J|A|C|S
ARTICLES
Published on Web 04/29/2006

Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction

Erik C. B. Johnson[†,§] and Stephen B. H. Kent*[,†,‡,§]

Contribution from the Department of Biochemistry and Molecular Biology, Department of Chemistry, Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637

Received December 8, 2005; E-mail: skent@uchicago.edu

Abstract: Native chemical ligation of unprotected peptide segments involves reaction between a peptide-α-thioester and a cysteine-peptide, to yield a product with a native amide bond at the ligation site. Peptide-α-thioalkyl esters are commonly used because of their ease of preparation. These thioalkyl esters are rather unreactive so the ligation reaction is catalyzed by in situ transthioesterification with thiol additives. The most common thiol catalysts used to date have been either a mixture of thiophenol/benzyl mercaptan, or the alkanethiol MESNA. Despite the use of these thiol catalysts, ligation reactions typically take 24−48 h. To gain insight into the mechanism of native chemical ligaton and in order to find a better catalyst, we investigated the use of a number of thiol compounds. Substituted thiophenols with pK_a > 6 were found to best combine the ability to exchange rapidly and completely with thioalkyl esters, and to then act as effective leaving groups in reaction of the peptide-thioester with the thiol side chain of a cysteine-peptide. A highly effective and practical catalyst was (4-carboxylmethyl)thiophenol ('MPAA'), a nonmalodorous, water-soluble thiol. Use of MPAA gave an order of magnitude faster reaction in model studies of native chemical ligation and in the synthesis of a small protein, turkey ovomucoid third domain (OMTKY3). MPAA should find broad use in native chemical ligation and in the total synthesis of proteins.

Introduction

Native chemical ligation[1] is widely used for the preparation of proteins by total or semi-synthesis. Ligation of chemically synthesized peptides enables the total synthesis of small proteins, and grants the researcher complete atom-by-atom control over the covalent structure of the protein molecule. To date, hundreds of proteins have been synthesized using native chemical ligation (for a review, see ref 2). Expressed protein ligation[3] is a form of protein semisynthesis[4] that uses native chemical ligation to react a peptide-thioester generated by recombinant means, with a Cys-peptide and is becoming widely used for the application of chemistry to the study of the molecular biology of proteins.[5]

Native chemical ligation involves the chemoselective reaction of two unprotected peptides in aqueous solution to give a single covalently linked ligation product. A peptide-thioester is reacted with a Cys-peptide to give a product polypeptide with a native peptide bond at the ligation site. While full details of the mechanism of native chemical ligation have not yet been elucidated, reaction is envisioned to occur according to the reaction scheme shown in Figure 1. A peptide-α-thioester is reacted with a peptide containing an N-terminal cysteine residue; transthioesterification with the side chain thiol of the N-terminal cysteine residue results in a thioester-linked intermediate, which spontaneously rearranges through a favorable five-membered ring intramolecular nucleophilic attack by the cysteine α-amino group to form a native amide bond at the Xxx-Cys ligation site. Previous work by our laboratory and by others has shown that the rate of ligation is dependent on the identity of the C-terminal amino acid in the peptide-thioester: ligation at less hindered C-terminal residues, such as Gly- or Ala-$^{\alpha}$COSR, proceeds more quickly than ligation at sterically hindered, β-branched C-terminal residues such as Ile- or Val-$^{\alpha}$COSR.[6] The reaction rate also depends on the nature of the thiol leaving group. Ligation with a highly activated 5-thio-2-nitrobenzoic acid thioester proceeds to completion within 10 min at pH 7.0, whereas ligation with an alkyl thioester is very slow and can take days to reach completion.[1] These and similar observations,[7] in conjunction with the absence of observable thioester-linked intermediate ligation product, suggest that the rate-limiting step in native chemical ligation is the transthioesterification with the thiol moiety of the side chain of the N-terminal cysteine.

Initially, synthetic peptide-thioalkylesters were prepared by alkylating peptide-thiocarboxylates;[1] more recently, peptide-thioalkyl esters have been prepared by direct synthesis.[6,8] These

[†] Department of Biochemistry and Molecular Biology.
[‡] Department of Chemistry.
[§] Institute for Biophysical Dynamics.

(1) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776−779.
(2) Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.
(3) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705−6710.
(4) Rees, A. R.; Offord, R. E. *Biochem. J.* **1976**, 159, 467−486.
(5) Muir, T. W. *Annu. Rev. Biochem.* **2003**, 72, 249−289.
(6) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 10068−10073.
(7) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* **1997**, *119*, 4325−4329.
(8) Aimoto, S. *Biopolymers* **1999**, *51*, 247−265.

10.1021/ja058344i CCC: $33.50

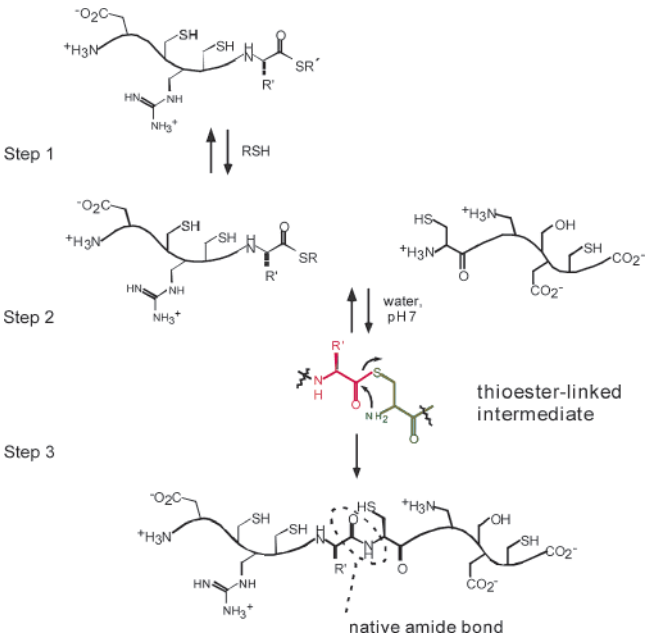
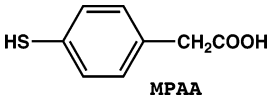


Figure 1. Proposed Mechanism for Native Chemical Ligation. *Step 1*: exogenous thiol-thioester exchange; *Step 2*: transthioesterification with the side chain thiol of N-terminal cysteine creates a thioester-linked intermediate ligation product; *Step 3*: the thioester-linked intermediate undergoes spontaneous intramolecular nucleophilic rearrangement to form an amide bond at the ligation site. Steps 1 & 2 are freely reversible, whereas Step 3 is irreversible, under the conditions used (aqueous solution, pH7, added thiol catalyst).

alkyl thioesters are straightforward to prepare by solid phase peptide synthesis and are easy to handle during work up and purification but are relatively unreactive under native chemical ligation conditions in water at pH 7. For that reason, exogenous thiols have been added to native chemical ligation reactions to promote the in situ formation of more active thioesters and thus increase the kinetics of ligation.[7] The added thiol also maintains the N-terminal cysteine in reduced form and reverses any nonproductive transthioesterification with the thiol moieties of side chains of internal cysteine residues. The most common thiol catalysts used for native chemical ligation are a 1% benzyl mercaptan/3% thiophenol mix for chemically synthesized peptide-thioesters,[7] or 2-mercaptoethanesulfonate sodium salt (MESNA) for recombinant peptide-thioesters.[5] However, even with these catalysts, ligations at unhindered sites still often require days to go to completion, during which prolonged time side reactions can occur.

We were therefore interested in further exploring the mechanism of native chemical ligation, with the goal of improving ligation rates through the use of better thiol catalysts. We examined a set of fourteen different thiols selected on the basis of their pK_a. We found that aryl thiols were the most effective catalysts, and that in particular (4-carboxymethyl)thiophenol (MPAA) is a highly effective catalyst for native chemical ligation, is freely water soluble, and



MPAA

does not have a strong or offensive odor. The application of this improved catalyst was exemplified by the synthesis of the OMTKY3 polypeptide chain by the native chemical ligation of two unprotected peptide segments an order of magnitude more rapidly than previously reported.

Table 1. Thiols Tested for Catalytic Activity[a]

thiol	pK_a	solubility in water
Benzyl mercaptan	9.67	sparingly soluble
2-mercaptoethanesulfonate, sodium salt (MESNA)	9.2[b]	very soluble
(4-amino)thiophenol	8.75	soluble
(4-hydroxy)thiophenol	6.8	soluble
(4-methyl)thiophenol	6.8	sparingly soluble
(4-methoxy)thiophenol	6.75	sparingly soluble
thiophenol	6.6	sparingly soluble
4-(carboxymethyl)thiophenol (MPAA)	6.6[c]	very soluble
(3-hydroxy)thiophenol	6.5	soluble
(4-chloro)thiophenol	6.11	sparingly soluble
(3-carboxy)thiophenol	5.78[d]	very soluble
(4-nitro)thiophenol	4.67	soluble
2-Mercaptopyrimidine	1.65	very soluble
2-Mercaptopyridine	−1	soluble

[a] pK_a values and solubility characteristics were obtained from "calculated properties" for each compound in SciFinder Scholar v.2006. [b] pK_a was assumed to be the same as the cysteine side chain sulfur. [c] pK_a was obtained from ref 9. [d] pK_a was obtained from ref 10.

Results

A set of thiols, both alkyl and aryl, were chosen based on the pK_a of the thiol moiety (Table 1). On the basis of the previously reported[1] much greater reactivity of the 5-thio-2-nitrobenzoic acid thioester compared with an alkyl thioester, our expectation was that thiols with lower pK_a would be better leaving groups and therefore better catalysts for native chemical ligation.

We tested these thiols for their ability to catalyze native chemical ligation reactions under standard conditions. A model ligation reaction test system was used, consisting of the peptide-thioester Leu-Tyr-Arg-Ala-Leu-*COSCH*$_2$CH$_2$COLeu (1 mM), and the N-terminal cysteine peptide Cys-Leu-Tyr-Leu-Ala-Ala (2 mM). Leucine was chosen as the C-terminal amino acid on the peptide thioester because ligations at this site have previously been demonstrated to be slow.[6]

Thiols (10 mM) were present in 10-fold molar excess relative to the starting peptide-$^{\alpha}$thioalkylester. Tris(2-carboxyethyl)-phosphine (TCEP) (20 mM)was included in the ligation reaction to maintain all thiols in reduced form. The kinetics of the ligation reactions, as well as the extent of thioester-thiol exchange, were determined using analytical HPLC; products were identified by electrospray mass spectrometry. The amounts of ligation and exchange products were determined by manual integration of peak areas, with detection at 214 nm; the extinction coefficients of all peptide species are very similar.

The results for the fourteen thiols and for the control reaction in the absence of catalyst are shown in Figure 2. Most of the added thiols increased the rate of ligation over the control, with the exception of 2-mercaptopyridine and (4-nitro)thiophenol. The extent of ligation after 2 h reaction is plotted in relation to the pK_a of the thiol catalyst in Figure 3. The thiol which gave the fastest reaction was (4-amino)thiophenol, with a pK_a of 8.75. To our initial surprise, thiols with *higher* pK_a increased the rate of native chemical ligation to a greater extent than those with *lower* pK_a. However, this relationship held true only for thiols where the mercaptan group was located directly on the phenyl ring, i.e., aryl thiols (aryl mercaptans). Thiols where the mercaptan was attached to an alkyl group, such as benzyl

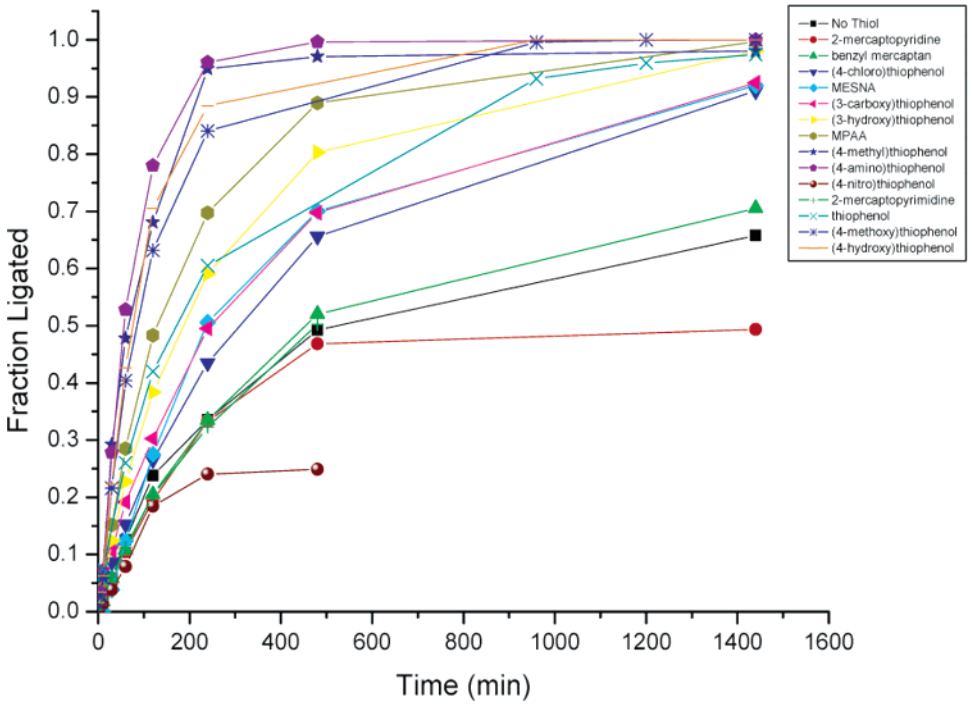


Figure 2. Effect of Thiols on Native Chemical Ligation. Two peptides, Leu-Tyr-Arg-Ala-Leu-$^{\alpha}$COS*CH$_2$CH$_2$COLeu* (1 mM) and Cys-Leu-Tyr-Leu-Ala-Ala (2 mM), were ligated at pH 7.0 in the presence of 10-fold molar excess (10 mM) of thiol and TCEP (20 mM). At each time point, the fraction ligated was determined by HPLC separation and integration of the ligated product as a fraction of the sum of {starting materials + thiol-thioester exchanged species + thioester hydrolysis (if any)}.

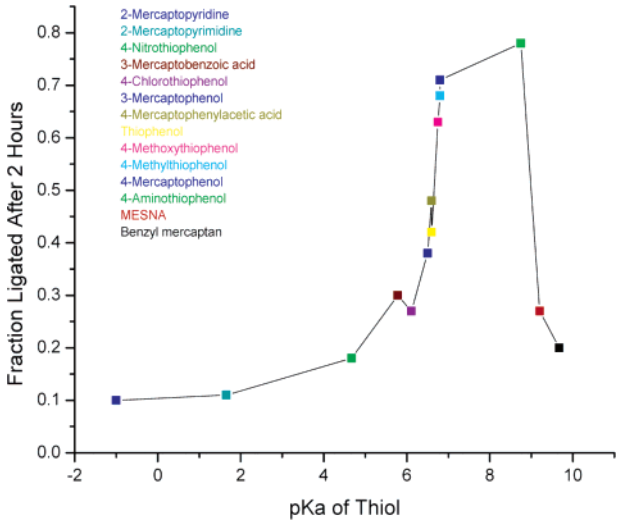


Figure 3. Fraction Ligated as a Function of Thiol pK_a. The fraction of ligated product in each reaction shown in Figure 2 was determined at 2 h and plotted as a function of the pK_a of the thiol used in the reaction. pK_a values are given in Table 1.

mercaptan and 2-mercaptoethanesulfonate (MESNA), were poor catalysts for native chemical ligation even though their pK_a values are higher than any of the phenyl thiol catalysts tested. Ligation with benzyl mercaptan and MESNA, with pK_a of 9.7 and 9.2, respectively, was substantially *slower* than for the substituted thiophenols with pK_a between 7 and 9.

Thiol-thioester exchange during the ligation reaction was observed only for those thiols with a pK_a of 6.5 or above, and increased with increasing pK_a of the mercaptan: benzyl mercaptan and MESNA showed the most thiol-thioester exchange during the course of the ligation reaction (data not shown). However, the results from the initial kinetic experiment indicated that phenyl thiols were superior ligation catalysts compared to alkanethiols. This suggests that the alkyl thioesters created by exchange with benzyl mercaptan and MESNA are less reactive than those of the phenyl thioesters. Further, the lack of accumulation of exchange products with aryl thiols suggests that thiol-thioester exchange was the rate-limiting step in catalysis; that is, any peptide-aryl thioester formed is immediately consumed in the native chemical ligation reaction.

An optimal catalyst would therefore be one that would balance leaving-group reactivity with the ability to undergo rapid and complete thiol-thioester exchange. We chose to further explore two phenyl thiols that met these criteria and that also gave clean reactions during the initial test: (3-hydroxy)-thiophenol and (4-carboxymethyl)thiophenol, abbreviated as MPAA for its trivial name, 4-mercaptophenylacetic acid. Model ligations of Leu-Tyr-Arg-Ala-Leu-$^{\alpha}$COS*CH$_2$CH$_2$COLeu* with Cys-Leu-Tyr-Leu-Ala-Ala were performed as was done in the initial kinetic experiment, using several concentrations of each thiol. Increasing the concentration of either thiol gave concomitant increases in ligation rate. MPAA was the more effective catalyst.

The results for MPAA are shown in Figure 4. At 250 or 400 mM added MPPA, the ligation rate approached the rate observed for the preformed Leu-Tyr-Arg-Ala-Leu-MPAA thioester. This suggested that thiol-thioester exchange was occurring rapidly at high concentrations of MPAA, and that even at 400 mM MPAA the thiol was not interfering with reaction of the Leu-Tyr-Arg-Ala-Leu-MPAA thioester with the thiol side chain of the N-terminal cysteine in the peptide Cys-Leu-Tyr-Leu-Ala-Ala, or with the subsequent intramolecular rearrangement of the initial thioester-linked product to form an amide bond at the ligation site.

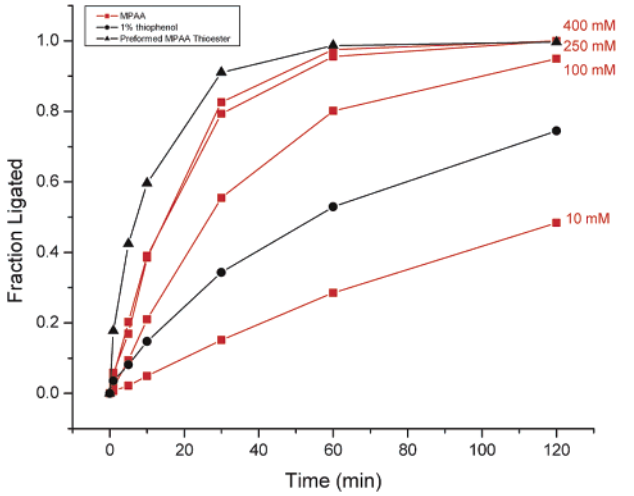


Figure 4. Ligation Rates at Increased Concentrations of the Thiol MPAA. (4-carboxymethyl)thiophenol (MPAA) was added at a concentration of 10, 100, 250, or 400 mM to the ligation reaction between Leu-Tyr-Arg-Ala-Leu-$^{\alpha}$COS*CH$_2$CH$_2$COLeu* and Cys-Leu-Tyr-Leu-Ala-Ala under the conditions used in Figure 2, and the yield of ligation product was determined by HPLC as a function of time (red dots). The reaction of the preformed Leu-Tyr-Arg-Ala-Leu-$^{\alpha}$MPAA thioester under the same conditions in the absence of added thiol is also shown (black triangles). The same reaction in the presence of 1%v/v thiophenol is shown for comparison (black circles).

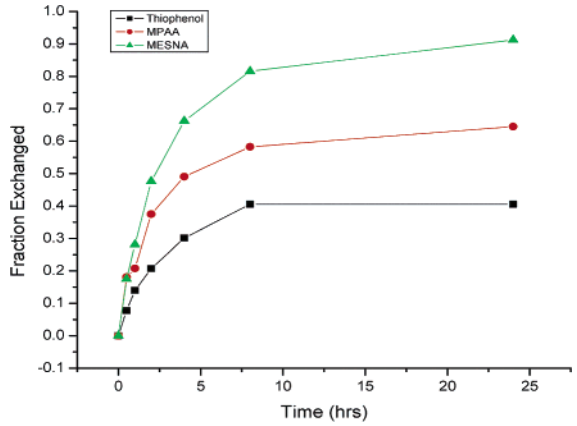


Figure 5. Thiol-thioester Exchange. The peptide Leu-Tyr-Arg-Ala-Leu-$^{\alpha}$COS*CH$_2$CH$_2$COLeu* (1 mM) was incubated at pH7 with 10 mM MESNA (green), MPAA (red), or thiophenol (black), and the fraction of thiol-thioester exchange as a function of time was determined by HPLC analysis.

To help dissect the catalytic mechanism, the thioester exchange reaction with added thiol was studied separately. The peptide-thioester Leu-Tyr-Arg-Ala-Leu-$^{\alpha}$*COSCH$_2$CH$_2$COLeu* (1 mM) was incubated with 10 mM MESNA, MPAA, or thiophenol, and the extent of thiol-thioester exchange was measured. The resulting data are shown in Figure 5.

The widely used water-soluble alkanethiol MESNA[5] gave the most exchange, about 90% after 24 h. In contrast, MPAA and thiophenol gave less exchange, although exchange with MPAA was both faster and more complete than with thiophenol at the same concentration (10 mM). Unlike thiophenol, however, MPPA is freely soluble in water and at higher concentrations MPAA-thioester exchange is rapid and complete (see Experimental Procedures).

To further dissect the catalytic mechanism, model ligations with *preformed* MPAA and thiophenyl peptide-thioesters were performed in the absence of added thiols (Figure 6).

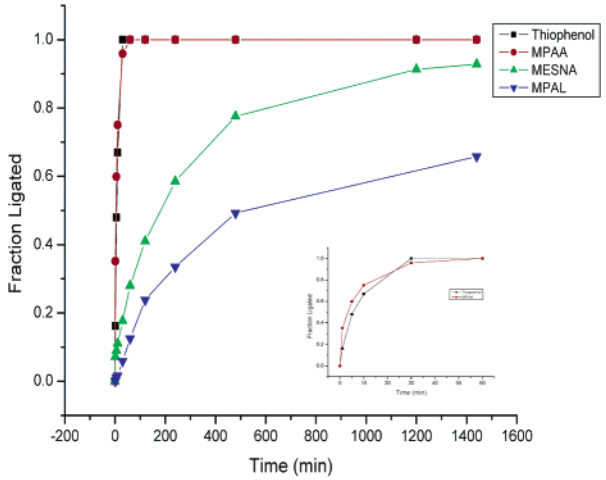


Figure 6. Ligation Rates with Preformed Thioesters. The peptide Leu-Tyr-Arg-Ala-Leu-αCO-X (1 mM) where X = S*CH$_2$CH$_2$COLeu*, MESNA, MPAA, or thiophenol, was reacted at pH7 with 2 mM Cys-Leu-Tyr-Leu-Ala-Ala with no thiols added to the reaction. Ligation yield as a function of time was determined by HPLC analysis. Inset: expanded view of rate of ligation with the preformed phenyl thioesters of MPAA and thiophenol.

Table 2. Effect of Various Thiols on the Rate of Ligation of OMTKY3(6−23)−$^{\alpha}$*COSCH$_2$CH$_2$COLeu* Thioester and (Cys^{24}-56)OMTKY3[a,b]

thiol	time (h)
200 mM MPAA	2
1% benzyl mercaptan	16
1% benzyl mercaptan, 3% thiophenol	16
2% thiophenol	24
100 mM MESNA	>36
No Thiol	>36

[a] OMTKY3(6−23) = Val-Asp-Cys-Ser-Glu-Tyr-Pro-Lys-Pro-Ala-Cys-Thr-Leu-Glu-Tyr-Arg-Pro-Leu. [b] OMTKY3(24−56) = Cys-Gly-Ser-Asp-Asn-Lys-Thr-Tyr-Gly-Asn-Lys-Cys-Asn-Phe-Cys-Asn-Ala-Val-Val-Glu-Ser-Asn-Gly-Thr-Leu-Thr-Leu-Ser-His-Phe-Gly-Lys-Cys.

Ligation with either the thiophenyl or MPAA preformed thioesters proceeded essentially to completion in 30 min, whereas the same ligation using the preformed MESNA alkyl thioester required more than 24 h to go to completion. Ligation with the canonical preformed *SCH$_2$CH$_2$COLeu* thioester was even slower (Figure 6). The rapid rates of reaction of the preformed MPAA and thiophenyl thioesters were presumed to reflect the rate of transthioesterification with the side chain thiol of the N-terminal cysteine, and were identical within experimental uncertainty (Inset, Figure 6).

To demonstrate the practical utility of (4-carboxymethyl)-thiophenol (MPAA) as a catalyst for the synthesis of proteins by native chemical ligation, we repeated the previously reported synthesis of the third domain of turkey ovomucoid trypsin inhibitor (OMTKY3), using the unprotected peptide segments 6-23-$^{\alpha}$COS*CH$_2$CH$_2$COLeu* and Cys^{24}-56. The amino acid sequences are shown in Table 2. Previous syntheses of this protein showed that ligation at the Leu^{23}-Cys^{24} ligation site required from 24 to 36 h before approaching completion.[11,12] Different thiol catalysts were analyzed for their ability to

(9) DeCollo, T. V.; Lees, W. J. *J. Org. Chem.* **2001**, *66*, 4244−4249.
(10) Eldin, S.; Jencks, W. P. *J. Am. Chem. Soc.* **1995**, *117*, 4851−4857.
(11) Lu, W. Y.; Qasim, M. A.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 8518−8523.
(12) Lu, W. Y.; Qasim, M. A.; Laskowski, M.; Kent, S. B. H. *Biochemistry* **1997**, *36*, 673−679.

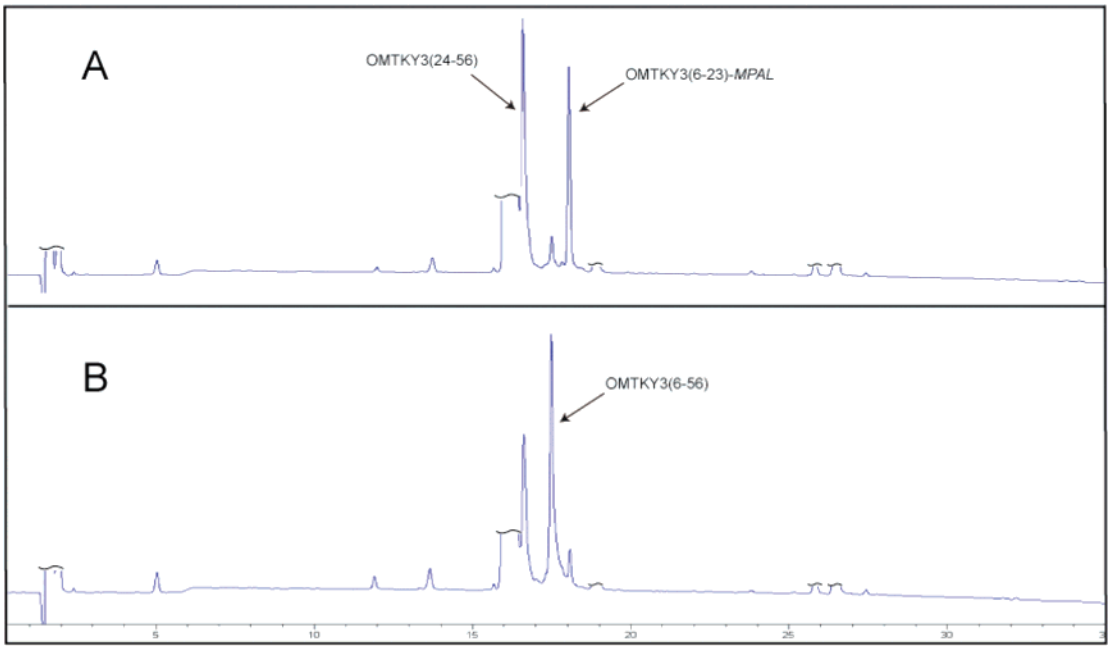


Figure 7. Ligation of OMTKY3(6−23)-$^{\alpha}$CO*CH$_2$CH$_2$COLeu* and (Cys^{24}-56)OMTKY3. (A) Ligation after 15 min. (B) Ligation after 2 h. Ligation was perfomed with a slight excess of OMTKY3(Cys^{24}-56), and was essentially complete at 2 h. Reversed-phase HPLC chromatography was performed using a linear gradient of buffer B in buffer A where B = acetonitrile + 0.08% TFA, and A = water + 0.1% TFA. Detection was at 214 nm. Nonpeptidic peaks corresponding to the thiol catalyst are truncated for clarity: if desired, the nonpeptidic thiol catalyst-related components can be removed by use of a short column to which they do not adsorb, or by isocratic elution of these components prior to initiating the gradient.

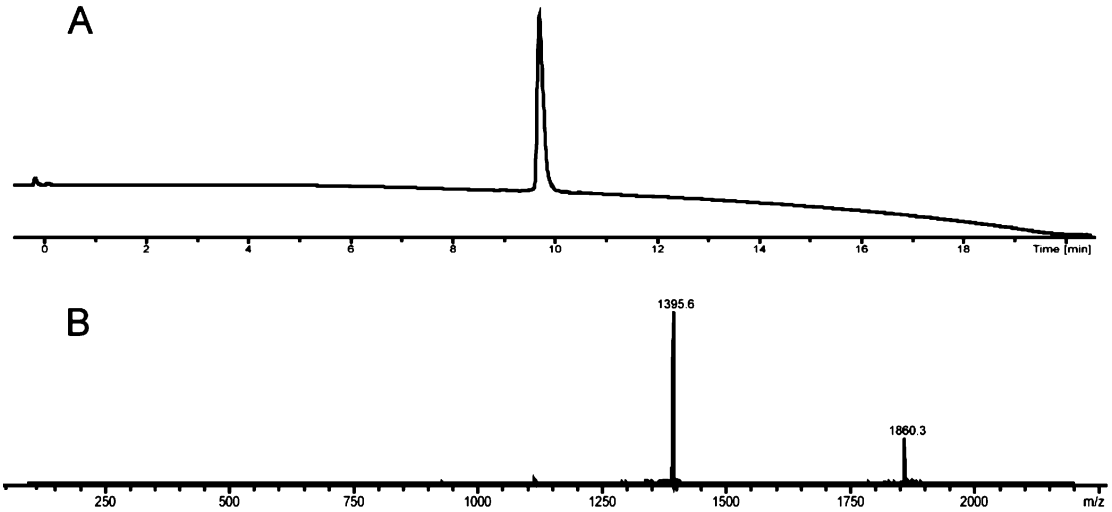


Figure 8. Ligation Product OMTKY3(6−56). (A) HPLC of purified OMTKY3(6−56), using a different linear gradient of buffer B in buffer A where B = acetonitrile + 0.08% TFA, and A = water + 0.1% TFA. (B) Eletrospray MS of the entire peak in (A) (Calculated mass (average isotopes) 5579.2 Da; Observed mass 5578.2 ± 0.6 Da).

accelerate this ligation, with time points taken at 30 min, 1, 2, 4, 12, 16, 24, and 36 h. The results are shown in Table 2.

Ligation in the presence of MPAA showed an order of magnitude (i.e., ∼10-fold) rate increase over the standard catalyst mix used for native chemical ligation, 1% benzyl mercaptan/3% thiophenol, and was essentially complete within 2 h using 200 mM MPAA. Addition of more nucleophilic thiols, such as MESNA or benzyl mercaptan, to MPAA substantially slowed the reaction, presumably by competing with MPAA for thiol-thioester exchange (data not shown).

HPLC chromatograms of the ligation using 200 mM MPAA are shown in Figure 7. The ligation proceeded cleanly, and the ligated product was purified by preparative HPLC and characterized by LC−MS as the full-length OMTKY3(6−56) (Figure 8).

Discussion

Our studies of the effects of a range of different thiol additives on the native chemical ligation reaction have given new insights into details of the mechanism and how it affects the practical use of this reaction. We expected that such understanding would allow us to improve upon the standard protocol(s) currently used for native chemical ligation reactions, to yield faster and more complete reactions. This, in turn, would lead to more facile synthetic access to proteins.

The data reported here support the understanding of the native chemical ligation reaction mechanism previously put forward,[7] which is shown in modified form in Figure 1. The reactant peptide-$^{\alpha}$COS*CH$_2$CH$_2$COLeu* alkyl thioester is converted in situ to a more reactive peptide-thioester by addition of a thiol catalyst

to the reaction mixture; this more reactive peptide-thioester then undergoes a transthioesterification reaction with the side chain of the N-terminal Cys in a Cys-peptide reactant, to give a thioester-linked initial product; the thioester linked intermediate undergoes a spontaneous and irreversible intramolecular rearrangement to give the final amide-linked peptide ligation product. Because the unrearranged species has never been observed, rearrangement of the thioester-linked intermediate is believed to be extremely rapid, and to never become rate-limiting. (This should be contrasted with the use of 1-amino,2-thiol auxiliaries in so-called extended native chemical ligation where the rearrangment of the thioester-linked intermediate is often rate-limiting. See for example: Canne, L. E.; Bark, S. J.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 5891−5896; Kawakami, T.; Akaji, K.; Aimoto, S. *Org. Lett.* **2001**, *3*, 1403−1405; Offer, J.; Boddy, C. N.; Dawson, P. E. *J. Am. Chem. Soc.* **2002**, *124*, 4642−4646; Macmillan, D.; Anderson, D. W. *Org. Lett.* **2004**, *6*, 4659−4662.)

This mechanism is consistent with all of the observations in our current work. Testing a number of different thiol compounds for their ability to catalyze native chemical ligation, we observed that for alkanethiols such as MESNA and benzyl mercaptan, thiol-thioester exchange was not rate limiting: these alkanethiols underwent rapid exchange with the peptide-$^{\alpha}$COS*CH*$_2$CH$_2$-COLeu thioester (Figure 5), but because alkanethiols are poor leaving groups the transthioesterifcation by the N-terminal cysteine (Step 2, Figure **1)** becomes rate limiting. In contrast, aryl thiols are good thioester leaving groups in the reaction with the side chain thiol of an N-terminal cysteine, and at low concentrations of added thiol the initial thiol-thioester exchange was the rate-limiting step. The comparative leaving group behavior in the transthioesterifcation reaction with the side chain thiol of an N-terminal cysteine can be inferred from the rates of ligation in the absence of added thiol catalyst: preformed peptide-COS-aryl thioesters react ∼100-fold faster than peptide-$^{\alpha}$COS*CH*$_2$CH$_2$COLeu, and ∼50-fold faster than peptide-COS-MESNA (Figure 6).

An optimal catalyst for native chemical ligation would therefore be a thiol that could undergo rapid and complete exchange with the reactant peptide-$^{\alpha}$COS*CH*$_2$CH$_2$COLeu alkyl thioester while maintaining high reactivity toward cysteine transthioesterification. Thiophenol meets these criteria to some extent, and has traditionally been used to catalyze native chemical ligation reactions.[7] Drawbacks of thiophenol include its poor solubility in aqueous ligation buffers[13] and its strong and offensive odor. Poor solubility limits the concentration of thiophenol available in solution for thiol-thioester exchange, which as a consequence is slow and incomplete, thereby limiting the catalytic utility of this thiol. In contrast, the aryl thiol (4-carboxymethyl)thiophenol (also named 4-mercaptophenylacetic acid, MPAA) meets the criteria for being a near-optimal catalyst for native chemical ligation. MPAA is highly water-soluble and undergoes thiol-thioester exchange more rapidly and more completely than thiophenol (Figure 5); at high concentrations of MPAA, thiol-thioester exchange is rapid and complete. MPAA is also an excellent leaving group for reaction with the side chain thiol of an N-terminal cysteine, as shown by rapid ligation of the preformed peptide-MPAA thioester (Figure 6).

The practical advantages of MPAA as a catalyst for native chemical ligation reactions are apparent in the studies reported here. Using 250 mM MPAA, ligation of a Cys-peptide with a model peptide containing a Leu-$^{\alpha}$COS*CH*$_2$CH$_2$COLeu alkyl thioester was essentially complete within 1 h (Figure 4), a rate that approached that observed with the preformed MPAA thioester (Inset, Figure 6). Using 200 mM MPAA as catalyst, ligation of a similar Leu-thioester in the longer OMTKY3(6−23)- α*COSCH*$_2$CH$_2$COLeu peptide required only 2 h to approach completion, in contrast to the 16−24 h using traditional thiol catalysts (Table 2). This represents a 10-fold rate enhancement over the use of thiophenol as a catalyst for this ligation. Additional advantages of MPAA are that, in contrast to thiophenol, it is a nonmalodorous, easily handled solid, which allows ligations to be performed outside a chemical fume hood.

We also investigated the use of a *preformed* peptide-MPAA aryl thioester. (Preformed peptide-thioarylesters have previously been used in native chemical ligation, with or without added thiol catalysts—see for example: Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776−779; von Eggelkraut-Gottanka, R.; Klose, A.; Beck-Sickinger, A. G.; Beyermann, M. *Tetrahedron Lett.* **2003**, 44, 3551−3554.) In this way, one could essentially avoid the rate limiting thiol-thioester exchange step during native chemical ligation as the peptide would, in essence, have already undergone complete exchange. Peptide-MPAA thioesters are readily prepared as isolated compounds by postsynthetic treatment of the conventional peptide-$^{\alpha}$COS*CH*$_2$*CH*$_2$*COLeu* thioesters with MPAA, using a procedure based on that recently reported for the preparation of peptide-thiophenyl esters.[14] Peptide-MPAA thioesters are more readily prepared as isolated compounds than peptide-thiophenyl thioesters by this method, because of the much higher solubility of MPAA over thiophenol. Peptide-MPAA thioesters are stable for prolonged periods in aqueous solution at room temperature and at neutral pH, and are not difficult to handle and store during postsynthesis workup. Alternatively, preformed peptide-MPAA aryl thioesters could be directly prepared by chemical synthesis. The 4-carboxymethyl group located on MPAA provides an obvious handle for synthesis of thiophenyl esters on resin using Boc chemistry solid phase peptide synthesis (Brad Pentelute, Duhee Bang, personal communication). Even with the use of preformed peptide-MPAA thioesters in native chemical ligation, it would probably still be necessary to use added MPAA during ligations with peptides that contain internal cysteines. The added thiol serves to reverse unproductive side reactions such as thiolactone formation, and to reverse transthioesterification with cysteines not located on the N-terminus of the Cys-peptide.

Conclusions

Alkanethiols such as MESNA are poor catalysts for the native chemical ligation reaction. By contrast, certain aryl thiols, such as (4-carboxymethyl)thiophenol (MPAA), are excellent catalysts. Using MPAA as a catalyst for native chemical ligation enables proteins to be synthesized rapidly—ligations are typically complete within less than an hour—and in high yields. We are

(13) Thiophenol is only sparingly soluble in water; in solubility tables it is listed as insoluble in water. It behaves as if soluble only to ∼20 mM (based on ∼10-fold rate difference between thiophenol and MPAA, in Table 2). The data in Figure 4 also support this—the ligation rate in the presence of 1% thiophenol (nominally ∼90 mM, but visibly not completely dissolved) is what would be expected for ∼20 mM MPAA.

(14) Bang, D.; Pentelute, B. L.; Kent, S. B. H. *Angew Chem.*, accepted for publication.

currently using the catalyst MPAA to extend total chemical synthesis to larger and more complex protein targets, such as the HIV-1 protease, RNase A, and erythropoietin. MPAA is such an effective catalyst for native chemical ligation that we expect to be able to routinely perform ligations at sites such as Ile-Cys, Val-Cys, and Thr-Cys that are prohibitively slow using thiophenol as a catalyst. We expect that MPAA will also prove to be a much more effective alternative to the alkanethiol MESNA for use in the semi-synthesis of proteins by expressed protein ligation.

Experimental Procedures

Materials. 2-(1*H*-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), *S*-trityl-mercaptopropionic acid, *p*-methylbenzhydrylamine (MBHA) resin, and Boc-amino acids were obtained from Peptides International, Kentucky. Side chain protecting groups used were: Arg(Tos), Cys(4MeBzl), Asp(cHex), Glu(cHex), His(Bom), Lys(2ClZ), Ser(Bzl), Thr(Bzl), Tyr(BrZ). *N*,*N*-diisopropylethylamine (DIEA) was obtained from Applied Biosystems, Foster City. *N*,*N*-Dimethylformamide (DMF), dichloromethane, methanol, diethyl ether, and HPLC-grade acetonitrile were purchased from Fischer. Trifluoroacetic acid (TFA) was obtained from Halocarbon Products, New Jersey. HF was purchased from Matheson. Tris(2-carboxyethyl)-phosphine hydrochloride (TCEP), *p*-cresol, and all thiols were purchased from Sigma-Aldrich, except benzyl mercaptan, which was purchased from Acros.

Peptide Synthesis. Peptide-$^{\alpha}COSCH_2CH_2$COLeu thioesters were synthesized on a $HSCH_2CH_2$COLeu-MBHA-resin, similar to that described previously.[6] Peptides were synthesized on a 0.2 mmol scale for model peptides and on a 0.4 mmol scale for the OMTKY3 peptides, using the manual Boc chemistry in situ neutralization/HBTU protocol described previously,[15] with slight modifications. Briefly, N^{α}-Boc removal was achieved with neat TFA, using one rapid wash followed by a 2 min batch treatment. The N^{α}-deprotected peptide-resin was subject to a single flow wash with DMF. Amino acids (2.2 mmol) were dissolved in 4 mL of 0.5 M HBTU (i.e., 2.0 mmol) and were activated by addition of DIEA (1 mL, 6.6 millimole) for 1 min, and then added to the peptide-resin; coupling was carried out for 12 min, followed by a single DMF flow wash to remove excess activated amino acid and soluble coproducts. After chain assembly was complete, the N^{α}-Boc group was removed as described above, and the peptide-resin was washed with dichloromethane and dried by aspiration. The peptide was cleaved from the resin and the side-chain protecting groups were simultaneously removed by treatment with anhydrous HF at 0 °C for 1 h, with 10% *p*-cresol added as a scavenger. After removal of HF by evaporation, the peptide was precipitated and washed with cold diethyl ether and dissolved in aqueous acetonitrile + 0.1% TFA, then lyophilized.

LC−MS Analysis. Analytical reversed-phase HPLC was performed on an Agilent 1100 system with either Microsorb C-18 or C-4 (5 μm 2.1 × 50 mm) silica columns packed in-house. Peptide masses were obtained using on-line electrospray MS detection. Semipreparative HPLC was performed on a Vydac C-18 (10 × 250 mm) column, except for purification of OMTKY3(6−56), which was performed on a Vydac C-8 (10 × 250 mm) column. Peptides were eluted from the column using an appropriate shallow gradient of acetonitrile/0.08% TFA versus water/0.1% TFA. Fractions containing the desired purified peptide product were identified by analytical LC-MS, then combined and lyophilized.

Preliminary Evaluation of Thiols as Catalysts for Native Chemical Ligation. Each thiol compound was separately dissolved to 20 mM in ligation buffer (6 M GdmHCl, 0.2 M phosphate, 40 mM TCEP) at pH 7.5, with heat and sonication if necessary. A 25-μL portion of a stock solution of Cys-Leu-Tyr-Leu-Ala-Ala (8 mM in 6M GdmHCl, 0.2M phosphate, pH 6.2) was added to 25 μL Leu-Tyr-Arg-Ala-Leu-$^{\alpha}COSCH_2CH_2$COLeu stock solution (4 mM in 6M GdmHCl, 0.2M phosphate, pH 6.7), and the ligation started by immediately adding 50 μL of ligation buffer containing the thiol catalyst. Final reaction conditions were as follows: 1 mM Leu-Tyr-Arg-Ala-Leu-$^{\alpha}COSCH_2CH_2$COLeu, 2 mM Cys-Leu-Tyr-Leu-Ala-Ala, 10 mM thiol, 20 mM TCEP, 6 M GdmHCl, 0.2 M phosphate, pH 7.0, 25 °C. For experiments with increased thiol concentrations, the final concentration of thiol was achieved as described above by diluting the concentrated thiol ligation buffer 1:1 with the peptide stock mix. For each reaction, at each time point 5 μL aliquots were withdrawn and the reaction quenched by adding 5 μL of 1:1 acetonitrile:H_2O + 1% TFA, and the solution was stored at −20 °C until analyzed. The extent of ligation was followed by analytical reversed-phase HPLC and quantified by manual integration of peaks detected at 214 nm.

Thiol-Thioester Exchange. Thiol-thioester exchange was assayed by dissolving the thiol to 20 mM in ligation buffer, pH 7.2. 50 μL of a solution of Leu-Tyr-Arg-Ala-Leu-$^{\alpha}COSCH_2CH_2$COLeu (2 mM in 6M GdmHCl, 0.2M phosphate, pH 6.7) was added to 50 μL thiol solution. Final reaction conditions were 1 mM Leu-Tyr-Arg-Ala-Leu-$^{\alpha}COSCH_2CH_2$COLeu, 10 mM thiol, 20 mM TCEP, 6 M GdmHCl, 0.2 M phosphate, pH 7.0, 25 °C. Aliquots were withdrawn at each time point, and the reaction quenched and analyzed as described above.

Ligation with Preformed Thioesters. Preformed MESNA and MPAA thioesters were obtained by dissolving 8 mg Leu-Tyr-Arg-Ala-Leu-$^{\alpha}COSCH_2CH_2$COLeu in 1 mL ligation buffer (9.6 mM, no TCEP) containing either 600 mM MPAA or 750 mM MESNA, and adjusting the solution to pH 7.0 and allowing the reaction to proceed for 1 h. Preformed thiophenyl thioester was obtained by dissolving 16 mg Leu-Tyr-Arg-Ala-Leu-$^{\alpha}COSCH_2CH_2$COLeu in 16 mL ligation buffer (no TCEP), adding 2% v/v thiophenol, and adjusting the suspension to pH 7.0. The reaction was stopped after 3 h and worked up, even though incomplete. Exchanged thioester peptides were purified by semipreparative reversed-phase HPLC after each reaction had gone to completion. Ligation kinetics with preformed thioesters were assayed using 1 mM Leu-Tyr-Arg-Ala-Leu-thioester, 2 mM Cys-Leu-Tyr-Leu-Ala-Ala, 20 mM TCEP, 6 M GdmHCl, 0.2 M phosphate, pH 7.0, 25 °C as described above.

OMTKY3 Ligations. Ligations were performed with 1 mM of OMTKY3(6−23)- $\alpha COSCH_2CH_2$COLeu and (Cys^{24}-56)OMTKY3 each, in ligation buffer, pH 7.0, 25 °C. Aliquots were taken at each time point as described previously, and the progress of the ligation assessed by analytical reversed-phase HPLC on a Zorbax C8 (4.6 × 150 mm) column. After the ligation using MPAA as catalyst had gone to completion, OMTKY3(6−56) was purified by semipreparative reverse-phase HPLC as described above, and characterized by LC−MS.

Acknowledgment. We thank Thomas Durek for providing the peptide Cys-Leu-Tyr-Leu-Ala-Ala. We gratefully acknowledge support from the Department of Energy Genomes to Life Genomics Program (Grant DE-FG02-04ER63786) and the MD/PhD Graduate Training in Growth and Development program at the University of Chicago (NIH T32 HD007009).

JA058344I

(15) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. *Int. J. Pept. Protein Res.* **1992**, *40*, 180−193.