Selectivity in the Post-Translational, Transglutaminase-Dependent Acylation of Lysine Residues[†]

S. N. Prasanna Murthy,[‡] Thomas J. Lukas,[§] Theodore S. Jardetzky,[||] and Laszlo Lorand*[,‡]

Department of Cell and Molecular Biology and Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, 303 East Chicago Avenue, Chicago, Illinois 60611, and Department of Structural Biology, Stanford University School of Medicine, Stanford, California 94305

Received December 19, 2008; Revised Manuscript Received February 12, 2009

ABSTRACT: Transglutaminases (TGs) are known to exhibit remarkable specificities not only for the Q (or Gln) sites but also for the K (or Lys) sites of proteins with which they react. To gain further insight into K-site specificity, we examined the reactions of dansyl-ε-aminocaproyl-GlnGlnIleVal with three chemically and structurally well-characterized proteins (bovine pancreatic ribonuclease A, bovine pancreatic trypsin inhibitor, and chicken egg white lysozyme), as catalyzed by TG2, a biologically important post-translational enzyme. The substrates represent a total of 20 potential surface sites for acylation by the fluorescent Gln probe, yet only two of the lysine side chains reacted with TG2. While the K1 site of ribonuclease and the K15 site of the trypsin inhibitor could be readily acylated by the enzyme, none of the lysines in lysozyme were modified. The findings lead us to suggest that the selection of lysine residues by TG2 is not encoded in the primary amino acid sequence surrounding the target side chain but depends primarily on its being positioned in an accessible segment of the protein structure.

Transglutaminases (TGs) are post-translational, protein-remodeling enzymes. TG2,[1] the historical prototype in the TG family of gene products chosen for this study, is found in many cells and also in the extracellular matrix. It is known to play important roles in a variety of biological processes (*1*−*3*).

TGs catalyze nucleophilic displacements, such as hydrolysis and transamidation, at γ-amides of select glutamyl residues in proteins and peptides (P). Earlier, it was assumed that the reaction occurred only at the primary amide of a Q site (P-$CH_2CH_2CO|NH_2$), but it is now known that secondary amides at E sites in branched γ:ε isopeptides (P-$CH_2CH_2CO|NHR$) may also be targeted (*4*).

TGs are homologous to the papain family of proteases with which they share considerable active site (CySH, His, Asp/Asn) sequence and structural similarities and also a similar acylation/deacylation kinetic pathway. There are, however, significant differences that uniquely characterize these protein-remodeling enzymes. The acyl intermediate stabilizing function of Gln19 of papain is carried out by a Trp residue conserved in all TGs, corresponding to Trp241 in TG2 (*5, 6*). Also, unlike with papain, TGs display remarkable specificities for the substituents attached to the nucleophilic, electron pair-donating second (or donor) substrate, H_2NR', which competes against the water molecules for the deacylation of the acylenzyme intermediate (P-$CH_2CH_2CO|$S-E). Hence, in TG-catalyzed reactions, pre-equilibration would occur between the acylenzyme and the second substrate prior to deacylation (*5*−*7*). The appearance of extra subsites (exosites) for binding amines into a Michaelis complex must have been a significant step in evolution, a major departure from the simple papain model of deacylation by direct bimolecular substitution (*8, 9*).

Among small amines, monosubstituted diaminopentanes, exemplified by dansylcadaverine, are most effective as second substrates (*10, 11*). They mimic the ε-NH_2Lys side chains in proteins (K sites) with which TGs react to cross-link macromolecules by γ:ε side chain bridges. However, of the multitude of Lys-containing proteins in the biological milieu, only very few with select subsets of Lys residues in their structures seem to be targeted by the resident TG when the latent enzyme becomes activated by Ca^{2+} ions (*12*).

In an attempt to further investigate the exceptional amine specificity of TG2 for the selection of K sites in proteins, we studied the TG2-dependent acylation of three chemically and structurally well-characterized substrates: RNaseA or bovine ribonuclease, BPTI or bovine pancreatic trypsin inhibitor, and chicken egg white lysozyme. There are 10 Lys residues in RNaseA (*13*), four in BPTI (*14*), and six in lysozyme (*15*), yet we found that the first two of these proteins each contained only a single K site, Lys1 in RNaseA and Lys15 in BPTI, which could be enzymatically modified by incorporation of the fluorescent, dansyl-ε-aca-GlnGlnIl-

[†] This work was supported by NIH Grant HL-02212.

* To whom correspondence should be addressed. E-mail: l-lorand@northwestern.edu. Telephone: (312) 503-0591. Fax: (312) 503-7912.

[‡] Department of Cell and Molecular Biology, Feinberg School of Medicine, Northwestern University.

[§] Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University.

[||] Stanford University School of Medicine.

[1] Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; dansyl, *N*-(5-aminopentyl)-5-dimethylaminonaphthalene-1-sulfonamide; dansyl-ε-aca-QQIV, dansyl-ε-aminocaproyl-Gln-Gln-Ile-Val; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; PMSF, phenylmethanesulfonyl fluoride; RNaseA, bovine pancreatic ribonuclease A; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; TPCK, *N*-*p*-tosyl-L-phenylalanine chloromethyl ketone; TG2, transglutaminase 2 (EC 2.3.2.13).

10.1021/bi802323z CCC: $40.75
Published on Web 02/17/2009

eVal, acylating probe (*12, 16, 17*); interestingly, lysozyme did not react with TG2.

EXPERIMENTAL PROCEDURES

Materials. Bovine pancreatic ribonuclease A (RNase A, type XII-A), chicken egg white lysozyme, Sephadex G-50, dithiothreitol, phenylmethanesulfonyl fluoride (PMSF), iodoacetic acid, and iodoacetamide were from Sigma (St. Louis, MO). Bovine pancreatic trypsin inhibitor (BPTI; also known as aprotinin) was purchased from Boehringer Mannheim and was further purified by being passed through Sephadex G-50 with 20 mM imidazole-HCl and 10% glycerol (pH 7.5). The K15A mutant of BPTI (*18*) was a gift from M. Laskowski, Jr., of Purdue University (West Lafayette, IN). Human erythrocyte transglutaminase (TG2) was prepared by affinity chromatography on an Affigel 15 column, using the gelatin-binding 42K fragment of human fibronectin as a ligand (*19, 20*). The fluorescent dansyl peptide, dansyl-ε-aca-QQIV (*16, 17*), employed as the acylating probe, was synthesized by K. N. Parameswaran, and an anti-dansyl affinity column (*12*) was also prepared in our laboratory. Ultrapure urea was purchased from Bio-Rad (Hercules, CA) and TPCK-treated trypsin from Worthington Biochemicals (Freehold, NJ).

Methods. The TG2-dependent acylation of RNaseA and BPTI, purification of the modified proteins, digestion by trypsin, isolation of labeled peptides, and sequencing were carried out by previously published procedures (*12, 21, 22*). Specifically, acylation of RNaseA was performed at 37 °C for 2 h in 2 mL of 20 mM imidazole-HCl buffer (pH 7.5) and 10% glycerol containing 73 μM RNase A (1 mg/mL), 1 mM dansyl-ε-aca-QQIV, 1 mM $CaCl_2$, 25 μM DTT, and 0.365 μM TG2 (30 μg/mL) and stopped by the addition of 40 μL of 0.1 M EDTA. Urea was added (to 8 M), and excess dansyl probe was removed from the labeled protein by gel filtration on Sephadex G-50 [2.5 cm × 48 cm, equilibrated with 0.2 M Tris-HCl (pH 8.2) and 4 M urea; 3.5 mL fractions were collected]. Reduction and carboxymethylation were carried out as recommended by Clark and Gurd (*23*). Labeled protein fractions from the Sephadex column (fractions 28−32) were pooled, mixed with 150 μL of 1 M dithiothreitol (final concentration of 10 mM, 37 °C, 30 min) and, for a further 30 min, with 300 μL of 1.5 M iodoacetic acid (final concentration of 30 mM) and dialyzed against 2 × 2 L of ammonium bicarbonate for 18 h at 4 °C. Digestion with TPCK-treated trypsin was carried out at 37 °C for 20 h in 20 mL of 50 mM ammonium bicarbonate at a trypsin:RNaseA ratio of 1:10 (w/w) and was stopped by addition of 2 μL of 200 mg/mL PMSF in DMSO. Peptides, acylated by the dansyl probe, were isolated by their affinity for a column containing an anti-dansyl antibody (2.5 cm × 4.5 cm; 3.6 mg of monoclonal anti-dansyl IgG per milliliter of Affigel 10 resin). The column was washed with 150 mL of 50 mM Tris-HCl and 0.1 M NaCl (pH 7.5), and bound peptides were eluted with 10% (v/v) acetic acid, concentrated by lyophilization to ∼1 mL, and further purified by reverse phase HPLC. This was carried out on an Ultrasphere octyl column (Altex, Beckman; Waters 2690 Separation Module, Waters 996 Photodiode Array Detector, and Varian 2070 Spectrofluorometer) with a gradient from 0 to 36% acetonitrile, containing 0.1% TFA. Fluorescent fractions (excitation at 338 nm and emission at 530 nm) were concentrated for reinjection and finally for amino acid sequence analysis.

Similarly, the enzymatic acylation of BPTI was performed at 37 °C for 2 h in 8 mL of 20 mM imidazole-HCl (pH 7.5), 100 mM NaCl, 10% glycerol, 0.09 mM DTT, 0.115 mM BPTI (0.75 mg/mL), 1 mM dansyl-ε-aca-QQIV, 0.625 μM TG2 (50 μg/mL), and 5 mM $CaCl_2$. The reaction was terminated by addition of EDTA (5 mM), and urea was added (to 4 M). Fluorescently labeled BPTI was isolated on Sephadex G-50 (2.5 cm × 48 cm), using 0.2 M Tris-HCl (pH 8.2), in 8 M urea as eluent; 3.5 mL fractions were collected. This was followed by dialysis against water, and concentrating the solution to ∼1 mL which was further purified by reverse phase HPLC (Beckman) on an Ultrasphere octyl column (Altex, Beckman) with a gradient from 10 to 80% acetonitrile in water containing 0.1% TFA over 70 min. Fluorescence from HPLC runs was monitored by a PGC Instrument (Frederick, MD) model 2 device using a 338 nm excitation filter and a 500 nm emission filter. Data were recorded with Hewlett-Packard 3390A integrators (Wilmington, DE). Fluorescent fractions were collected on a Foxy 200 instrument (ISCO, Lincoln, NE) set to detect peaks by slope based on absorbance at 240 nm, and relevant fractions were pooled and dried by lyophilization. The protein was dissolved in 3 mL of 4 M urea in 0.2 M Tris-HCl buffer (pH 8.2) and 60 μL of 1 M DTT (20 mM) and incubated at 37 °C for 30 min; 150 μL of 1 M iodoacetamide (50 mM) was then added and incubation continued for an additional 30 min. This was followed by dialysis against 4 L of 0.1 M ammonium bicarbonate for 20 h at 4 °C. The ca. 1.2 mg of reduced, carboxamidomethylated protein was digested with TPCK-treated trypsin at 37 °C in 4 mL of 0.1 M ammonium bicarbonate [at a trypsin:substrate ratio of 1:10 (w/w)]. Twenty hours later, the trypsin was inactivated with 10 μL of 200 mg/mL PMSF in DMSO.

Separation of the acylated, dansyl-ε-aca-QQIV decorated peptides was accomplished by passage through an anti-dansyl IgG-Affigel 10 column and reverse phase HPLC. An Ultrasphere octyl column was employed with a linear gradient from 16 to 35% acetonitrile and 0.1% TFA at a flow rate of 1 mL/min over a period of 60 min. Fluorescent fractions were collected, concentrated, and reapplied to the HPLC column for further purification with either 17% acetonitrile (for peptide I) or 23% acetonotrile (for peptide II) in 0.1% TFA. Fluorescent fractions were collected, concentrated, and sequenced.

Protein concentrations were measured with a BCA assay (Pierce, Rockford, IL) according to the manufacturer's instructions using bovine serum albumin as a standard. The extent of dansyl-ε-aca-QQIV incorporation was estimated by using a molar extinction coefficient of 4670 at 327 nm for the dansyl moiety (*24*).

Electrophoresis on polyacrylamide gels (16% acrylamide in the separating gel at an acrylamide:bisacrylamide ratio of 20:1) in the presence of SDS (0.2%) was carried out with the discontinuous buffer system of Laemmli (*25*). Prior to electrophoresis, the samples were treated with 2% SDS and 40 mM DTT at 37 °C for 30 min. Fluorescent bands on the gel were visualized by illumination with UV light (UVL-56, Ultraviolet Oroducts Inc., SanGabriel, CA). Proteins were stained with Coomassie Brilliant Blue R (Sigma) in 10% acetic acid and 25% isopropyl alcohol.

Amino acid sequence analyses of RNaseA fragments by standard automated Edman chemistry on Applied Biosystems (Foster City, CA) instruments were carried out at Northwestern University Medical School (Chicago, IL); those of BPTI were performed first at the Harvard Microchemistry facility and also confirmed here.

TG2-catalyzed incorporation of dansyl-ε-aca-QQIV into wild-type BPTI and its K15A mutant was carried out at 37 °C up to 60 min, in 20 mM imidazole-HCl buffer (pH 7.5), 100 mM NaCl, 10% glycerol, 0.09 mM DTT, 0.5 mg/mL BPTI (76.9 μM), 1 mM fluorescent probe dansyl-ε-aca-QQIV, 10 μg/mL TG2 (0.125 μM; enzyme:substrate ratio of 1:615), and 1 mM $CaCl_2$. At appropriate intervals of time, TG2 activity was terminated by addition of EDTA and SDS−PAGE buffer.

RESULTS

There is a long history of investigating TG-dependent modifications by incorporation of fluorescent dansylated compounds into proteins. Procedures were developed in this laboratory to employ amine substrates, exemplified by dansylcadaverine, for identifying TG-reactive Q sites (*21, 22, 26*−*28*) and for using dansylated Gln peptides as acylating probes for marking the TG-reactive K sites of proteins (*12, 21, 22, 29, 30*). The dansyl tag makes for easy recognition of products not only because of fluorescence but also because readily available anti-dansyl antibodies allow for the direct affinity purification and identification of labeled sequences. We chose three well-characterized proteins [RNaseA (*31*), BPTI (*32*), and lysozyme (*33*)] as potential substrates for acylation (Figure 1A−C) by TG2 to identify the sites which could be modified by incorporation of the dansyl-ε-aca-QQIV probe.

Reactions with the first two substrates were conducted to near completion, and the small amounts of unmodified RNaseA and BPTI were separated from the labeled proteins by gel filtration which also removed the excess acylating probe (Figure 2). Following digestion by trypsin, dansyl-tagged fragments were purified from other peptides by affinity chromatography, using a monoclonal anti-dansyl IgG as a ligand. The unlabeled material passed through the column unhindered, while the dansyl-bearing fragments were retained but could be eluted with 10% acetic acid (Figure 3).

The fluorescent peptides were further processed by reverse phase HPLC and sequenced by the Edman procedure. Lys residues acylated with the dansyl-ε-aca-QQIV probe were not recognized as known Edman products; hence, they are marked with an X in Figure 4. In sequencing the large "peptide I" fragment from modified BPTI, we identified Cys residues in their carboxymethylated form (marked C between brackets in Figure 4) but did not quantify them. The prominent peak on the HPLC profile of modified RNaseA (representing ca. 88% of dansyl fluorescence) yielded a sequence of XETAAAK, corresponding to residues 1−7 of the protein and pointing to K1 as the TG2-reactive Lys. Modification of the α-amino group of this N-terminal Lys could be excluded on the grounds that the Edman sequencing reaction progressed down the peptide chain in a normal fashion.

The fragmentation pattern of modified BPTI revealed two fluorescently labeled products. The minor peak, termed

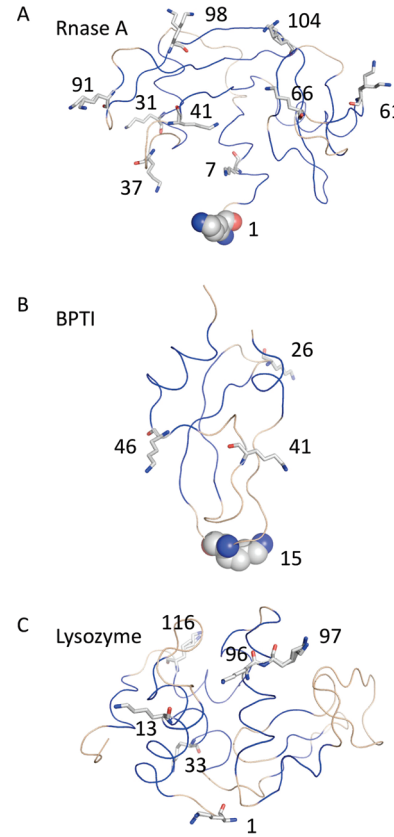

FIGURE 1: Positions of Lys residues in RNaseA, BPTI, and chicken egg lysozyme. (A) RNase A lysines are shown using Protein Data Bank (PDB) entry 7RSA (*31*). (B) BPTI lysines are shown using PDB entry 1BPI (*32*). (C) Hen egg white lysozyme lysines are shown using PDB entry 2VB1 (*33*). Reactive lysine residues are shown in CPK sphere representations, and nonreactive lysine residues are shown in stick representations. Alternate conformations are shown when present in the PDB entry. The backbone ribbon traces are colored blue for α-helical and β-sheet secondary structural elements in each protein. Though there are 10 Lys surface side chains in RNaseA (K1, K7, K31, K37, K41, K61, K66, K91, K98, and K104), six in lysozyme (K1, K13, K33, K96, K97, and K116), and four in BPTI (K15, K26, K41, and K46), in RNaseA and BPTI only a single Lys residue (K1 and K15, respectively) could be modified in the TG2-catalyzed reaction with the dansyl-ε-aca-QQIV probe. None of the Lys residues of lysozymes reacted with TG2.

"peptide I" in Figure 4, proved to be a tripeptide with a sequence of XAR, whereas the major peak "peptide II" material (ca. 75% of dansyl fluorescence) yielded the sequence RPDFCLEPPYTGPCXAR which corresponds to the N-terminal segment of residues 1−17 of BPTI, identifying K15 as the TG2-acylated Lys. We can offer no explanation for the unconventional cleavage by trypsin, apparently between a carboxamidomethylated Cys and dansyl-ε-aca-QQIV-derivatized K15 residue, to produce the minor fragment represented by peptide I on the HPLC profile, i.e., residues 15−17 of modified BPTI.

It is most interesting that, among the 10 Lys side chains on the surface of RNaseA, i.e., K1, K7, K31, K37, K41, K61, K66, K91, K98, and K104 (Figure 1A), only residue K1 could be acylated by TG2 and that, of the four side chain ε-amines on the surface of BPTI, i.e., K15, K26, K41, and K46 (Figure 1B), only residue K15 could be modified (Figure 4; a preliminary report of this finding appears in ref *34*). To further investigate whether K15 is merely a favored substrate of TG2 which outcompetes the other lysines in BPTI (K26

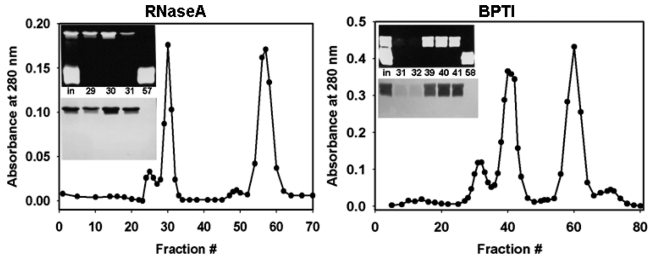

FIGURE 2: Purification of acylated RNaseA and BPTI. Following the TG2-catalyzed reaction with dansyl-ε-aca-QQIV (see Experimental Procedures), unreacted materials were separated from the modified proteins when the samples were passed through a Sephadex G-50 column. The peak eluting around fraction 25 in the left panel represents unreacted RNaseA, and that around fraction 32 in the right panel unmodified BPTI, whereas the peaks at fractions 58 (left) and 60 (right) are excess dansyl peptide. Acylated RNaseA was eluted around fraction 30 (left) and acylated BPTI around fraction 40 (right). Fractions which exhibited dansyl fluorescence and which were positively stained by Coomassie blue; i.e., fractions 28−32 (left) and 38−43 (right) were pooled for further processing. Insets present SDS−PAGE profiles photographed under UV light (top) and after staining with Coomassie blue (bottom). Reaction mixtures applied to the column for exclusion chromatography are marked "in". Approximately 20 μL of fractions was applied for SDS−PAGE analysis.

and K46) with similar solvent accessibility and identical basicity/nucleophilicity (*35, 36*) or whether TG2 exhibits true discrimination toward K15, we compared the reactivity of the K15A protein mutant with that of the wild type. As seen in Figure 5, the absence of a Lys residue at position 15 which was identified earlier by the direct labeling approach as the sole target for TG2 (Figure 4) eliminated the ability of the protein to react with the enzyme altogether.

Chicken egg white lysozyme, which has six Lys residues in its structure [K1, K13, K33, K96, K97, and K116 (Figure 1C)], could not be acylated by TG2 under the conditions employed for RNaseA.

DISCUSSION

Unlike with the papain family of enzymes, in TG2 catalysis, in general [with the exception of the breaking of the γ−ε isopeptide bond (unpublished results)], transamidation rather than hydrolysis seems to be the preferred kinetic path. This is why, historically, the reactions incorporating small amines into proteins and cross-linking protein units by N^{ε}-(γ-glutamyl)lysine side chain bridges have been regarded as the main functions of the enzyme (E) under physiological conditions. The ratio for partitioning the thiolester acylenzyme intermediate (P-COS-E) between an amine substrate and water by aminolysis or hydrolysis depends on the (i) effective concentration, (ii) nucleophilicity, and (iii) binding affinity of H_2NR', the second substrate:

As mentioned in the introductory section, in case of the Q site-containing first substrates, monosubstituted diaminopentanes, with an alkyl group equivalent in length to a Lys residue in proteins, are the most effective. While this alone represents a quite unique selection of Lys targets, the specificity of the enzyme is far more stringent. There are numerous data to indicate that, of the large number of Lys residues in proteins, only a handful might react with TG2 (*12, 21, 22, 29, 37, 38*), and the results presented here substantiate these observations. As seen in panels A and B of Figure 1, there are 10 lysines in RNaseA (K1, K7, K31, K37, K41, K61, K66, K91, K98, and K104) and four in BPTI (K15, K26, K41, and K46). In the reactions of the proteins with TG2, the potential amine site concentrations are constant. For BPTI, there is also convincing evidence that the ε-ammonium ions of K15, K26, and K46 have identical pK_a values (*35, 36*), so that these Lys residues also represent the same effective nucleophilic concentrations. Thus, the finding that there is only a single TG2-reactive K site in BPTI would seem to indicate that it is the binding affinity of this specific Lys residue in being able to form a Michaelis complex with the acylenzyme intermediate that is mainly responsible for substrate selection. This special character of the K15 side chain in BPTI was confirmed by showing that the K15A mutant did not react with TG2 (Figure 5).

The currently available crystal structures of TG2 in the inactive (*39−41*) and Ca^{2+}-activated form, blocked by an active center-directed inhibitor (*42*), illustrate that the catalytic triad of the inactive enzyme is initially buried in an interdomain interface in an environment that would not allow substrate side chains or extended polypeptide chains access to the active site. It could be that the binding of the first Q substrate to the enzyme induces a significant conformational change near the active site and that these structural changes may dictate the K site selectivity we observed with the RNaseA and BPTI substrates. It has also been proposed that, in Ca^{2+}-activated TG2—in the course of a major conformational change known to accompany the loss of bound nucleotide (*43−45*)—a "tunnel" is formed by the interaction of Trp241 and Trp332 into which the primary γ-amide of the Q residue of the first substrate would fit for proper positioning to the active center (*42*). Indeed, as expected, in this recently determined structure of Ca^{2+}-activated human TG2, blocked by the covalent inhibitor, a large rearrangement of the TG2 domains and exposure of the active site is observed. The inhibitor mimics the position of the glutamine substrate, which is observed to lie on one side of the active site tunnel. The lysine donor could gain access to the active site on the opposite side of this tunnel, which could sterically restrict the access and reactivity of lysine residues displayed in the context of a protein structure. A neighboring Thr360 residue, adjacent to the tunnel bounded by Trp241 and Trp332, when mutated to Trp or Ala, affects the catalytic efficiency of TG2. However, such a static conformation might not fully account for the behavior of functional TG2. The Trp241/Trp332 "jaws of the tunnel" may not be fixed in the active enzyme but might move to open and shut in step with each cycle of the steady state. Opening the "vise" would seem to be required for the release of the cross-linked products of transamidation, and the dynamic model would also be more suited to accommodate γ:ε branched isopeptide substrates (*4*) for which hydrolysis, rather than transamidation, appears to be favored (unpublished results), perhaps because the leaving group could interfere with the binding of incoming amines.

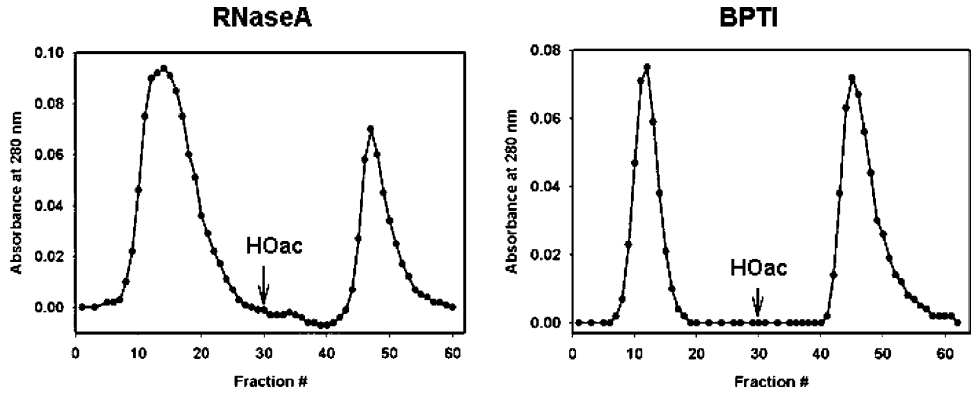

FIGURE 3: Separation of dansyl-ε-aca-QQIV-modified tryptic fragments by affinity for an anti-dansyl antibody column. The mixtures of tryptic peptides obtained from enzymatically acylated RNaseA (left panel) and BPTI (right panel) were applied to the affinity column (Experimental Procedures). Unmodified peptides passed through unhindered, whereas the retained, TG2-modified, and dansyl-tagged fragments could be eluted with 10% acetic acid [HOac (arrows)].

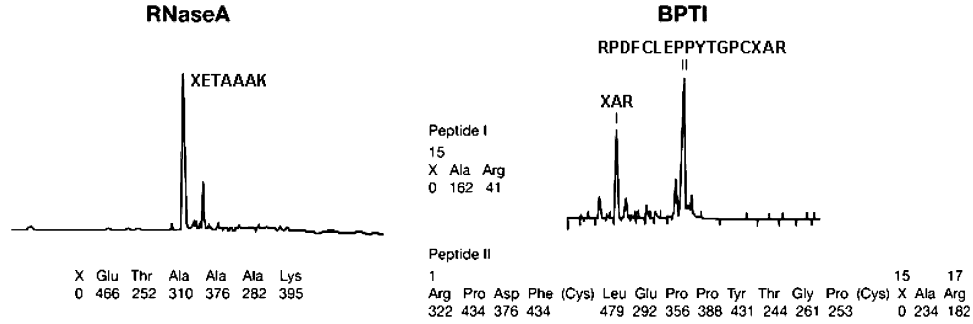

FIGURE 4: Isolation of dansyl-tagged peptides by HPLC and sequencing. The labeled peptides from RNaseA (left) and BPTI (right) were further fractionated by reverse phase HPLC on Ultrasphere octyl columns, eluted with a linear gradient from 0 to 60% acetonitrile in 0.1% TFA (60 min; left panel) and a gradient from 16 to 35% acetonitrile in 0.1% TFA (60 min; right) at flow rates of 1 mL/min (Experimental Procedures). Tracings of dansyl fluorescence (excitation at 338 nm and emission at 500 nm) exhibited a predominant peak from a RNaseA fragment (left; ca. 88% of total fluorescence), which was sequenced (quantities of Edman products are given in picamoles). An X indicates a position in which no known amino acid was recovered in the Edman cycle and is, thus, assumed to be a Lys residue derivatized by the dansyl-ε-aca-QQIV acyl probe, i.e., a K site for TG2. Analysis of the BPTI fragments produced two fluorescent peaks (right; marked as peptides I and II, with ca. 75% of total fluorescence), each of which was reprocessed on the same column, prior to sequencing, by elution with 17% acetonitrile (for peptide I) and 23% acetonitrile (for peptide II) in 0.1% TFA.

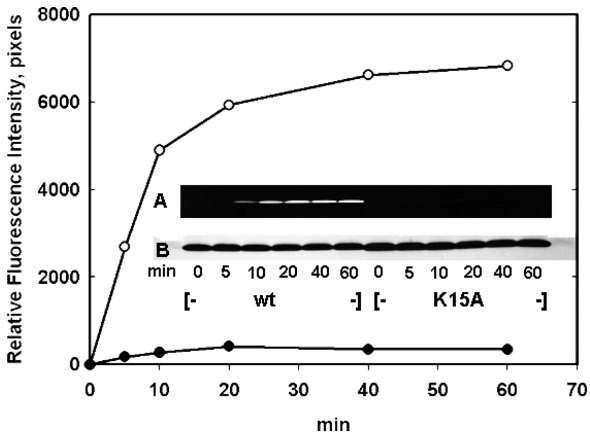

FIGURE 5: K15A mutant of BPTI that cannot be acylated by TG2. Samples were removed for analysis by SDS−PAGE at various times during the enzymatic reaction (abscissa, in minutes), run in parallel with the mutant (K15A) and wild-type (wt) BPTI (Experimental Procedures). Gels were photographed under UV illumination (inset, panel A), and the negatives were scanned to record fluorescence intensities due to the covalent decoration of the protein by the acylating dansyl-ε-aca-QQIV marker (ordinate, pixels). Coomassie blue-stained gel bands are shown in panel B.

Our findings clearly show that there are no amino acid sequence similarities between RNaseA and BPTI, flanking the enzyme-reactive K sites (Figure 1). While steric access to the active site tunnel may provide significant specificity for substrates, inspection of the structures of RNaseA, BPTI, and lysozyme suggests that there may be, at least, two features of this. One determinant is likely governed by access to the lysine side chain that is the target of modification, but this alone does not appear to be sufficient to explain the findings as many of the lysines in these model substrate proteins are surface-exposed and their relative reactivity does not appear to correlate with observed hydrogen bond or salt bridge interactions of the side chain. The second determinant of specificity likely involves overall access of the enzyme to the targeted lysine residue. TG2 may require specific interactions with main chain atoms of the targeted lysine and its neighboring residues, potentially involving the induction of an optimal conformation of the substrate polypeptide chain within the TG2 active site. The overall reactivity of each lysine may also be influenced by steric clashes of the protein substrate with TG2 that would prevent insertion of the lysine side chain into the active site. While most RNaseA and BPTI lysines appear to have similar side chain exposure and accessibility (Figure 1A,B), many of these are located in constrained secondary structures or turns that would restrict access to the peptide backbone. It is notable that the two reactive lysines, K1 of RNase A and K15 of BPTI, are not located in canonical secondary structure elements but in regions of extended polypeptide chain. However, other nonreactive lysines, such as K26 and K46 of BPTI, are also

not located in canonical secondary structures. Thus, the positioning of a lysine residue relative to canonical helical, sheet, or turn structures in substrate proteins, by itself, is not a simple predictor of reactivity. K26 is in a turn that would very likely restrict the flexibility of the main chain, and K46 is also in a region of the protein that appears more restricted than K15, although it is not a canonical turn. In addition, NMR data show that K46 forms a salt bridge with D50 in BPTI (*46*), which is not seen in the crystal structure, and that the K46−D50 interaction is with the main chain of K46. In the crystal structure, the backbone amine of K41 is involved in a hydrogen bond network mediated by two buried water molecules and involving residues N44 and Y10, which may hinder the accessibility of TG2 to K41. Taken together, these considerations provide a satisfactory explanation for why only the K15 residue of BPTI is selected to react with TG2.

In lysozyme, none of the exposed lysines are substrates for TG2 (Figure 1C), although most of the side chains are exposed to solvent. In this protein, K13, K33, K96, and K97 are positioned in helical secondary structures, K116 is at the end of a short helical segment, and K1 forms bidentate backbone hydrogen bonds with residue Thr40 that engage both its amino and carbonyl groups. For all of these residues, it is possible that the disposition of the lysine main chain atoms and their local interactions may block access to the TG2 active site. In contrast, K1 in RNaseA and K15 in BPTI exhibit greater exposure of both the side chain and local main chain regions. K1 of RNaseA is located at the extended N-terminal region preceding the first α-helix, and K15 in BPTI inserts into the trypsin active site, forming backbone interactions with the protease. Moreover, electrostatically, K15 of BPTI is in a highly positive environment of the protein (*47*), and the K1 residue of RNaseA is part of a positively charged band around the molecule (*48*). Thus, we suggest that the flexible structural features of these substrate sites, with proximity to charged or polar residues which would ensure an increased level of solvent exposure, explain their being selected for modification by TG2. This conclusion is supported by a statistical regression analysis of literature data compiled on TG2 [albeit obtained under a variety of different reaction conditions (*36*)] and by earlier reports on the reactions of the enzyme with fibrinogen (*29, 49*), vimentin (*21*), tau protein (*22*), αB crystallin (*12, 37*), and GAPDH (*50*). In all of these substrates, TG2 was found to preferentially target K sites located in flexible, or even unstructured segments of the polypeptide chain, often in the N- and C-terminal domains of the proteins. A more detailed understanding of K site specificity may require a structural analysis of an authentic, trapped intermediate of TG2 bound to one of its protein targets. While there is clearly an important role for local structure in determining lysine reactivity, overall specificity appears to be a multidimensional problem likely involving structure, dynamics, sequence, shape, and other features in an as yet to be understood hierarchy. Nevertheless, the findings in this article indicate that an overriding feature of K selection by TG2 is that the lysine residue be located in a highly flexible portion of the protein substrate.

ACKNOWLEDGMENT

It is a pleasure to thank Drs. Martin Egli, the late Michael Laskowski, Jr., Frederic M. Richards, Andrew D. Robertson, Celia A. Schiffer, Brian Schoichet, and Merridee A. Wouters for their valuable help and advice on various aspects of this project and, last but not least, our colleague, the late James H. Wilson, for the isolation of dansyl-tagged protein fragments.

REFERENCES

1. Fesus, L., and Piacentini, M. (2002) Transglutaminase 2: An enigmatic enzyme with diverse functions. *Trends Biochem. Sci. 27*, 534–539.
2. Griffin, M., Casadio, R., and Bergamini, C. M. (2002) Transglutaminases: Nature's biological glues. *Biochem. J. 368*, 377–396.
3. Lorand, L., and Graham, R. M. (2003) Transglutaminases: Crosslinking enzymes with pleiotropic functions. *Nat. Rev. Mol. Cell Biol. 4*, 140–156.
4. Parameswaran, K. N., Cheng, X. F., Chen, E. C., Velasco, P. T., Wilson, J. H., and Lorand, L. (1997) Hydrolysis of γ:ε isopeptides by cytosolic transglutaminases and by coagulation factor XIIIa. *J. Biol. Chem. 272*, 10311–10317.
5. Curtis, C. G., Stenberg, P., Brown, K. L., Baron, A., Chen, K., Gray, A., Simpson, I., and Lorand, L. (1974) Kinetics of transamidating enzymes. Production of thiol in the reactions of thiol esters with fibrinoligase. *Biochemistry 13*, 3257–3262.
6. Parameswaran, K. N., and Lorand, L. (1981) New thioester substrates for fibrinoligase (coagulation factor XIIIa) and for transglutaminase. Transfer of the fluorescently labeled acyl group to amines and alcohols. *Biochemistry 20*, 3703–3711.
7. Stenberg, P., Curtis, C. G., Wing, D., Tong, Y. S., Credo, R. B., Gray, A., and Lorand, L. (1975) Transamidase kinetics. Amide formation in the enzymic reactions of thiol esters with amines. *Biochem. J. 147*, 155–163.
8. Bender, M. L. (1971) *Mechanisms of Homogeneous Catalysis from Protons to Proteins*, Wiley-Interscience, New York.
9. Jencks, W. P. (1969) *Catalysis in Chemistry and Enzymology*, McGraw-Hill Book Co., New York.
10. Lorand, L., Rule, N. G., Ong, H. H., Furlanetto, R., Jacobsen, A., Downey, J., Oner, N., and Bruner-Lorand, J. (1968) Amine specificity in transpeptidation. Inhibition of fibrin cross-linking. *Biochemistry 7*, 1214–1223.
11. Lorand, L., Parameswaran, K. N., Stenberg, P., Tong, Y. S., Velasco, P. T., Jonsson, N. A., Mikiver, L., and Moses, P. (1979) Specificity of guinea pig liver transglutaminase for amine substrates. *Biochemistry 18*, 1756–1765.
12. Lorand, L., Velasco, P. T., Murthy, S. N., Wilson, J., and Parameswaran, K. N. (1992) Isolation of transglutaminase-reactive sequences from complex biological systems: A prominent lysine donor sequence in bovine lens. *Proc. Natl. Acad. Sci. U.S.A. 89*, 11161–11163.
13. Smyth, D. G., Stein, W. H., and Moore, S. (1963) The sequence of amino acid residues in bovine pancreatic ribonuclease: Revisions and confirmations. *J. Biol. Chem. 238*, 227–234.
14. Kassell, B., Radicevic, M., Ansfield, M. J., and Laskowski, M., Sr. (1965) The Basic Trypsin Inhibitor of Bovine Pancreas. IV. The Linear Sequence of the 58 Amino Acids. *Biochem. Biophys. Res. Commun. 18*, 255–258.
15. Canfield, R. E. (1963) The Amino Acid Sequence of Egg White Lysozyme. *J. Biol. Chem. 238*, 2698–2707.
16. Parameswaran, K. N., Velasco, P. T., Wilson, J., and Lorand, L. (1990) Labeling of ε-lysine crosslinking sites in proteins with peptide substrates of factor XIIIa and transglutaminase. *Proc. Natl. Acad. Sci. U.S.A. 87*, 8472–8475.
17. Lorand, L., Parameswaran, K. N., and Velasco, P. T. (1991) Sorting-out of acceptor-donor relationships in the transglutaminase-catalyzed cross-linking of crystallins by the enzyme-directed labeling of potential sites. *Proc. Natl. Acad. Sci. U.S.A. 88*, 82–83.
18. Castro, M. J., and Anderson, S. (1996) Alanine point-mutations in the reactive region of bovine pancreatic trypsin inhibitor: Effects on the kinetics and thermodynamics of binding to β-trypsin and α-chymotrypsin. *Biochemistry 35*, 11435–11446.
19. Radek, J. T., Jeong, J. M., Murthy, S. N., Ingham, K. C., and Lorand, L. (1993) Affinity of human erythrocyte transglutaminase for a 42-kDa gelatin-binding fragment of human plasma fibronectin. *Proc. Natl. Acad. Sci. U.S.A. 90*, 3152–3156.

20. Murthy, S. N., Velasco, P. T., and Lorand, L. (1998) Properties of purified lens transglutaminase and regulation of its transamidase/crosslinking activity by GTP. *Exp. Eye Res. 67*, 273–281.
21. Clement, S., Velasco, P. T., Murthy, S. N., Wilson, J. H., Lukas, T. J., Goldman, R. D., and Lorand, L. (1998) The intermediate filament protein, vimentin, in the lens is a target for cross-linking by transglutaminase. *J. Biol. Chem. 273*, 7604–7609.
22. Murthy, S. N., Wilson, J. H., Lukas, T. J., Kuret, J., and Lorand, L. (1998) Cross-linking sites of the human tau protein, probed by reactions with human transglutaminase. *J. Neurochem. 71*, 2607–2614.
23. Clark, J. F., and Gurd, F. R. (1967) Effect of alkylation of sperm whale myoglobin on response to extremes of temperature and pH. *J. Biol. Chem. 242*, 3257–3264.
24. Deranleau, D. A., and Neurath, H. (1966) The combination of chymotrypsin and chymotrypsinogen with fluorescent substrates and inhibitors for chymotrypsin. *Biochemistry 5*, 1413–1425.
25. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*, 680–685.
26. Lorand, L., and Ong, H. H. (1966) Labeling of amine-acceptor cross-linking sites of fibrin by transpeptidation. *Biochemistry 5*, 1747–1753.
27. Lorand, L., Chenoweth, D., and Gray, A. (1972) Titration of the acceptor cross-linking sites in fibrin. *Ann. N.Y. Acad. Sci. 202*, 155–171.
28. Murthy, S. N., Wilson, J., Zhang, Y., and Lorand, L. (1994) Residue Gln-30 of human erythrocyte anion transporter is a prime site for reaction with intrinsic transglutaminase. *J. Biol. Chem. 269*, 22907–22911.
29. Lorand, L. (1991) Identification of ε-Lysine Cross-Linking Sites in the α Chains of Fibrin and the Tissue Transglutaminase Carrying Role of Fibronectin in Plasma. 20th Linderstrom-Land Conference, Vingsted, Denmark.
30. Lorand, L. (2001) Factor XIII: Structure, activation, and interactions with fibrinogen and fibrin. *Ann. N.Y. Acad. Sci. 936*, 291–311.
31. Svensson, L. A., Sjolin, L., Gilliland, G. L., Finzel, B. C., and Wlodawer, A. (1986) Multiple conformations of amino acid residues in ribonuclease A. *Proteins 1*, 370–375.
32. Parkin, S., Rupp, B., and Hope, H. (1996) Structure of bovine pancreatic trypsin inhibitor at 125 K definition of carboxyl-terminal residues Gly57 and Ala58. *Acta Crystallogr. D52*, 18–29.
33. Wang, J., Dauter, M., Alkire, R., Joachimiak, A., and Dauter, Z. (2007) Triclinic lysozyme at 0.65 Å resolution. *Acta Crystallogr. D63*, 1254–1268.
34. Murthy, S. N. P., Wilson, J., Lukas, T., and Lorand, L. (1998) How do transglutaminases select the reactive lysine residues in their protein substrates? *FASEB J. 12* (8), A1432, Abstract 702.
35. Brown, L. R., De Marco, A., Wagner, G., and Wuthrich, K. (1976) A study of the lysyl residues in the basic pancreatic trypsin inhibitor using ^{1}H nuclear magnetic resonance at 360 MHz. *Eur. J. Biochem. 62*, 103–107.
36. March, K. L., Maskalick, D. G., England, R. D., Friend, S. H., and Gurd, F. R. (1982) Analysis of electrostatic interactions and their relationship to conformation and stability of bovine pancreatic trypsin inhibitor. *Biochemistry 21*, 5241–5251.
37. Groenen, P. J., Bloemendal, H., and de Jong, W. W. (1992) The carboxy-terminal lysine of αB-crystallin is an amine-donor substrate for tissue transglutaminase. *Eur. J. Biochem. 205*, 671–674.
38. Csosz, E., Bagossi, P., Nagy, Z., Dosztanyi, Z., Simon, I., and Fesus, L. (2008) Substrate Preference of Transglutaminase 2 Revealed by Logistic Regression Analysis and Intrinsic Disorder Examination. *J. Mol. Biol. 383*, 390–402.
39. Liu, S., Cerione, R. A., and Clardy, J. (2002) Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity. *Proc. Natl. Acad. Sci. U.S.A. 99*, 2743–2747.
40. Ahvazi, B., and Steinert, P. M. (2003) A model for the reaction mechanism of the transglutaminase 3 enzyme. *Exp. Mol. Med. 35*, 228–242.
41. Ahvazi, B., Boeshans, K. M., and Steinert, P. M. (2004) Crystal structure of transglutaminase 3 in complex with GMP: Structural basis for nucleotide specificity. *J. Biol. Chem. 279*, 26716–26725.
42. Pinkas, D. M., Strop, P., Brunger, A. T., and Khosla, C. (2007) Transglutaminase 2 undergoes a large conformational change upon activation. *PLoS Biol. 5*, e327.
43. Murthy, S. N., Lomasney, J. W., Mak, E. C., and Lorand, L. (1999) Interactions of G(h)/transglutaminase with phospholipase Cδ1 and with GTP. *Proc. Natl. Acad. Sci. U.S.A. 96*, 11815–11819.
44. Murthy, S. N., Iismaa, S., Begg, G., Freymann, D. M., Graham, R. M., and Lorand, L. (2002) Conserved tryptophan in the core domain of transglutaminase is essential for catalytic activity. *Proc. Natl. Acad. Sci. U.S.A. 99*, 2738–2742.
45. Begg, G. E., Carrington, L., Stokes, P. H., Matthews, J. M., Wouters, M. A., Husain, A., Lorand, L., Iismaa, S. E., and Graham, R. M. (2006) Mechanism of allosteric regulation of transglutaminase 2 by GTP. *Proc. Natl. Acad. Sci. U.S.A. 103*, 19683–19688.
46. Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., and Wuthrich, K. (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol. 196*, 611–639.
47. Schiffer, C. A., and van Gunsteren, W. F. (1996) Structural stability of disulfide mutants of basic pancreatic trypsin inhibitor: A molecular dynamics study. *Proteins 26*, 66–71.
48. Fedorov, A. A., Joseph-McCarthy, D., Fedorov, E., Sirakova, D., Graf, I., and Almo, S. C. (1996) Ionic interactions in crystalline bovine pancreatic ribonuclease A. *Biochemistry 35*, 15962–15979.
49. Murthy, S. N., Wilson, J. H., Lukas, T. J., Veklich, Y., Weisel, J. W., and Lorand, L. (2000) Transglutaminase-catalyzed crosslinking of the Aα and γ constituent chains in fibrinogen. *Proc. Natl. Acad. Sci. U.S.A. 97*, 44–48.
50. Orru, S., Ruoppolo, M., Francese, S., Vitagliano, L., Marino, G., and Esposito, C. (2002) Identification of tissue transglutaminase-reactive lysine residues in glyceraldehyde-3-phosphate dehydrogenase. *Protein Sci. 11*, 137–146.

BI802323Z