU-UEUEUEU*C(sCy3-MH)*-OH (KalbTG K-tag-Cy3, 5,863.9 g/mol) and Z-RSKLG*O2Oc*EUEUEUEUEUEUEUEUEU-EUEUEUEUEUEUEUEUEUEUEU*C(Cy5-MH)*-OH (MTG K-tag-Cy5, 5,723.9 g/mol), where U is β-alanine, *O2Oc* is 8-amino-3,6-dioxaoctanoic acid, and *C(sCy3-MH)* and *C(Cy5-MH)* represent a C-terminal cysteine modified post-synthesis by sulfo-Cy3 maleimide and Cy5 maleimide, respectively.

For the orthogonal labeling experiment, a molecule containing both KalbTG and MTG Q-tags was chemically synthesized, with a primary chemical structure of Z-GGGYRYRQGGG*PE-G27*GIEGRG*PEG27**PEG27*GGGDYALQGG-OH (6,620.6 g/mol).

All peptides were synthesized via standard Fmoc-based solid phase peptide synthesis in a 0.25-mmol scale using commercially available building blocks. After solid-phase synthesis, peptides were cleaved with TFA/triisopropylsilane/water (95:2.5:2.5) and precipitated with diisopropylether followed by purification via RP18-HPLC using a water/TFA acetonitrile gradient. Dye labeling was achieved by reaction of the peptides with sulfo-Cy3 maleimide (Lumiprobe) and sulfo-Cy5 maleimide (GE Healthcare), respectively. Purification of dye-labeled peptides was achieved by RP18-HPLC using a water/TFA acetonitrile gradient. Identity of the peptides was confirmed by LC-MS (Thermo Scientific RSLC-MSQplus system), applying a Kinetex C18 2.6-μm, 50 $\times$ 3-mm column (Phenomenex).

If not noted otherwise, labeling reactions were performed for 15 min at 37 °C in the presence of 72 μM substrate protein, 720 μM label peptide, and 1 μM transglutaminase in 200 mM MOPS, pH 7.2, and 1 mM EDTA. For the pH-dependent labeling profile, experiments were performed in 200 mM MOPS buffer adjusted to pH 6.2, 6.8, and 7.4 with NaOH or HCl or 200 mM Tris buffer adjusted to pH 8.0, 8.5, and 9.0 HCl. For the orthogonal labeling experiment, 1.5 μM KalbTG was added to a volume of 20 μl containing 100 μM substrate peptide and 1 mM KalbTG K-tag-Cy3. After incubation for 30 min at 37 °C, 1 mM MTG K-tag-Cy5 and 1.5 μM MTG were added and incubated

---

for an additional 15 min at 37 °C. The reaction was stopped by the addition of 50 mM TCA. Samples were taken between incubation steps and analyzed by SDS-PAGE and in-gel fluorescence (Bio-Rad ChemiDoc gel documentation system, Cy3 and Cy5 LED, and filter sets).

Crystallization and structure determination of KalbTG

KalbTG in PBS was crystallized at 22 °C using the sitting drop (200-nl) vapor diffusion method by 1:1 mixing of 8 mg/ml protein with an unbuffered reservoir consisting of 0.2 M ammonium tartrate, 20% PEG 3350. Crystals were cryoprotected in reservoir solution containing 20% ethylene glycol before flash-cooling in liquid nitrogen. Data were collected at 100 K at SLS beamline PX-II using a Pilatus 6M detector and integrated and scaled in space group P3 with XDS (47). The $l = 3n$ reflections have I/σ of >9, rendering the presence of a screw axis unlikely. Self-Patterson and twinning analyses did not reveal suspicious data pathologies. The cell volume is consistent with two or three KalbTG molecules in the asymmetric unit, with Matthews parameters of 3.5 Å^3/Da and 2.3 Å^3/Da, respectively. Whereas the $\kappa = 180°$ section of the self-rotation function did not indicate a 2-fold NCS axis, a peak in the $\kappa = 164°$ section at $\omega = 0°$, $\phi = 0°$ indicated that at least two molecules in the asymmetric unit are related by a 164° rotation, which turned out to be correct after molecular replacement. Data collection statistics are summarized in supplemental Table 1.

The structure of KalbTG (226 residues) was determined by molecular replacement using the *S. mobaraensis* transglutaminase (354 residues, PDB entry 3IU0) as the search model. The first attempts using the complete *S. mobaraensis* TG were unsuccessful, probably because the enzymes are of very different sizes. The two transglutaminases share 28.2% sequence identity and 38.9% sequence similarity over the entire length of KalbTG. A variant of *S. mobaraensis* TG devoid of loop regions and trimmed to the hydrophobic core resulted in a potential solution with PHASER (48) when searching for two molecules in the asymmetric unit in space group P3 with a log-likelihood gain of 213. Trigonal space groups P31 and P32 did not yield solutions, consistent with the high intensities of the $l = 3n$ reflections. The molecular replacement model was refined in BUSTER (49) to an initial R_{free} of 46%. Some secondary structure elements were visible in the electron density maps and were included in the model, which was then submitted to 10 cycles of automatic model building and refinement in CBUC-CANEER and REFMAC5 (50). The resulting model included the entire KalbTG catalytic domain and had an R_{free} of 30%. The structure was completed in COOT (51) and refined with PHENIX (52) to an R_{free} value of 23% at 1.9 Å resolution with excellent stereochemistry. There are two molecules in the asymmetric unit that are virtually identical (root mean square deviation 0.26 Å over all atoms) and exhibit excellent electron density (supplemental Fig. 3) and stereochemistry (supplemental Table 2). Model refinement statistics are collected in supplemental Table 2. The first 19 N-terminal amino acids (MGGG-STTAQAAAVAAPTPR) and the C-terminal artificial GGGS-His$_8$ tag are disordered in the structure.

DSC

Measurements were performed with a starting temperature of 20 °C and a final temperature of 90 °C on a VP-capillary DSC instrument (MicroCal/GE Healthcare) using PBS as a reference. A scanning rate of 90 °C/h was applied. The mature and active KalbTG enzyme with the 2× SlyD removed was measured at a protein concentration of 0.7 mg/ml in PBS. Data analysis was performed with Origin version 7, SW 2.0.

Antibody expression and purification

Heavy and light chains of TU1.20 (mouse monoclonal antibody against TSH) were cloned into standard mammalian expression vectors featuring a CMV promotor (46) and including different conjugation tags (LLQGA and GGGSYR-YRQGGGS) at the heavy chain C terminus. Both plasmids encoding heavy and light chain were co-transfected into suspension-adapted human embryonic kidney HEK293-F cells (Life Technologies/Thermo Fischer Scientific). HEK293-F cells were cultured in shaker flasks at 37 °C in FreeStyle 293 expression medium (Thermo Fisher Scientific) under serum-free medium conditions. The cells were transfected at ~2 × 10^6 vital cells/ml with the expression plasmids (0.5 mg/liter of cell culture) complexed by PEIpro (Polyplus) transfection reagent (1.3 ml/liter of cell culture) in PBS buffer. The culture supernatant was collected at day 7 post-transfection by centrifugation. IgG was purified via one-step protein A affinity purification (HiTrap MabSelect SuRe, GE Healthcare) according to the supplier's instructions.

Biotinylation

TU1.20 tagged at the C terminus of each heavy chain with LLQGA (26.7 µM) was conjugated for 16 h at 37 °C with a 100-fold excess of biotin-dPEG(23)-NH$_2$ (Iris Biotech) via microbial transglutaminase (Zedira) (0.1 units/µl) or incubated (100 µM TU1.20) for 16 h at 37 °C with a 100-fold excess of biotin-dPEG(23)-NH$_2$ and KalbTG (0.5 µM) as a control. To achieve quantitative biotinylation, TU1.20 tagged at the C terminus of each heavy chain with YRYRQGGGS (100 µM) was conjugated 16 h at 37 °C with a 25-fold excess of biotin-dPEG(23)-NH$_2$ via KalbTG (0.5 µM). To assess amino-donor specificity, the YRYRQGGGS-tagged IgG was conjugated with either a 5-fold excess of biotin-dPEG(23)-NH$_2$ (Iris Biotech) or K-tag-biotin label (Z-RYESKG-PEG27-K(biotin)) via KalbTG (0.5 µM). All conjugates were purified via Superdex 200 to remove enzyme and unconjugated label.

TSK GFC300 SA-FLUO analytics

IgG biotinylation was assessed by complex formation with SA-FLUO. SA-FLUO–IgG-biotin complexes were analyzed by analytical size-exclusion chromatography (TOSOH GFC300). In detail, equal volumes (50 µl) of biotinylated IgG ($c = 0.6$ mg/ml) and SA-FLUO ($c = 0.3$ mg/ml) were mixed and incubated for 5 min at room temperature. 20 µl of 100 mM biotin were added. 25 µl of each sample were injected on a TSK GFC300 SW 7.8 × 150-mm column, and extinction profiles were monitored at 280 and 494 nm. Successful biotinylation is

SASBMB

J. Biol. Chem. (2017) 292(38) 15622–15635 **15633**

monitored by the existence of SA-FLUO–IgG-Biotin complexes having extinction profiles at 280 and 494 nm.

Elecsys TSH immunoassay

Site-specifically labeled IgG-biotin conjugates of TU1.20 were used at a concentration of 2.5 µg/ml in original TSH buffer replacing the original reagent in the R1 compartment of the Elecsys immunoassay rackpack (Roche Diagnostics). Cal1 and Cal2 from the TSH CalSet (Roche Diagnostics) were analyzed by the TSH Elecsys immunoassay (Roche Diagnostics) on a Cobas E170 module.

References

1. Rashidian, M., Dozier, J. K., and Distefano, M. D. (2013) Enzymatic labeling of proteins: techniques and approaches. *Bioconjug. Chem.* **24,** 1277–1294
2. Strop, P., Delaria, K., Foletti, D., Witt, J. M., Hasa-Moreno, A., Poulsen, K., Casas, M. G., Dorywalska, M., Farias, S., Pios, A., Lui, V., Dushin, R., Zhou, D., Navaratnam, T., Tran, T. T., *et al.* (2015) Site-specific conjugation improves therapeutic index of antibody drug conjugates with high drug loading. *Nat. Biotechnol.* **33,** 694–696
3. Zhou, Q., and Kim, J. (2015) Advances in the development of site-specific antibody-drug conjugation. *Anticancer Agents Med. Chem.* **15,** 828–836
4. Ando, H., Adachi, M., Umeda, K., Matsuura, A., Nonaka, M., Uchio, R., Tanaka, H., and Motoki, M. (1989) Purification and characteristics of a novel transglutaminase derived from microorganisms. *Agric. Biol. Chem.* **53,** 2613–2617
5. Martins, I. M., Matos, M., Costa, R., Silva, F., Pascoal, A., Estevinho, L. M., and Choupina, A. B. (2014) Transglutaminases: recent achievements and new sources. *Appl. Microbiol. Biotechnol.* **98,** 6957–6964
6. Kieliszek, M., and Misiewicz, A. (2014) Microbial transglutaminase and its application in the food industry: a review. *Folia Microbiol.* **59,** 241–250
7. Strop, P. (2014) Versatility of microbial transglutaminase. *Bioconjug. Chem.* **25,** 855–862
8. Folk, J. E., and Cole, P. W. (1966) Transglutaminase: mechanistic features of the active site as determined by kinetic and inhibitor studies. *Biochim. Biophys. Acta* **122,** 244–264
9. Jeger, S., Zimmermann, K., Blanc, A., Grünberg, J., Honer, M., Hunziker, P., Struthers, H., and Schibli, R. (2010) Site-specific and stoichiometric modification of antibodies by bacterial transglutaminase. *Angew. Chem. Int. Ed. Engl.* **49,** 9995–9997
10. Sugimura, Y., Yokoyama, K., Nio, N., Maki, M., and 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
11. Lee, J. H., Song, E., Lee, S. G., and Kim, B. G. (2013) High-throughput screening for transglutaminase activities using recombinant fluorescent proteins. *Biotechnol. Bioeng.* **110,** 2865–2873
12. Malešević, M., Migge, A., Hertel, T. C., and Pietzsch, M. (2015) A fluorescence-based array screen for transglutaminase substrates. *Chembiochem* **16,** 1169–1174
13. Forsström, B., Axnäs, B. B., Stengele, K. P., Bühler, J., Albert, T. J., Richmond, T. A., Hu, F. J., Nilsson, P., Hudson, E. P., Rockberg, J., and Uhlen, M. (2014) Proteome-wide epitope mapping of antibodies using ultra-dense peptide arrays. *Mol. Cell Proteomics* **13,** 1585–1597
14. Rebets, Y., Tokovenko, B., Lushchyk, I., Rückert, C., Zaburannyi, N., Bechthold, A., Kalinowski, J., and Luzhetskyy, A. (2014) Complete genome sequence of producer of the glycopeptide antibiotic Aculeximycin *Kutzneria albida* DSM 43870T, a representative of minor genus of Pseudonocardiaceae. *BMC Genomics* **15,** 885
15. Katsuya, S., Noriki, N., and Masao, M. (1996) Some characteristics of a microbial protein cross-linking enzyme: transglutaminase. In *Macromolecular Interactions in Food Technology*, pp. 271–280, American Chemical Society, Washington, D. C.
16. Pasternack, R., Dorsch, S., Otterbach, J. T., Robenek, I. R., Wolf, S., and Fuchsbauer, H. L. (1998) Bacterial pro-transglutaminase from *Streptoverticillium mobaraense*: purification, characterisation and sequence of the zymogen. *Eur. J. Biochem.* **257,** 570–576
17. Zotzel, J., Pasternack, R., Pelzer, C., Ziegert, D., Mainusch, M., and Fuchsbauer, H. L. (2003) Activated transglutaminase from *Streptomyces mobaraensis* is processed by a tripeptidyl aminopeptidase in the final step. *Eur. J. Biochem.* **270,** 4149–4155
18. Sarafeddinov, A., Arif, A., Peters, A., and Fuchsbauer, H. L. (2011) A novel transglutaminase substrate from *Streptomyces mobaraensis* inhibiting papain-like cysteine proteases. *J. Microbiol. Biotechnol.* **21,** 617–626
19. Geertsma, E. R., and Dutzler, R. (2011) A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry* **50,** 3272–3278
20. Scholz, C., Schaarschmidt, P., Engel, A. M., Andres, H., Schmitt, U., Faatz, E., Balbach, J., and Schmid, F. X. (2005) Functional solubilization of aggregation-prone HIV envelope proteins by covalent fusion with chaperone modules. *J. Mol. Biol.* **345,** 1229–1241
21. Oteng-Pabi, S. K., and Keillor, J. W. (2013) Continuous enzyme-coupled assay for microbial transglutaminase activity. *Anal. Biochem.* **441,** 169–173
22. Kamiya, N., Doi, S., Tanaka, Y., Ichinose, H., and Goto, M. (2007) Functional immobilization of recombinant alkaline phosphatases bearing a glutamyl donor substrate peptide of microbial transglutaminase. *J. Biosci. Bioeng.* **104,** 195–199
23. Andres, H., Duefel, H., Gerg, M., Kowalewsky, F., Scholz, C., and Schraeml, M. (August 11, 2012) An antibody specifically binding to insulin-like growth factor-1. International Patent WO/2012/150321
24. Kashiwagi, T., Yokoyama, K., Ishikawa, K., Ono, K., Ejima, D., Matsui, H., and Suzuki, E. (2002) Crystal structure of microbial transglutaminase from *Streptoverticillium mobaraense*. *J. Biol. Chem.* **277,** 44252–44260
25. Fernandes, C. G., Plácido, D., Lousa, D., Brito, J. A., Isidro, A., Soares, C. M., Pohl, J., Carrondo, M. A., Archer, M., and Henriques, A. O. (2015) Structural and functional characterization of an ancient bacterial transglutaminase sheds light on the minimal requirements for protein cross-linking. *Biochemistry* **54,** 5723–5734
26. Yang, M. T., Chang, C. H., Wang, J. M., Wu, T. K., Wang, Y. K., Chang, C. Y., and Li, T. T. (2011) Crystal structure and inhibition studies of transglutaminase from *Streptomyces mobaraense*. *J. Biol. Chem.* **286,** 7301–7307
27. Josten, A., Haalck, L., Spener, F., and Meusel, M. (2000) Use of microbial transglutaminase for the enzymatic biotinylation of antibodies. *J. Immunol. Methods* **240,** 47–54
28. Dennler, P., Chiotellis, A., Fischer, E., Brégeon, D., Belmant, C., Gauthier, L., Lhospice, F., Romagne, F., and Schibli, R. (2014) Transglutaminase-

based chemo-enzymatic conjugation approach yields homogeneous antibody-drug conjugates. *Bioconjug. Chem.* **25,** 569–578

29. Mori, Y., Goto, M., and Kamiya, N. (2011) Transglutaminase-mediated internal protein labeling with a designed peptide loop. *Biochem. Biophys. Res. Commun.* **410,** 829–833

30. Date, M., Yokoyama, K., Umezawa, Y., Matsui, H., and Kikuchi, Y. (2004) High level expression of *Streptomyces mobaraensis* transglutaminase in Corynebacterium glutamicum using a chimeric pro-region from *Streptomyces cinnamoneus* transglutaminase. *J. Biotechnol.* **110,** 219–226

31. Du, K., Liu, Z., Cui, W., Zhou, L., Liu, Y., Du, G., Chen, J., and Zhou, Z. (2014) pH-dependent activation of *Streptomyces hygroscopicus* transglutaminase mediated by intein. *Appl. Environ. Microbiol.* **80,** 723–729

32. Lin, S. J., Hsieh, Y. F., Wang, P. M., and Chu, W. S. (2007) Efficient purification of transglutaminase from recombinant *Streptomyces platensis* at various scales. *Biotechnol. Lett.* **29,** 111–115

33. Liu, S., Zhang, D., Wang, M., Cui, W., Chen, K., Du, G., Chen, J., and Zhou, Z. (2011) The order of expression is a key factor in the production of active transglutaminase in *Escherichia coli* by co-expression with its pro-peptide. *Microb. Cell Fact.* **10,** 112

34. Marx, C. K., Hertel, T. C., and Pietzsch, M. (2008) Random mutagenesis of a recombinant microbial transglutaminase for the generation of thermostable and heat-sensitive variants. *J. Biotechnol.* **136,** 156–162

35. Pfleiderer, C., Mainusch, M., Weber, J., Hils, M., and Fuchsbauer, H. L. (2005) Inhibition of bacterial transglutaminase by its heat-treated proenzyme. *Microbiol. Res.* **160,** 265–271

36. Yang, H. L., Pan, L., and Lin, Y. (2009) Purification and on-column activation of a recombinant histidine-tagged pro-transglutaminase after soluble expression in *Escherichia coli. Biosci. Biotechnol. Biochem.* **73,** 2531–2534

37. Sommer, C., Volk, N., and Pietzsch, M. (2011) Model based optimization of the fed-batch production of a highly active transglutaminase variant in *Escherichia coli. Protein Expr. Purif.* **77,** 9–19

38. Sommer, C., Hertel, T. C., Schmelzer, C. E., and Pietzsch, M. (2012) Investigations on the activation of recombinant microbial pro-transglutaminase: in contrast to proteinase K, dispase removes the histidine-tag. *Amino Acids* **42,** 997–1006

39. Wang, K., Wang, B., Yang, H. L., and Pan, L. (2013) New strategy for specific activation of recombinant microbial pro-transglutaminase by introducing an enterokinase cleavage site. *Biotechnol. Lett.* **35,** 383–388

40. Rickert, M., Strop, P., Lui, V., Melton-Witt, J., Farias, S. E., Foletti, D., Shelton, D., Pons, J., and Rajpal, A. (2016) Production of soluble and active microbial transglutaminase in *Escherichia coli* for site-specific antibody drug conjugation. *Protein Sci.* **25,** 442–455

41. Thiele, A., Stangl, G. I., and Schutkowski, M. (2011) Deciphering enzyme function using peptide arrays. *Mol. Biotechnol.* **49,** 283–305

42. Jawa, V., Cousens, L. P., Awwad, M., Wakshull, E., Kropshofer, H., and De Groot, A. S. (2013) T-cell dependent immunogenicity of protein therapeutics: preclinical assessment and mitigation. *Clin. Immunol.* **149,** 534–555

43. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T. L. (2008) NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36,** W5–W9

44. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7,** 539

45. Duckert, P., Brunak, S., and Blom, N. (2004) Prediction of proprotein convertase cleavage sites. *Protein Eng. Des. Sel.* **17,** 107–112

46. Green, M. R., and Sambrook, J. (2012) *Molecular Cloning: A Laboratory Manual,* 4th Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

47. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132

48. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674

49. Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M., and Bricogne, G. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2210–2221

50. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242

51. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501

52. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221

53. Vita, R., Overton, J. A., Greenbaum, J. A., Ponomarenko, J., Clark, J. D., Cantrell, J. R., Wheeler, D. K., Gabbard, J. L., Hix, D., Sette, A., and Peters, B. (2014) The immune epitope database (IEDB) 3.0. *Nucleic Acids Res.* **34,** D405–D412, 10.1093/nar/gku938

54. Gerber, P. (1986) *Moloc molecular modeling package,* Gerber Molecular Design, Basel, Switzerland