Photolabeling of Tissue Transglutaminase Reveals the Binding Mode of Potent Cinnamoyl Inhibitors[†]

Christophe Pardin, Isabelle Roy, Roberto A. Chica, Eric Bonneil, Pierre Thibault, William D. Lubell, Joelle N. Pelletier, and Jeffrey W. Keillor*

Département de chimie, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, PQ H3C 3J7, Canada

Received October 29, 2008; Revised Manuscript Received February 11, 2009

ABSTRACT: We have recently developed a new class of cinnamoyl derivatives as potent tissue transglutaminase (TG2) inhibitors. Herein, we report the synthesis of a diazirine derivative of these inhibitors and its application to the photolabeling of its binding site on guinea pig liver transglutaminase. Two novel homology models were generated for this commonly studied TG2, which differ in the conformational state they represent. Tryptic digest and mass spectrometric analysis of the photolabeling experiment showed that only residue Cys230 was labeled, and our homology models were used to visualize these results. This visualization suggested that Cys230 is somewhat more solvent-exposed in the "closed" conformation of TG2, compared to the "open" conformation. Docking experiments suggested binding modes consistent with the labeling pattern that would block access to the tunnel leading to the active site, consistent with the observed mode of inhibition. However, while these modeling simulations favored the closed conformation as the target of our cinnamoyl inhibitors, native PAGE experiments indicated the open conformation of the enzyme in fact predominates in the presence of our photolabeling derivative. These results are important for understanding the binding modes of TG2 inhibitors in general and will be critical for the structure-based design of future inhibitors.

Transglutaminases (TGases,[1] EC 2.3.2.13) make up a family of Ca^{2+}-dependent enzymes that catalyze the transfer of an acyl group from the γ-carboxamide group of a peptide-bound glutamine residue to the ε-amino group of a peptide-bound lysine residue, resulting in an isopeptide bond that may cross-link peptides or proteins (*1*−*3*). In mammals, TGases found in tissue, plasma, and epidermis have been extensively characterized. Tissue TGases are involved in various different biological processes such as endocytosis (*4, 5*), apoptosis (*6*), and cell growth regulation (*7*). However, unregulated, high TGase activities have been implicated in physiological disorders involved in disease states such as acne (*8*), the formation of cataracts (*9*), immune system diseases (*10*), psoriasis (*11*), Alzheimer's disease (*12, 13*), Huntington's disease (*14, 15*), Celiac disease (*16*), and cancer metastasis (*17*). The development of TGase inhibitors has thus been pursued to provide lead compounds for therapeutic development and probes for better understanding their roles in normal and abnormal physiologies.

A number of TGase inhibitors have been developed to regulate excess TGase activity. These include irreversible inactivators, such as dihydroisoxazole derivatives (*17*), gluten peptide analogues (*18*), and dipeptide-bound 1,2,4-thiadiazoles (*19*), α,β-unsaturated amides, and epoxides (*20, 21*). Furthermore, thieno[2,3-*d*]pyrimidin-4-one acylhydrazide derivatives represent a new class of reversible TGase inhibitors (*22*).

We have recently discovered *trans*-cinnamoyl derivatives that serve as potent reversible inhibitors of guinea pig liver TGase (*23, 24*). These inhibitors were shown to be competitive with acyl donor substrates and exhibited K_i values as low as 174 nM (*24*). Their selectivity for TG2 was demonstrated by their ineffectiveness against similar enzymes such as Factor XIIIa and caspase 3 (*23*).

Detailed analysis of the binding mode of TG2 inhibitors has been hindered by the lack of structural data. Few X-ray structures of TG2 have been published (*25*−*27*), and for many years, the only two structures available (*25, 26*) lacked any bound substrate or synthetic inhibitor. Our preliminary study (*23*) of the binding mode of the new family of cinnamoyl inhibitors was therefore based on computational analysis, using the structure of one homologous TG2 apoenzyme (*25*). Since the cinnamoyl inhibitors were determined to be competitive with respect to donor substrate, they were docked into the shallow groove at the mouth of the active site tunnel, previously hypothesized to be the donor substrate binding site (*28*). However, steady-state kinetic data alone cannot distinguish between inhibition through interaction at this substrate binding site and inhibition through binding at the tightly coupled nucleotide binding site (*29*). Additional structural information was therefore required to unambiguously identify where our competitive inhibitors were bound. Furthermore, a third TG2 structure was recently

[†] We thank the Natural Sciences and Engineering Research Council (NSERC) and the Canadian Institutes of Health Research (CIHR) for financial support. In addition, C.P. thanks the Université de Montréal for a graduate bursary, and I.R. and R.A.C. thank NSERC for graduate scholarships.

* To whom correspondence should be addressed. Telephone: (514) 343-6219. Fax: (514) 343-7586. E-mail: jw.keillor@umontreal.ca.

[1] Abbreviations: DMF, dimethylformamide; EDTA, ethylenediaminetetraacetic acid; MOPS, morpholinopropanesulfonic acid; TGase, transglutaminase; TG2, tissue transglutaminase.

10.1021/bi802021c CCC: $40.75
Published on Web 03/09/2009

published (*27*) that stands out notably from the first two structures in that it features a ligand bound in the active site and a remarkably different, "open" conformation. This raised the additional question of which form of the enzyme may bind our inhibitors. In an effort to resolve these issues, we performed labeling experiments to provide structural evidence regarding TG2 ligand binding and to determine the mode of inhibition of the cinnamoyl inhibitors.

MATERIALS AND METHODS

Synthesis. (i) Materials. Dicyclohexylcarbodiimide (DCC) was obtained from Sigma-Aldrich. Triethylamine (Et_3N) was purchased from ACP. Buffer salts were from Sigma Chemical Co. Dichloromethane was obtained from EMD. Substrate *N*-Cbz-Glu(γ-*p*-nitrophenyl ester)Gly (*30*), (*E*)-3-(4-aminophenyl)-1-(pyridin-3-yl)prop-2-en-1-one (*24*), and 4,4-azopentanoic acid (*31*) were synthesized in our laboratory according to published procedures.

^1H and ^{13}C NMR spectra were recorded on a Brüker 300 MHz spectrometer in the solvent indicated in the text. Chemical shifts are reported in parts per million with internal reference to TMS. High-resolution mass spectra (HRMS) were recorded on a LC-MSD-TOF instrument from Agilent technologies in positive electrospray mode. Either protonated molecular ions $(M + H)^+$ or sodium adducts $(M + Na)^+$ were used for empirical formula confirmation.

*(ii) Methods. 3-(3-Methyl-3H-diazirin-3-yl)-N-{4-[(E)-3-oxo-3-(pyridin-3-yl)prop-1-enyl]phenyl}propanamide (***3***).* 3,3-Azobutyric acid **1** (*31*) (0.5 mmol) was dissolved in 2 mL of CH_2Cl_2, treated with DCC (0.25 mmol), and stirred for 30 min at 0 °C. Then the mixture was filtered to remove the dicyclohexylurea (DCU) and washed with 0.5 mL of CH_2Cl_2. Azachalcone **2** (*23*) (0.1 mmol) and NEt_3 (0.25 mmol) were added to the filtrate and stirred overnight at room temperature. Then CH_2Cl_2 was removed under reduced pressure, and the reaction mixture was diluted with 30 mL of EtOAc and washed with 3 × 5 mL of 0.1 N NaOH. The organic layer was dried with $MgSO_4$, filtered, and evaporated. The crude solid was purified by flash chromatography (100% EtOAc) followed by precipitation with an EtOAc/Et_2O mixture to produce amide **3** as a pale yellow solid in 62% isolated yield: mp 163−165 °C; ^1H NMR ($CDCl_3$) δ 9.18 (s, 1H, 2-pyridyl), 8.76 (d, 1H, J = 3.6 Hz, 4-pyridyl), 8.27 (m, 2H, NH and 6-pyridyl), 7.75 (d, 1H, J = 15.6 Hz, C*H*CHCO), 7.58 (m, 4H, Ar), 7.45 (dd, 1H, J = 7.9 Hz, 4.9 Hz, 5-pyridyl), 7.35 (d, 1H, J = 15.6 Hz, CHC*H*CO), 2.18 (t, 2H, J = 7.6 Hz, CH_2CH_2CO), 1.82 (t, 2H, J = 7.6 Hz, CH_2CH_2CO), 1.01 (s, 3H, Me); ^{13}C NMR (d_6-DMSO) δ 187.75 (CO-pyridyl), 169.72 (CONH), 152.71 (C6-pyridyl), 149.13 (C2-pyridyl), 144.07 (*C*HCHCO), 141.21 (C4 Ar), 135.40 (C4-pyridyl), 132.50 (C3-pyridyl), 129.65 (C1 Ar), 128.70 (C2 and C6 Ar), 123.44 (C5-pyridyl), 119.49 (CH*C*HCO), 118.42 (C3 and C5 Ar), 30.39 (C diaziryl), 28.83 (CH_2CH_2CO), 25.30 (Me), 18.93 (CH_2CH_2CO); ^{13}C NMR (d_6-DMSO) δ 187.75, 169.72, 152.71, 149.13, 144.07, 141.21, 135.40, 132.50, 129.65, 128.70, 123.44, 119.49, 118.42, 30.39, 28.83, 25.30, 18.93; HRMS (FAB) calcd for $C_{19}H_{19}N_4O_2$ ($[M + H]^+$) 335.1503, found 335.1510.

Kinetic Methods. The change in absorbance at 405 nm was followed on a Cary 100 BIO UV−visible spectrophotometer at 25 °C, in a buffer composed of 50 mM $CaCl_2$, 50 μM EDTA, and 0.1 M MOPS (pH 7.0). All aqueous solutions were prepared using deionized water purified with a Millipore BioCell system. All kinetic assays were conducted using 900 μL of buffer, 50 μL of 0.15 mg/mL TGase, and 25 μL of a 2.2 mM stock solution of substrate *N*-Cbz-Glu(γ-*p*-nitrophenyl ester)Gly, in the presence of 0−25 μL of an anhydrous DMF stock solution of the inhibitor. DMF was added to complete the reaction volume (1 mL). Inhibition by synthetic analogue **3** was studied for the reaction of 54.4 μM (∼2 K_M) chromogenic TGase substrate *N*-Cbz-Glu(γ-*p*-nitrophenyl ester)Gly with 1.0 milliunit of recombinant guinea pig liver TGase as previously reported (*30*). Initial rates were measured in the presence of 0−15 μM **3**, plotted against inhibitor concentration, and fit by linear regression. The IC_{50} value of 28 ± 0.7 μM was calculated from extrapolation to 50% of the rate observed in the absence of inhibitor.

Photolabeling. Photolabeling was carried out in an UV oven in buffer composed of 180 mM Tris-acetate (pH 7.0) and 4.5 mM $CaCl_2$. All aqueous solutions were prepared using deionized water purified from the Millipore BioCell system. The photolabeling reaction mixture was composed of 375 μL of buffer, 100 μL of 1.25 mg/mL TGase, and 25 μL of an anhydrous DMF stock solution of inhibitor **3** (1 mg/mL). The sample was incubated for 30 min at 37 °C and then irradiated for 4 min in a Luzchem model LZC-5 UV oven with six 350 nm, 8 W bulbs.

Mass Spectrometric Analyses. For ESI-MS analysis, samples were solubilized in 50 mM ammonium bicarbonate and 1 M urea (EM Science, Mississauga, ON). Chromatographic separation was effected using a 40 to 90% gradient of 0.1% formic acid in acetonitrile on a Poroshell CP column with an 1100 LC system coupled to an ESI-MSD-TOF mass spectrometer (Agilent Technologies) operated in positive electrospray mode.

For MS/MS analysis, samples were prepared as described above. Reduction of disulfide bonds was achieved by incubating the protein samples in 25 mM tris(2-carboxyethyl)phosphine (TCEP) (Pierce Biotechnology, Rockford, IL) at 37 °C for 1 h. Samples were digested using modified porcine trypsin (Promega, Madison, WI) for 8 h at 37 °C and acidified with formic acid (EM Science); 200 ng samples of tryptic digests were injected onto a homemade 300 μm × 5 mm C18 precolumn (Jupiter 3 μm, 300 Å, Phenomenex, Torrance, CA). NanoLC separations were carried out on a Waters nano-ACQUITY LC system using a custom 150 μm × 10 cm C18 column (Jupiter 5 μm, 300 Å, Phenomenex). Peptide elution was performed using a gradient from 10 to 60% in 56 min. The nano-LC instrument was interfaced to a Q-TOF Premier apparatus (Waters, Millford, MA), and the mass spectrometer was calibrated daily with Glu-fibrinopeptide B (Sigma-Aldrich, Oakville, ON). Tandem mass spectra were acquired in the data-dependent acquisition mode. Three iterative exclusion runs were performed for each sample. Precursors sequenced in the preceding run were excluded in subsequent nanoLC−MS/MS analyses to reduce the number of redundant MS/MS spectra and obtain more comprehensive sequence coverage. Correlation of fragment ions with the predicted transglutaminase sequence was performed using Mascot version 2.1 (Matrix Science, London, U.K.). Searches were performed using variable modifications (deamidation N/Q, oxidation M, cinnamoyl moiety, and semitryptic

```
  1 MAEDLILERC DLQLEVNGRD HRTADLCRER
 31 LVLRRGQPFW LTLHFEGRGY EAGVDTLTFN
 61 AVTGPDPSEE AGTMARFSLS SAVEGGTWSA
 91 SAVDQQDSTV SLLLSTPADA PIGLYRLSLE
121 ASTGYQGSSF VLGHFILLYN PWCPADAVYM
151 DSDQERQEYV LTQQGFIYQG SAKFINGIPW
181 NFGQFEDGIL DICLMLLDTN PKFLKNAGQD
211 CSRRSRPVYV GRVVSAMVNC NDDQGVLQGR
241 WDNNYSDGVS PMSWIGSVDI LRRWKDYGCQ
271 RVKYGQCWVF AAVACTVLRC LGIPTRVVTN
301 FNSAHDQNSN LLIEYFRNES GEIEGNKSEM
331 IWNFHCWVES WMTRPDLEPG YEGWQALDPT
361 PQEKSEGTYC CGPVPVRAIK EGHLNVKYDA
391 PFVFAEVNAD VVNWIRQKDG SLRKSINHLV
421 VGLKISTKSV GRDEREDITH TYKYPEGSEE
451 EREAFVRANH LNKLATKEEA QEETGVAMRI
481 RVGQNMTMGS DFDIFAYITN GTAESHECQL
511 LLCARIVSYN GVLGPVCSTN DLLNLTLDPF
541 SENSIPLHIL YEKYGDYLTE SNLIKVRGLL
571 IEPAANSYVL AERDIYLENP EIKIRVLGEP
601 KQNRKLIAEV SLKNPLPVPL LGCIFTVEGA
631 GLTKDQKSVE VPDPVEAGEQ AKVRVDLLPT
661 EVGLHKLVVN FECDKLKAVK GYRNVIIGPA
```

FIGURE 1: LC−MS/MS peptide sequencing and identification of modified residues of guinea pig liver TGase, after photoinactivation with **3**. Bold letters indicate peptides observed in the analysis; highlighting color indicates the corresponding residue shown in Figure 3.

digestion). Parent ion and fragment ion mass tolerances were both set at ±0.6 Da.

Nondenaturing Gel Assays. Samples of 6 μg of TGase were incubated for 15 min (5 min in the presence of calcium) at 25 °C in buffer composed of 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, and 0.5 mM EDTA. To this buffer were added the following ligands for incubation: 5 mM $CaCl_2$, 1 mM $MgCl_2$ and 25 μM GTP, or 28.5 μM **3** (5-fold K_i) and 5% DMF. Loading buffer [0.5 M Tris, 20% glycerol, and bromophenol blue (pH 6.8 at 4 °C)] was added to ice-chilled samples in a 3:1 sample:buffer ratio. Nondenaturing PAGE was performed on 10% resolving gels (100 V at 4 °C) by using Laemmli buffers (*32*) without SDS, with the pH adjusted at 4 °C. For analysis of incubated TGase, the appropriate ligands were also added to the resolving gels and running buffers in the following concentrations: 0.5 mM $CaCl_2$, 0.5 mM $MgCl_2$ and 25 μM GTP, or 28.5 μM **3** and 5% DMF. Gels were stained with 0.25% Coomassie brilliant blue R-250 in a 40% methanol/10% acetic acid mixture and destained in a 25% methanol/10% acetic acid mixture. Six to eight samples of each incubation mixture were analyzed, of which representative examples are shown in Figure 5.

Homology Modeling. Two homology models of guinea pig liver TGase were prepared: the first of the well-known closed-form structures and the second of the recent open-form structure. The closed form used the Protein Data Bank (PDB) coordinates of the crystal structures of human TG2 (1KV3) (*25*), red sea bream TG2 (1G0D) (*26*), human epidermal TGase (1L9M) (*33*), and human factor XIIIa (1EVU) (*34*) as templates, while the open form used the coordinates of the only crystal structure of the open form of human TG2 (2Q3Z) (*27*) as the template. Crystallized water molecules, heterologous molecules, and extra monomers were removed. Primary sequence alignment was derived using either the Tcoffee server (*35*) (closed form) or ClustalW (*36*) (open and closed forms), using default parameters. Models were generated using Modeler 8v1 (closed form) or Modeler 9v2 (open form) (*37*) with default parameters for loop conformations, using the crystal structure coordinates.

For the closed form, model quality was assessed using Molprobity (*38*), Whatcheck (*39*), and ProsaII (*40*). Models were refined using an all-atom tether restraint of 100 kcal/Å^2 through 1000 steps of steepest descents followed by a conjugate gradient minimization until convergence of 0.001 kcal mol^{-1} Å^{-1}, with a distance-dependent dielectric constant of 4 to mimic the interior of proteins (Discover, InsightII). Following refinement, model T2 was retained (see the Supporting Information) and was subjected to an unconstrained molecular dynamics simulation at 300 K consisting of a 1 ps equilibration of the molecular system, followed by simulated exploration of conformational space for 300 ps. A nonbond cutoff of 15 Å was applied. The radius of gyration of the structure during the simulation was measured using the Decipher module of InsightII. For the open form, models were refined without restraint through 1000 steps of steepest descents followed by a conjugate gradient minimization until convergence of 0.001 kcal mol^{-1} Å^{-1} under a CVFF force field, with a distance-dependent dielectric constant of 80 to mimic the hydrophilic environment of proteins in water (Discover and InsightII). Model quality was assessed using Molprobity (*38*) and ProsaII (*40*).

Docking Simulations. The geometry of ligand **3** was minimized semiempirically (AM1) using Hyperchem 6.0. Docking of the optimized structure was performed against the closed-form homology model of guinea pig liver TGase (vide supra) using AutoDock version 3.0.5, with grid dimensions of 90 × 70 × 80 npts and a grid spacing of 0.375 Å. The grid was centered on the homology model to include the photolabeled residue identified herein through mass spectrometry as well as the residues of the donor substrate binding site identified by molecular modeling (*28*). The genetic algorithm implemented in AutoDock 3.0.5 (*41*) was applied with a starting population size of 50. Fifty runs were performed with a maximum number of energy evaluations of 250000 and a maximum number of generations of 27000. Auto-Tors was used to define the available torsions for the ligands; seven possible torsions for the ligand were thus obtained. Of the 50 conformations obtained, 49 had very similar binding energies corresponding to −6.5 ± 1 kcal/mol. Docking was repeated using a larger grid (126 × 80 × 126 npts, grid spacing of 0.375 Å) for 100 runs. Similar results were obtained, confirming the quality of the docking protocol. The majority of conformations obtained were consistent with either the observed photolabeling or the observed competitive TG2 inhibition. Conformations consistent with both these empirical observations were retained as potential binding models, of which two are shown in Figure 4. For the open-form homology model, docking simulations were performed as described above for the larger grid.

RESULTS AND DISCUSSION

Photolabile Inhibitor. Photoreactive azachalcone **3** was designed on the basis of structure−activity studies on our new class of azachalcone TG2 inhibitors (*23*). Diazirine **1** (*31*) was selected as the photolabile moiety because its small volume is not likely to interfere with enzyme binding and its rapid photoactivation is initiated at wavelengths

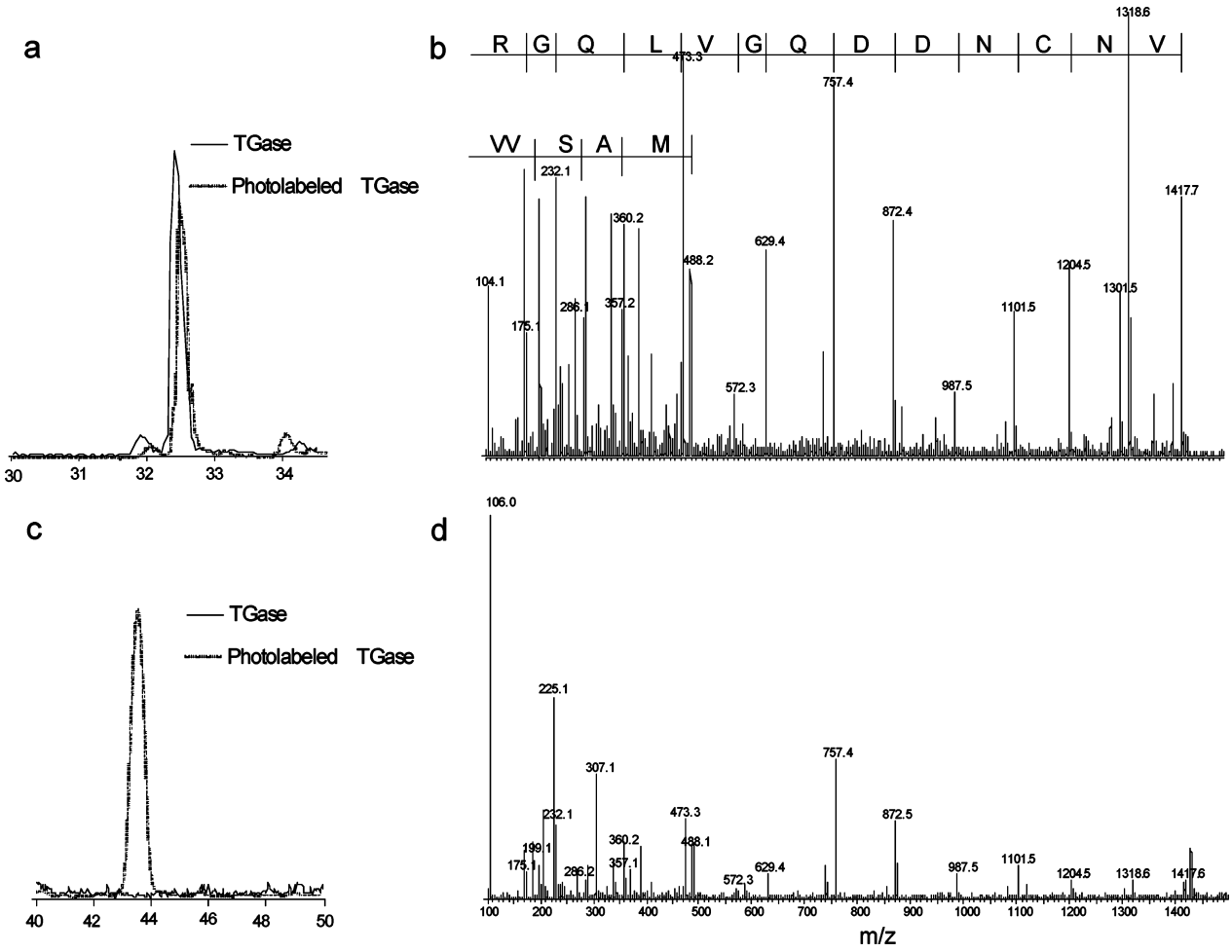
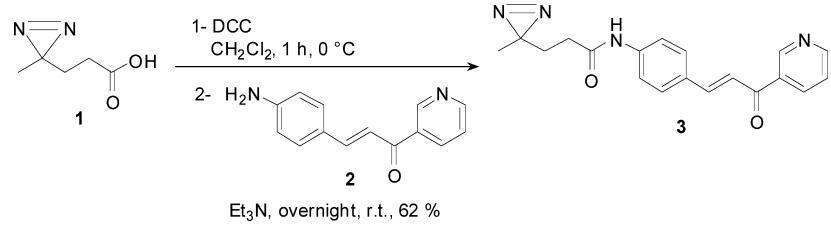


FIGURE 2: Identification of photolabeled peptide ions by LC−MS/MS. (a) Reconstructed ion chromatogram for m/z 952.9^{2+} from unlabeled tryptic peptide V_{223}VSAMVNCNDDQGVLQGR$_{240}$ of TGase present in both control and photolabeled samples. (b) MS/MS spectrum of the m/z 952.9^{2+} precursor ion confirming the sequence assignment. (c) Reconstructed ion chromatogram for the m/z 1120.0^{2+} ion from modified tryptic peptide V_{223}VSAMVNCNDDQGVLQGR$_{240}$ of TGase uniquely detected in a photolabeled sample. (d) MS/MS spectrum of the m/z 1120.0^{2+} precursor ion showing the characteristic fragment ion at m/z 307.1 corresponding to the azachalcone moiety ($C_{19}H_{19}N_2O_2^+$). Y-Type fragment ions are shown in panel b for the sake of convenience.

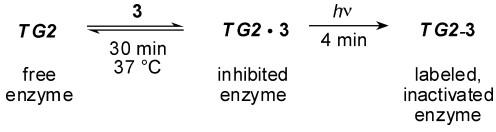
Scheme 1: Synthesis of Photolabile Inhibitor **3**



benign to protein structure (*42*). Photoactivation of alkyl diazirines generates diazo and carbene derivatives (*43, 44*) that react rapidly to covalently label residues in their proximity (*45*). Carboxylic acid **1** was coupled to azachalcone **2** (*23*) using dicyclohexylcarbodiimide (DCC) and triethylamine to provide photoactive inhibitor **3** in 62% isolated yield (Scheme 1).

Azachalcone **3** was determined to be a reversible TGase inhibitor having an IC_{50} value of 28 μM in direct continuous colorimetric assays (see Materials and Methods). It is interesting to note that the IC_{50} value of **3** is lower than that of the parent azachalcone **2** (148 μM) but similar to that of its *N*-acetylated derivative (28 μM) (*23*). Incubation of **3** with guinea pig liver TGase for 30 min at 37 °C and irradiation for 4 min led to a partial decrease in TGase activity. This

Scheme 2: Photoinactivation of TG2



solution was filtered over a 30 kDa molecular mass cutoff membrane and re-evaluated for activity, relative to control TGase that was irradiated in the absence of **3**. This confirmed that the activity loss was due to irreversible inhibition (Scheme 2). The covalent incorporation of photolabel in inactivated TGase was confirmed by ESI-MS.

Mass Spectrometry Analysis. LC−MS/MS analyses were performed on tryptic digests of TGase from control and photolabeled samples, representing 85% coverage of the

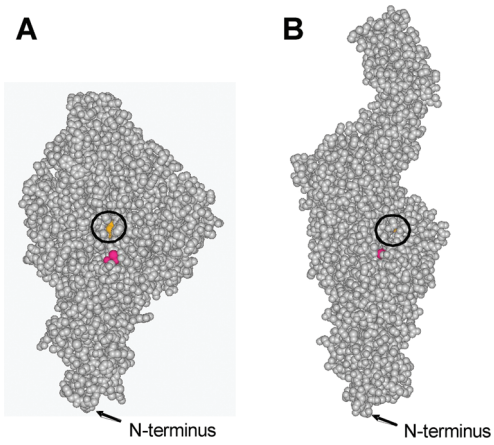


FIGURE 3: Homology models used to visualize the position of labeled residue Cys230, colored magenta. The active site is encircled, and the catalytic Cys277 is colored orange: (A) closed-form space-filling model and (B) open-form space-filling model.

peptide sequence (Figure 1), to identify residues modified with inhibitor **3**. Identification of a tryptic peptide bearing modified residues was facilitated through the observation of characteristic fragment ions corresponding to the cleavage of the azachalcone moiety. The unmodified tryptic peptide V_{223}VSAMVNCNDDQGLQGR$_{240}$ was observed at *m*/*z* 952.9^{2+} in both control and photolabeled samples (Figure 2a). The product ion spectrum of *m*/*z* 952.9^{2+} confirmed the sequence assignment of the unmodified peptide ion (Figure 2b). However, a doubly protonated peptide ion at *m*/*z* 1120.0^{2+} was observed only in the photolabeled sample (Figure 2c). MS/MS of the corresponding precursor ion yielded y-type fragment ions consistent with the backbone peptide sequence V_{223}VSAMVNCNDDQGLQGR$_{240}$ (Figure 2d) together with specific fragment ions associated with the photolabel group. More specifically, the presence of the azachalcone moiety was readily identified by abundant fragment ions at *m*/*z* 307.1 and 225.1 corresponding to the cleaved photolabeling group ($C_{19}H_{19}N_2O_2^+$) and its amide bond cleavage fragment ($C_{14}H_{13}N_2O_1^+$), respectively. The localization of the modified residue was assigned to the Cys230 residue. Indeed, the abundant peak at *m*/*z* 106.0 is present in only the MS/MS spectrum of the peptide ion *m*/*z* 1120.0^{2+}. This fragment was attributed to C_3H_7NOS (±4 ppm of theoretical value, compared to >40 ppm for all other empirical formulas) generated by collision-induced cleavage of the photolabeling group and subsequent rearrangement of the cysteine residue. Comparison of the relative abundance of the *m*/*z* 952.9 peptide ions from both samples (Figure 2a) suggested that in the photolabeled sample approximately 15% of the TGase was labeled with photolabile inhibitor **3** on this peptide. This yield is typical for diazirine-based photolabeling (*46*) and corresponds roughly to the observed decrease in enzymatic activity.

Homology Modeling. Guinea pig liver TGase shares significant sequence identity (homology) with TGases from other vertebrates: human TG2, 83% (90%); red sea bream TG2, 43% (59%); human epidermal TGase, 39% (56%); and human factor XIIIa, 36% (51%). Guinea pig liver TGase is one of the most thoroughly characterized TGases, making it a common choice for inhibition studies. However, its structure has never been determined. Therefore, the interpretation of our labeling results, in the context of an inhibitor binding mode, required the generation of the first structural homology models for this enzyme, based on the closed (*25, 26*) and more recent open (*27*) conformations of homologous TG2 structures. As a result of the high degree of homology, sequence alignments obtained using two algorithms (see Materials and Methods) were identical except in the extremities of gaps, which should have little effect on homology modeling. Five three-dimensional models for guinea pig liver TGase were thus derived for each sequence alignment file.

For the closed-form models, evaluation of model quality identified a low percentage (≤2.0%) of residues having backbone torsional angles in the disallowed regions of the Ramachandran plot, consistent with properly folded models. Furthermore, the percentages obtained for the models were in the same range as the values for the template crystal structures. No deleterious residue contacts were detected that were not already present in the template crystal structures. All models had regions with unfavorable residue interaction energies (Table S1 of the Supporting Information). For the open-form model, a higher percentage (between 13.8 and 15.9%) of residues in the disallowed regions of the Ramachandran plot was obtained, and those percentages were higher than for the template crystal structure (4.6%) (Table S2 of the Supporting Information). This poorer result is likely due to the fact that a single template was available for modeling the open form. Nonetheless, low rmsd values between all models (between 2.99 and 5.59 Å for the main chain atoms) demonstrated coherence. All models were refined and analyzed to select the final closed-form and open-form models, according to total energies of each structure, rmsd values, rotamer and Ramachandran outliers, and clash score (Supporting Information).

The retained, closed-form model T2 was subjected to an unconstrained molecular dynamics simulation (Figure S1 of the Supporting Information). The total energy of model T2 decreased rapidly in the first 25 ps and then decreased slowly for the remainder of the simulation. Its radius of gyration decreased rapidly by >1 Å in parallel with the energy and then by <1 Å during the remainder of the simulation. Together, these results indicate that the model was near a minimum energy structure. While this simulation does not allow us to establish that the global minimum has been approximated, it confirms that the closed-form model T2 of guinea pig liver TGase is stable at room temperature. The quality of this closed-form model was validated according to backbone conformation, residue contacts, residue interactions, and overall stability. All values fell within the limits of statistically valid structures. The overall structure of model T2 was very similar to that of human TG2 (backbone rmsd = 1.44 Å) as well as red sea bream tissue TGase (backbone rmsd = 4.03 Å) (Figure S2 of the Supporting Information). These data were consistent with guinea pig liver TGase having a higher degree of sequence homology with human TG2 than with red sea bream TG2. The structural similarity between model T2 and the template structures is not surprising, as the four templates are themselves structurally similar. The active site region is negatively charged for all the closed-form structures (Figure S2 of the Supporting Information), and the overall charge distribution is similar, the donor substrate binding groove being largely hydrophobic (*26*). These observations validate the use of the widely studied guinea pig liver TGase for testing inhibitors.

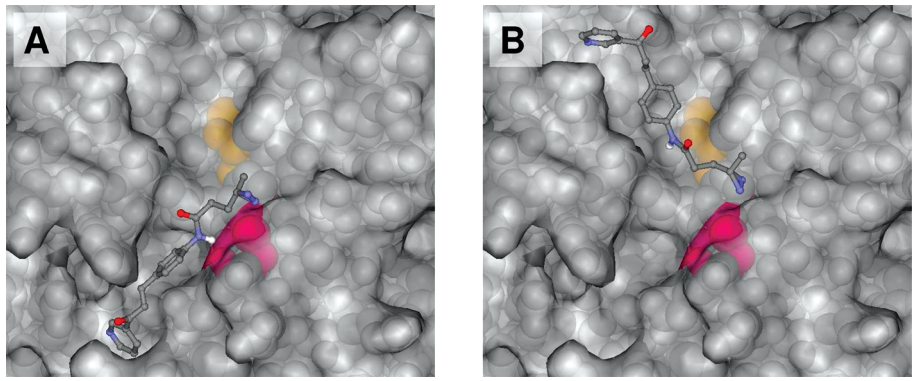


FIGURE 4: Two potential binding modes of photolabile inhibitor **3** in the donor substrate binding site of guinea pig liver TGase, suggested by simulated docking experiments with the closed-form homology model. Residue Cys230 is colored magenta, and catalytic Cys277 at the bottom of the active site tunnel is colored orange.

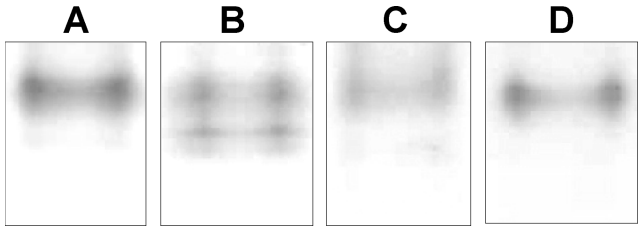


FIGURE 5: Nondenaturing PAGE investigation of the conformational effect of ligand binding. (For concentrations and conditions, see Materials and Methods.) (A) TGase as purified herein, without incubation. (B) TGase incubated in the presence of Mg·GTP, showing both faster and slower migrating species. (C) TGase incubated in the presence of calcium, showing predominantly the slower migrating species. (D) TGase incubated in the presence of inhibitor **3**, showing predominantly the slower-migrating form.

Having been generated from a single template crystal structure, the open-form model was of lower quality. Nonetheless, the high degree of sequence homology between the human TG2 template and the modeled guinea pig liver TGase is reflected by a high level of overall structural similarity and charge distribution in the open form as well (Figure S3 of the Supporting Information).

We used our homology models to visualize the position of labeled residue Cys230. In the closed-form model (Figure 3A), it is clear that Cys230 is located in a surface-accessible location within the hydrophobic groove that was previously identified as a potential donor substrate binding site (*28*). Binding in this area is clearly consistent with competitive inhibition kinetics due to blocking access to the active site. The labeled residue Cys230 does not appear to be as surface-accessible in the open-form model (Figure 3B). While these modeling results suggest that labeling by photoreactive azachalcone **3** occurs on the closed form of guinea pig liver TGase, it is important to acknowledge that the open-form model is based on only one empirical structure, and the solvent accessibility of residue Cys230 in this model is determined solely by side chain conformations.

Docking Simulations. Simulated docking experiments were then performed with our closed-form and open-form homology model, using the semiempirically (AM1) minimized structure of **3** as a ligand, to see if it could be bound as a competitive inhibitor in orientations consistent with the observed labeling patterns. Over the course of 150 simulations on the closed-form model, two predominant binding modes were readily distinguished (Figure 4). The resulting structures were not energy-minimized and should not be interpreted as representing final coordinates of bound **3**. However, these modes both demonstrate the affinity of the rigid aromatic scaffold of our inhibitors for the generally hydrophobic substrate binding groove (*47*). In both cases, the reactive diazirine is proximal to Cys230 while the remainder of the inhibitor establishes multiple contacts within the active site area. Both conformations are consistent with the observed competitive mode of inhibition.

Simulated docking with the open-form model yielded conformational clustering of poor quality only and was not successful in identifying any predominant binding mode. Furthermore, none of the binding models thus generated were consistent with the observed labeling, but this is not surprising, given that the labeled Cys230 appears to be less surface-accessible in this model (Figure 3B). However, the significance of these docking simulations is strictly dependent on the reliability of the available homology model (see above).

Conformational Considerations. Our modeling results suggest that the initial binding of **3** occurs on the closed form of TGase. However, the crystal structure of TGase in the open conformation is the only one that features a ligand bound in its active site (*27*). This peptidic irreversible inhibitor closely resembles a high-affinity donor substrate, so the open-form crystal structure represents, in the very least, the best model that is known for the acyl−enzyme intermediate. According to the well-known ping-pong kinetic mechanism that characterizes the catalytic cycle of TGase, it is the acyl−enzyme intermediate, not the free enzyme, that binds the second acyl acceptor substrate, and the open structure provides the first clues for the location of this substrate binding site (*27*). Further, that work and other structural studies (*48*−*50*) suggest that this open conformation may be representative of the active conformation of the free enzyme, which binds the acyl donor substrate.

To provide more direct empirical evidence indicating which conformation of TGase is involved in the binding of **3**, nondenaturing PAGE experiments were performed. As shown in Figure 5A, TGase purified for these studies gave rise to one intense band on a nondenaturing gel assessed in the absence of added ligands. After incubation of purified TGase with a subsaturating (*51*) concentration of Mg·GTP (25 μM), two distinct bands were observed (Figure 5B). This is consistent with previous observations (*27, 50*) that the binding of GTP shifts the conformational equilibrium to favor the faster migrating, presumably closed form of TGase. In

contrast, incubation in the presence of calcium favors the slower migrating, presumably open-form conformation (Figure 5C), which has been observed previously (*27*). Incubation of TGase with inhibitor **3** gave one intense band consistent with the open-form conformation (Figure 5D). The incubation of TGase with **3** in the presence of added calcium gave similar results, as expected (results not shown).

The experimental data from the native gels support the binding of **3** to the open form of the enzyme. However, our results from docking simulations with the open-form homology model (vide supra) were inconclusive, suggesting that this model does not accurately represent the binding-competent form of the enzyme. The proximity of the labeled residue, Cys230, to the surface of the open-form homology model suggests that potential inaccuracies as subtle as side chain conformations may be sufficient to render the model inconsistent with our empirical binding and labeling results. Alternatively, a conformational change, smaller in magnitude than that relating the open and closed conformational states, may be required to provide the open form with an environment suitable for inhibitor binding and subject to subsequent labeling.

Summary. A photolabile derivative of a potent reversible inhibitor was prepared and used to photolabel guinea pig liver TGase in the region of its inhibitor binding site. The residue modified by this labeling reaction was identified by MS/MS to be Cys230. The first homology models for the closed and open forms of guinea pig liver TGase were generated and used to locate Cys230 on and near the surface of the enzyme, respectively, in the shallow hydrophobic groove at the mouth of the tunnel leading to the active site. Nondenaturing gels suggest the inhibitor is bound by the open-form conformation. No predominant binding model was identified through docking simulations with the open-form homology model, suggesting that this model must differ from the form of the enzyme that binds the inhibitor, at least at the level of side chain conformations. The empirical evidence provided herein will be critical for the subsequent structure-based design of inhibitors for tissue TGase.

ACKNOWLEDGMENT

We are grateful to our departmental colleague, Prof. Andreea Schmitzer, for the use of her UV oven.

SUPPORTING INFORMATION AVAILABLE

Refinement parameters, molecular dynamics data, and structures of the guinea pig liver transglutaminase homology models. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Folk, J. E., and Cole, P. W. (1996) Transglutaminase: Mechanistic features of the active site as determined by kinetic and inhibitor studies. *Biochim. Biophys. Acta 22*, 244.
2. Achyuthan, K. E., Slaughter, T. F., Santiago, M. A., Enghild, J. J., and Greenberg, C. A. (1993) Factor XIIIa-derived peptides inhibit transglutaminase activity. Localization of substrate recognition sites. *J. Biol. Chem. 268*, 21284.
3. Greenberg, C. S., Birchbichler, P. J., and Rice, R. H. (1991) Transglutaminases: Multifunctional cross-linking enzymes that stabilize tissues. *FASEB J. 5*, 3071.
4. Levistzki, A., Willingham, M., and Pastan, I. H. (1980) Evidence for participation of transglutaminase in receptor-mediated endocytosis. *Proc. Natl. Acad. Sci. U.S.A. 77*, 2706–2710.
5. Davies, P. J., Davies, D. R., Levitzki, A., Maxfield, F. R., Milhaud, P., Willingham, M. C., and Pastan, I. H. (1980) Transglutaminase is essential in receptor-mediated endocytosis of α2-macroglobulin and polypeptide hormones. *Nature 283*, 162–167.
6. Fesus, L., Thomazy, V., and Falus, A. (1987) Induction and activation of tissue transglutaminase during programmed cell death. *FEBS Lett. 224*, 104–108.
7. Birckbichler, P. J., Orr, G. R., Patterson, M. K., Conway, E., Carter, H. A., and Maxwell, M. D. (1983) Enhanced transglutaminase activity in transformed human lung fibroblast cells after exposure to sodium butyrate. *Biochim. Biophys. Acta 763*, 27–34.
8. De Young, L., Ballaron, S., and Epstein, W. (1984) Transglutaminase Activity in Human and Rabbit Ear Comedogenesis: A Histochemical Study. *J. Invest. Dermatol. 82*, 275–279.
9. Azari, P., Rahim, J., and Clarkson, D. P. (1981) Transglutaminase activity in normal and hereditary cataractous rat lens and its partial purification. *Curr. Eye Res.,* 463–469.
10. Fesus, L. (1982) Transglutaminase activation: Significance with respect to immunologic phenomena. *Surv. Immunol. Res. 1*, 297–304.
11. Schroeder, W. T., Thatcher, S. M., Stewart-Galetka, S., Annarella, M., Chema, D., Siciliano, M. J., Davies, P. J. A., Tang, H. Y., Sowa, B. A., and Duvic, M. (1992) Type I keratinocyte transglutaminase: Expression in human skin and psoriasis. *J. Invest. Dermatol. 99*, 27–34.
12. Selkoe, D. J., Abraham, C., and Ihara, Y. (1982) Brain transglutaminase: In vitro crosslinking of human neurofilament proteins into insoluble polymers. *Proc. Natl. Acad. Sci. U.S.A. 79*, 6070–6074.
13. Norlund, M. A., Lee, J. M., Zainelli, G. M., and Muma, N. A. (1999) Elevated transglutaminase-induced bonds in PHF tau in Alzheimer's disease. *Brain Res. 851*, 154–163.
14. Dedeoglu, A., Kubilus, J. K., Jeitner, T. M., Matson, S. A., Bogdanov, M., Kowall, N. W., Matson, W. R., Cooper, A. J. L., Ratan, R. R., and Beal, M. F. (2002) Therapeutic Effects of Cystamine in a Murine Model of Huntington's Disease. *J. Neurosci. 22*, 8942–8950.
15. Mastroberardino, P. G., Iannicola, C., Nardacci, R., Bernassola, F., De Laurenzi, V., Melino, G., Moreno, S., Pavone, F., Oliverio, S., and Fesus, L. (2002) 'Tissue' transglutaminase ablation reduces neuronal death and prolongs survival in a mouse model of Huntington's disease. *Cell Death Differ. 9*, 873–880.
16. Piper, J. L., Gray, G. M., and Khosla, C. (2002) High selectivity of human tissue transglutaminase for immunoactive gliadin peptides: Implications for celiac sprue. *Biochemistry 41*, 386–393.
17. Choi, K., Siegel, M., Piper, J. L., Yuan, L., Cho, E., Strnad, P., Omary, B., Rich, K. M., and Khosla, C. (2005) Chemistry and biology of dihydroisoxazole derivatives: Selective inhibitors of human transglutaminase 2. *Chem. Biol. 12*, 469–475.
18. Hausch, F., Halttunen, T., Mäki, M., and Khosla, C. (2003) Design, synthesis, and evaluation of gluten peptide analogs as selective inhibitors of human tissue transglutaminase. *Chem. Biol. 10*, 225.
19. Marrano, C., de Macédo, P., and Keillor, J. W. (2001) Synthesis and Evaluation of Dipeptide-Bound 1,2,4-Thiadiazoles as Irreversible Transglutaminase Inhibitors. *Bioorg. Med. Chem. 9*, 3231.
20. Marrano, C., de Macédo, P., and Keillor, J. W. (2001) Evaluation of Novel Dipeptide-Bound α,β-Unsaturated Amides and Epoxides as Irreversible Inhibitors of Guinea Pig Liver Transglutaminase. *Bioorg. Med. Chem. 9*, 1923.
21. de Macédo, P., Marrano, C., and Keillor, J. W. (2002) Synthesis of Dipeptide-Bound Epoxides and α,β-Unsaturated Amides as Potential Irreversible Transglutaminase Inhibitors. *Bioorg. Med. Chem. 10*, 355.
22. Duval, E., Case, A., Stein, R. L., and Cuny, G. D. (2005) Structure-activity relationship study of novel tissue transglutaminase inhibitors. *Bioorg. Med. Chem. 15*, 1885.
23. Pardin, C., Pelletier, J. N., Lubell, W. D., and Keillor, J. W. (2008) Cinnamoyl inhibitors of tissue transglutaminase. *J. Org. Chem. 73*, 5766–5775.
24. Pardin, C., Roy, I., Lubell, W. D., and Keillor, J. W. (2008) Reversible and competitive triazole-based inhibitors of tissue transglutaminase. *Chem. Biol. Drug Des. 72*, 189–196.

25. Nogushi, K., Ishikawa, K., Yokoyama, K.-I., Ohtsuka, T., Nio, N., and Suzuki, E.-I. (2001) Crystal Structure of Red Sea Bream Transglutaminase. *J. Biol. Chem. 276*, 12055–12059.
26. 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.
27. 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.
28. Chica, R. A., Gagnon, P., Keillor, J. W., and Pelletier, J. N. (2004) Tissue transglutaminase acylation: Proposed role of conserved active site Tyr and Trp residues revealed by molecular modelling of peptide substrate binding. *Protein Sci. 13*, 979.
29. Case, A., and Stein, R. L. (2007) Kinetic analysis of the interaction of tissue transglutaminase with a nonpeptidic slow-binding inhibitor. *Biochemistry 46*, 1106–1115.
30. Leblanc, A., Gravel, C., Labelle, J., and Keillor, J. W. (2001) Kinetic Studies of Guinea Pig Liver Transglutaminase Reveal a General-Base-Catalyzed Deacylation Mechanism. *Biochemistry 40*, 8335.
31. Church, R. F. R., and Weiss, M. J. (1970) Diazirines. II. Synthesis and properties of small functionalized diazirine molecules. Observations on the reaction of a diaziridine with the iodine-iodide ion system. *J. Org. Chem. 35*, 2465.
32. Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature 227*, 680–685.
33. Ahvazi, B., Kim, H. C., Kee, S. H., Nemes, Z., and Steinert, P. M. (2002) Three-dimensional structure of the human transglutaminase 3 enzyme: Binding of calcium ions changes structure for activation. *EMBO J. 21*, 2055.
34. Garzon, R. J., Pratt, K. P., Bishop, P. D., Le Trong, I., Stenkamp, R. E., and Teller, D. C. (2000) Human Factor XIII with calcium bound in the ion site. Protein Data Bank, doi 10.2210/pdb1evu/pdb.
35. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A Novel Method for Fast and Accurate Multiple Sequence Alignment. *J. Mol. Biol. 302*, 205.
36. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673.
37. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779.
38. Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Cα geometry: Φ, ϕ and Cβ deviation. *Proteins: Struct., Funct., Genet. 50*, 437.
39. Hooft, R. W. W., Vriend, G., Sander, C., and Abola, E. E. (1996) Errors in protein structures. *Nature 381*, 272.
40. Sippl, M. J. (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins 17*, 355.
41. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem. 19*, 1639.
42. Nassal, M. (1984) 4′-(1-Azi-2,2,2-trifluoroethyl)phenylalanine, a photolabile carbene-generating analog of phenylalanine. *J. Am. Chem. Soc. 106*, 7540.
43. Brunner, J. (1993) New photolabeling and crosslinking methods. *Annu. Rev. Biochem. 62*, 483–514.
44. Ziebell, M. R., Nirthanan, S., Husain, S. S., Miller, K. W., and Cohen, J. B. (2004) Identification of Binding Sites in the Nicotinic Acetylcholine Receptor for [^{3}H]Azietomidate, a Photoactivatable General Anesthetic. *J. Biol. Chem. 279*, 17640–17649.
45. Das, J., Zhou, X., and Miller, K. W. (2006) Identification of an alcohol binding site in the first cysteine-rich domain of protein kinase Cδ. *Protein Sci. 15*, 2107–2119.
46. Addona, G. H., Husain, S. S., Stehle, T., and Miller, K. W. (2002) Geometric Isomers of a Photoactivable General Anesthetic Delineate a Binding Site on Adenylate Kinase. *J. Biol. Chem. 277*, 25685–25691.
47. Keillor, J. W. (2005) Tissue Transglutaminase Inhibition. *Chem. Biol. 12*, 410.
48. Di Venere, A., Rossi, A., De Matteis, F., Rosato, N., Finazzi, A., and Mei, G. (2000) Opposite Effects of Ca^{2+} and GTP Binding on Tissue Transglutaminase Tertiary Structure. *J. Biol. Chem. 275*, 3915–3921.
49. Casadio, R., Polverini, E., Mariani, P., Spinozzi, F., Carsughi, F., Fontana, A., Polverino de Laureto, P., Matteucci, G., and Bergamini, C. M. (1999) The structural basis for the regulation of tissue transglutaminase by calcium ions. *Eur. J. Biochem. 262*, 672–679.
50. 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.
51. Lai, T.-S., Slaughter, T. F., Peoples, K. A., Hettasch, J. M., and Greenberg, C. S. (1998) Regulation of Human Tissue Transglutaminase Function by Magnesium-Nucleotide Complexes. *J. Biol. Chem. 273*, 1776–1781.

BI802021C