Pergamon

Tetrahedron 56 (2000) 3449–3491

TETRAHEDRON

Tetrahedron Report Number 527

Peptide Segment Coupling by Prior Ligation and Proximity-Induced Intramolecular Acyl Transfer

Don M. Coltart*

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

Received 15 July 1999

Contents

* Tel.: +1-780-492-3428; fax: +1-780-492-8231.

PII: S0040-4020(00)00147-2

1. Introduction

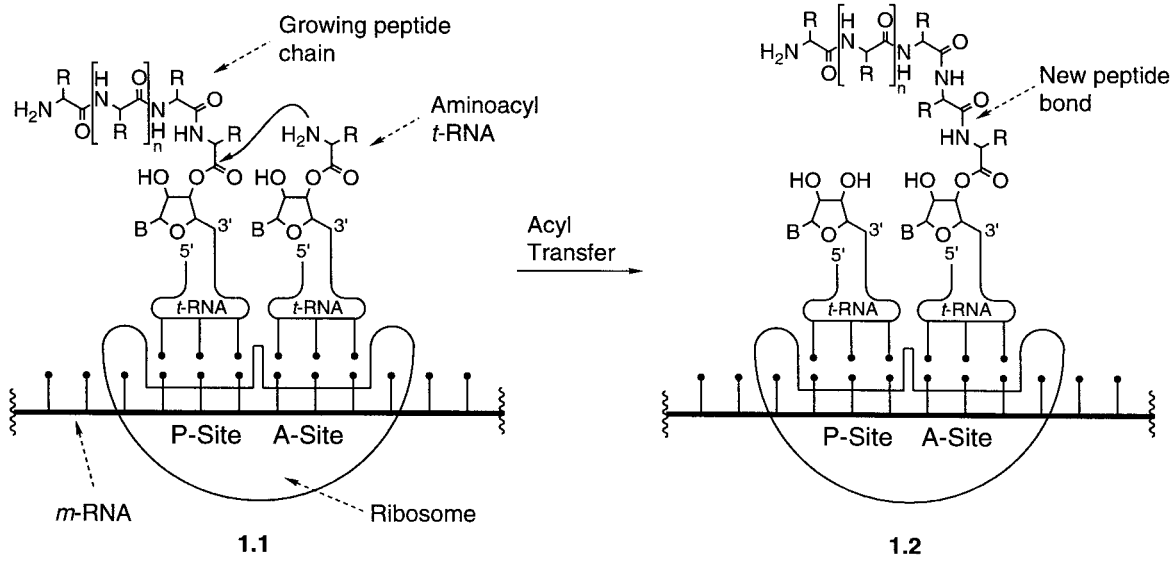
Central to the understanding and use of peptides is the extent to which one can manipulate the primary sequence. Complete control—through synthesis—of the primary sequence of a peptide of up to 50 amino acid residues is possible using chemically based techniques.[1] For larger peptides, however, chemical synthesis is a much more arduous task, and in these cases site-directed mutagenesis[2] is the approach normally used to modify the primary sequence. This technique is not without its shortcomings since, with few exceptions, only naturally occurring amino acids can be incorporated into the peptide chain. To some extent this restriction can be avoided by biosynthetic site-specific incorporation[3] of unnatural amino acids, but this and site-directed mutagenesis are typically used to alter only the sequence of an existing peptide, and so do not offer ready access to peptides that have no known, naturally occurring counterpart. Therefore, the development of purely chemical synthetic procedures would provide a valuable complementary method for controlling the primary sequence of peptides. An approach based on chemical synthesis would offer an opportunity to incorporate directly non-genetically coded amino acids, as well as structural and functional modifications of the peptide backbone and/or its side-chains, and would allow linking of discrete protein domains.

Of fundamental importance to the synthesis of peptides by chemical methods is the formation of peptide bonds between the constituent amino acid residues. Peptide bond formation was reported in 1881 by Curtius,[4] and later (1902) by Fischer.[5] The process typically relies on enthalpic activation of the α-carboxyl group of one amino acid, which then undergoes nucleophilic attack by the α-amino group of a second amino acid. In order to avoid side reactions that might result from the presence of electrophilic and nucleophilic centers in the amino acid side-chains, protecting groups are usually employed. The application of these basic principles to peptide synthesis in the solid phase was reported by Merrifield in 1963.[6] His method has undergone many improvements,[7–10] and it is currently possible to carry out routine stepwise solid phase synthesis of peptides of up to about 50[11] amino acids. The practicality of solid phase synthesis of peptides greater than about 50 residues in length is marred by the accumulation of byproducts that result during each cycle of the synthetic sequence; these byproducts eventually increase to a level that renders purification of substantial amounts of the desired product difficult.[7,8]

Chemical synthesis of large peptides by intermolecular coupling of smaller, pre-assembled peptides, using conventional peptide bond-forming techniques has been studied.[12] Conceptually, an approach of this type would combine the power of solid phase peptide synthesis to generate moderately sized segments with a technique that allows rapid assembly of the segments into a large peptide. On basis of the ease of discriminating between the relatively small starting materials and the products, it should also be easier to purify and characterize intermediates after each peptide bond-forming step. In practice, however, these potential advantages are not fully realized, and such intermolecular segment couplings have proven unsatisfactory.[8]

The above approach relies ultimately on coupling between the free α-amino group of one peptide segment and the activated α-carboxyl of another. Recently, considerable progress has been made by use of a less conventional procedure for generating a peptide bond between two segments. The approach relies on highly specific and efficient ligation of two peptide segments prior to formation of the peptide bond that will link them in the desired way. Once the two subunits have been ligated, the carboxyl terminus of one and the amino terminus of the other are close together, so that a highly efficient, proximity-driven intramolecular acyl transfer can occur to form a new peptide bond.

Advances in the area of peptide segment coupling that utilize this novel approach have been the result mainly of the independent research carried out by the groups of Kemp, Kent, and Tam. Although their approaches share the



Scheme 1. B=nucleotide base.

common features of prior ligation and intramolecular acyl transfer, they each employ a different method to incorporate these features.

The present review describes these approaches to peptide segment coupling, and is limited to those methods that involve *both prior ligation and intramolecular acyl transfer*.

2. Significance of Prior Ligation and Proximity-Induced Intramolecular Acyl Transfer

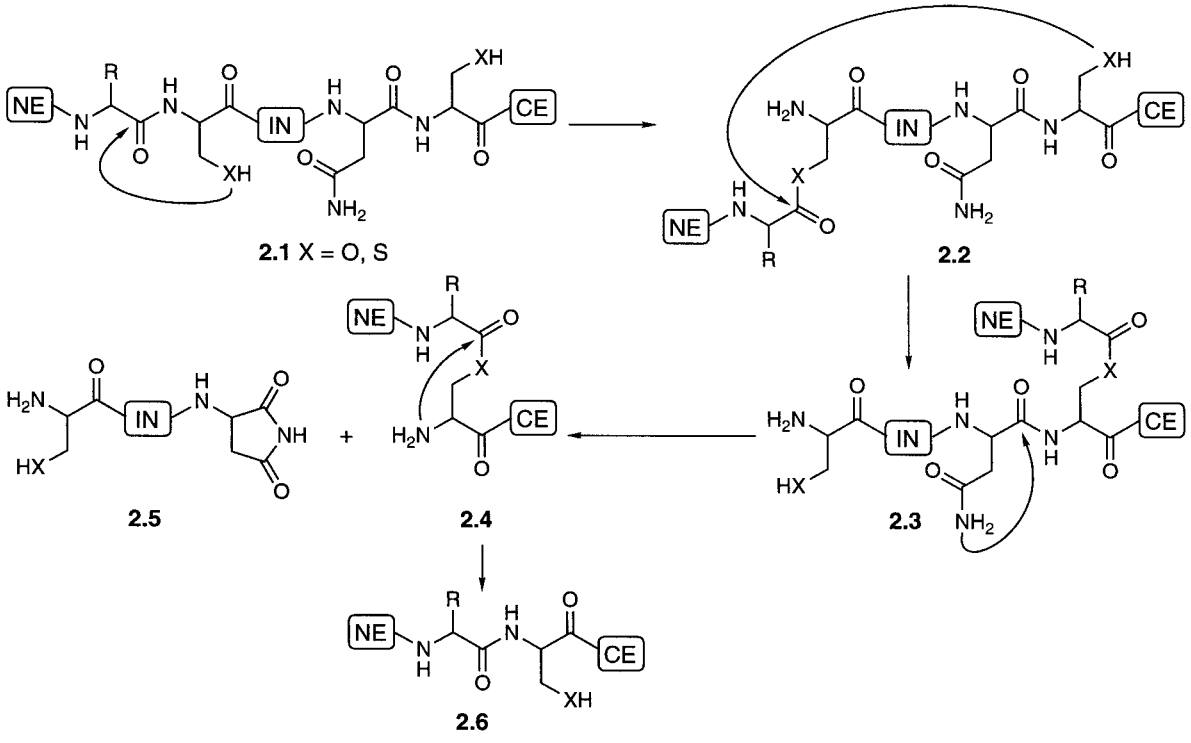
It has long been recognized that intramolecular reactions often occur much more readily than corresponding intermolecular processes.[13] The exact source of the rate acceleration is controversial,[14] but it appears to result largely from the fact that the reacting functional groups are held close together, thereby increasing their effective local concentration.

Peptide bond formation by intramolecular acyl transfer was demonstrated by Wieland in the early 1950s,[15] and shortly afterwards by Brenner.[16] The concept of entropic activation, as well as its significance as it pertains to peptide bond formation by intramolecular acyl transfer, was first formally described by Brenner in a publication[17] that adumbrated much of the future developments in the area of peptide segment coupling via intramolecular acyl transfer, and briefly mentioned the prior ligation aspect of this approach.

Nature takes advantage of entropic proximity effects in the formation of peptide bonds during protein synthesis. In vivo peptide bond formation occurs when the *C*-terminal acyl group of a growing peptide chain is transferred to the α-amino group of an aminoacyl *t*-RNA (**1.1→1.2**, Scheme 1). Prior to the transfer the *C*-terminal acyl group and the α-amino group of the aminoacyl *t*-RNA are held, through a complex series of associations, in such a way that they are in proximity. Thus, acyl transfer can proceed despite the fact that the acyl group is in the form of an ordinary ester (linked to a *t*-RNA molecule) and, therefore, has no significant enthalpic activation. This description is of course an oversimplification of the actual cellular processes involved.

It is reasonable to assume that proximity effects also play a role during the in vivo post-translational splicing[18] of proteins. In this type of process an intervening peptide segment (the intein) is excised while its flanking regions (the exteins) are coupled by formation of a new peptide bond.[18] The sequence is initiated (Scheme 2) by an intramolecular *N→O* or *N→S* acyl transfer (**2.1→2.2**) which gives an intermediate (**2.2**) that then undergoes a second intramolecular acyl transfer (**2.2→2.3**) and cleavage (**2.3→2.4**), producing a transient intermediate (**2.4**). This rearranges either by intramolecular *O→N* or *S→N* acyl transfer (**2.4→2.6**) to give the coupled product (**2.6**).

Proximity effects have also been exploited in vitro[19] for protein semisynthesis and, in this approach, two peptide segments derived from a protein (with or without subsequent modification) are associated (either covalently[19e] or non-covalently) in such a way that the reactive *C*-terminal α-carboxyl of one segment and *N*-terminal α-amino group of the other are close together, thus facilitating acyl transfer. Self-association can occur due to the propensity of large peptide fragments to form complexes approximating the native conformation of the original protein but, given that the degree of complementarity necessary for self-association is quite specific, this approach to segment coupling is



Scheme 2. NE=*N*-terminal extein; CE=*C*-terminal extein; IN=intein.

unlikely to be generally useful. However, as an outcome of the intein/extein studies summarized in Scheme 2, a new type of in vitro semisynthesis is being developed which may prove to be more general.[20] In this approach, the essential *N*-terminal portion of a particular intein (having an extein at its *N*-terminus) and the essential *C*-terminal portion of the same intein (having an extein attached to its *C*-terminus) are allowed to associate. This association results in an active splicing system, which ultimately links the exteins, forming a new peptide and excising the intein-like complex. The generality of this approach would arise from the ability to attach various exteins to the truncated intein segments.

In terms of developing a generalized method for coupling peptide segments by taking advantage of intramolecular acyl transfer processes, the ligation-based strategy mentioned above is extremely important. This approach offers several advantages[21] over conventional techniques for peptide segment coupling, as these have a number of characteristics that cause problems when applied to the coupling of large peptides.[12e,22,23] Firstly, given the large size of the peptide segments, an entropic barrier exists that reduces coupling efficiency. This comes about because the terminal α-amino group of the amine component and the activated terminal acyl group of the acyl component do not have a high enough reactivity to make the intermolecular reaction sufficiently rapid at the high dilutions necessary to solubilize large peptide fragments; competing intramolecular side reactions then occur. Secondly, protection of the side-chain functional groups is necessary in order to prevent undesired reactions between the activated acyl group and peptide nucleophiles other than the intended amino group. This situation requires that a large number of protecting groups be removed in the final step of the synthesis—an operation which also leads to problems of product purification. Moreover, the required activation of the carboxyl terminus of the *N*-terminal peptide renders coupling of the segments susceptible to epimerization at the center adjacent to the activated acyl group.[7b] Finally, in order to prevent solvolysis of the activated terminal acyl group during the coupling process, dipolar aprotic solvents are required. Unfortunately, peptides are often poorly soluble in these solvents and tend to associate significantly in them—a factor which further inhibits the coupling process.

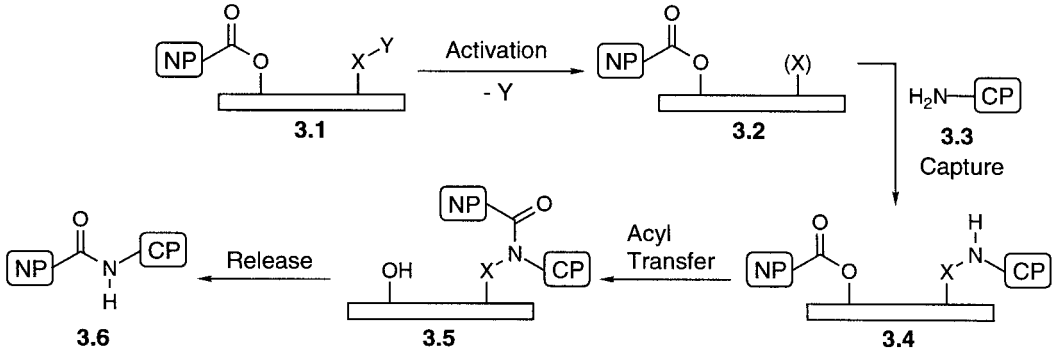
The advantages of a ligation-based approach can be outlined as follows. Since the relatively weak mutual reactivity of the terminal acyl carbon of one fragment and the terminal amino nitrogen of the other cannot be relied on to join the peptide chains, a separate and efficient ligation reaction is used to link the fragments and bring the two coupling sites into proximity. Once the fragments have been ligated, formation of the peptide bond will follow first-order kinetics and, consequently, factors such as steric hindrance by a bulky α-substituent on either of the amino acid residues involved in the coupling should exert less significant retarding effects compared with corresponding intermolecular reactions. Additionally, placing the terminal amino and the terminal acyl groups close together results in a high local concentration of these species, and this imparts a strong entropic advantage to the system, thereby removing the need for enthalpic activation of the acyl group. Given that acyl transfer now involves a relatively unactivated acyl group, the likelihood of epimerization or reaction with nucleophiles other than the closely positioned terminal amine is decreased and, consequently, the requirement for protecting groups is diminished or even removed. It may be possible, therefore, to use unprotected peptide segments in aqueous solution, and this approach would most likely avoid solubility problems, for example, and facilitate purification.

3. Prior Amine Capture Strategies

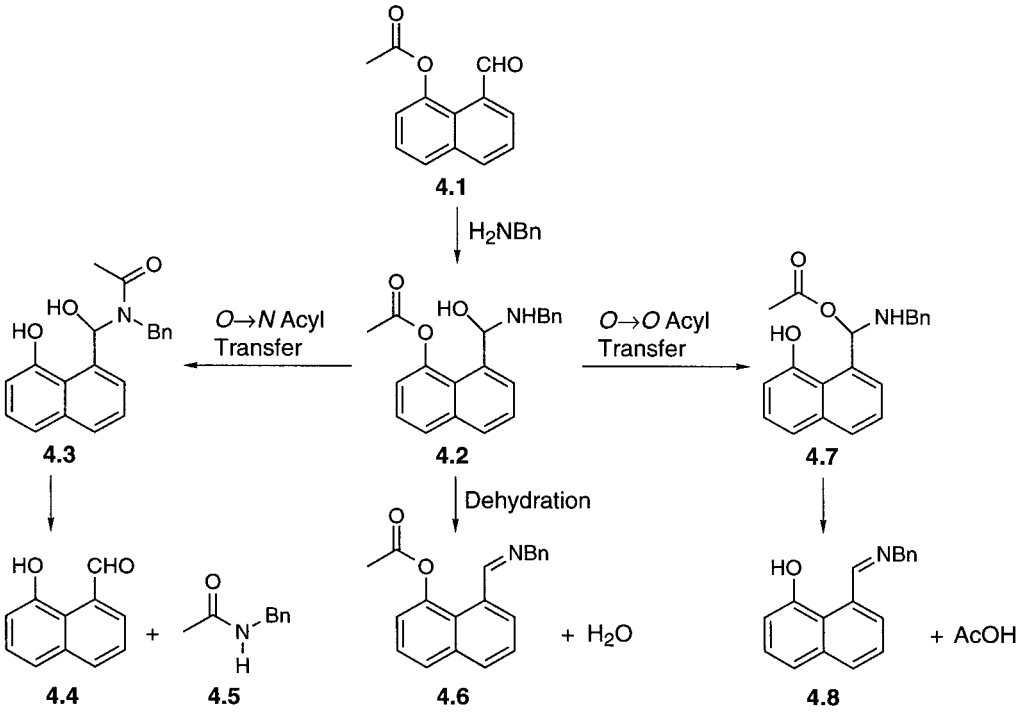
The first general approach to a ligation-based coupling technique is the *Prior Amine Capture Strategy* reported by Kemp in 1975.[21,24]

The principles of the Prior Amine Capture Strategy can be outlined in a general way by reference to Scheme 3. Structure **3.1** represents the *N*-terminal peptide chain that is ultimately to undergo coupling with the *C*-terminal chain, and is derivatized at its *C*-terminus in the form of a (weakly activated) ester. The alcohol portion of this ester is a template which, in the ligated product (cf. **3.4**), will serve to position the appropriate amino and acyl groups in proximity so that rapid intramolecular acyl transfer can occur. The group represented as X–Y in **3.1** is the protected form of one of the functional groups involved in the initial ligation and is referred to as the *capture site*. In this case X is the capture atom and Y is its protecting group. The capture site is capable of bonding to the terminal α-amino group of the *C*-terminal peptide (**3.3**).

The ligation is initiated by removal of the protecting group (**3.1**→**3.2**) so as to activate the capture site.



Scheme 3. NP=*N*-terminal peptide; CP=*C*-terminal peptide.



Scheme 4.

Capture (**3.2**+**3.3**→**3.4**) then occurs in the presence of the *C*-terminal segment to give the system represented as **3.4**. At this point, intramolecular acyl transfer takes place giving the rearranged system **3.5**. Finally, release of the template from **3.5** by cleavage of the X–N bond (**3.5**→**3.6**) results in formation of the coupled peptide **3.6**, which contains a new peptide bond linking what were originally two separate peptides.

In order to reduce the above plan to practice, Kemp recognized that several important criteria had to be met. First, a versatile, clean, and reliable method of attaching a template (containing a masked capture group) to the *C*-terminus of the *N*-terminal peptide (cf. **3.1**) would have to be found and, secondly, the template itself would have to meet certain criteria. For instance, the template would need to withstand all reaction conditions met during the coupling process but still be removable without damage to the newly formed peptide. The template must, of course, also correctly position the acyl and amino components close together so as to facilitate intramolecular acyl transfer.

Another prerequisite for implementing the Prior Amine Capture Strategy is that a suitable capture step be developed. This step would have the crucial role of bringing the peptide segments together before peptide bond formation; it would have to occur rapidly and efficiently, and also in a solvent that inhibits self-association of peptides—a phenomenon that can, for example, decrease solubility and/or restrict access to reaction sites.

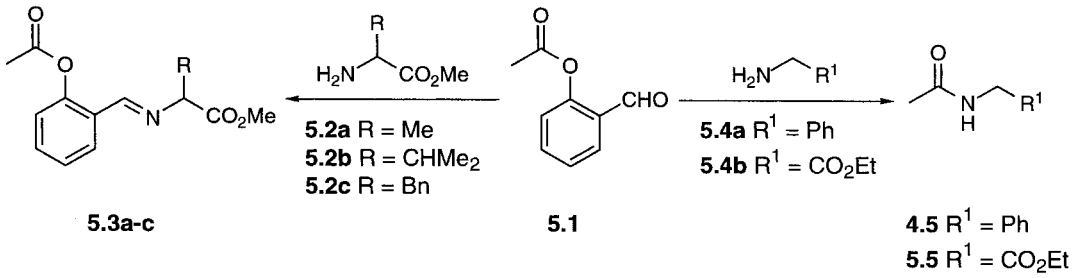
With the above objectives in mind, Kemp examined two general approaches for amine capture. In one, a carbonyl group is the electrophilic capture site (see Sections 3.1 and 3.2) and, in the other, a nitro olefin is used (see Section 3.3).

3.1. Ligation by hemiaminal formation

Initial studies of the Prior Amine Capture Strategy involved examination of 8-hydroxy-1-naphthaldehyde (**4.4**, Scheme 4) as a potential candidate both for promoting amine capture and for serving as a template to allow acyl transfer.[25] It was thought that initially a hemiaminal would be formed and that this species would act as an intermediate through which intramolecular acyl transfer could occur. In order to test the idea, aldehyde **4.1** was allowed to react with benzylamine in DMSO; *N*-benzylacetamide (**4.5**), the *O*→*N* acyl transfer product, was isolated in 70% yield, along with acetic acid (30% yield). This result suggested that the hemiaminal did indeed form (**4.1**→**4.2**), and that it then underwent both *O*→*N* acyl transfer (**4.2**→**4.3**) and *O*→*O* acyl transfer (**4.2**→**4.7**). The ratio of products was solvent dependent (Table 1), but no conditions were found that led exclusively to formation of **4.5**. In DMF the ratio of **4.5** to acetic acid was the same as in DMSO. When 1:1

Table 1. Effect of solvent on the product distribution for reaction of **4.1** with benzylamine[25]

Solvent	Relative product ratios		
	O→*N* Acyl transfer product (**4.5**)	*O*→*O* Acyl transfer product (AcOH)	Dehydration product (**4.6**)
DMSO	7	3	0
DMF	7	3	0
DMSO–H_2O (1:1)	3	7	0
MeCN–H_2O (1:4)	1	9	0
MeCN	1	0	1
Benzene	0	7	3
CCl_4	0	7	3
$CHCl_3$	0	1	4



Scheme 5.

Table 2. Product distribution when **5.1** reacts with different amines[25]

Amine	*O*→*N* Acyl transfer product (%)	Dehydration product (%)
$BnNH_2$	100	0
H_2N, CO_2Et	90	0
H_2N, CO_2Me	0	100
H_2N, CO_2Me	0	100
H_2N, CO_2Me, Bn	0	100

DMSO–water was used, a product ratio of 3:7 was observed, and with 1:4 acetonitrile–water, the product ratio was 1:9. Competing imine formation (**4.2→4.6**) was significant in certain other solvents. For example, use of acetonitrile resulted in formation of **4.5** and **4.6** in a 1:1 ratio. In some instances no product corresponding to *O*→*N* acyl transfer was detected. This was the case when the reaction was carried out in benzene, carbon tetrachloride, or chloroform; **4.5** was not observed, whereas both acetic acid and **4.6** were, in ratios of 7:3, 7:3 and 1:4, respectively.

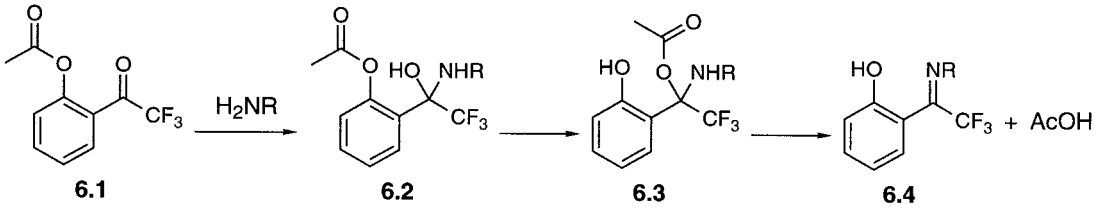
In this system (Scheme 4), the intramolecular acyl transfer proceeds through a seven-membered transition state, and rate constants for the combined *O*→*N* and *O*→*O* acyl transfer processes in acetonitrile, DMF and 1:4 acetonitrile–water were found to be 0.1, 0.2 and 15 $M^{-1} s^{-1}$, respectively. An estimation of the rate of the intermolecular reaction between benzylamine and the acetyl group of **4.1** was made by measuring the rate constant for reaction between benzylamine and 8-acetoxy-1-nitronaphthalene. The value was $1\times10^{-3} M^{-1} s^{-1}$ in acetonitrile. Since reaction of **4.1** with benzylamine (also in acetonitrile) had a rate constant two orders of magnitude greater (i.e. ca. $1\times10^{-1} M^{-1} s^{-1}$), the intramolecularity of the acyl transfer via **4.2** was implied, although not unambiguously proven.

The hemiaminal approach to ligation was further explored[25] by examining two other templates. The first was 2-acetoxybenzaldehyde (**5.1**, Scheme 5) which, when treated with benzylamine in any of the above solvents, gave **4.5** in quantitative yield, with rate constants close to those observed for **4.1** in DMF, acetonitrile, and 1:4 acetonitrile–water. The assumption that formation of **4.5** was, in fact, the result of an intramolecular process was based on the observation that the corresponding reaction using the isomeric aldehyde 4-acetoxybenzaldehyde proceeded about 150 times slower. Although the reaction with **5.1** was very efficient, experiments with other amines revealed that the product composition depended on the amine used (Table 2). For example, reaction of ethyl glycinate (**5.4b**) with 2-acetoxybenzaldehyde in benzene resulted in a 90% yield of the *O*→*N* acyl transfer product (**5.5**). However, the methyl esters of alanine (**5.2a**), valine (**5.2b**), and phenylalanine (**5.2c**) each gave exclusively the imine derivatives (**5.3a–c**).[26]

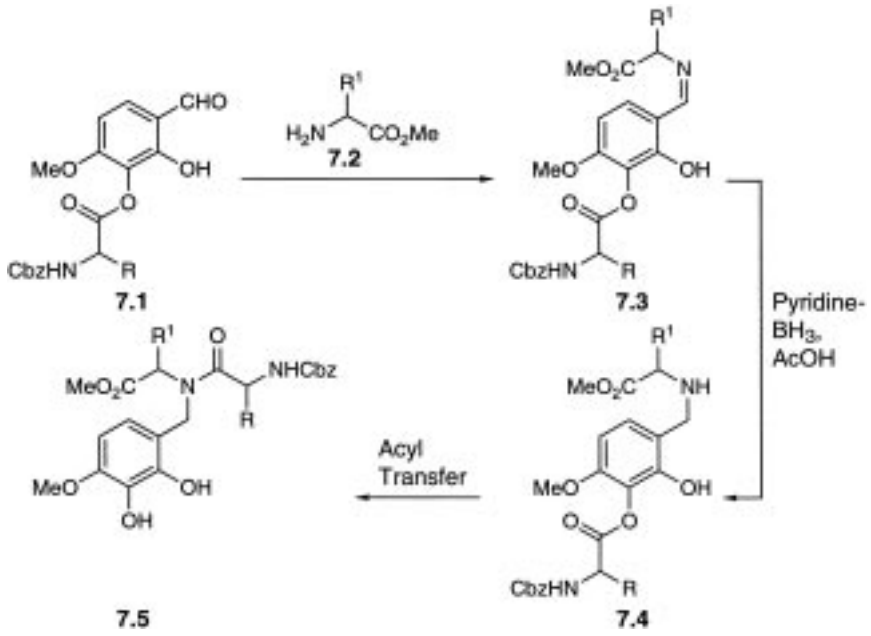
2-Acetoxytrifluoroacetophenone (**6.1**) was also examined (Scheme 6) as a template for acyl transfer, in the hope that the intermediate hemiaminal (cf. **6.2**) would not be prone to dehydration.[25] When **6.1** was treated with any of the above-mentioned amines, however, only products corresponding to *O*→*O* acyl migration (**6.1→6.2→6.3→6.4**) were observed and, in each case, acetic acid was isolated in quantitative yield. Evidently, direct dehydration of the intermediate hemiaminal was indeed avoided using **6.1**—but not with the intended result.

3.2. Ligation by imine formation

A slightly modified form of the hemiaminal approach—but



Scheme 6.



Scheme 7.

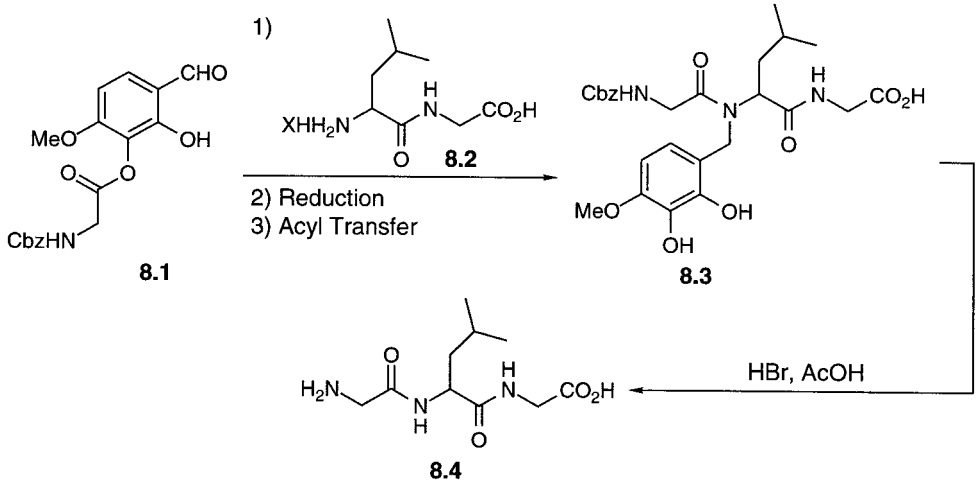
Table 3. Half-times for intramolecular acyl transfer reactions of **7.4**[27]

Acylating component (**7.1**)	Amine component (**7.2**)	Acyl transfer half-time (min)
R=H	R^1=Bn	70
H	CH_2CHMe_2	40
H	Me	15
Me	Me	70
Bn	Me	120

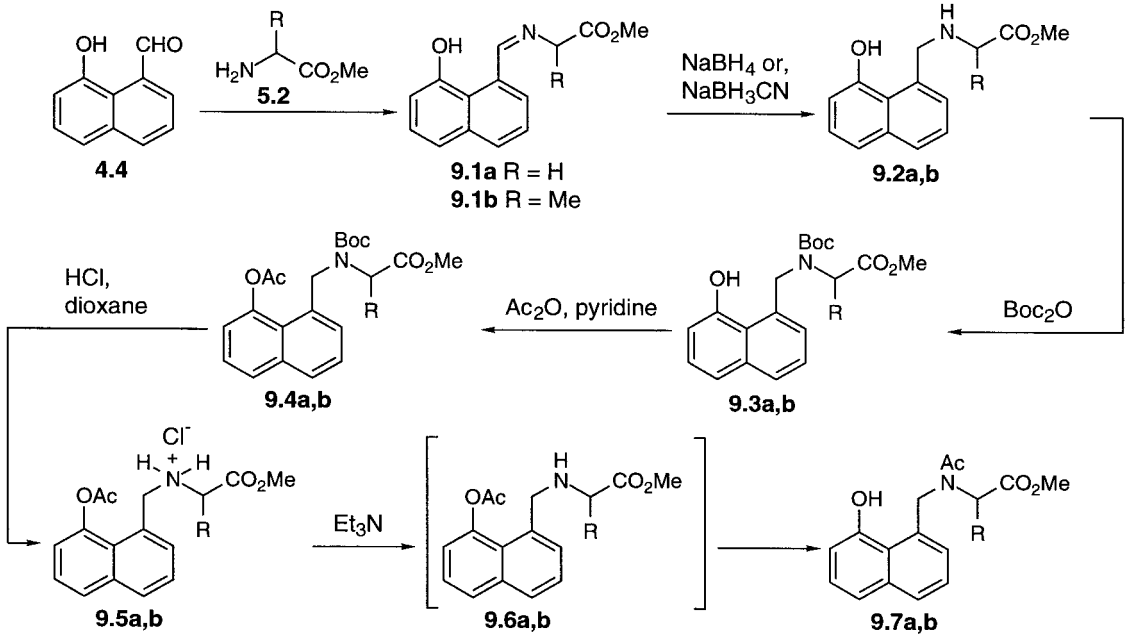
still involving nucleophilic addition of an amine to a carbonyl group—was also examined.[27] This version avoids dehydration as an unwanted side reaction by actually incorporating the loss of water into the capture process.

The idea was tested using **7.1** (Scheme 7) as the amine capture device.[27] Here, an imine is formed (**7.1**+**7.2**→**7.3**) by dehydration of the hemiaminal generated from the amine and the aldehyde. Direct and facile formation of the imine was expected, as such behavior is characteristic of salicylaldehydes.[28] Both R and R^1 were varied (R=H, R^1=Me, CH_2CHMe_2, Bn; R=Me, R^1=Me; R=Bn, R^1=$CHMe_2$) to give the corresponding imines; in all cases imine formation was efficient (**7.1**+**7.2**→**7.3**). The imines were then reduced (**7.3**→**7.4**), using pyridine–borane in acetic acid, to give a secondary amine that underwent intramolecular acyl transfer (**7.4**→**7.5**).[29] The transfer was slowed somewhat by the use of polar solvents such as DMSO, but proceeded readily in other solvents. When R=H, the half-times for acyl transfer to captured AlaOMe, LeuOMe, and PheOMe in deuterochloroform were 15, 40, and 70 min, respectively (Table 3).

Variation of the acyl fragment also influenced the rate; with R=Me or Bn half-times to captured AlaOMe were 70 and 120 min, respectively. Yields for the acyl transfer reaction



Scheme 8. XH=tetramethylguanidine.



Scheme 9.

were good. For example, when **7.1** (R=H) was treated with (±)-PheOEt, the corresponding acyl transfer product (cf. **7.5**) was obtained in 85% yield. The method could also be used to make a tripeptide.[27] Thus, reaction of **8.1** (=**7.1**, R=H) with LeuGly tetramethylguanidine salt (**8.2**, Scheme 8), followed by imine reduction and acyl transfer, gave **8.3** in 92% yield.

Cleavage of the 4-methoxy-2,3-dihydroxybenzyl moiety from the rearranged products was easily and efficiently achieved by treatment either with HBr–AcOH or with TFA in the presence of resorcinol.[27] For example, cleavage from the tripeptide derivative **8.3**, with concomitant removal of the amine protecting group, by HBr–AcOH, gave GlyLeuGly (**8.4**) in 84% yield.

Given that the acyl group in the procedures of Schemes 7 and 8 is only weakly activated as an *O*-aryl ester, epimerization at the adjacent chiral center was unlikely.[30] However, the stereochemical integrity with respect to the α-carbon of the captured amino acid, once it had been derivatized as an imine, was of concern, and so the extent, if any, of epimerization at this center was examined, using an isotopic dilution assay.[30] When **8.1**, which was ^{14}C-labeled at the glycine α-carbon, and L-PheGlyOEt were allowed to react in acetonitrile for periods of 1.5 and 14 h, the *racemic* tripeptide was obtained in 0.1 and 0.3% yield, respectively, after the complete sequence of imine formation, imine reduction, intramolecular acyl transfer, and cleavage with TFA. This result indicated that the extent of epimerization was very low.

The approach of Scheme 7 is a practical example of peptide bond formation using a ligation-based technique. However, although the amine capture step worked well, the low rates of acyl transfer showed that refinement of the template was still required.

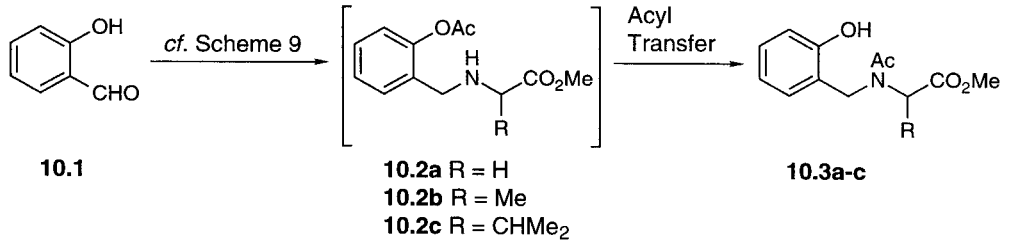
An efficient intramolecular acyl transfer is one in which a clean and rapid reaction that is not overly sensitive to steric factors is observed[31] and, in an effort to better satisfy these criteria, two other templates (see later, Schemes 9 and 10) were examined as candidates for facilitating acyl transfer. These templates differ from one another in the size and shape of the cyclic transition state through which the acyl transfer must occur.

Compounds **9.6a**,**b** (Scheme 9), derived from 8-formyl-1-naphthol, provided a second opportunity to examine acyl transfer through a seven-membered transition state.[32] The corresponding hydrochloride salts (**9.5a**,**b**) were derived from 8-formyl-1-naphthol (**4.4**) as follows. Treatment of the naphthol with either GlyOMe or AlaOMe gave an imine (**4.4**→**9.1a**,**b**), which was subsequently reduced and protected as its *N*-Boc derivative (**9.1a**,**b**→**9.2a**,**b**→**9.3a**,**b**). The phenolic hydroxyl was acylated with acetic anhydride in the presence of pyridine, and deprotection of the secondary amine, using HCl in dioxane, gave the hydrochloride salts of **9.6a**,**b** (**9.3a**,**b**→**9.4a**,**b**→**9.5a**,**b**). Treatment of each salt with triethylamine served to generate **9.6a**,**b** in situ, and the liberated amines underwent acyl transfer, giving amides **9.7a**,**b** (**9.5a**,**b**→**9.6a**,**b**→**9.7a**,**b**).

The acyl transfer reactions of **9.6a**,**b** were studied in a variety of solvents. Half-times for the glycine derivative

Table 4. Half-times for intramolecular acyl transfer reactions of **9.6a**,**b**[32]

Acyl transfer substrate	Solvent	Acyl transfer half-time
9.6a	MeCN	3.6 min
	$CHCl_3$	9 min
	DMF	50 min
	DMSO	36 min
9.6b	MeCN	3.3 h
	$CHCl_3$	3.5 h



Scheme 10.

Table 5. Half-times for intramolecular acyl transfer reactions of **10.2a–c**[32]

Acyl transfer substrate	Solvent	Acyl transfer half-time
10.2a	MeCN	36 s
	DMSO	6 min
	$CHCl_3$	7 min
	DMF	10 min
10.2b	H_2O (pH 10)	24 s
	MeCN	50 min
	$CHCl_3$	2.2 h
	DMSO	11.5 h
	DMF	12.5 h
10.2c	H_2O (pH 10)	3.3 min
	MeCN	4.5 h

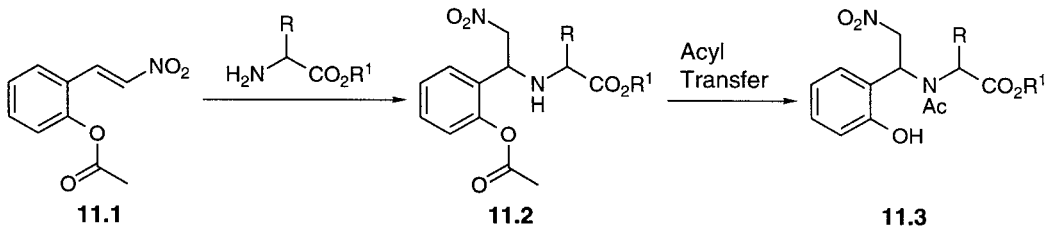
9.6a in acetonitrile, chloroform, DMF, and DMSO were 3.6, 9, 50 and 36 min, respectively (Table 4), but a much slower reaction was observed with the more sterically hindered alanine derivative **9.6b**, which exhibited half-times of 3.3 and 3.5 h in acetonitrile and chloroform, respectively. Hence, the rate of acyl transfer via a seven-membered transition state with the naphthalene template is sensitive to steric effects, as shown by the Ala/Gly half-time ratios of 55 and 23 in acetonitrile and chloroform, respectively.

The related acyl transfer via a six-membered transition state was also examined further by studying **10.2a–c** (Scheme 10),[32] obtained from 2-hydroxybenzaldehyde (**10.1**) in a manner analogous to the formation of **9.6a,b**. The rates of acyl transfer for **10.2a–c** were measured in several dipolar aprotic solvents and, for **10.2b** and **10.2c**, also in water (pH 10) (Table 5). In general, the acyl transfer reactions proceeded most readily in water, having half-times of 24 s for **10.2b** and 3.3 min for **10.2c**. When aprotic dipolar solvents were used, the highest rate of acyl transfer for each of the compounds was observed in acetonitrile, where half-times of 36 s, 50 min and 4.5 h were measured for **10.2a–c**, respectively. Acyl transfer was slowest in DMF in the case of **10.2a** and **10.2b** [acyl transfer for **10.2c** was measured only in water (pH 10) and acetonitrile]. In comparison to the seven-membered acyl transfer systems **9.6a,b**, the present templates (**10.2a–c**) exhibited consistently greater acyl transfer rates and, with the exception of the reaction in chloroform, where the rate for **10.2a** was only 1.3 times as fast as for **9.6a**, the rate of acyl transfer for **10.2a** was on average 5.7 times as fast as that for **9.6a** in the solvents examined. The alanine derivatives also showed a similar trend, with **10.2b** reacting 1.6 times as fast as **9.6b** in chloroform, and 4.0 times as fast in acetonitrile. Like its seven-membered counterpart, the acyl transfer system based on **10.1** was sensitive to the steric bulk of the captured amine, and the effect was even more pronounced in this case. In acetonitrile, the Ala/Gly half-time ratio was 83 and the Val/Gly half-time ratio was 450.

3.3. Ligation by Michael addition

As an alternative to the carbonyl-containing templates described above (Sections 3.1 and 3.2), the nitrostyrene derivative **11.1** (Scheme 11) was evaluated for its ability to act as an amine capture device, and also as a template to mediate intramolecular acyl transfer.[31] Amine capture in this case occurs by Michael addition of the primary amine function to the nitro olefin (**11.1→11.2**) and positions the acyl group and the amine nitrogen in such a way that intramolecular acyl transfer (**11.2→11.3**) can occur. It was expected that this amine capture system should have certain inherent advantages compared to the templates discussed above. First, unlike approaches involving a hemiaminal intermediate, problems of competing elimination or *O→O* acyl transfer cannot arise. The required Michael addition was also expected to be efficient, as reaction of β-nitrostyrene with primary amines is known to proceed rapidly and in nearly quantitative yield.[33] Secondly, the capture reaction gives a secondary amine directly, thus avoiding the reduction step required in systems where capture results in an imine.

Michael addition between **11.1** and the amino acid esters GlyOEt, AlaOMe, PheOMe, and ValOMe was easily carried



Scheme 11.

Table 6. Rates of acyl transfer in **11.2**[31]

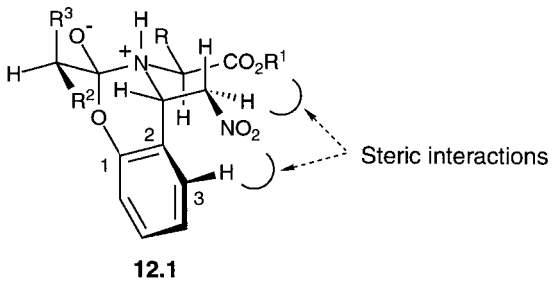
Acyl transfer substrate (**11.2**)	Acyl transfer rate (min^{-1})
R=H, R^1=Et	0.02
R=Me, R^1=Me	2×10^{-4}
R=Bn, R^1=Me	7×10^{-5}
R=$CHMe_2$, R^1=Me	2×10^{-5}

out in acetonitrile, and the adducts were then tested for their ability to undergo acyl transfer. Despite the fact that the required intramolecular transfers were occurring through a presumably favorable six-membered transition state, slow reactions were observed in all cases, and the rate constants are listed in Table 6. The data reveal a Gly/Ala rate ratio of 100, a Gly/Phe rate ratio of 286, and a Gly/Val rate ratio of 1000, suggesting a strong dependency of the rate on steric factors. Replacement of the acyl fragment (MeCO) with $CbzHNCH_2CO$ showed no significant changes in the rate of acyl transfer.

The steric effects observed in this study were rationalized by analogy to a previously established model that dealt with steric effects on aminolysis of peptide *p*-nitrophenyl esters.[34] Based on this model, the structure of the transition state leading to acyl transfer in compounds **11.2** was proposed to be **12.1** (Scheme 12), or a diastereomer. From the diagram, it can be seen that steric interactions exist between the nitromethylene group (CH_2NO_2) and either the ester (CO_2R^1) or α-alkyl substituent (R) of the amino acid, as well as between the nitromethylene group and H(3) of the aromatic template. Structure **12.1** is not capable of undergoing any stabilizing conformational changes that relieve these interactions without introducing new ones.

The above model suggests that sensitivity to steric factors should be expected in any derivative of **3.4** (Scheme 3) except where the capture site X is small and, especially, if it has no substituents (cf. the CH_2NO_2 group of **12.1**). For instance, minimal steric factors would be expected if X where sulfur, a methylene group, or an sp^2 atom. Consideration of model **12.1** seemed to emphasize the fact that unfavorable steric interactions can nullify entropically favorable ring sizes for the transition state. However, a way to avoid such steric interactions is to modify the link—and its attachments—that join the captured amine to the template.

The case where the link is a simple methylene group had been dealt with to some extent by examination of systems



12.1

Scheme 12. R^2=*N*-terminal peptide; R^1=*C*-terminal peptide; R, R^3=amino acid side-chains.

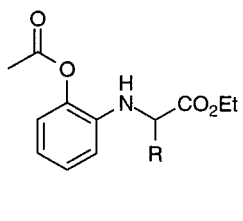
7.4, **9.6** and **10.2** (see above), and an attempt[31] was then made to test the effect of having X as an sp^2 atom. This was done[31] by examining acyl transfer rates for the ethyl esters of *N*-(2-acetoxyphenyl)glycine (**13.1a**), *N*-(2-acetoxyphenyl)alanine (**13.1b**), and *N*-(2-acetoxyphenyl)-valine (**13.1c**). Intramolecular *O*⟶*N* acyl transfer, which proceeded through a five-membered transition state in these compounds, was very slow. The poor nucleophilicity of the nitrogen atom (it is part of an aniline system), and the development of strain in the transition state, were regarded as factors responsible for the low rates. The Gly/Ala and Gly/Val rate ratios were comparable to those observed in corresponding intermolecular reactions.

Finally, the effect of using a sulfur atom as the link was studied briefly[31] by examination of **13.2** (Scheme 13), but even for the unhindered case shown, the rate of intramolecular acyl transfer was very low (3×10^{-4} min^{-1})—for reasons that were not identified.

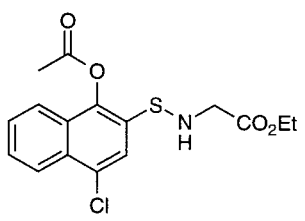
4. Prior Thiol Capture Strategies

As a result of the investigations described above, it had become apparent that a coupling strategy in which acyl transfer proceeds through a relatively small transition state of five to seven members is not effective, due to steric crowding. The suspicion had also developed that the amino group might not be nucleophilic enough to achieve rapid and clean capture at high dilution. On the matter of steric crowding, it appeared that in the ligation product the segment linking the amino and acyl components was either too inflexible, or failed to position them properly. Consequently, there were limited opportunities for avoiding unfavorable nonbonded interactions in reaching a suitable transition state for acyl transfer. Kemp reasoned that the best way to avoid these unfavorable interactions was simply to increase the length of the spacer arm while maintaining its rigidity. The problem of low nucleophilicity of the amine nitrogen, on the other hand, had to be dealt with by finding an entirely new way of ligating the peptide segments.

An extremely elegant approach to these apparently separate problems was developed[35] through a single modification of the initial strategy. The modification required that cysteine be the *N*-terminal residue of the *C*-terminal peptide chain. Of course, this imposes the limitation that the coupling site between the two peptides has to be between that cysteine and another amino acid. The approach, which is referred to as the *Prior Thiol Capture Strategy*, is outlined in Scheme

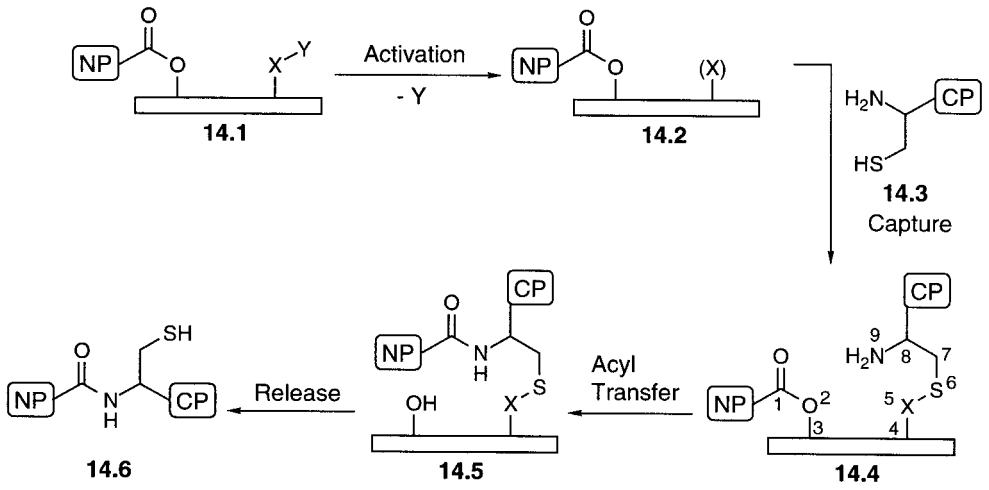


13.1a R = H
13.1b R = Me
13.1c R = $CHMe_2$



13.2

Scheme 13.



Scheme 14. NP=*N*-terminal peptide; CP=*C*-terminal peptide.

14.[35] In the first step, the masked capture site (X–Y) of the template is activated (**14.1→14.2**); it then undergoes reaction with the thiol function of the *N*-terminal cysteine to give the ligated product (**14.2**+**14.3→14.4**). Intramolecular acyl transfer (**14.4→14.5**) ensues, forming the rearranged system **14.5**, which contains a new peptide bond. Finally, cleavage of the template (**14.5→14.6**) liberates the newly formed peptide **14.6**.

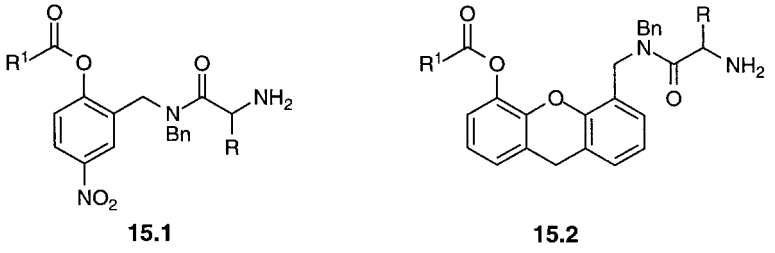
The potential of this strategy to overcome the limitations of the Prior Amine Capture Strategy can be explained as follows. The thiol function of the terminal cysteine is strongly nucleophilic and, in contrast to the amino group that is involved in the Prior Amine Capture Strategy, this difference should greatly increase the likelihood of effecting smooth ligation of the peptide segments prior to acyl transfer, even at high dilution. Moreover, the ensuing acyl transfer transition state would necessarily consist of at least nine atoms (cf. **14.4**) and, therefore, could provide ample conformational freedom to overcome unfavorable interactions resulting from steric crowding. Before trying the Prior Thiol Capture Strategy, however, two matters required attention. The first was to determine what type of ligation process would best take advantage of the reactivity of the thiol group, and the second was to establish if acyl transfer could proceed through a transition state as large as that necessitated by this approach.

The consequences of a larger acyl transfer transition state had already been evaluated to some extent during earlier work[32] in which intramolecular acyl transfer in **15.1** and **15.2** (Scheme 15) had been examined. The transfer in these systems proceeds via transition states of nine and twelve members, respectively. In each system, R and R^1 were varied extensively, and the rates of acyl transfer were measured in a variety of solvents. These studies showed that acyl transfer could, indeed, proceed efficiently through transition states involving medium to large rings, as indicated by Scheme 14.

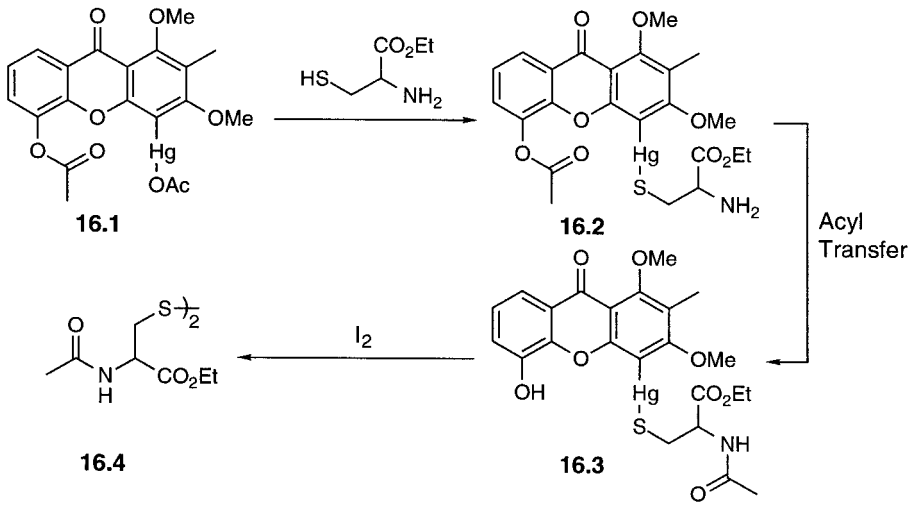
The question of what type of ligation process would best take advantage of the reactivity of the thiol group was examined by studying both mercaptide formation (Section 4.1) and disulfide formation (Section 4.2).

4.1. Ligation by mercaptide formation

Kemp's early work on ligation by thiol capture involved the use of an organomercury derivative.[36] Compounds of this type were chosen because they have a high affinity for thiol groups and react rapidly with them.[37] The C–Hg–S linkage that would be produced by ligation has a linear geometry and, in order to test acyl transfer via the resulting large ring that is necessarily formed, compound **16.2** was studied. This compound was generated in situ (Scheme 16) by treating **16.1** with the ethyl ester of cysteine at concentrations ranging from 10^{-2} to 10^{-3} M in either DMF or DMSO. Half-times for the acyl transfer (**16.2→16.3**) were found to be 24 and 8 h in DMF and DMSO, respectively, and were independent of concentration, suggesting that transfer was intramolecular. The template was detached from the rearranged product by iodine-mediated oxidation (**16.3→16.4**), so as to release the disulfide **16.4**.



Scheme 15.



Scheme 16.

4.2. Ligation by disulfide exchange

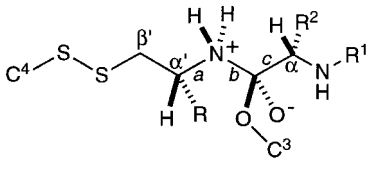
Capture by disulfide formation was also examined[38,39] as a means of ligating two peptide segments prior to their coupling by acyl transfer. In this approach, which has also been explored in a modified form (see later, Scheme 66), (X) in structure **14.2** (Scheme 14) takes the form of a thiol group, and ligation occurs by disulfide formation between this template thiol and the cysteine thiol at the *N*-terminus of the *C*-terminal peptide. Intramolecular acyl transfer from **14.4** (X=S) generates the peptide bond, and then cleavage of the resulting disulfide **14.5** (X=S) liberates the newly formed peptide **14.6**. Several practical points had to be considered in order to develop this version of the Prior Thiol Capture Strategy, and Kemp has carried out extensive research on each of these points.

4.2.1. Template design. Implementation of the Prior Thiol Capture Strategy requires a template that allows rapid intramolecular acyl transfer of a weakly activated acyl function to the weakly nucleophilic amino group of a cysteine residue.[38,40] As indicated in Scheme 14, the transfer necessarily involves a ring of at least nine atoms (cf. **14.4**, X=S). In general, cyclization of medium rings is entropically unfavorable.[13c,41] For this reason, the template would have to be constructed in such a way that, not withstanding the ring size of the transition state, the amine and the acyl group are easily positioned in a manner that confers a proximity-based entropic advantage, thus making the lack of enthalpic activation at the acyl substituent irrelevant.

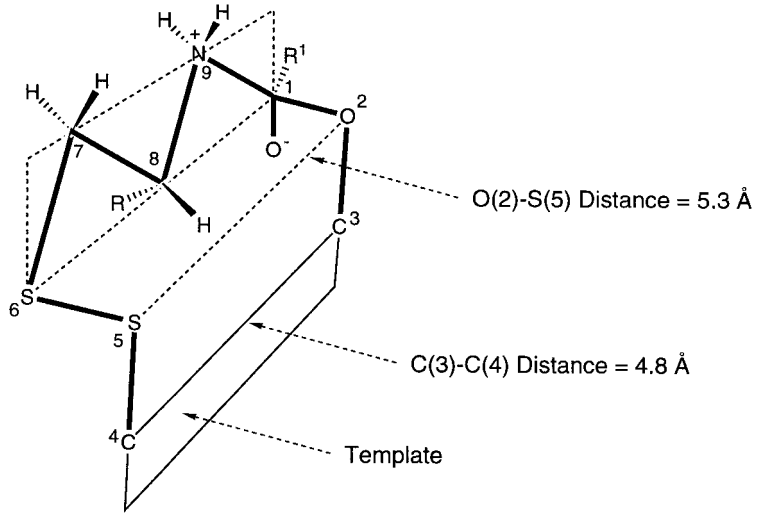
4.2.1.1. Effective molarity. The entropic advantage for an intramolecular reaction, which results as a consequence of restraining the nucleophilic and the electrophilic components in proximity, can be expressed quantitatively in terms of *effective molarity* (EM).[13b,c] The EM for a system can be estimated when a corresponding intermolecular version of a particular reaction exists. In such a case, the EM of the system is expressed as the ratio of the first-order rate constant (for the intramolecular reaction) to the second-order rate constant (for the intermolecular one). In practice, EM is equal to the concentration of the external electrophile or nucleophile that must be added to the intramolecular reaction mixture so that both intra- and intermolecular reactions proceed at the same rate.

Kemp has used the concept of EM to guide the development of an optimal template for acyl transfer.[40,42] This was done by measuring the EM with various templates that were deemed to possess the potential to function in the required manner, and then excluding from further study those with an unacceptably low EM value, the minimum value being set at EM=1 M. An effective template would be associated with both a high EM and a high rate of intramolecular acyl transfer.

4.2.1.2. Transition state model. In order to aid the design of suitable templates for the disulfide approach, a transition state model was proposed[32,40] on the basis of the following considerations. Information from earlier studies[43] justified the assumption that the transition state has nearly tetrahedral geometry at both the acyl carbon and the amine nitrogen, a fully formed C–N bond, and partially broken N–H and C–O bonds (Scheme 17). Furthermore, by analogy to a previously reported transition state model for intermolecular aminolysis of *p*-nitrophenyl esters,[34] it was assumed for the present purposes that the bond which connects the α-carbon (see Scheme 17) and the acyl substituent (bond *c*), and the bond which connects the α′-carbon and the amino substituent (bond *a*) have an antiperiplanar relationship about the C–N bond (bond *b*) in the acyl transfer transition state. Further refinement[40] of the transition state model was carried out by applying assumptions regarding bond angles and the conformational preferences of the bonds not already defined. The result of these considerations was the transition state model **18.1** (Scheme 18).[40] This structure implies that an ideal



Scheme 17. R^1=*N*-terminal peptide; R=*C*-terminal peptide; R^2=amino acid side-chain.



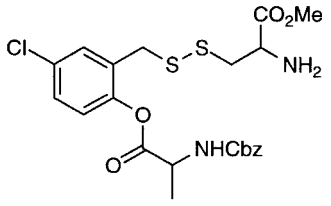
18.1

Scheme 18. R^1=*N*-terminal peptide; R=*C*-terminal peptide.

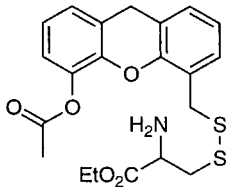
template would be one that can accommodate the particular geometric arrangement of atoms shown and provide a framework connecting C(4) and C(3) by a distance of 4.8 Å, and do so in a way that the resulting C(4)–S(5) and C(3)–O(2) bonds are oriented to give an O(2)–S(5) distance of 5.3 Å. It was also deemed important that the template be rigid, so as to decrease the number of random conformations that the system could sample before reaching the required transition state conformation. Finally, in order to avoid any major van der Waals interactions with the α-hydrogen [i.e. with H(8)] of the cysteine residue, the template should be free of substituents that protrude towards H(8); planar structures satisfy this requirement to some extent. Given that the model shown in Scheme 17 has rotational degrees of freedom in the C(α′)–C(β′)–S–S region, other low energy conformations must be accessible and, indeed, as indicated below, subsequent experiments indicated that **18.1** is not the only model that corresponds to a low-energy transition state for acyl transfer, but it was nevertheless adopted as the working model.

4.2.1.3. Evaluation of templates. While model **18.1** emerged from the theoretical considerations given above, Kemp also included related structures as candidates for evaluation, thus broadening the survey and providing an opportunity to identify the effects of template flexibility and transition state conformation.

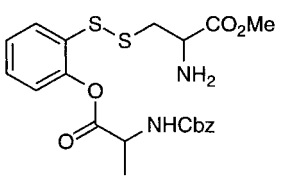
The first set of template-bound amino acids corresponding to model **18.1** that were examined are those shown in Scheme 19.[35] When dissolved in DMSO[44] for 20–40 h, compounds **19.1** and **19.2** each gave products corresponding to acyl transfer. However, the rates of transfer were low, and products resulting from disulfide exchange were also detected. Notwithstanding these undesirable features, each compound displayed a high EM. An accurate value could



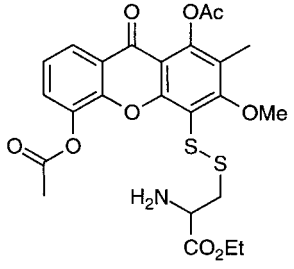
19.1



19.2

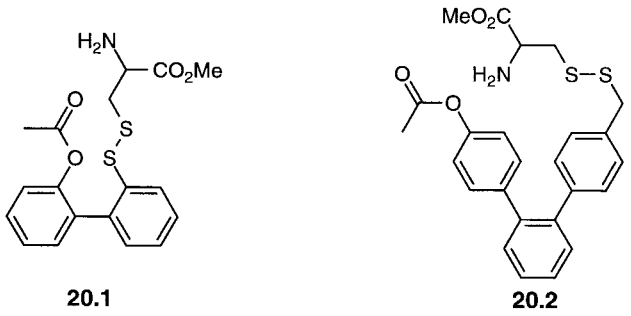


19.3



19.4

Scheme 19.



20.1 20.2

Scheme 20.

not be obtained due to the formation of byproducts, but an EM in the range of 3–14 M was estimated for **19.2**,[42,45] and 0.5–10 M[42] for **19.1**. These compounds have C(3)–C(4) distances (cf. Scheme 18) of 2.53 and 4.16 Å, and O(2)–S(5) distances of 2.5–4.2 and 2.6–5.4 Å, respectively,[42] and so only **19.2** approaches the values specified by the transition state model, but neither of them provides a sufficiently rigid framework, as indicated by the range of O(2)–S(5) distances that they can accommodate.

Compounds **19.3** and **19.4** were found[35] to undergo concentration-independent acyl transfer in both DMSO and DMF. The half-time for reaction of **19.4** in DMSO was 2.7 h,[35] and in DMF it was ca. 28 h.[35] An estimate of the EM of **19.4** was made[35] by determining the rate of intermolecular aminolysis for the reaction between ethyl *S*-benzyl cysteine and 1,3-dimethoxy-2-methyl-5-acetoxyxanthone. From this experiment, an EM of 0.5–0.7 M was determined[35,42] for the system. For **19.3**, an EM value between 0.08 and 1.3 M was estimated.[42] The C(3)–C(4) distances for compounds **19.3** and **19.4** were 1.40 and 4.76 Å, respectively, and the O(2)–S(5) distances were 2.98 and 4.81 Å, respectively.[42] Thus, although each of these compounds provided adequate rigidity, only **19.4** provided distances required of the transition state model **18.1**.

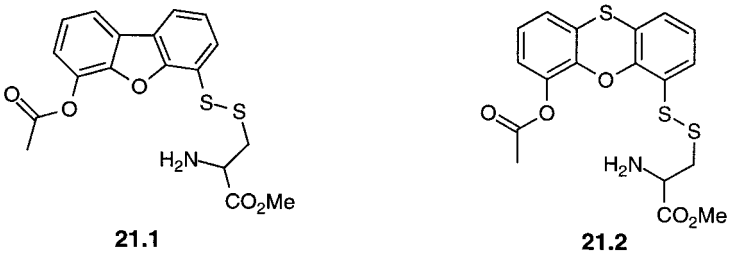
Compounds **20.1** and **20.2** were next examined for their ability to facilitate intramolecular acyl transfer (Scheme 20).[42,46] In the case of **20.1**, both the C(3)–C(4) and the O(2)–S(5) distances fell short of the optimal values, and for **20.2**, both were larger than desired.[42] In addition, each compound showed a range of values for both distances and lacked the required rigidity. When either **20.1** or **20.2** was kept in DMF, no products corresponding to acyl transfer could be detected, even after five days; only products of disulfide exchange were observed.[46] This result with **20.2** is noteworthy in the light of later work[42] which suggested that acyl transfer could also proceed through a more extended transition state than the one shown by **18.1**; such a transition state should be accessible to **20.2**, yet acyl transfer did not occur—presumably because of the flexibility of the system.[46]

The dibenzofuran **21.1** (Scheme 21) provided a promising example of a potential template for the Prior Thiol Capture Strategy, having a C(3)–C(4) distance of 4.82 Å and an O(2)–S(5) distance of 5.45 Å.[42] This template was examined along with the phenoxythiin system **21.2**,[40,42] which exhibited corresponding distances of 4.35 and 3.90 Å, respectively.[42] Efficient acyl transfer was not observed in the case of **21.2**,[40] and an EM of less than 0.1 M was determined.[40,42] Lack of acyl transfer in this system was attributed primarily to its inability to achieve a suitable O(2)–S(5) distance and, to some extent, to the flexibility of the system.[40,42]

As was evident by the C(3)–C(4) and O(2)–S(5) distances given above, the dibenzofuran template (**21.1**) very effectively accommodated the structural requirements of the proposed transition state. The half-times for acyl transfer in this system were found[40] to be solvent dependent, with values of 23, 23, 2 and 2 h being observed in DMF, 5:1 hexafluoroisopropyl alcohol (HFIP)–DMSO, DMSO, and 1:5 DMSO–HMPA, respectively. In acetonitrile acyl transfer was not detected, even after 30 h. The half-time for the corresponding intermolecular reaction in DMSO was obtained by treating 4-acetoxydibenzofuran with ethyl *S*-benzyl-L-cysteinate. From this experiment, the EM of compound **21.1** was determined to be 4.6 M.

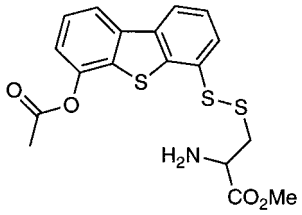
An examination of the effect of substituents attached to the phenolic ring of the dibenzofuran template on the rate of acyl transfer was undertaken.[40] As expected, an electron withdrawing substituent *para* to the phenolic oxygen caused a rate increase; in the case of a nitro group, for example, an increase of more than 3000 was observed.[40]

The sulfur analog **22.1** (Scheme 22) presumably has a degree of structural rigidity comparable to the dibenzofuran system, but the critical dimensions and the hypothesized transannular steric interactions with the cysteine α-hydrogen are quite different.[42] The C(3)–C(4) distance in **22.1** is 5.23 Å, whereas the O(2)–S(5) distance is 6.30 Å. This system did not undergo efficient acyl transfer[42]—a fact which lends support to the belief that the spatial relationships and nonbonded interactions implied by the transition state model **18.1** are important.



21.1 21.2

Scheme 21.



22.1

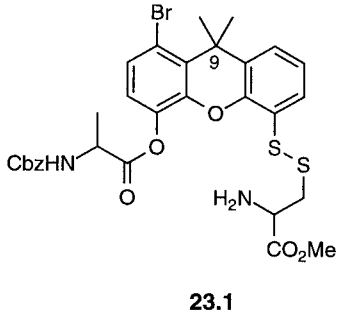
Scheme 22.

The next set of templates studied included those shown in Scheme 23. The 9,9-dimethylxanthene derivative **23.1**[47] is structurally related to compound **19.2**, but possesses methyl substituents at C(9), which enforce a folded conformation on the ring system. The bromo substituent *para* to the acyl group was necessary because the reactivity of the unsubstituted system was too low to permit accurate monitoring of the intra- and intermolecular reactions. The EM calculated for this system was approximately 0.1 M, indicating that intramolecular acyl transfer, if it occurred at all, was not facilitated relative to the corresponding intermolecular process.

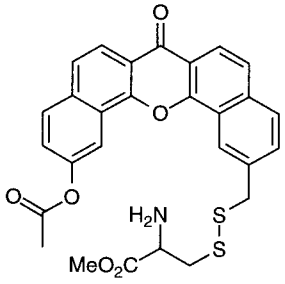
Compound **23.2**[47] provided an opportunity to test the assumption mentioned earlier in connection with compound **20.2**—that the acyl transfer could occur through a transition state with a more extended conformation than shown by model **18.1**. In the case of **20.2**, the absence of acyl transfer was attributed, in part, to its lack of rigidity.[46] Hence, studies with **23.2** would identify what benefit a rigid system might provide. In the event, the rate constant for the intramolecular process in DMSO was 0.02 h^{-1}, which corresponds to an EM of 0.3 M for **23.2**. Although the EM is low, the fact that acyl transfer occurred was regarded[47] as significant in that it further supported[48] the hypothesis that acyl transfer can take place through a more extended transition state, different from the one predicted by the model **18.1**.

Compound **24.1** and its nitro analog **24.2** were also evaluated (Scheme 24).[47] These are fundamentally different from the previously studied systems, in that they provided an opportunity to examine intramolecular acyl transfer using a template containing a chiral center. The compounds embody the distances of Scheme 18 and the spanning atomic linkages of the dibenzofuran **21.1**, but their nonplanar conformations potentially reduce transition state crowding at the cysteine α-hydrogen. Both **24.1** and **24.2** were obtained as 1:1 mixtures of diastereomers, and the acyl transfer reaction of each diastereomeric mixture was examined.[47] The two diastereomers within each mixture reacted at different rates, and the rate constant for the more reactive isomer in **24.1** and **24.2** was found to be 0.013 and 2.3 h^{-1}, respectively. The rates of the corresponding intermolecular reactions were determined; the respective EMs for **24.1** and **24.2** (faster diastereomer in each case) were calculated as 6 and 1.3 M.

Another report,[49] published at the same time, described the examination of compound **25.1** (Scheme 25). It showed relatively low rates of transfer, with a corresponding EM of 0.3 M. Conversion of **25.1** into its sulfoxide derivatives **25.2** not only provided another opportunity to examine intramolecular acyl transfer across a template possessing a chiral center, but also offered the possibility of exploring what effect oxidation at the sulfur atom of **25.1** would have on the rate of reaction, since the presence of the sulfoxide group might facilitate acyl transfer in a manner similar to that of DMSO.[49] As expected, each of the (separated) diastereomers of **25.2** underwent acyl transfer at a different rate, and the higher rate constant was 7.0 h^{-1}. This corresponds to an EM of 0.22 M. Thus, compound **25.2** appears to be even less efficient as an acyl transfer system than **25.1**. The slower-reacting diastereomer of **25.2** had an EM of ca. 0.0015 M.

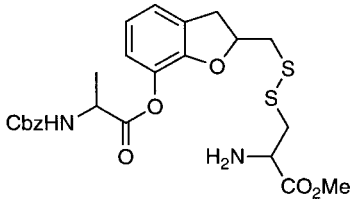


23.1

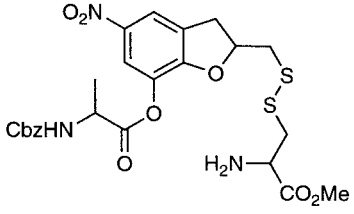


23.2

Scheme 23.

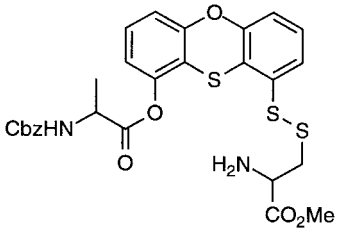


24.1

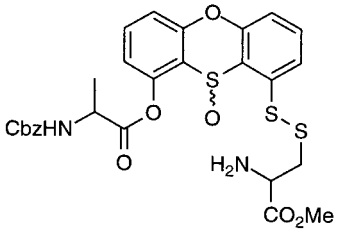


24.2

Scheme 24.



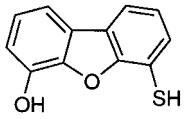
25.1



25.2

Scheme 25.

During the course of the above research it had become apparent that 4-hydroxy-6-mercaptodibenzofuran (**26.1**, Scheme 26) was the best template of those examined, and was worthy of further study. In this template the C(3)–C(4) distance was found to be 4.82 Å, which compares well to the distance (4.8 Å) required by the transition state model **18.1**. The O(2)–S(5) distance of 5.45 Å was also close to the value of 5.3 Å specified by the model. In addition, the template provided the rigid framework that was needed, and was expected to show only one moderately weak van der Waals interaction between the furan oxygen and the cysteine α-hydrogen in the acyl transfer transition state.



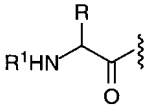
26.1

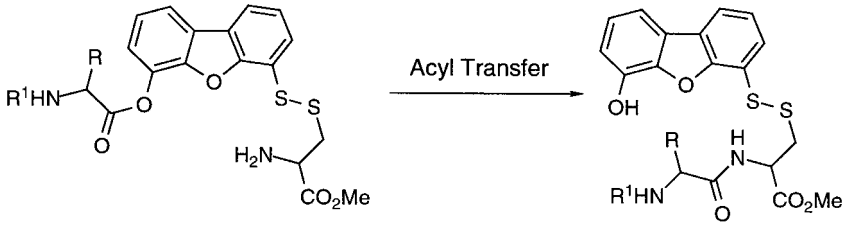
Scheme 26.

4.2.2. Effect of amino acid side-chains on the rate of intramolecular acyl transfer. Since compound **26.1** emerged as a promising template for acyl transfer by the Prior Thiol Capture Strategy, Kemp undertook an investigation into the effect that the side-chain of the acylating component would have on transfer rates in this system.[50] It has been argued[34,50] that, for unassociated peptide segments, the rate of peptide bond formation is largely determined by the two substituents that neighbor the new bond. On this basis, it was recognized that coupling reactions giving dipeptides by intramolecular acyl transfer across the dibenzofuran template (cf. **27.1→27.2**, Scheme 27) should serve as suitable models for corresponding reactions in which actual peptide segments are coupled.

The rates of acyl transfer for the formation of several dipeptides using system **27.1** were measured in DMSO (Table 7)[50] and, with four exceptions [L-Pro, L-Val, L-Asp, L-Asp(*t*-Bu)], the transfer half-times were found to be between 2 and 4 h. The low rate of transfer in the case of the valine derivative was assumed to result from β-branching of the side-chain, a situation that is known to slow acyl transfer in intermolecular reactions as well.[34,51] In the case of the aspartic acid derivative, the high rate of transfer was rationalized on the basis of a stabilizing intramolecular hydrogen-bonding interaction (in the transition state) between the side-chain carboxyl group and the attacking amino group; this interpretation is consistent with the concept that an ideal template may well incorporate a hydrogen-bonding site.[50] The retarding effect associated with the proline residue was attributed to steric interactions in the transition state, but was deemed to be unique to that residue. Thus, intramolecular acyl transfer in **27.1** was tolerant of a range of substituent changes, but was significantly retarded by extreme steric interactions.

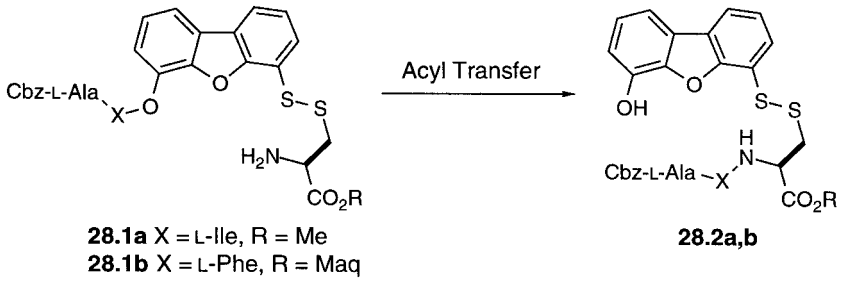
Table 7. Half-times for intramolecular acyl transfer reactions of **27.1** (Mbh=3,3′-dimethoxybenzhydryl; Ans=9-anthracenesulfonyl)[50]

	Acyl transfer half-time (h)
L-AlaGly	2.0
L-Ala	3.0
L-Leu	4.0
L-Pro	34
L-Val	51
L-Lys(Cbz)	3.2
L-Asn	3.1
L-Asn(Mbh)	2.4
L-Asp(*t*-Bu)	1
L-Asp	~0.1
L-Arg(Ans)	1.8
L-Thr(*t*-Bu)	2.5



27.1 27.2

Scheme 27.



Scheme 28. Maq=2-oxymethylanthraquinone.

4.2.3. Extent of epimerization. The possibility of epimerization at the α-carbon of the acyl fragment during acyl transfer, using the dibenzofuran template **26.1**, was examined with the model systems **28.1a**,**b**.[52] Given that the acylating component in these systems is only weakly activated as an aromatic ester, the epimerization levels were expected to be low, and this was indeed the case. Compound **28.1a** was allowed to react (Scheme 28) in DMSO for three days, in order to ensure complete acyl transfer, giving **28.2a**. This was then hydrolyzed in the presence of 12 M HCl-propionic acid to liberate the individual amino acid components. The L-isoleucine/D-*allo*-isoleucine ratio was measured and from this the extent of epimerization was determined to be 0.20±0.27%, thus indicating that, within the limits of accuracy of the method, no epimerization occurs during the coupling process.

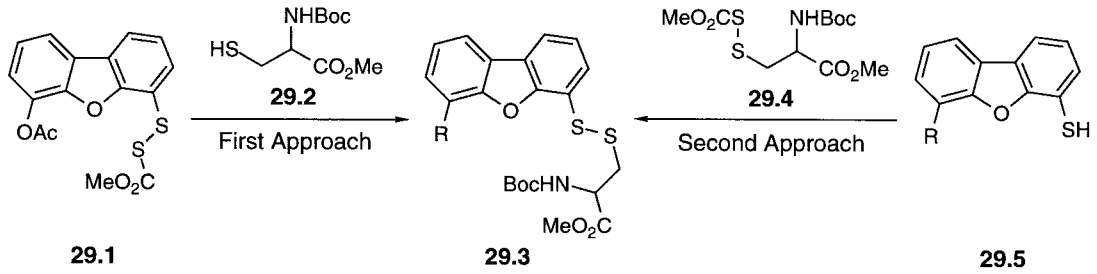
Further support for the absence of epimerization was obtained[52] by examining the acyl transfer products from **28.1b**. It had been established that the L–L–L form of **28.2b** was separable from its L–D–L epimer using HPLC, and that the separation could be achieved with baseline resolution. Hence, once the acyl transfer reaction for **28.1b** was complete, the products were analyzed using this HPLC technique. Comparison of the product mixture with reference standards indicated that the ratio of L–L–L to L–D–L epimers was greater than 99.9:0.1.

4.2.4. Disulfide formation. In the Prior Thiol Capture Strategy, initial ligation of the two peptide segments occurs by formation of a mixed disulfide and, in order to optimize this process, Kemp investigated the formation of disulfides involving dibenzofuran systems.

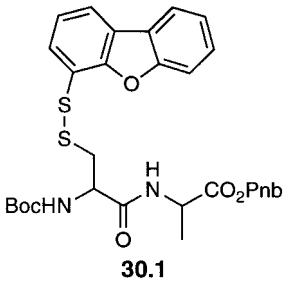
Two general approaches to mixed disulfides were explored.[22] In each, the disulfide was generated by reaction of a free thiol with a sulfenylcarbomethoxy-derivatized thiol (-SScm), a type of reaction that was known[53] to be specific and efficient. In the first approach[22] (Scheme 29), the template thiol was derivatized with ScmCl (cf. **29.1**), and was then allowed to react with a free cysteine thiol [**29.1**+**29.2**→**29.3** (R=OAc)]. In the second approach[22] (Scheme 29), it was the cysteine thiol that was derivatized with ScmCl (cf. **29.4**), and it underwent reaction with the free template thiol [**29.4**+**29.5** (R=OAc)→**29.3** (R=OAc)]. Of the two routes, the second was found to be very much more effective, and typically gave the desired disulfide in near quantitative yields.

Important features of the ligation step were identified by studying the kinetics of the reaction of **29.5** (R=H) with **29.4** to give **29.3** (R=H)[22] in mixtures of HFIP–acetonitrile, with and without water, and in the presence and absence of various tertiary amines. From these experiments, two general factors became apparent. First, the amine exhibited a catalytic effect on the reaction—an observation that is consistent with the view that the thiolate anion of **29.5** (R=H) acts as the nucleophilic species. The catalytic effect was independent of the base strength or concentration, suggesting that the actual catalyst was the conjugate base of the solvent (HFIP). The second factor revealed by the kinetic study was that an increase in solvent polarity, through the addition of water, caused a dramatic increase (ca. 10^4) in reaction rate.

It had been established[54] that a reaction time of 12 h was generally required to achieve approximately 90% conversion of the ligated product to the acyl transfer product and, as a result of this long reaction time, it was necessary to determine to what extent disulfide exchange occurs during the acyl transfer. To this end,[54] **30.1** (Scheme 30) was stored in DMSO for 19 h. Analysis of the product composition showed that the compound was largely (98%) unchanged and that only 2% had been converted into the corresponding symmetrical disulfides. Hence DMSO was deemed to be a suitable solvent for avoiding the problem of disulfide exchange.

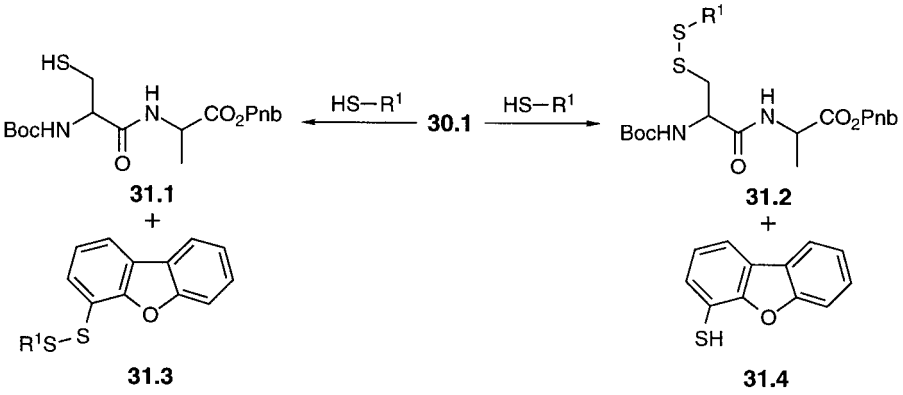


Scheme 29.

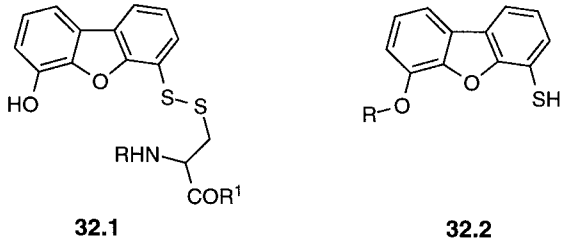


Scheme 30. Pnb=*p*-nitrobenzyl.

If the initial ligation leading to the required disulfide is not quantitative, then free thiol in the form of the template-linked *N*-terminal peptide [cf. **14.2** (X)=SH, Scheme 14] would be present in the reaction mixture. The effect that this thiol might have on disulfide exchange was also examined,[54] by separately treating **30.1** with thiophenol or benzyl thiol in DMSO (Scheme 31). In each case an excess of thiol ranging from 10–100 fold was used. Under these conditions, rapid disulfide exchange occurred to give the mixed disulfides **31.3** and **31.2** in a ratio of 1:99.



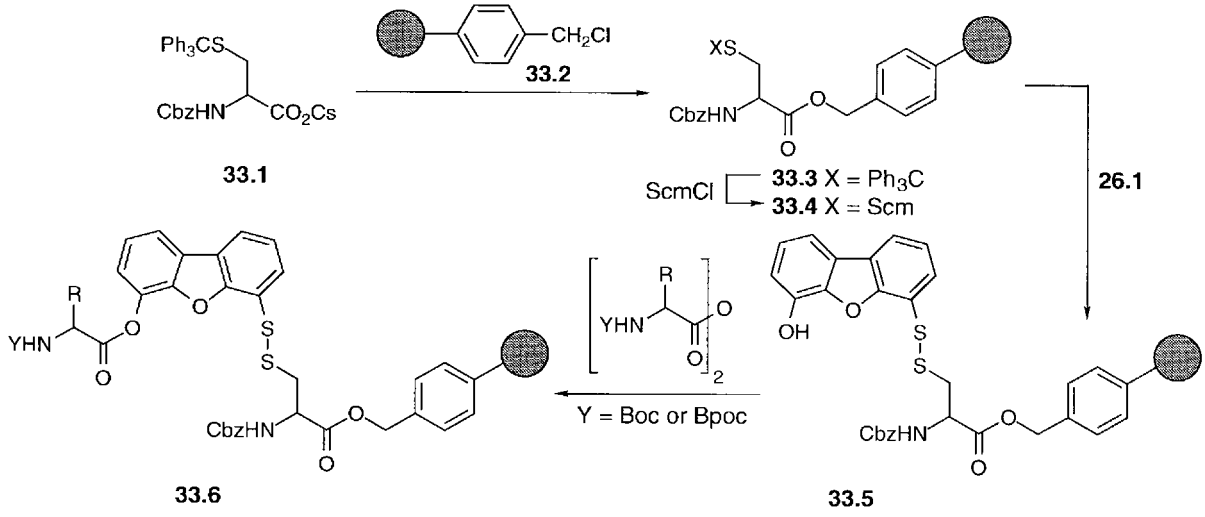
Scheme 31. R^1=Ph or Bn; Pnb=*p*-nitrobenzyl.



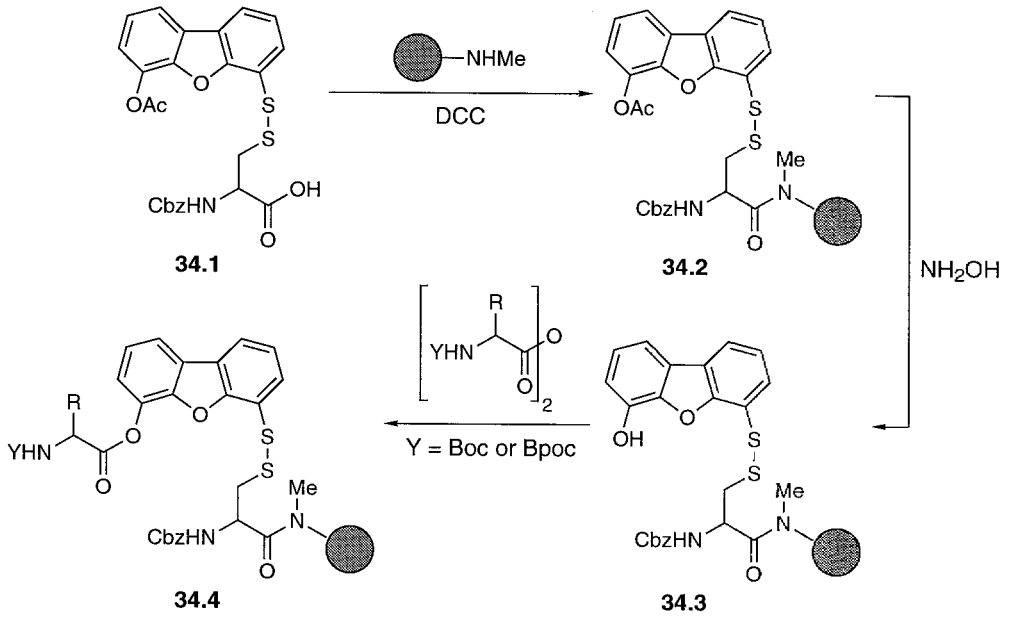
Scheme 32. R=CbzPFT(*t*-Bu); R^1=GGAOPnb; Pnb=*p*-nitrobenzyl.

In a subsequent experiment,[54] it was established that unwanted thiol-catalyzed disulfide exchange between the product **32.1** (Scheme 32) of an acyl transfer reaction and the capture thiol **32.2** used in that reaction could be suppressed almost completely by addition of the thiol scavenger silver nitrate, before acyl transfer.

4.2.5. Synthesis of the dibenzofuran-linked *N*-terminal peptide segments. Application of the Prior Thiol Capture Strategy requires that the *N*-terminal peptide (cf. **14.1**, Scheme 14) be generated in such a way that its *C*-terminus



Scheme 33. ● =polystyrene resin; Scm=SCO_2Me.



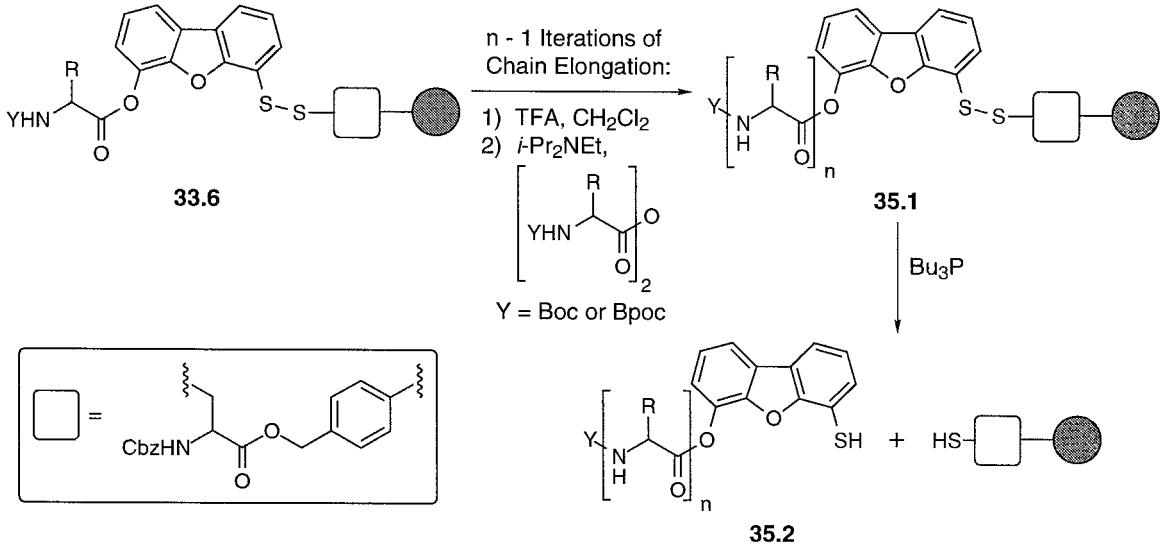
Scheme 34. =polystyrene resin.

is linked to the template. For this purpose, the 4-hydroxy-6-mercaptodibenzofuran-linked solid support **33.6** was developed.[38] This support was synthesized by the route outlined in Scheme 33. Commercially available chloromethylated polystyrene resin (**33.2**) was treated with the derivatized cysteine cesium salt **33.1**, to give **33.3**. Exchange of the trityl protecting group for an Scm group (**33.3**→**33.4**) was achieved by treatment with ScmCl, and exposure of **33.4** to the mercaptodibenzofuran **26.1** under standard conditions led to disulfide **33.5**. Finally, acylation of the phenolic hydroxyl with the symmetrical anhydride of an N^{α}-blocked amino acid gave the required resin-bound disulfide **33.6**. As described below, the resin was evaluated for its ability to survive the conditions of solid phase synthesis.

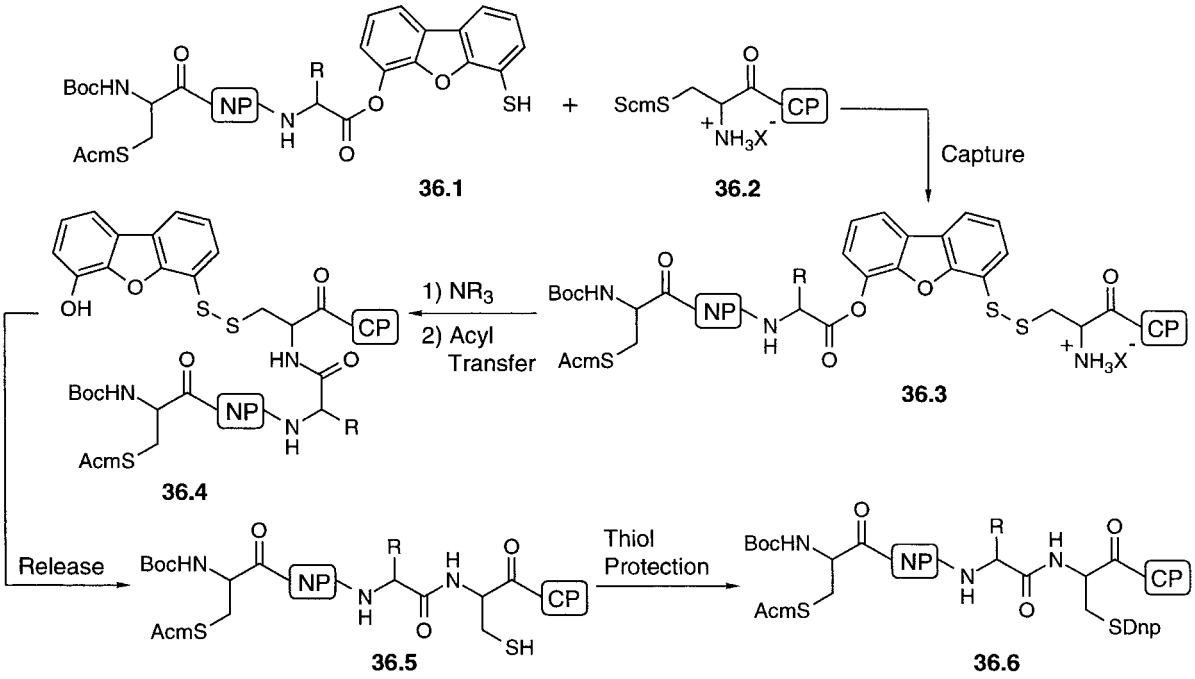
An alternative method was also developed[22] for linking the dibenzofuran template to a solid support, and involves the use of an aminomethyl polystyrene resin. In this case (Scheme 34), compound **34.1**[55] was coupled to the resin, using DCC, to form **34.2**. Removal of the acetyl group with hydroxylamine then generated the phenolic species **34.3**, which underwent *O*-acylation (**34.3**→**34.4**) in a manner analogous to that involved in the conversion of **33.5** into **33.6**.

The phenolic ester that anchors the peptide chain in **33.6** and **34.4** precludes the use of the fluorenylmethoxycarbonyl group (Fmoc) for N^{α}-protection, because of the requirement of a secondary amine as the deblocking reagent, but two other standard solid phase synthesis protocols are suitable for chain extension.[22,38]

The first is the combination of Boc protection for the



Scheme 35. =polystyrene resin.



Scheme 36. NP=*N*-terminal peptide; CP=*C*-terminal peptide; X=CF_3CO_2; Acm=$CH_2NHCOMe$; Scm=SCO_2Me; Dnp=2,4-dinitrophenyl.

α-amino group[38] and benzyl protection[38] for the side-chain functions. Chain elongation (Scheme 35) was initiated by treatment of **33.6** (Y=Boc) with 50% TFA in CH_2Cl_2 to give the corresponding TFA salt. This salt was then treated with an appropriately protected aminoacyl anhydride in the presence of *i*-Pr_2NEt, giving a product elongated by one residue. The sequence was repeated until the desired peptide (**35.1**) had been assembled. Treatment with tributylphosphine[38] then served to reductively cleave the disulfide bond, giving the protected *N*-terminal peptide **35.2**, already derivatized at its *C*-terminus as a 4-hydroxy-6-mercaptodibenzofuran ester. The peptide could also be obtained in fully unprotected form by simply effecting global deprotection of the side-chains prior to resin cleavage.

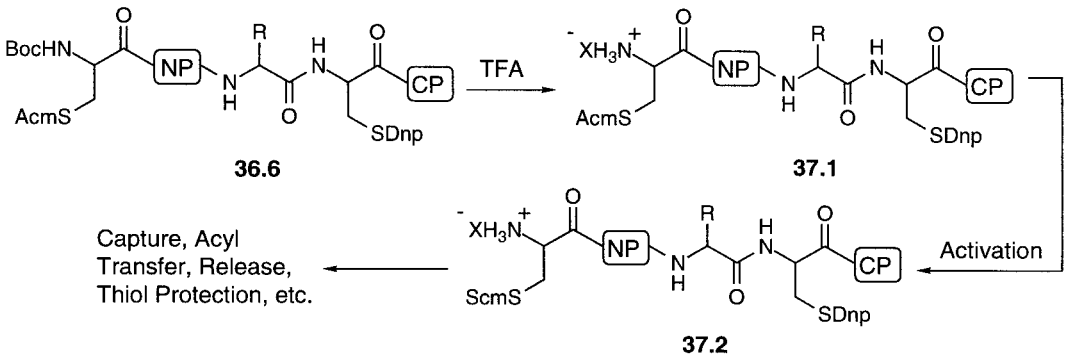
The second solid phase protocol[38] employed Bpoc (*p*-$PhC_6H_4CMe_2OCO$–) protection of the α-amino group (Y=Bpoc, Scheme 35), combined with *t*-butyl[22] and Boc protection[38] for the side-chain functions. In this case, chain elongation was carried out as described above, with the exception that the Bpoc protecting group was removed using 0.5% TFA in the presence of 1% thioanisole.

4.2.6. Cysteine thiol protection and a modified Prior Thiol Capture Strategy. In order to implement the Prior Thiol Capture Strategy, it was necessary to have available a variety of protecting groups for cysteine thiol residues, both internal and terminal.[22] Protection of this type is required because there is no general method[56] for forming a disulfide selectively from one thiol in the presence of others. Aside from the cysteine side-chains, no other protection of substituents in either peptide segment is required.[57]

The approach that was developed[22,56,58] to deal with selective protection of cysteine residues is summarized in Scheme 36.[56] The thiol group of the *N*-terminal cysteine of the *C*-terminal peptide **36.2** is protected with a sulfenyl-carbomethoxy (Scm) group, and the *N*-terminal cysteine thiol of the dibenzofuran-linked peptide **36.1** is protected as an acetamidomethyl (Acm) thioether.[59] Compound **36.2** undergoes ligation with the *N*-terminal peptide, by nucleophilic displacement of the Scm moiety by the template thiol, to generate the acyl transfer intermediate (**36.1**+**36.2**→**36.3**). Treatment with an amine initiates acyl transfer, so as to couple the two peptide segments and give compound **36.4**. Cleavage of the template from **36.4** is then followed by thiol protection (treatment with 2,4-dinitrofluorobenzene in a bicarbonate buffer) as the 2,4-dinitrophenylsulfenyl (Dnp) derivative **36.6** (**36.4**→**36.5**→**36.6**).

The above series of protecting groups used for the cysteine thiol functions was chosen for the following reasons. First, both the Dnp- and the Acm-protected thiols are inert to attack by the nucleophilic template thiol, and so no complications would result from formation of disulfides other than the one required for acyl transfer. The second reason is that, unlike the Dnp-protected thiol, the Acm-protected thiol is readily convertible into the corresponding Scm form,[60] which *is* susceptible to nucleophilic attack by the template thiol. Discrimination between the protected thiol groups is required so that, once the coupled product has been liberated from the template (**36.4**→**36.5**), the free thiol group can be converted into its Dnp-protected form (**36.5**→**36.6**) and, following removal of the N^{α}-Boc group (**36.6**→**37.1**, Scheme 37), the Acm-protected *N*-terminal cysteine thiol can then be activated by conversion into its Scm derivative (**37.1**→**37.2**)—a sequence of steps that would set the stage for a second ligation between **37.2** and a new template-linked *N*-terminal peptide (cf. **36.1**), marking the beginning of a new cycle of peptide coupling. In principle, the whole sequence can be repeated as many times as desired.

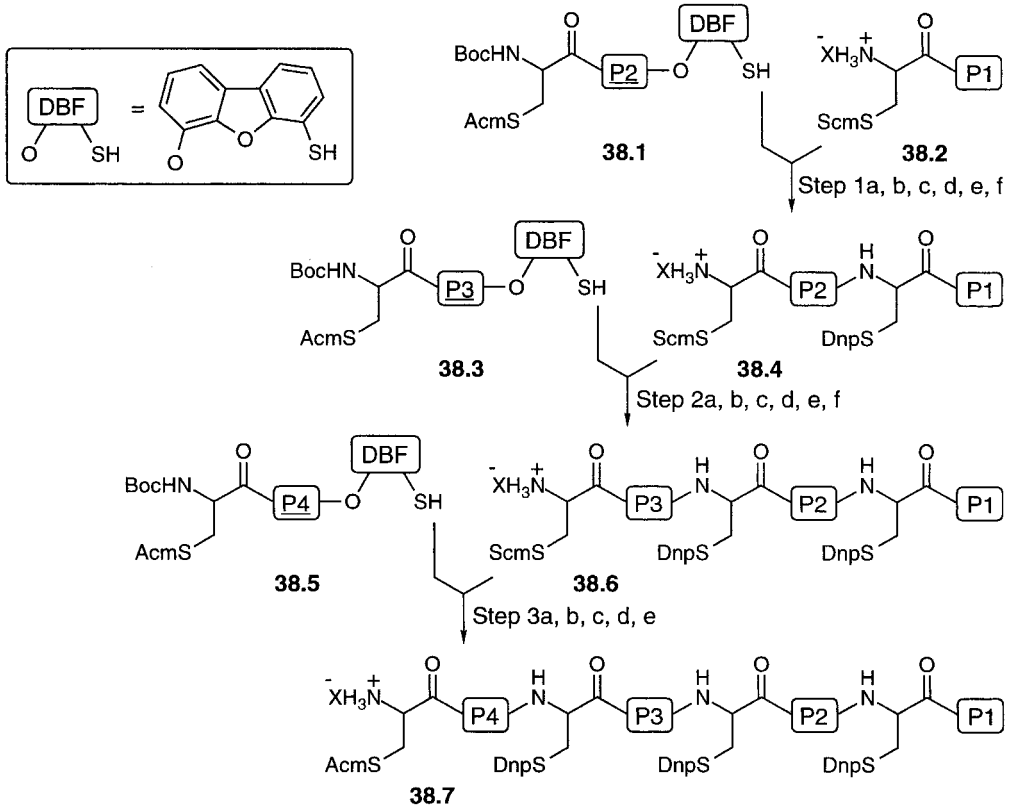
4.2.7. Application to peptide synthesis. The Prior Thiol Capture Strategy has been applied to the synthesis of



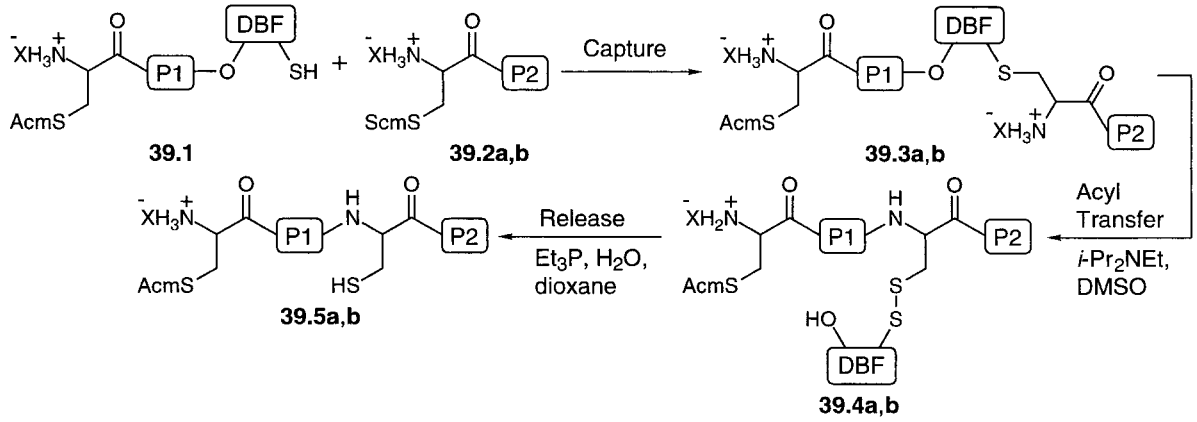
Scheme 37. NP=*N*-terminal peptide; CP=*C*-terminal peptide; X=CF_3CO_2; Acm=$CH_2NHCOMe$; Scm=SCO_2Me; Dnp=2,4-dinitrophenyl.

several medium-sized peptides. One of these is the 29-residue *C*-terminal segment of the protein *basic pancreatic trypsin inhibitor* (BPTI).[22] This segment corresponds to residues 30–58 of the native protein. The synthesis of this peptide was the first major test of the Prior Thiol Capture Strategy, and was effected by carrying out three cycles of the process outlined in Schemes 36 and 37, but with one modification. Instead of using *N*-terminal peptide segments with protecting groups only on the *N*-terminal nitrogen and cysteine residues (cf. **36.1**), fully protected segments were used. The non-cysteinyl protecting groups were all TFA-labile and so, after Dnp protection of the new thiol group (cf. **36.5→36.6**), treatment with TFA also removed the non-cysteine side-chain protecting groups (both SAcm and SDnp are stable under these conditions). This reaction gave a new *C*-terminal peptide, protected only on its cysteine side-chains, which could undergo activation (cf. **37.1→37.2**) and enter into another coupling cycle.

The synthesis of the peptide in question was carried out as summarized in Scheme 38. In the first step, the activated *C*-terminal peptide **38.2** underwent capture by the fully protected segment **38.1** (step 1a), to give the ligation intermediate. Acyl transfer (step 1b) then occurred, and was followed by trialkylphosphine-mediated template cleavage (step 1c), Dnp protection of the liberated thiol (step 1d), deprotection of the *N*-terminal nitrogen and non-cysteine amino acid side-chains (step 1e) and, finally, activation of the *N*-terminal cysteine thiol by conversion into its Scm derivative (step 1f). The product of this coupling cycle



Scheme 38. a=capture; b=acyl transfer (*i*-Pr_2NEt); c=template cleavage (Et_3P); d=thiol protection (SH→SDnp); e=deprotection (TFA); f=activation (SAcm→SScm); P1=GGA; P2=MR(Pmc)T(*t*-Bu); P2=deprotected form of P2; P3=R(Pmc)AK(Boc)R(Pmc)NNFK(Boc)S(*t*-Bu)AE(*t*-Bu)D(*t*-Bu); P3=deprotected form of P3; P4=QT(*t*-Bu)FVY(Dnp)GG; P4=QTFVY(Dnp)GG; X=CF_3CO_2; Acm=$CH_2NHCOMe$; Scm=SCO_2Me; Dnp=2,4-dinitrophenyl; Pmc=pentamethylchromane.



Scheme 39. P1=LNELDADEQADL; P2=LARFGDDGENL (for **a** series); P2=ESLHDHADELYRSC(S*t*-Bu)LARFGDDGENL (for **b** series); P2=ESLHDHADELYRSCLARFGDDGENL (for **39.5b**); X=CF_3CO_2.

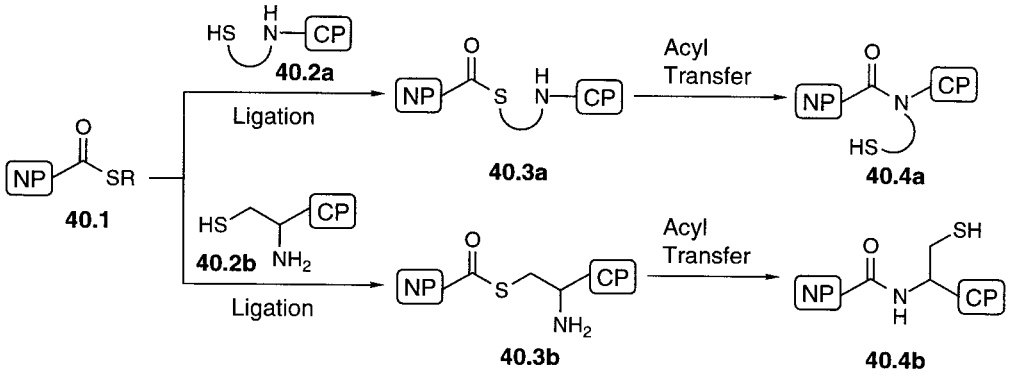
was the activated, cysteine-protected octapeptide **38.4**, which corresponds to residues 51–58 of native BPTI. During the second cycle of the synthesis, this segment was coupled to the template-derivatized 13-residue segment (**38.3**), corresponding to residues 38–50 of the native protein. The product of this second coupling cycle was the 21-residue peptide **38.6**. In the third and final cycle, the activated segment **38.6** was combined with the octapeptide **38.5**, to give the target peptide **38.7**, protected only on its cysteine side-chains. Overall yields for each of the six-step cycles were between 50% and 75%. At high pH a second product was obtained resulting from intramolecular acyl transfer to the ϵ-amino group of the lysine residue at site 3 of peptide P3, but no acyl transfer to the second lysine of P3 at site 8 was detected. It is not unexpected that at high pH an ϵ-amino group is acylated since, under these conditions, the lysine ϵ-amino group is as abundant as the *N*-terminal cysteine α-amino group and, as reflected in the higher pK_a of its conjugate acid, it should be a better nucleophile. The fact that the more distal ϵ-amino group is not acylated indicates an upper limit to the distance over which intramolecular acyl transfer can occur, and a model study has shown that, within this distance, selectivity of transfer can be very largely controlled by proper pH adjustment.[62]

Subsequent to the above synthesis, the Prior Thiol Capture Strategy was applied to the synthesis of a 25- and a 39-residue peptide.[58] In these cases, however, peptide segments that were fully deprotected, except on their cysteine residues, were used. The 39-residue peptide corresponded to residues 26–63 of the 63-amino acid *ColE1 repressor of primer protein*, but had isoleucine-37 replaced by leucine, and had a cysteine residue attached to its *N*-terminus. This peptide was synthesized by coupling a 13-residue segment corresponding to amino acids 26–37 of the native protein (but with the additional *N*-terminal cysteine) to a 27-residue segment corresponding to amino acids 27–63 of the native protein and having the leucine-37 substitution. The 25-residue peptide was synthesized by coupling the same 13-residue *N*-terminal fragment used for the 39-residue peptide with a segment corresponding to residues 52–63 of the native protein.

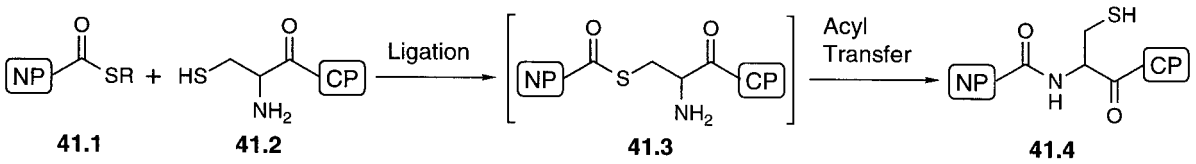
Each of the *C*-terminal segments used for the synthesis of these two peptides was generated using standard solid phase procedures. Synthesis of the common *N*-terminal fragment, however, made use of the special dibenzofuran-linked solid phase resin (**34.3**) described above (see Scheme 34).

The synthesis of the 25- and 39-residue peptides is outlined in Scheme 39. In each case, the coupling cycle was initiated by reacting the template-linked *N*-terminal peptide **39.1** with the appropriate activated *C*-terminal peptide (**39.2a**,**b**), to give the ligated product (**39.1**+**39.2a**→**39.3a**; **39.1**+**39.2b**→**39.3b**).

The ligated product so obtained was then treated with base



Scheme 40. NP=*N*-terminal peptide; CP=*C*-terminal peptide.



Scheme 41. NP=*N*-terminal peptide; CP=*C*-terminal peptide.

in order to initiate acyl transfer, and this process gave rise to compounds **39.4a**,**b**. At this stage reductive cleavage was carried out in order to remove the template and generate the desired peptide segments **39.5a**,**b**. (In the case of template cleavage from **39.4b**, the internal cysteine *t*-butyldisulfide protecting group was also cleaved.) In this way, the 25-residue peptide **39.5a** and the 39-residue peptide **39.5b** were obtained in 82% and 80% yield, respectively.

5. Native Chemical Ligation Strategies

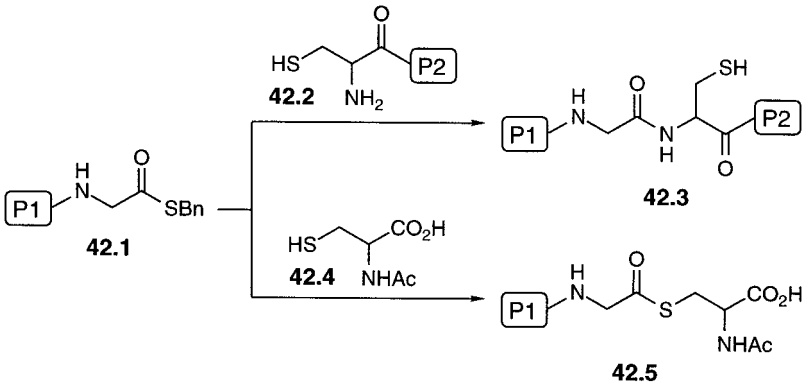
A third ligation-based approach to peptide segment coupling is the *Native Chemical Ligation Strategy*, which was developed by Kent.[63] In this strategy, ligation of the segments prior to peptide bond formation occurs through a thioester link between the two segments to be coupled. The thioester is formed by exchange between a thiol and a thioester, a type of reaction that has been studied extensively[64] and is known to proceed in a highly chemoselective manner. Kent's approach can be outlined in a general way by reference to Scheme 40. Here, the thiol is part of the *N*-terminal amino acid residue of the *C*-terminal peptide segment (**40.2a**,**b**), and the thioester is simply a derivatized form of the carboxyl terminus of the *N*-terminal peptide (**40.1**). Ligation occurs when the two components are mixed in solution, and the reaction generates an intermediate (**40.1**+**40.2a**→**40.3a**; **40.1**+**40.2b**→**40.3b**) that is capable of undergoing intramolecular *S*→*N* acyl transfer. The transfer in this case gives rise to a coupled product (**40.4a**,**b**) which, depending on the specific approach used, may or may not require further manipulation.

From Kent's original ideas on Native Chemical Ligation two related procedures have emerged. The main distinction between them is the source of the nucleophilic thiol involved in the ligation step. This ligation determines the site at which coupling occurs between the two peptide segments. In one approach, the source of the thiol is an *N*-terminal cysteine residue, and coupling occurs between that cysteine and some other amino acid. In the other approach, an *N*-ethanemercapto- (*N*-CH_2CH_2SH) or *N*-oxyethanemercapto- (*N*-OCH_2CH_2SH) derivatized glycine or alanine residue provides the nucleophilic sulfur and, as a result, coupling occurs between that glycine or alanine and some other amino acid.

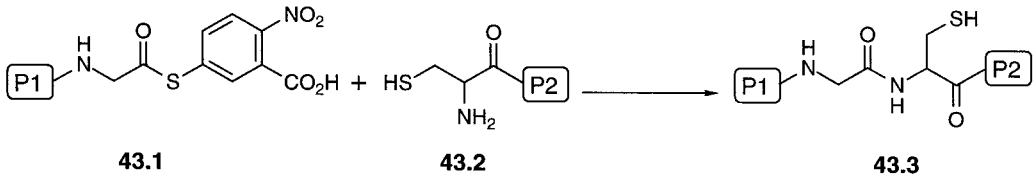
5.1. Ligation by thioester exchange involving cysteine

In the first of the Native Chemical Ligation approaches,[63] an *N*-terminal cysteine on the *C*-terminal peptide provides the thiol substituent that is required for ligation. The ligation is chemoselective and so no protecting groups should be required on either peptide segment. The general strategy, which has also been applied in another context (see Scheme 62), is outlined in Scheme 41. The principle is that ligation occurs when the thioester moiety of the *N*-terminal peptide (**41.1**) undergoes nucleophilic attack by the thiol group of the *N*-terminal cysteine of the other peptide (**41.1**+**41.2**→**41.3**). This process gives rise to an acyl transfer intermediate (**41.3**) which is not isolated, but which undergoes spontaneous intramolecular acyl transfer to generate the coupled product directly (**41.3**→**41.4**). In this case no further manipulation is required.

5.1.1. Mechanistic and practical aspects. Intramolecular acyl transfer in the cysteine-based Native Chemical Ligation Strategy must proceed via a five-membered transition state from an intermediate (cf. **41.3**) that is not isolated. Indirect evidence has been obtained for the proposed intermediate in two separate experiments. In the first, the thioester-derivatized *N*-terminal peptide **42.1** (Scheme 42) was treated with *N*-acetylcysteine (**42.4**).[63] Ligation



Scheme 42. P1=LYRA; P2=RAEYS.



Scheme 43. P1=MscCTETLQNAHSMALPALEPSTRYWARVRVRT; P2=TGYNGIWSEWSEARSWDT; Msc=[2-(methanesulfonyl)ethyloxy]carbonyl.

proceeded to give **42.5**, but this compound cannot undergo acyl transfer because the amino group of the cysteine is blocked, and so **42.5** was isolated in the form shown. When the same *N*-terminal peptide (**42.1**) was treated with a fully deprotected *C*-terminal peptide having an *N*-terminal cysteine, the expected coupling product was obtained (**42.1**+**42.2**→**42.3**).

Additional evidence for the proposed intermediate was obtained indirectly from an experiment that was also used to established the requirement for cysteine at the *N*-terminus of the *C*-terminal peptide.[63] In this work, a thioester-derivatized *N*-terminal peptide (cf. **41.1**) was combined with a 10-fold molar excess of the *C*-terminal peptide Leu-enkephalin[65] which has, at its *N*-terminus, a tyrosine residue. No reaction occurred. However, when the same *N*-terminal peptide was treated with a different *C*-terminal peptide, having a cysteine residue at its *N*-terminus (cf. **41.2**), rapid coupling took place. These observations implied not only that the proposed intermediate did indeed form, but also that the ligation proceeds in a highly chemoselective manner.

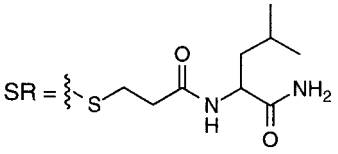
Coupling between **42.1** and **42.2** to give **42.3** (Scheme 42) proceeds readily in solution at pH 6.8, but very slowly when the pH is below 6.0.[63] In addition, when coupling between **43.1**, made as described below, and **43.2** (Scheme 43) was carried out at pH 7.0, reaction was essentially complete after only 5 min. However, when the same components were coupled at pH 5.0, the reaction was only about 50% complete after 10 min.[63] These observations are qualitatively in accord with the plausible mechanism that the thiolate is the reactive nucleophile.

The coupling of **43.1** and **43.2** showed the expected influence of the sulfur leaving group on the ligation process. The former compound was made from the corresponding thioacid by treatment with 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman's reagent). This transformation is assumed to involve the intermediacy of a short-lived disulfide [-C(O)SSC$_6$H$_3$(NO$_2$)CO$_2$H], which rapidly affords the thioester **43.1**.[66] The coupling experiment confirmed that the nature of the thioester leaving group influences the rate of reaction; the initial investigations into coupling by Native Chemical Ligation had employed benzyl thiol as the leaving group (cf. **42.1**) but, in the present case, 5-mercapto-2-nitrobenzoic acid serves in that role, and the result was much faster coupling.[63] In a subsequent experiment, addition of thiophenol to a coupling reaction that used a benzyl thiol leaving group also enhanced the rate of the coupling.[67] Presumably, the increase is due to the fact that the PhS group replaces BnS prior to ligation and, therefore, provides a better leaving group. This technique of in situ conversion to a more reactive species has been used in the synthesis of peptides with over 100 amino acid residues.[67]

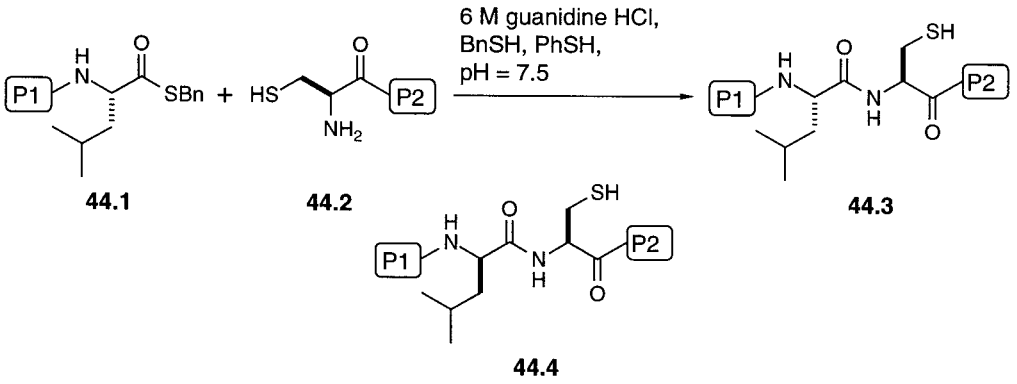
A practical aspect of the coupling reaction is the need to suppress oxidation of free cysteine thiol groups. Oxidation of this type renders the ligation step ineffective, since it decreases the nucleophilic character of the sulfur atom on the *N*-terminal cysteine, and prevents ligation. It was established,[63] fortunately, that this problem could easily be circumvented by carrying out the ligation in the presence of an excess of the thiol corresponding to the thioester leaving group.

Despite the fact that coupling in the Native Chemical Ligation Strategy is carried out with completely unprotected peptide segments, no incompatibility with the presence of other cysteine residues was apparent.[63] Presumably, this results from the fact that inter- or intramolecular reaction of cysteine residues, other than the intended one, with the thioester is unproductive and also reversible. The necessary juxtaposition of the required groups for acyl transfer can realistically be achieved only by reaction of the *N*-terminal

Table 8. Effect of *N*-terminal residue (X) of *C*-terminal segment on coupling of LYRAX-SR and CRANK by Native Chemical Ligation[68]



X	Coupling Time (h)	X	Coupling Time (h)	X	Coupling Time (h)	X	Coupling Time (h)
Gly	≤4	Ala	≤9	Arg	≤24	Ile	≥48
Cys	≤4	Met	≤9	Asn	≤24	Leu	≥48
His	≤4	Phe	≤9	Asp	≤24	Pro	≥48
		Trp	≤9	Gln	≤24	Thr	≥48
		Tyr	≤9	Glu	≤24	Val	≥48
				Lys	≤24		
				Ser	≤24		



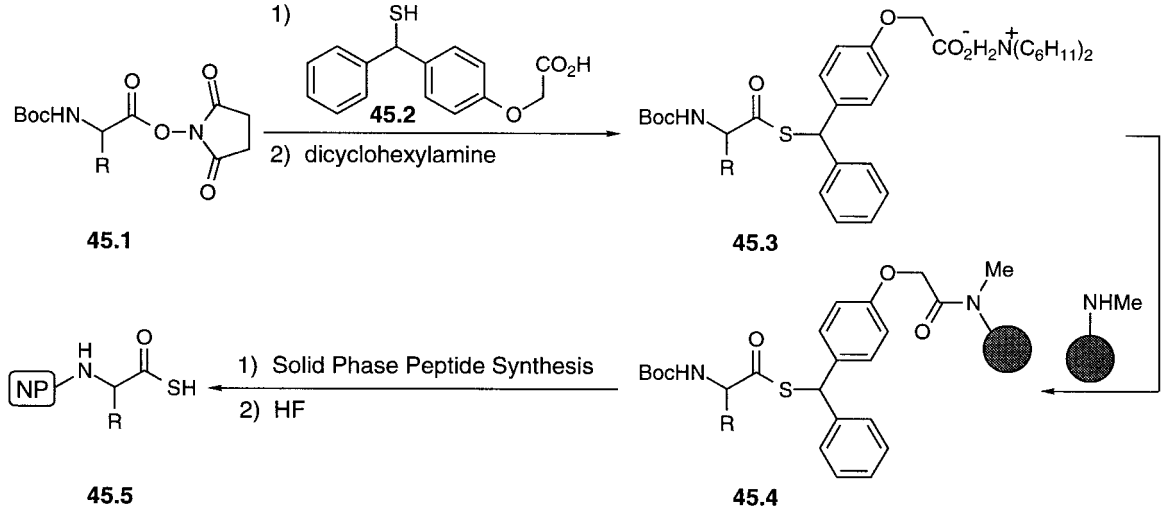
Scheme 44. P1=PACTLEYRP; P2=GSDNKTYGN.

cysteine of the *C*-terminal peptide (cf. **41.2**) with the thioester substituent of the *N*-terminal peptide (cf. **41.1**). Reaction of other cysteines with the thioester will undoubtedly occur, but does not result in a system in which an amino group and the acyl group are sufficiently close to permit rapid acyl transfer. Instead, the products of such undesired ligation would be expected to react with the excess of the thiol corresponding to the leaving group, and revert back to starting material. Eventually, a productive ligation occurs, and is followed by the desired acyl transfer.

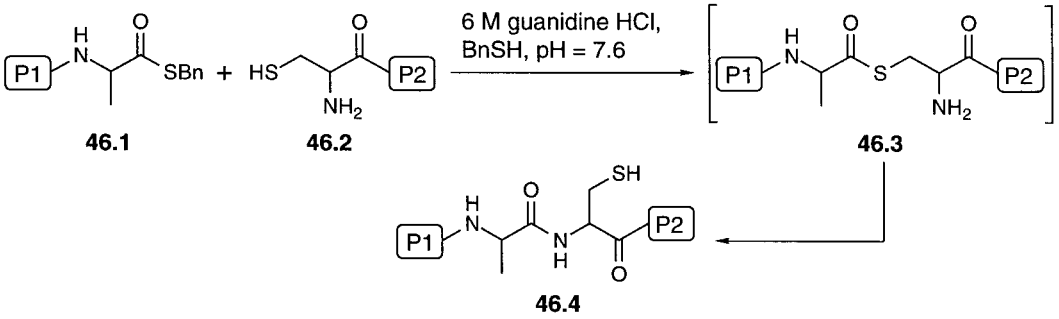
An investigation into the effect of the *C*-terminal amino acid of the *N*-terminal segment on the rate of coupling in the Native Chemical Ligation Strategy has been carried out. This was done by examining coupling of the *N*-terminal segment LYRAX-SR and the *C*-terminal segment CRANK.[68] In this case, X-SR is an amino acid thioester[69] corresponding to one of the 20 naturally occurring amino acids. Each of the 20 possible thioesters was examined and the amino acids were then grouped according to the time required for the coupling to reach completion (see Table 8). Ligation proceeds efficiently in all cases, except when the thioester-derivatized amino acid is β-branched or is proline. Interestingly, when the thioester was a histidine or cysteine residue, coupling occurred at approximately the same rate as with the sterically unhindered glycine residue. It was suggested that this result might be due to a catalytic effect displayed by the side-chain (thiol or imidazole) on the rate-limiting thioester exchange.

5.1.2. Extent of epimerization. The extent, if any, of epimerization during the coupling process was monitored in the following way.[70] The model peptides **44.1** and **44.2** were prepared and allowed to react in 6 M guanidine hydrochloride at pH 7.5, in the presence of 1% benzyl thiol and 3% thiophenol (Scheme 44), to give **44.3**. The leucine epimer (**44.4**) of this compound was also made, but by standard solid phase peptide synthesis. HPLC conditions were found that allowed baseline resolution of the two epimers (**44.3** and **44.4**) and, when the product of the ligation reaction was analyzed, no epimerization product was detected; and if any had occurred it must have been to the extent of $<1\%$.

5.1.3. Synthesis of peptide segments. The peptide segments used for the cysteine-based Native Chemical Ligation Strategy were obtained by solid phase peptide synthesis. The synthesis of the *C*-terminal peptides was straightforward, given that they possessed no unusual features. However, synthesis of the *N*-terminal peptides—which are derivatized at their carboxyl termini as thioesters—required some developmental work. The thioesters could be obtained readily from the corresponding thioacids



Scheme 45. NP=*N*-terminal peptide; ● =polystyrene resin.



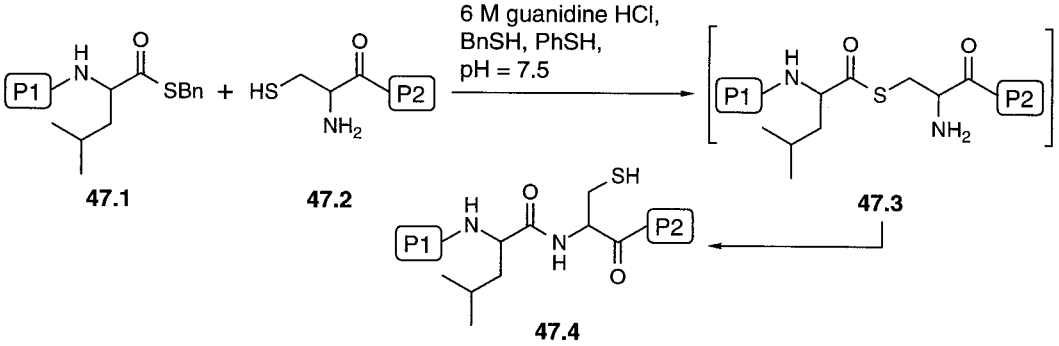
Scheme 46. P1=SAKELRCQCIKTYSKPFHPKFIKELRVIESGP; P2=ANTEIIVKLSDGRELCLDPKENWVQRVVEKFLKRAENS.

by reaction with an excess of benzyl bromide in 6 M guanidine hydrochloride at pH 4.6 for the benzyl esters, or with 5,5′-dithiobis(2-nitrobenzoic acid) in 8 M urea at pH 4.0 for the 5-mercapto-2-nitrobenzoic acid thioesters.[63] In order to obtain the starting thioacid-derivatized peptide, however, an optimized form of solid phase synthesis, based on the thioester-linked solid support **45.4** (Scheme 45), had to be used.[70–72] This support was generated[71] by reacting an appropriately protected *N*-hydroxysuccinimide amino acid ester (**45.1**) with the benzylic thiol **45.2**[73] to give, after treatment with dicyclohexylamine, salt **45.3**. This salt was then coupled to an aminomethyl polystyrene resin,[72] generating the required solid support (**45.3→45.4**). At this point, standard solid phase peptide synthesis was carried out in order to produce the required peptide segment, and then treatment of the resulting resin-bound peptide with HF cleaved the product from the support, and released the required peptidyl thioacid (**45.5**).

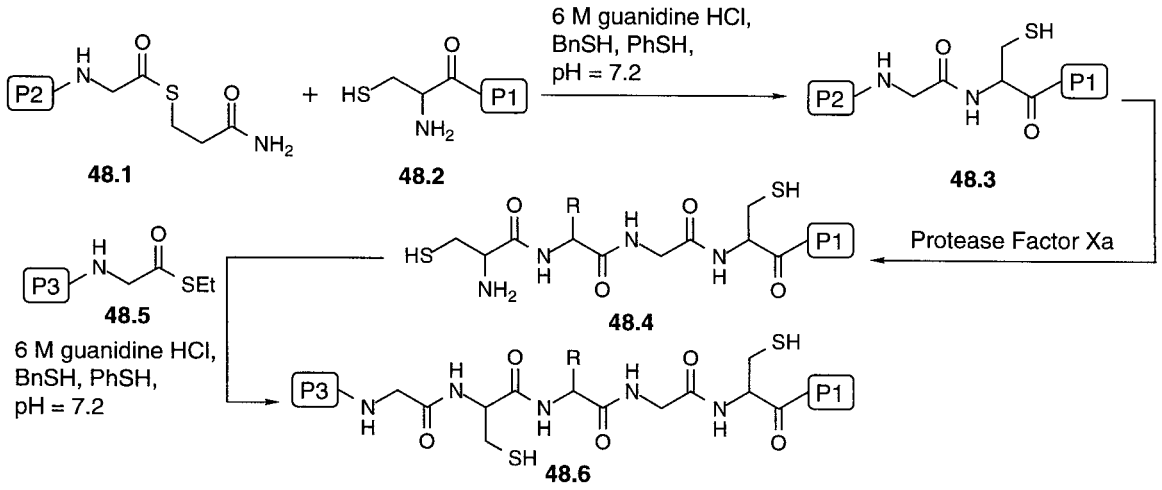
5.1.4. Application to peptide synthesis.[74] One of the first applications of the cysteine-based Native Chemical Ligation Strategy was in the synthesis of a *human interleukin 8* mutant (**46.4**, Scheme 46).[63] In this particular mutant, the histidine residue at position 33 was substituted by an alanine residue.[63] The protein was 72 amino acids in size, and contained a total of four cysteine residues. Its synthesis was achieved[63] by coupling the 33-residue *N*-terminal peptide segment **46.1** and the 39-residue *C*-terminal peptide segment **46.2**. Significantly, a total of 18 out of the 20 genetically encoded amino acids are represented in these segments. The site of coupling was chosen to be between cysteine 34 and alanine 33. The two peptide segments were allowed to react in 6 M guanidine hydrochloride at pH 7.6, in the presence of an excess of benzyl thiol, and underwent ligation to generate the acyl transfer intermediate (**46.1**+**46.2→46.3**). Intramolecular *S→N* acyl transfer (**46.3→46.4**) then gave the coupled product in approximately 60% yield after 48–72 h.

A second application of the cysteine-based Native Chemical Ligation Strategy was in the synthesis of the peptide corresponding to residues 6–56 of *turkey ovomucoid third domain*.[70] The synthesis of this peptide was achieved by coupling the 18-residue *N*-terminal peptide **47.1** (Scheme 47) and the 33-residue *C*-terminal peptide **47.2**. The coupling site was chosen to be between cysteine-24 and leucine-23. Peptides **47.1** and **47.2**, the former made[70] by the procedure summarized in Scheme 45, and the latter manually synthesized by an optimized[70] solid phase method based on Boc chemistry, underwent smooth ligation in 6 M guanidine hydrochloride at pH 7.5, in the presence of 1% benzyl thiol and 3% thiophenol, to give the acyl transfer intermediate **47.3**, which then rearranged to the coupled product **47.4**. Peptide **47.4** was isolated from the mixture after 36 h, and purified by reverse-phase HPLC. The isolated yield was typically about 56%.

One final noteworthy application of the Native Chemical Ligation Strategy is the formation of the 164-residue peptide **48.6** (Scheme 48).[76] In this example, the peptide segments were recombinantly derived, and so the approach serves to demonstrate the versatility, as well as the power of Native Chemical Ligation for coupling large segments. Since two sequential couplings were used, the potential for iteration was also demonstrated. The first coupling was between the 105-residue *C*-terminal peptide **48.2**, and the thioester-derivatized decapeptide **48.1**, and was achieved by storing the two components for 96 h in the presence of 6 M



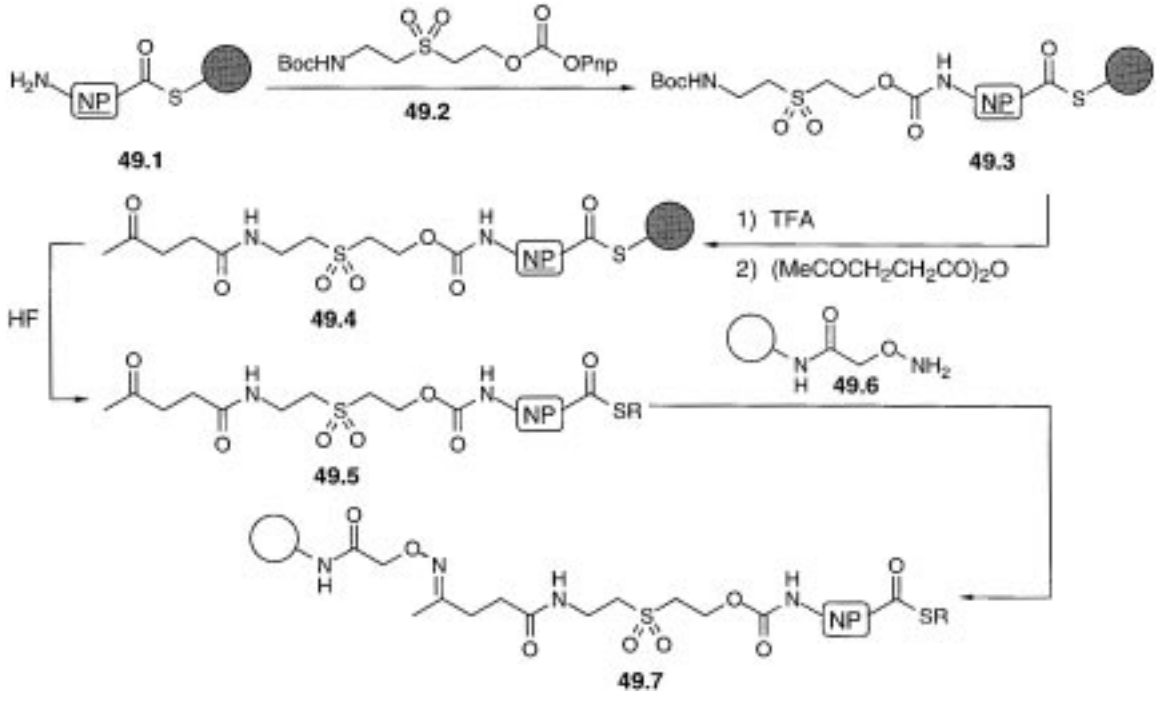
Scheme 47. P1=VDCSEYPKPACTLEYRP; P2=GSDNKTYGNKCNFCNAVVESNGTLTLSHFGKC.



Scheme 48. P1=LEKHSWYHGPVSRNAAEYLLSSGINGSFLVRESESSPGQRSISLRYEGRVYHYRINTASDGKLYVSSESRFNTLAELVHHHSTVADGLITTLHYPAPKRGIHRD; P2=RGKIEGRCK(Dns); P3=MLFVALYDFVASGDNTLSITKGEKLRVLGYNHNGEWAEAQTKNGQGWVPSNYITPV; R=$(CH_2)_4$NH-Dns; Dns=[5-(dimethylamino)-1-naphthyl]sulfonyl.

guanidine hydrochloride at pH 7.2, along with 1.5% benzyl thiol and 1.5% thiophenol (**48.1**+**48.2**→**48.3**). In this case, the decapeptide served as a masked form of the thioester-derivatized tripeptide CK(Dns)GSCH$_2$CH$_2$CONH$_2$. The additional seven residues of the decapeptide functioned as a protecting group for the *N*-terminal cysteine, and were removed enzymatically after the first coupling, by exposure to protease factor Xa (**48.3**→**48.4**). At this point, the stage was set for the second coupling, which was achieved by treating **48.4** with the 56-residue, thioester-derivatized peptide **48.5** under the same conditions used for the first coupling (**48.4**+**48.5**→**48.6**). The overall yield for the entire process was 28%.

5.1.5. Application to the solid phase (Solid Phase Chemical Ligation). The Native Chemical Ligation Strategy has recently been extended to allow coupling in the solid phase.[77] This can be accomplished using either of two separate approaches which Kent calls *Solid Phase Chemical Ligation* techniques. In the first approach, the thioester-derivatized *N*-terminal peptide is anchored to a resin through its *N*-terminus and is treated with a *C*-terminal peptide which has at its *N*-terminus a cysteine residue. Ligation and acyl transfer then proceed in the usual manner, with the exception that the coupled product remains attached to the solid support. Thus, peptide synthesis occurs on the solid support in the *N* to *C* direction. In the second



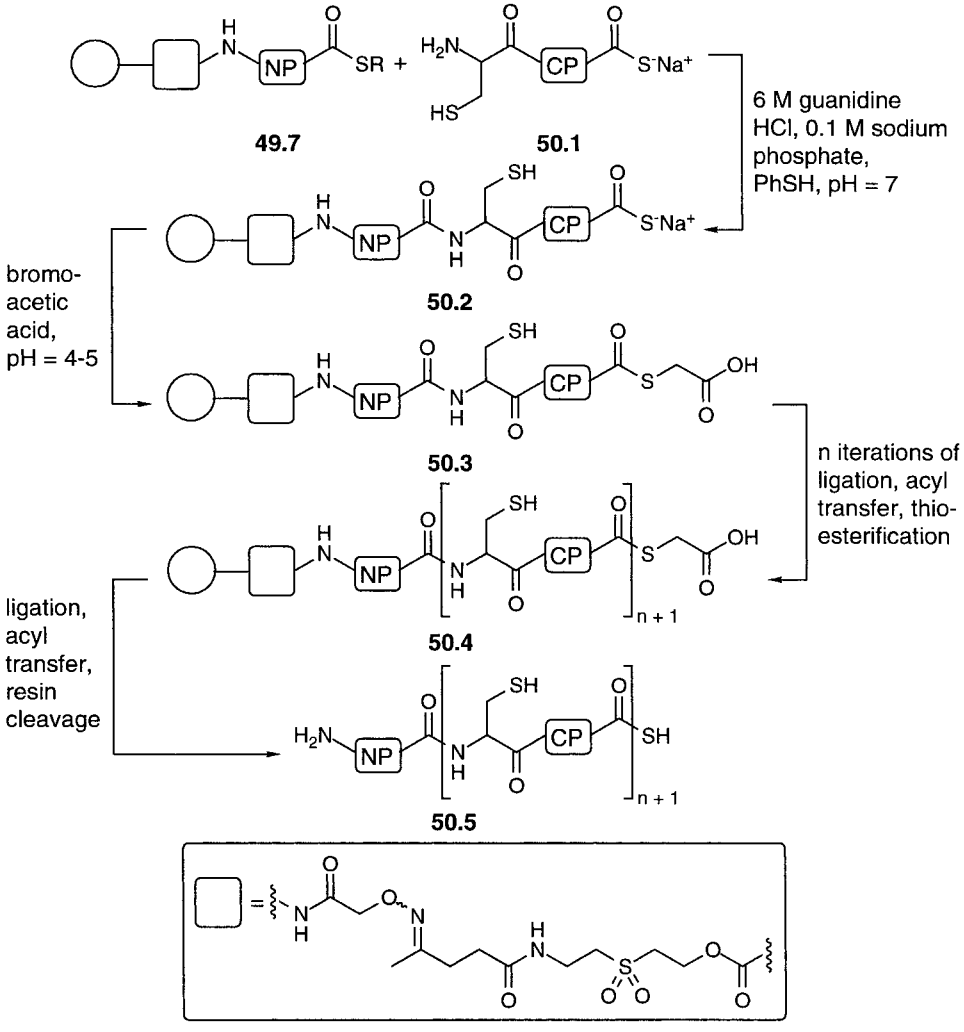
Scheme 49. NP=protected *N*-terminal peptide; NP=unprotected *N*-terminal peptide; Pnp=*p*-nitrophenyl; ●=polystyrene resin; ○=cellulose resin.

approach, coupling is carried out in the *C* to *N* direction by first attaching a *C*-terminal peptide bearing an *N*-terminal cysteine residue to the solid support via its *C*-terminus, and then treating the product with a thioester-derivatized *N*-terminal peptide. Again, ligation and acyl transfer proceed as usual, but result in a resin-bound coupled product.

During the development of this solid phase segment coupling approach, Kent was able to incorporate into the methodology the option of carrying out the process in an iterative fashion. In this case, the product of the first coupling reaction becomes the resin-bound peptide used in the second coupling cycle, and the product of that cycle, in turn, is then the resin-bound component in the next cycle. The iterative coupling requires that the non resin-bound peptide, which possesses both an *N*-terminal cysteine residue and a *C*-terminal thioester, be introduced in a masked form in which one or the other of its termini (which one would depend on which of the two approaches was being used) is suitably masked. Protection of this type is necessary to avoid unwanted cyclization or, possibly, polymerization of the non resin-bound peptide. Following ligation and acyl transfer, the protecting group is removed, and the product is ready to enter into another coupling cycle (see later).

5.1.5.1. Solid phase chemical ligation in the *N* to *C* direction.[77] Solid Phase Chemical Ligation in the *N* to *C* direction requires that the original *N*-terminal peptide be attached to the solid support. To achieve this, routine stepwise solid phase peptide synthesis was used to generate the desired *N*-terminal peptide (**49.1**, Scheme 49) but, prior to deprotection and resin cleavage, the unprotected α-amino residue was derivatized with *p*-nitrophenyl carbonate **49.2**[78] to give the Boc-protected amine **49.3**. Removal of the Boc group was followed by acylation,[79] and gave the keto derivative **49.4**. This species was subjected to global deprotection and resin cleavage (**49.4→49.5**), resulting in a free unprotected peptide possessing a cleavable linker at its amino terminus.[80] The ketonic group of the linker was employed, via oxime formation, as the site of attachment to a water-compatible, cellulose-based aminooxy acetic acid-derivatized resin (**49.6**), resulting in formation of **49.7**.

At this point Native Chemical Ligation could be carried out between the resin-bound thioester **49.7** and the first



Scheme 50. NP=*N*-terminal peptide; CP=*C*-terminal peptide; ◯=cellulose resin.

C-terminal segment (**50.1**, Scheme 50). In this process unwanted side reactions involving the *C*-terminus of the non resin-bound peptide were suppressed by introducing this peptide in the form of a thiocarboxylate ion. The *C*-terminal carbonyl in this form was found[77] to be sufficiently unreactive towards the *N*-terminal cysteine thiol, provided the thioester-containing *N*-terminal peptide (**49.7**) was present. However, in the absence of the thioester, detectable cyclization of the non resin-bound peptide did occur.[77] Ligation and acyl transfer (**49.7**+**50.1**→**50.2**) were carried out at pH 7 in the presence of 1% thiophenol. Once coupling was complete, the pH was lowered to 4–5 and the thiocarboxylate was converted into a thioester by treatment with bromoacetic acid (**50.2**→**50.3**). Excess reagent was washed away, and the pH was returned to 7 in preparation for another coupling cycle. Following the desired number of iterations, the peptide was liberated from the solid support by cleavage with aqueous sodium hydroxide at pH 12–14 (**50.3**→**50.4**→**50.5**), and then purified by HPLC.

Solid Phase Chemical Ligation in the *N* to *C* direction has been applied to the synthesis of three peptides ranging in size from 68 to 115 amino acid residues.[77] In each case, two iterations of the coupling cycle were employed to reach the target, and so the process involved the use of one resin-bound *N*-terminal peptide and two non resin-bound *C*-terminal peptides.

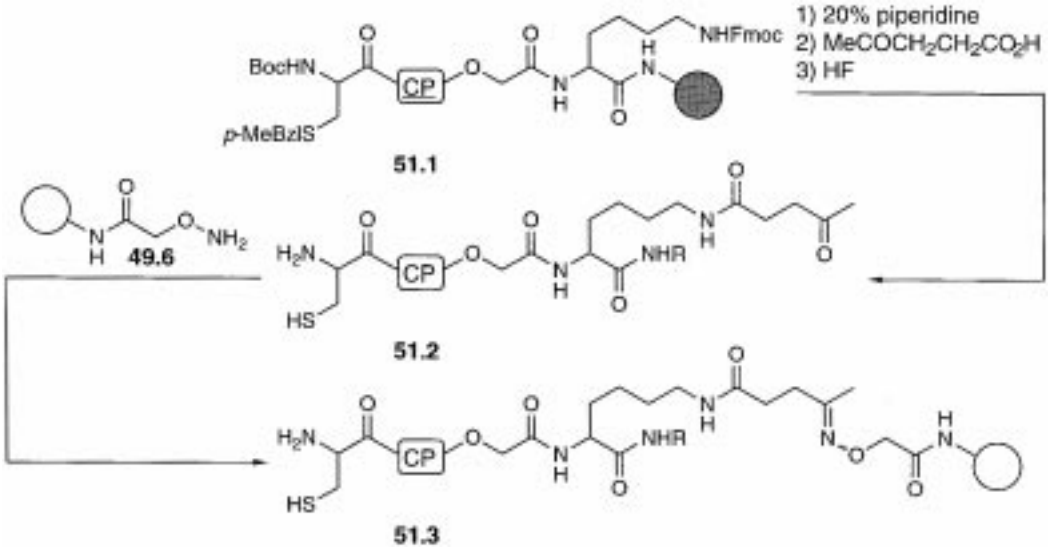
5.1.5.2. Solid phase chemical ligation in the *C* to *N* direction.[77]Preparation of the resin-linked *C*-terminal peptide for use in Solid Phase Chemical Ligation in the *C* to *N* direction is outlined in Scheme 51. The desired peptide segment was built up using standard stepwise solid phase techniques from a support incorporating a linker which contained a carboxyamidomethyl group attached to an ϵ-amino Fmoc-protected lysine residue. Once the desired peptide (**51.1**) had been generated, the Fmoc group was removed by treatment with 20% piperidine, and the liberated ϵ-amino group was acylated with levulinic acid. Global deprotection using HF was then effected and the resulting *C*-terminal-derivatized peptide **51.2** was isolated and purified by HPLC. The ketone functional group was used to anchor the newly formed *C*-terminal peptide to a water-compatible, cellulose-based aminooxy acetic acid-derivatized resin via formation of an oxime (**51.2**+**49.6**→**51.3**).

The non resin-bound, thioester-derivatized *N*-terminal peptide **52.1** (Scheme 52) used in the ligation/acyl transfer reaction was obtained by standard solid phase techniques and contained an *N*-terminal Acm-protected cysteine to prevent undesired side reactions. Ligation and acyl transfer (**52.1**+**51.3**→**52.2**) were carried out at pH 7 in the presence of 1% thiophenol and, once complete, excess reagents were washed away and the Acm group was removed by treatment with $Hg(OAc)_2$ and AcOH (**52.2**→**52.3**). The product so obtained could then be used in subsequent coupling cycles and, once the final peptide had been generated, it could be cleaved from the solid support by treatment with aqueous sodium hydroxide at pH 12–14 (**52.3**→**52.4**) and purified by HPLC.

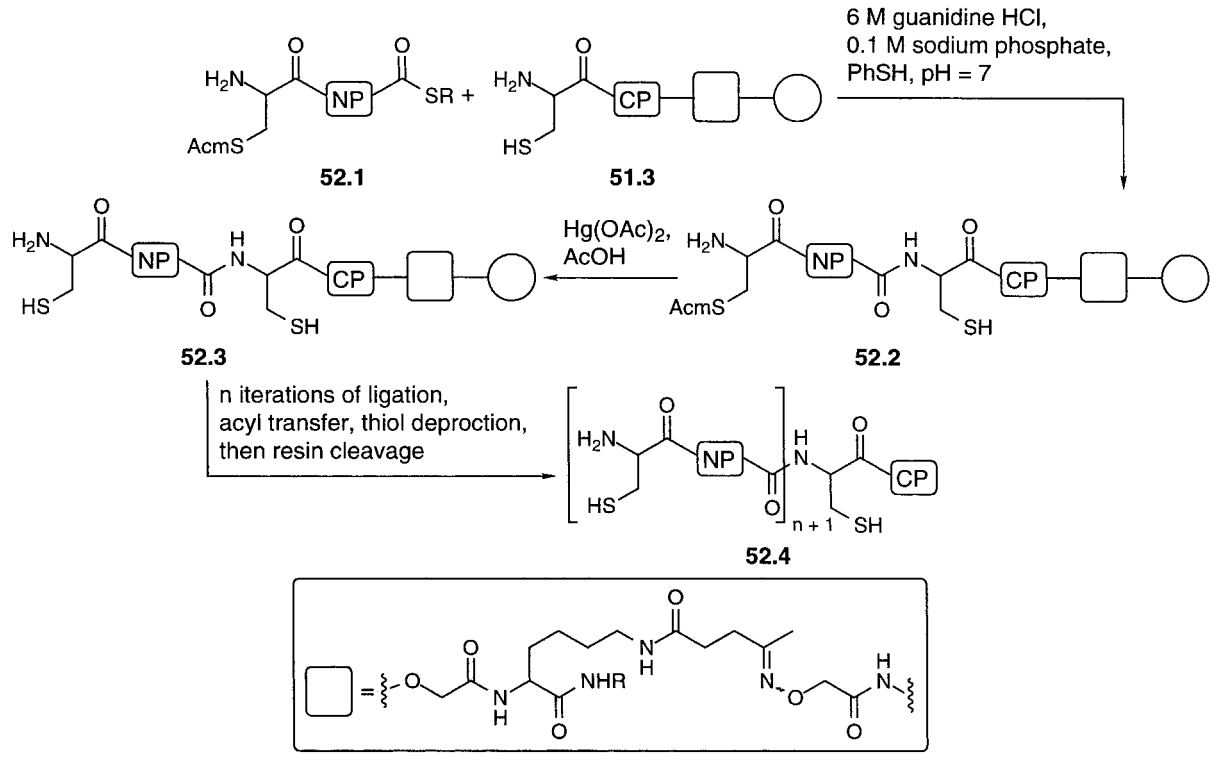
Solid Phase Chemical Ligation in the *C* to *N* direction has been applied to the synthesis of a 27-residue model peptide using two iterations of the coupling sequence, as well as to the synthesis of a 118-residue protein using three iterations of the sequence.[77]

5.2. Ligation by thioester exchange involving derivatized glycine or alanine

The second approach to peptide segment coupling by Native Chemical Ligation avoids the strict requirement for an *N*-terminal cysteine on the *C*-terminal peptide segment, and the scope of the process was expanded to permit coupling between either glycine or alanine in one segment and some other amino acid in the other segment. This version[75] of Native Chemical Ligation is outlined in Scheme 53. As in the cysteine-based approach, ligation occurs by nucleophilic attack by the thiol component (**53.2a**,**b**) on a thioester (**53.1**). The thiol component in this case is a glycine or alanine residue that is derivatized as either an



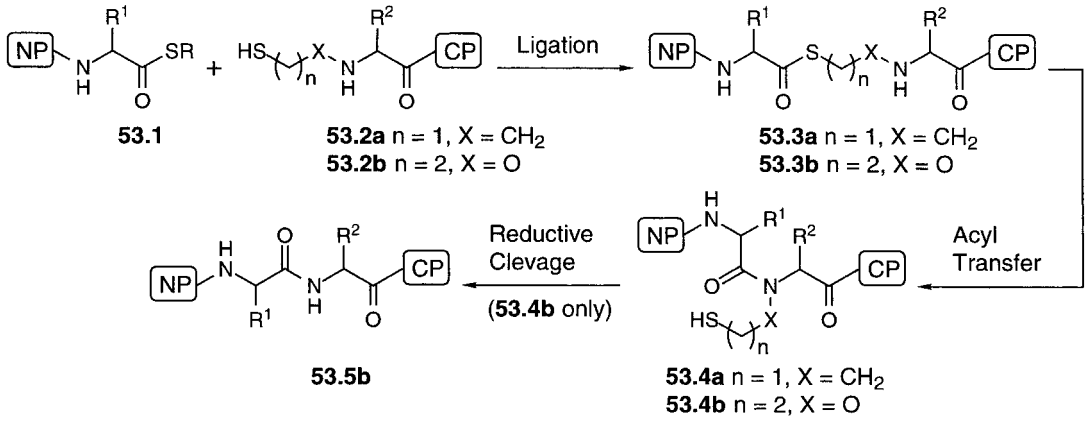
Scheme 51. CP=protected *C*-terminal peptide; CP=unprotected *C*-terminal peptide; =polystyrene resin; =cellulose resin.



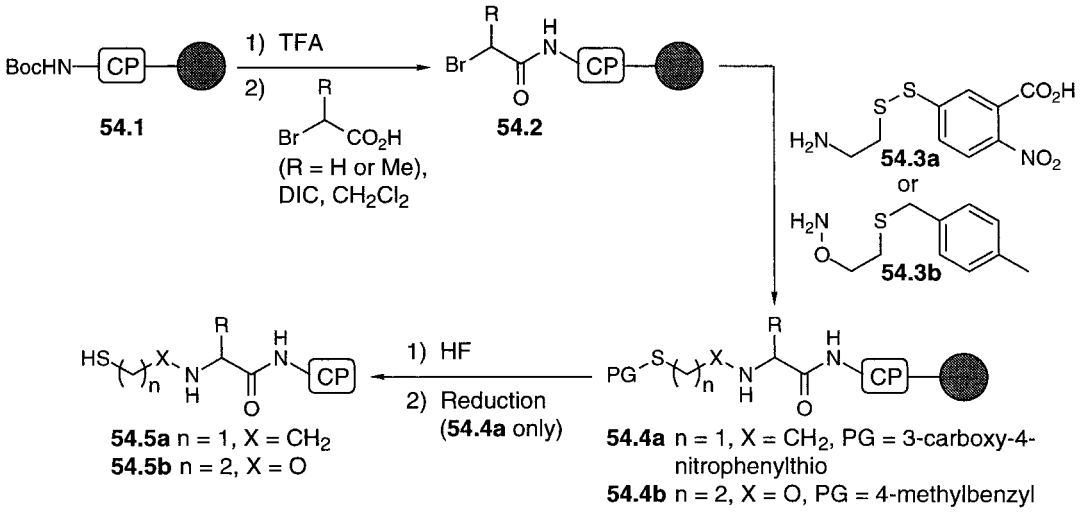
Scheme 52. NP=*N*-terminal peptide; CP=*C*-terminal peptide; ◯=cellulose resin.

N-ethanethiol (**53.2a**) or an *N*-(oxyethane)thiol (**53.2b**). These derivatized residues constitute the *N*-terminal amino acids of the corresponding *C*-terminal peptide chains. The product of the ligation (**53.3a** or **53.3b**) is a new thioester in which the thioacyl and α-amino groups are close to one another. Thus, a proximity induced *S*→*N* acyl transfer takes place to give the coupled peptide (**53.4a**,**b**). In the case of **53.4b**, the linking element used to correctly position the acyl and amino groups is then removed by reductive cleavage (**53.4b**→**53.5b**). However, for compound **53.4a**, no simple method for removing the linking element is available, and so a coupled product containing a non-standard amino acid is obtained.

Synthesis of the thioester-derivatized peptides **53.1** was achieved using the same approach as for the cysteine-based coupling strategy described above (cf. Scheme 45). However, the *C*-terminal peptides in the present case were made using a special approach that is summarized in Scheme 54. Standard solid phase peptide synthesis was used to generate the resin-bound peptide **54.1**. This was deprotected at its amino terminus, and then coupled either with bromoacetic acid (**54.1**→**54.2**, R=H) or with (±)-2-bromopropanoic acid (**54.1**→**54.2**, R=Me). The bromide so obtained was then displaced using either amine **54.3a** or **54.3b** to obtain compounds **54.4a** and **54.4b**, respectively. Treatment of **54.4b** with HF then served both to deprotect the peptide and cleave it from the solid support (**54.4b**→**54.5b**). In the case of **54.4a** the same process was applied, but a final reductive deprotection step had to be performed in order to liberate the terminal thiol, which had been protected as a disulfide (**54.4a**→**54.5a**). In the



Scheme 53. R=Bn, or 5-thio-2-nitrobenzoic acid; R^1=amino acid side-chain; R^2=H, or Me; NP=*N*-terminal peptide; CP=*C*-terminal peptide.



Scheme 54. CP=*C*-terminal peptide; ⬤=polystyrene resin.

Table 9. Model peptides used to study the approach of Scheme 53[75]

Model ligation	*N*-Terminal peptide (**53.1**)	*C*-Terminal peptide (**53.2**)
1	**53.1** R=Tnb,[a] R^1=H, NP=LYRA	**53.2a** R^2=H, CP=AGPAGD-NH_2
2	**53.1** R=Tnb, R^1=H, NP=LYRA	**53.2b** R^2=H, CP=RNTATIMMQRGNFR-NH_2
3	**53.1** R=Bn, R^1=Bn, NP=LYRA	**53.2b** R^2=H, CP=RNTATIMMQRGNFR-NH_2
4	**53.1** R=Bn, R^1=H, NP=LYRA	**53.2b** R^2=Me, CP=ARHTVHQRHLHG
5	**53.1** R=Bn, R^1=Bn, NP=LYRA	**53.2b** R^2=Me, CP=ARHTVHQRHLHG

[a] Tnb=5-thio-2-nitrobenzoic acid.

case where (±)-2-bromopropanoic acid was used to generate **54.2**, a mixture of peptides, epimeric at the *N*-terminal alanine reside, was obtained.[81]

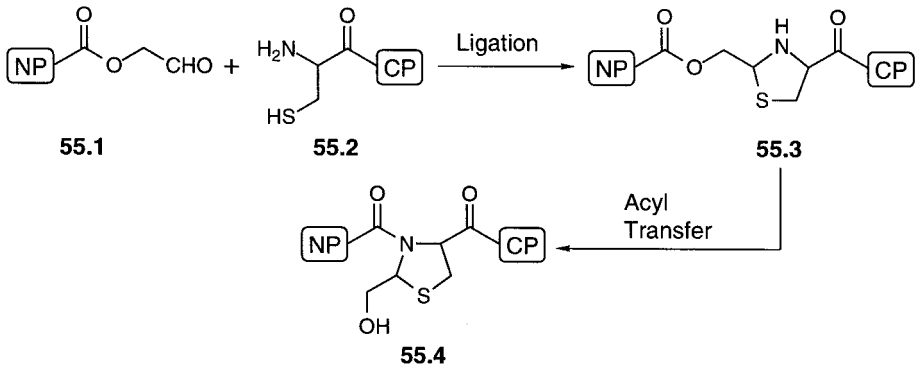
In order to test the efficiency of the coupling process outlined in Scheme 53, several model studies were undertaken. The peptides used for this purpose are shown in Table 9, and the results of the coupling reactions are collected in Table 10. Ligations 1 and 2 (see Table 10) established that straightforward coupling could be achieved using either the *N*-ethanethiol (**53.2a**) or the *N*-oxyethanethiol (**53.2b**) derivative. The reactions proceeded in good yield and within a reasonable time. The third model ligation provided an example of coupling of a more sterically hindered system and, compared to ligations 1 and 2, the rate of coupling was noticeably slower and gave poorer yields. A similar rate and yield decrease was seen for the fourth model ligation which, like the third, proceeded through a more sterically hindered transition state than either of the first two examples. Ligation 5 involved a coupling in which both the amino acids to be linked were substituted at their respective α-carbons. In this case the degree of steric hindrance was too severe and no coupling product was observed.

Notable features of the data in Tables 9 and 10 are the low rates of acyl transfer with sterically hindered systems, compared to the cysteine-based Native Chemical Ligation Strategy described earlier. In fact, the acyl transfer rates were so low that in each of the final three model ligations

Table 10. Results of model study for the approach of Scheme 53[75]

Model ligation	pH	Time (h)	Temperature (°C)	Approximate yield (%)	
				Unrearranged product (**53.3**)	Rearranged product (**53.4**)
1	7.0	4	25	Not detected	90
2	7.5	16[a]	25	Not detected	75
3	7.5 then	11.5	37	30	35
	4.5	10		0	64
4	7.5 then	17.5	37	39	52
	4.5	6.5		20	69
5	7.5 then	19	37	58	Not detected
	4.5	22		52	Not detected

[a] >80% complete after 1 h; extended reaction time was to ensure complete acyl transfer.



Scheme 55. NP=*N*-terminal peptide; CP=*C*-terminal peptide.

the unrearranged product was stable enough to be isolated. Thus, for practical purposes, the present method would appear to be limited to glycine at the ligation site; in the case of alanine, it has not yet been established that the key displacement (**54.2→54.4a**,**b**) can be done in a stereocontrolled way.

6. Orthogonal Ligation Coupling Strategies

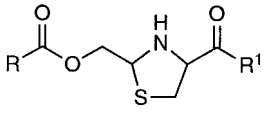
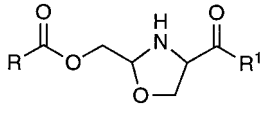
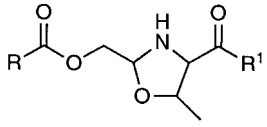
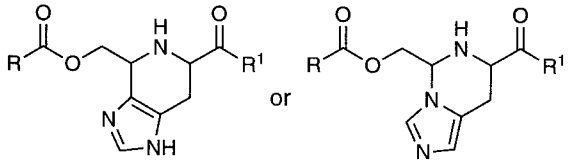
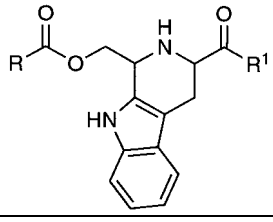
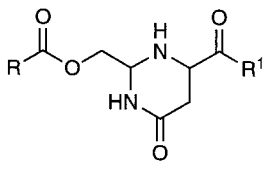
Several methods for coupling peptide segments—collectively referred to as *Orthogonal Ligation Coupling Strategies*—have been developed by Tam. Each of these methods is characterized by the type of reaction used to ligate the peptide segments prior to the peptide bond-forming acyl transfer reaction and, in all but two examples, the site of segment coupling occurs at a cysteine, a modified cysteine, or at a residue that becomes a cysteine in the final product, as described below.

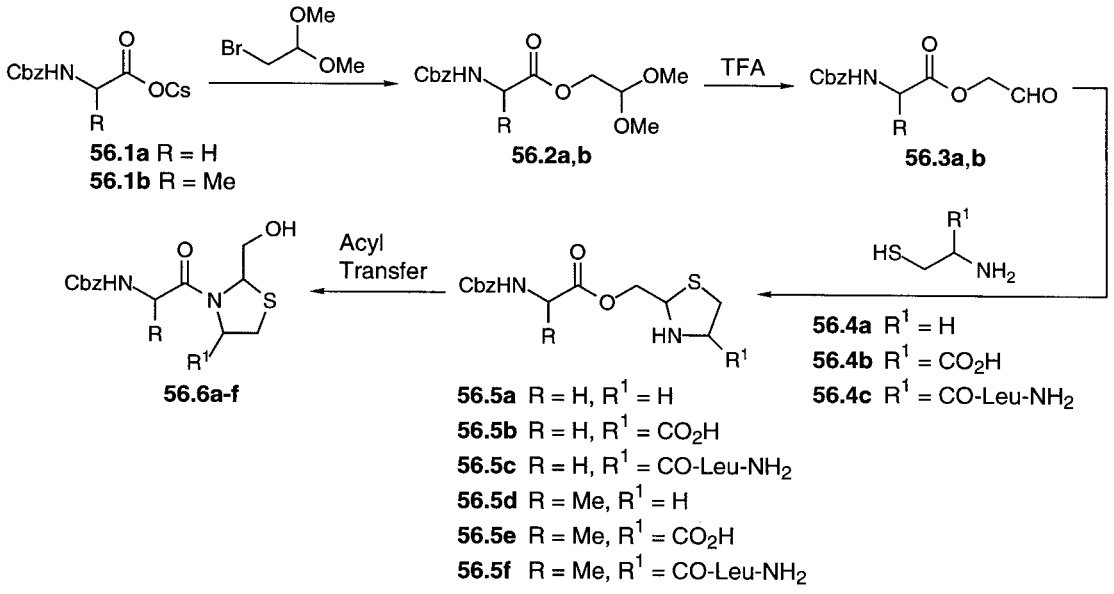
6.1. Ligation by thiazolidine formation

The first approach developed by Tam is one in which ligation of the peptide segments occurs by generation of a thiazolidine, and the general principle is outlined in Scheme 55.[23,82] The *C*-terminus of the *N*-terminal peptide segment is derivatized as a glycoaldehyde (**55.1**). The aldehyde group condenses with a cysteine residue located at the *N*-terminus of the other segment (**55.2**). The condensation is highly selective and gives the acyl transfer intermediate containing a thiazolidine (**55.1**+**55.2→55.3**). Acyl transfer proceeds through a five-membered transition state from this species to give the rearranged system **55.4**, which possesses a new peptide bond. However, unlike most of the segment coupling approaches described so far, the new peptide bond is not a part of a native peptide backbone because the peptide chain incorporates a thiazolidine at the site of coupling.

6.1.1. Mechanistic and model studies. Ligation of the peptide segments by thiazolidine formation occurs in a highly specific manner even though unprotected peptides are used. This results from the fact that, of the potentially many different ligation products that could form from nucleophilic attack on the aldehyde, only the thiazolidine (cf. **55.3**) is sufficiently stable in the aqueous environment in which the coupling is done and so is present in a significant amount. Reactions leading to other ligation products, such as imines or acetals, would be readily

Table 11. Heterocyclic ligation products resulting from reaction between glycoaldehyde-derivatized peptides and various *N*-terminal residues[83]

N-Terminus of *C*-terminal peptide	Ligation product	*N*-Terminus of *C*-terminal peptide	Ligation product
Cysteine		Serine	
Threonine		Histidine	
Tryptophan		Asparagine	



Scheme 56.

reversible under the conditions used, and none of these species would be expected to be present in a significant concentration.

The thiazolidine unit can be formed only by reaction between the aldehyde group of the glycoaldehyde-derivatized peptide (cf. **55.1**) and the amino and thiol groups of a cysteine residue at the *N*-terminus of the other peptide (cf. **55.2**). This high specificity requires that the *N*-terminus of the glycoaldehyde-derivatized peptide should be incapable of forming a stable heterocycle (cf. Table 11) by reacting with the glycoaldehyde unit. When this requirement is satisfied, no residue other than the *N*-terminal cysteine of the *C*-terminal peptide would have two nucleophilic groups suitably disposed to allow formation of a stable heterocycle.

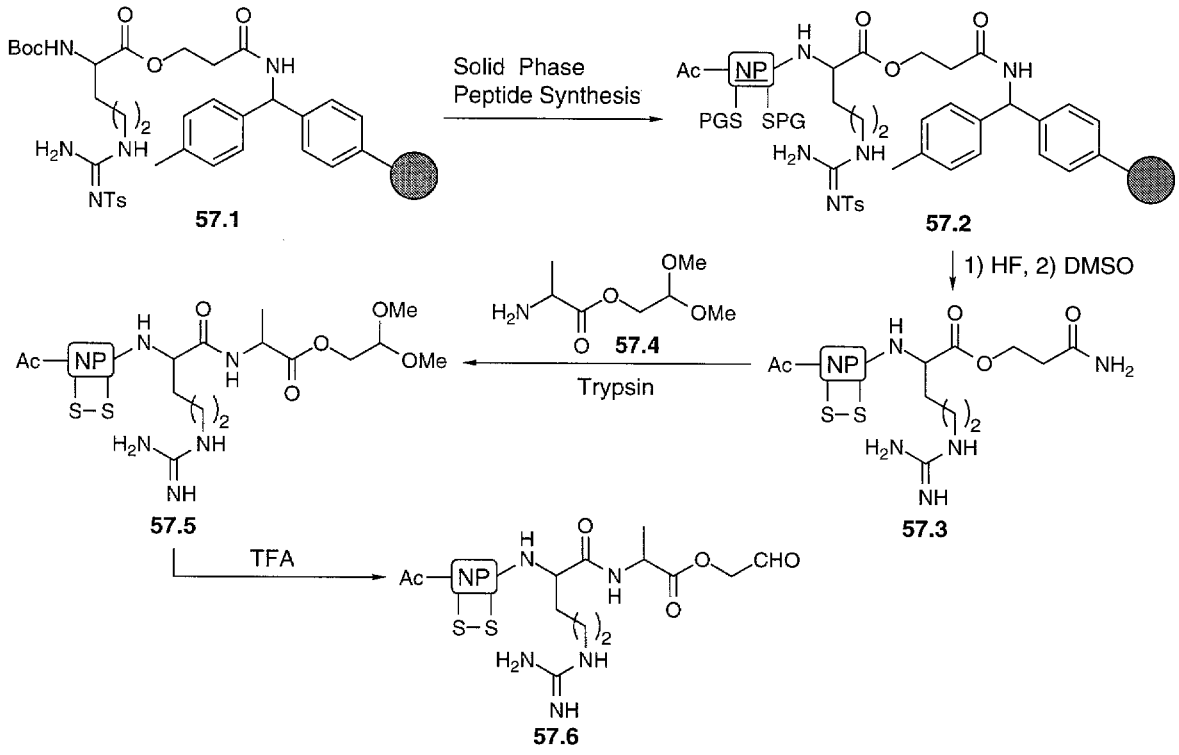
The possibility of effecting ligation using an *N*-terminal residue other than cysteine, but still having a nucleophilic side-chain, has also been explored.[83] A stable heterocyclic ligation product can be envisioned when either threonine, serine, tryptophan, histidine or asparagine is the *N*-terminus of the *C*-terminal peptide (see Table 11), and a special technique[83] was developed to establish whether or not such heterocycles do actually form under ligation conditions. The experiments revealed that each of the expected heterocycles was indeed produced, but that cysteine was the most effective in terms of rapidity and completeness of reaction. Next to cysteine was threonine, which did react completely, but at a significantly lower rate. Tryptophan, histidine and serine were found to react even more slowly than threonine, and only the reaction involving tryptophan went to completion. When asparagine was used as the *N*-terminal residue, almost no reaction was observed.

Initial model studies[23,82] (Scheme 56) for the thiazolidine-based ligation involved reacting a glycoaldehyde derivative of glycine or alanine (**56.3a**,**b**) with certain 1,2-aminothiols (**56.4a**–**c**). In order to generate the required glycoaldehyde derivatives, the cesium salts **56.1a**,**b** of the corresponding *N*α-Cbz-protected amino acids were treated with bromoacetaldehyde dimethyl acetal, to give compounds **56.2a**,**b**, from which aldehydes **56.3a**,**b** were obtained by the action of TFA. At pHs between 5 and 6, the ligation products **56.5a**–**f** were obtained within approximately 15 min.[23] The reaction could be carried out at a pH as low as 2, but in such cases several hours were required for completion.[82] At either neutral or basic pH, ligation was complete in less than 5 min,[23] but competing ester hydrolysis was observed at basic pH. It was desirable to carry out the ligation at acidic pH (generally pH 4–5[82]), not only to avoid ester hydrolysis, but also to avoid complicating reactions with other nucleophiles, such as those on the side-chains of lysine and arginine.[23,82,83]

In general, once the ligation product had formed, the acyl transfer reaction (**56.5a**–**f**→**56.6a**–**f**) could be initiated by increasing the pH of the solution.[23] Half-times for the acyl transfer with three different systems were obtained at various pHs (see Table 12). From these data, it can be seen that for the 2-aminoethanethiol derivatives **56.5a**,**d** an increase in pH gave a corresponding increase in the acyl transfer rate. In the case of the acyl transfer intermediate **56.5e**, however, this trend was not observed, and the highest rate occurred at approximately pH 7. Likewise, compounds **56.5c** and **56.5f** were found to undergo acyl transfer most readily[23,82] at the relatively low pH of 4–5 (data not shown in Table 12), but the reactions were slow (half-time ca. 20 h[23] and 24 h,[82] respectively).

Table 12. Half-times for acyl transfer of **56.5a**,**d**,**e** at various pHs[23]

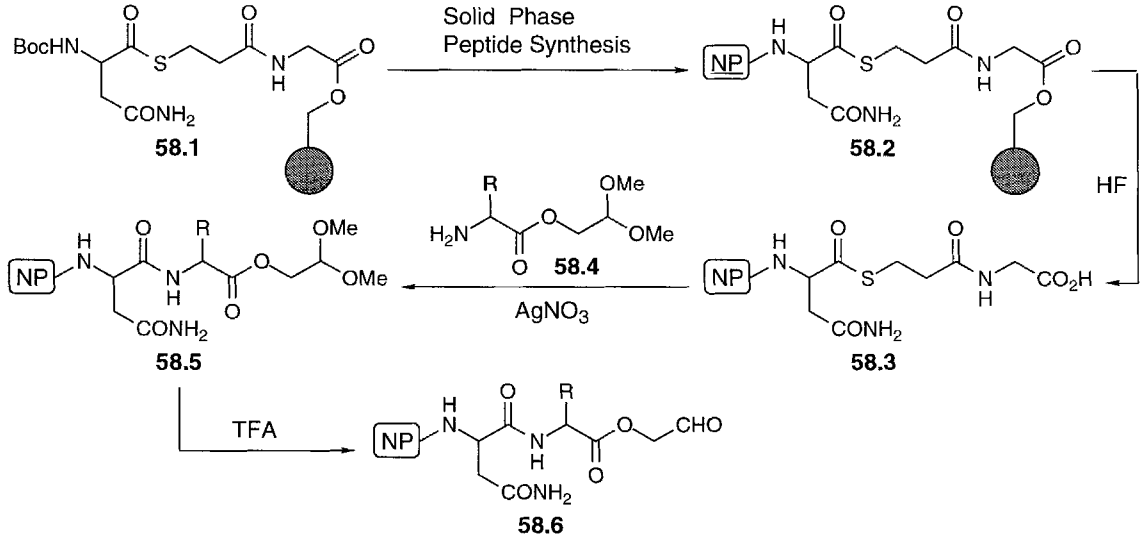
Substrate	Acyl transfer half-time (h)				
	pH 6	pH 7	pH 7.4	pH 8	pH 9
56.5a	21.5	8.8	–	3.4	2.8
56.5d	37.5	22.2	20.2	9.9	3.4
56.5e	55.0	8.4	9.5	11.0	9.3



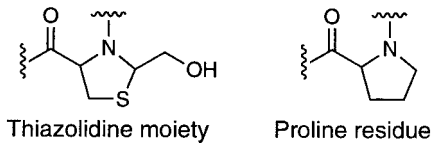
Scheme 57. PG=4-methylbenzyl; NP=protected *N*-terminal peptide; NP=unprotected *N*-terminal peptide; ⬤=polystyrene resin.

6.1.2. Synthesis of peptide segments. In order to apply the thiazolidine-based approach in a practical way, an effective means of attaching the glycoaldehyde moiety to the carboxyl terminus of the *N*-terminal peptide had to be found, and two methods were devised. In the first,[23,82] an enzymatic technique was used to couple a glycoaldehyde-derivatized amino acid to a peptide segment that was protected only at its *N*-terminus and at cysteine side-chains, and had its carboxyl terminus in the form of an ester. The synthesis[82] of this particular segment was based on solid phase techniques starting from a support such as **57.1**[84] (Scheme 57), which allows the required peptide segment to be liberated from the solid support with its carboxyl terminus already in the form of an ester — in the case of **57.1**, in the form of a 3-propylamido ester. Once the required synthetic peptide had been built up (**57.1**→**57.2**), side-chain deprotection and resin cleavage could be effected by treatment with HF, and this step was followed by oxidation of the cysteine residues to the corresponding disulfides (**57.2**→**57.3**). At this point, the minimally protected ester was subjected to trypsin-catalyzed coupling to the dimethoxyethyl ester of alanine (**57.3**+**57.4**→**57.5**). The species obtained from this reaction was then treated with TFA to give the required glycoaldehyde-derivatized peptide (**57.5**→**57.6**).

The second approach to glycoaldehyde-derivatized peptides



Scheme 58. NP=protected *N*-terminal peptide; NP=unprotected *N*-terminal peptide; ⬤=PAM resin.



Scheme 59.

relied on purely chemical means,[85] thereby avoiding possible limitations imposed by substrate specificity of enzymatic reactions. The *N*-terminal peptide was synthesized by a solid phase procedure, as before, but using the thioester-containing support **58.1** (Scheme 58), with the asparagine residue being required as part of the target peptide.[85] Once the required peptide had been obtained (**58.1→58.2**), treatment with HF served both to deprotect the side-chains and cleave the resin (**58.2→58.3**). The unprotected thioester-derivatized peptide was treated with a large excess of an appropriately derivatized amino acid (**58.4**) in the presence of silver ion, to give the masked form of the required glycoaldehyde-containing peptide (**58.5**).[85] Under these conditions, acylation of nucleophilic side-chains was not observed. As in the previous method, treatment with TFA effectively removed the protecting group, liberating the free aldehyde (**58.5→58.6**).

6.1.3. Application to peptide synthesis. The thiazolidine-based *Orthogonal Ligation Coupling Strategy* has been applied to the synthesis of a 15-residue peptide,[23] a 50-residue peptide,[82] and two HIV-1 protease analogs, each 99 residues in length.[85] As mentioned previously, the coupling products obtained from these reactions do not possess a native peptide backbone, as a thiazolidine moiety is incorporated into the backbone at the site of coupling. This moiety bears a strong structural resemblance to a proline residue (see Scheme 59) and, therefore, could potentially serve as a surrogate for proline.[86] It was this observation that suggested the synthesis of the HIV-1 protease analogs, since assessment of their biological activity could provide a means of testing how well the thiazolidine unit can serve in place of proline.

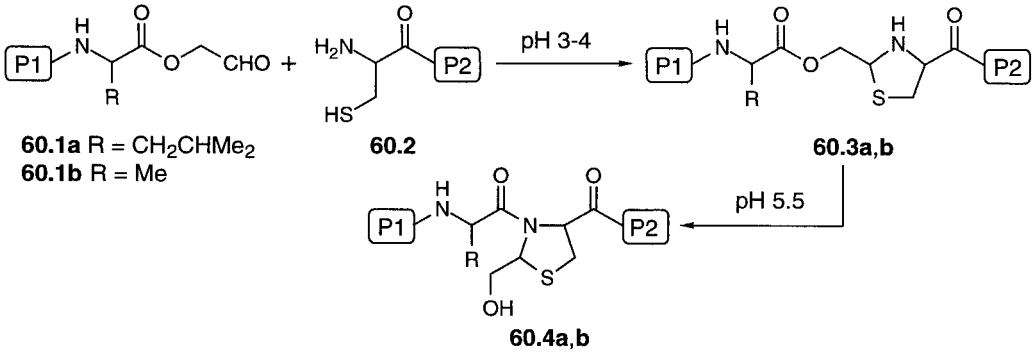
HIV-1 protease contains several proline residues, two of which are suitably positioned to serve as the coupling site. The one that was ultimately chosen for the coupling reactions was equivalent to proline-39 of the native protein. Consequently, each of the two analogs **60.4a**,**b** (Scheme 60) had a thiazolidine unit substituted for proline-39. In addition, cysteine-67 and cysteine-95 were replaced by α-aminobutyric acid. These substitutions were made in order to avoid potentially complicating reactions and, since the residues were known not to be involved in disulfide formation in the native protein, it was assumed that their replacement would not profoundly effect enzymatic activity. In one of the analogs (**60.4b**), leucine-38 of the native protein was substituted by alanine.

The *C*-terminal peptides required for assembly of the analogs corresponded to residues 40–99 of the native peptide, with an additional cysteine at their *N*-termini. These peptides (with the indicated substitutions) were synthesized using conventional solid phase methods and were used in a fully deprotected form.

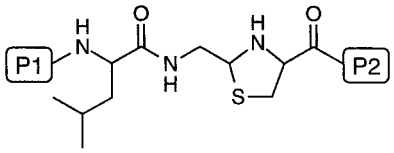
The *N*-terminal peptide segments used for the synthesis of the two HIV-1 protease analogs were each 38 residues in length and had their carboxyl termini in the required glycoaldehyde ester form, but were otherwise unprotected. These peptide segments were generated using the non-enzymatic approach (see Scheme 58).

In each case the ligation was initiated by combining the amino component (**60.2**) and the acyl component (**60.1a**,**b**) in a mixture of acetonitrile and water at pH 3-4.[85] Under these conditions ligation (**60.1a**,**b**+**60.2→60.3a**,**b**) was approximately 60–80% complete within 5–10 h. Although faster ligation would presumably have occurred at a higher pH, an increase in the pH beyond 4 resulted in precipitation of the peptide segments from solution.[85] The initial ligation had to be carried out in the presence of an aspartic protease inhibitor in order to prevent enzymatic cleavage (i.e. self-destruction) of the peptides. Once ligation was complete, the products (**60.3a**,**b**) were isolated, purified by HPLC, and then dissolved in a denaturing solution of guanidine hydrochloride and glycerol at pH 5.5, in order to effect acyl transfer (**60.3a**,**b→60.4a**,**b**). After 3–4 days of incubation, the coupled products were obtained in greater than 90% yield.

The enzymatic activity (at 22°C) of the two synthetic analogs (**60.4a**,**b**) was compared to that of both the native protein and a third analog (**61.1**, Scheme 61), which represented a stable form of the acyl transfer intermediates leading to **60.4a**,**b**. The kinetic data, which are listed in



Scheme 60. P1=PQITLWQRPLVTIRIGGQLKEALLDTGADDTVLEEMN; P2=GKWKPKMIGGIGGFIKVRQYDQIPVEIαGHKAIGTVLVGPTPVNIIGRNLLTQIGαTLNF; α=α-aminobutyric acid.



61.1

Scheme 61. P1=PQITLWQRPLVTIRIGGQLKEALLDTGADDTVLEEMN; P2=GKWKPKMIGGIGGFIKVRQYDQIPVEIαGHKAIGTVLVGPTPV-NIIGRNLLTQIGαTLNF; α=α-aminobutyric acid.

Table 13, show that each of the synthetic peptides has a binding affinity (K_m) similar to that of the native enzyme, but each peptide has a different catalytic activity. The analog with only the proline and cysteine substitutions (**60.4a**) retained essentially complete activity, compared to the native enzyme, and the one with the additional leucine substitution (**60.4b**) showed approximately 70% of the activity. On the other hand, the analog representing the non-rearranged acyl transfer system (**61.1**), had only 30% of the activity of the wild type enzyme.

6.2. Ligation by thioester exchange

The second approach used by Tam to couple peptide segments is closely related to the cysteine-based Native Chemical Ligation Strategy discussed above (cf. Scheme 41), but the thioester leaving group (SR in Scheme 41) in the present case is derived from 3-mercaptopropanoic acid.[87] Using a thioester of this type (cf. **62.2**, Scheme 62), an extensive investigation was carried out in order to determine optimal conditions for coupling two peptide segments. The model system used for this study consisted of peptide **62.3** and the *N*-protected glycine thioester **62.2**. These components were allowed to react in solution at various pHs and, in each case, four different main products were identified (**62.1**, **62.4**, **62.5** and **62.6**, Scheme 62). The relative amounts at each pH are shown in Table 14. The observed decrease in hydrolysis of **62.2** with increasing pH is consistent with mechanisms in which, for this pH window (5.6–7.6), each of the following conditions is met: (i) the hydrolysis shows no dependence on pH; (ii) the pK_a of the cysteine thiol lies above the highest pH studied; and (iii) the thiolate anion acts as the nucleophile.

Table 13. Kinetic data of synthetic HIV-1 protease analogs compared to wild type[85]

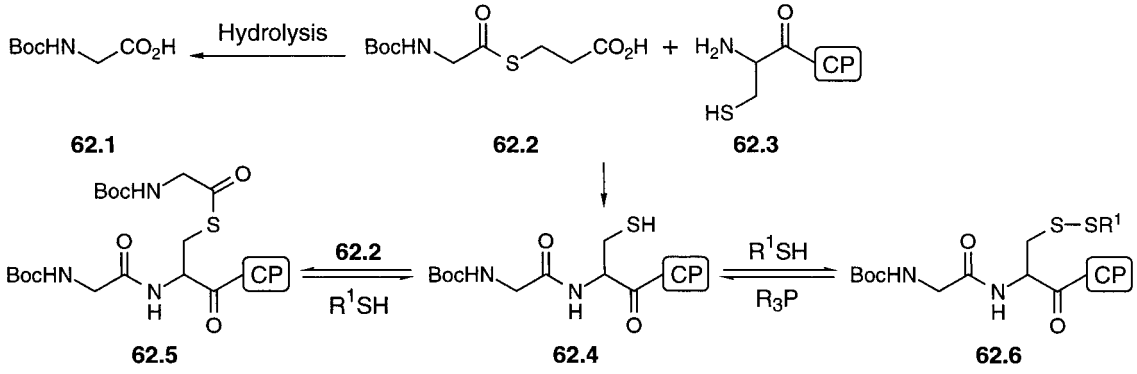
Substrate	K_m (μM)	V_{max} (μmol min^{-1} mg^{-1})	V_{max}/K_m (μmol min^{-1} mg^{-1} M^{-1})
Wild type	10.1	3.43	3.40×10^5
60.4a	11.9	3.96	3.33×10^5
60.4b	8.2	2.26	2.76×10^5
61.1	11.4	1.14	1.00×10^5

Further attempts were made to optimize the coupling under the basic conditions that minimize thioester hydrolysis.[87] This was achieved through an examination of the effect that a reducing environment would have on the product distribution for the reaction between **62.2** and **62.3** at pH 7.2. The reducing agents examined were 3-mercaptopropanoic acid and tris(2-carboxyethyl)phosphine. These were added in various amounts either alone or in combination, and it was found that almost exclusive formation of the desired product (**62.4**) could be achieved in the presence of a 2:10 mixture of tris(2-carboxyethyl)phosphine and 3-mercaptopropanoic acid. Under these conditions, formation of byproducts was almost completely suppressed. These optimized conditions were used[87] to synthesize several peptides ranging in size from 9–54 amino acids and, in each case, the coupling reactions proceed effectively, giving the acyl transfer products in yields of 60–88%.

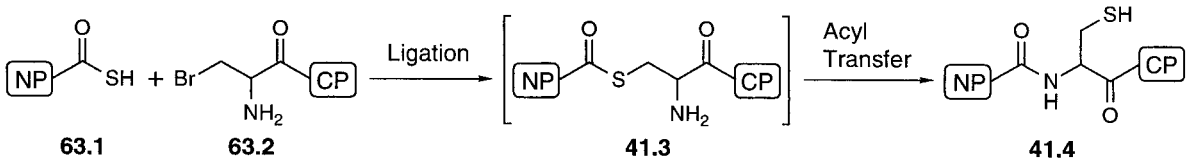
Tam has also explored segment coupling based on ligation by thioester exchange with a homocysteine residue[88] that takes the place of cysteine at the *N*-terminus of the *C*-terminal segment in his previous examples (cf. **62.3**, Scheme 62). Once ligation has occurred, acyl transfer takes place, this time via a six-membered ring, and gives rise to a product that has a homocysteine residue at the site of coupling. At this point, the homocysteine residue can be methylated,[89] so that a methionine unit is now present at the coupling site. The obvious limitation of this type of segment coupling, however, is that any cysteine thiols present in either of the segments would also be methylated.

6.3. Ligation by thioester formation

The Tam group has also explored the ligation of peptide segments, prior to peptide bond formation, by reacting a thioacid and a primary bromide (Scheme 63).[87] The thioacid in this case is a derivatized form of the carboxyl terminus of the *N*-terminal peptide (**63.1**), and the primary bromide is a bromoalanine residue at the *N*-terminus of the other peptide (**63.2**). The idea underlying this approach is that, once the



Scheme 62. R^1=$CH_2CH_2CO_2H$ or **62.3**; R_3P=$(HO_2CCH_2CH_2)_3P$; CP=FKA.



Scheme 63. NP=*N*-terminal peptide; CP=*C*-terminal peptide.

Table 14. Product distribution for reaction between **62.2** and **62.3**[87]

pH	Product yield (%)			
	62.1	**62.4**	**62.5**	**62.6**
5.6	78	5	2.5	12
6.6	34	43	3	20
7.2	38	44	3	16
7.6	15	56	4	25

two components have reacted, an acyl transfer intermediate (**41.3**) identical to that obtained using either of the thioester exchange ligation procedures already described (cf. Schemes 41 and 62) would be obtained. In this case, however, the ligation product is not formed by thioester exchange, but rather by nucleophilic substitution in which the bromide is displaced by the thioacid (**63.1**+**63.2**→**41.3**). At this point, acyl transfer should proceed, as in the thioester exchange strategies, to give the rearranged, coupled product with a new peptide bond (**41.3**→**41.4**) and, just as in the thioester exchange approaches, the product will have a cysteine residue at the site of coupling corresponding to what was the *N*-terminus of the *C*-terminal peptide chain, even though cysteine was not initially present at that position.

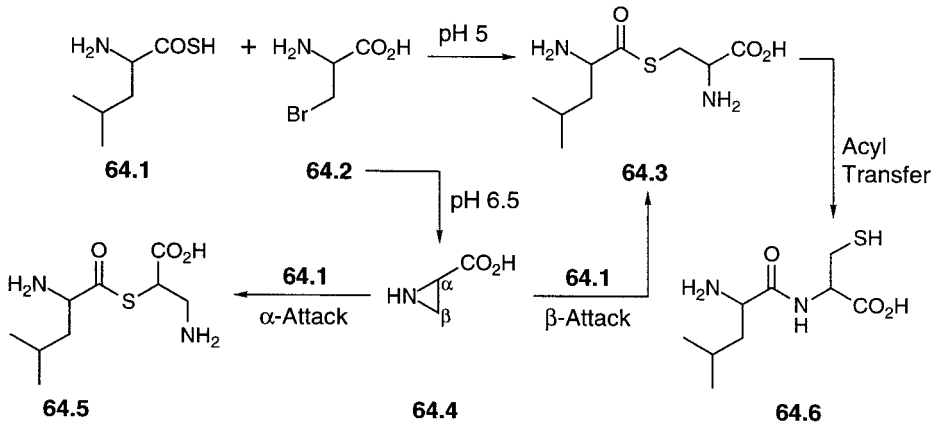
A model study of this approach was carried out using 2-amino-4-methylpentanethioic acid (**64.1**, Scheme 64) and 3-bromoalanine (**64.2**). The reaction was studied at several pHs and the results led to the following conclusions. When the reaction was tried at a pH greater than 6, the coupling process did not proceed smoothly to the desired product, but gave instead a 6:4 mixture of the desired product (**64.6**) and compound **64.5**, respectively. Compound **64.5** was judged to result from the formation of an aziridine (**64.2**→**64.4**), followed by attack at the α-position by the thioacid (**64.4**+**64.1**→**64.5**). Attack at the β position of the aziridine would give the desired ligation product (**64.3**) and, ultimately, **64.6**. If the pH of the solution was lowered below 5, however, it was found that very little (<3% yield) byproduct was formed. This outcome was suggested to result from the fact that at lower pH aziridine formation was not favored, and so the normal course of nucleophilic displacement of the bromide by the thioacid (**64.1**+**64.2**→**64.3**) could proceed, eventually giving the desired coupled product.

The possibility that the ligation was occurring by Michael addition of the thioacid to dehydroalanine, formed by HBr elimination from **64.2**, was excluded as follows. First, it was established that HBr elimination did not occur to an appreciable extent unless the pH of the solution was raised to about 11. Secondly, the product obtained from reaction of **64.1** and **64.2** was compared to appropriate reference samples, using HPLC. There was no evidence for epimerization, and so it was judged unlikely that ligation involved Michael addition.

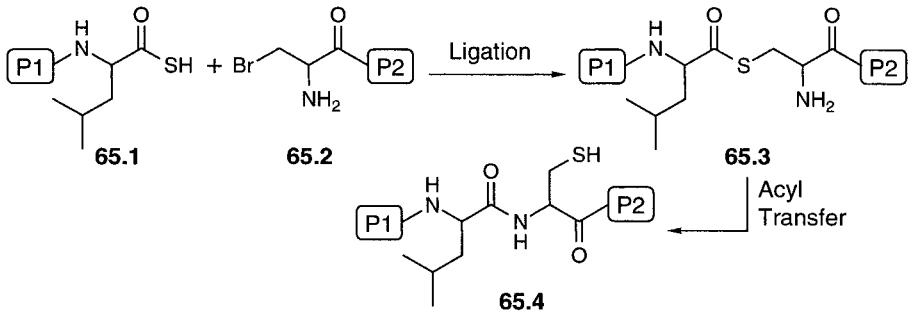
The method was tested (Scheme 65) by synthesizing a 12-residue peptide (**65.4**). This was generated by coupling the fully unprotected 4-residue thioacid segment **65.1** and the bromoalanine-containing 8-residue segment **65.2**, which was protected only at an internal cysteine residue. When combined in solution, the two components reacted to give the acyl transfer intermediate **65.3**, which subsequently rearranged to the desired product in 85% yield.

6.4. Ligation by disulfide exchange

Another peptide segment coupling approach that the Tam group has developed is based on ligation by a disulfide bond, which is formed by attack of a thioacid on an activated thiol.[90] Like Kemp's Prior Thiol Capture Strategy, the activated thiol is located at the *N*-terminus of the



Scheme 64.



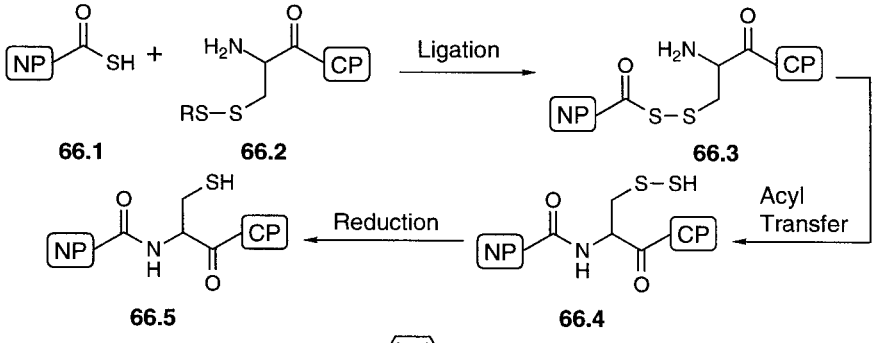
Scheme 65. P1=SAK; P2=PGGNAC(Acm)V; Acm=$CH_2NHCOMe$.

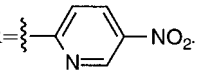
C-terminal peptide (**66.2**, Scheme 66) and is part of a cysteine residue. Unlike Kemp's approach, however, the activating group is a 2-mercapto-5-nitropyridyl, rather than an Scm group. Also different from Kemp's approach is the fact that the attacking nucleophilic sulfur atom is part of a thioacid corresponding to the carboxyl terminus of the other chain (**66.1**), rather than a thiol attached to a template. In fact, no template is used in the present case; instead, the two components react in solution to form an acyl transfer intermediate in situ (**66.1**+**66.2**→**66.3**), and this species then undergoes rearrangement through a six-membered transition state to give the coupled product, in which the cysteine thiol is derivatized as a hydrodisulfide (**66.3**→**66.4**). This is readily reduced to generate the desired coupled product (**66.5**).

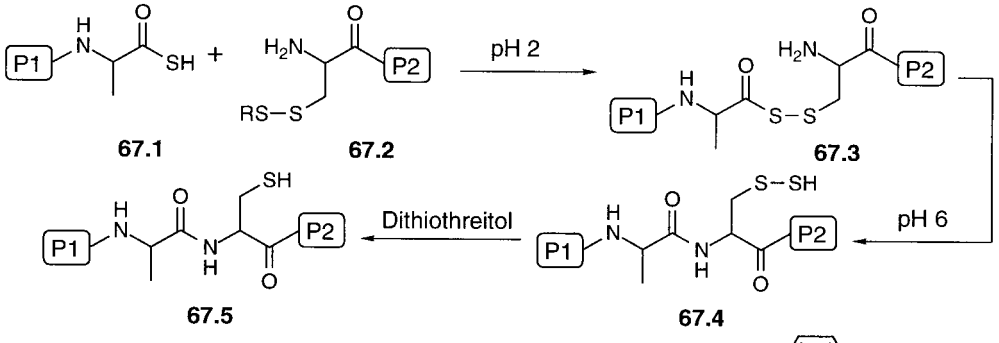
This disulfide ligation-based route to peptide segment coupling has been applied[90] to the synthesis of a 32-residue peptide, generated by coupling a 17-residue *C*-terminal segment and a 15-residue *N*-terminal segment. The latter was made[90] by solid phase synthesis, and its *C*-terminus was in the form of a thioacid. The coupling process was initiated by activating the *N*-terminal cysteine thiol of the *C*-terminal segment, which was fully unprotected except at its internal cysteine. Activation was accomplished by treatment with 2,2′-dithiobis(5-nitropyridine). The product of this reaction (**67.2**, Scheme 67) was then combined with the thioacid-derivatized *N*-terminal peptide **67.1** in a solution of acetonitrile, water, and TFA at pH 2. Under these conditions the ligation product was formed (**67.1**+**67.2**→**67.3**). The pH was then adjusted to 6, which induced efficient acyl transfer, giving the hydrodisulfide **67.4**. Acyl transfer via the six-membered transition state accessible from **67.3** was so efficient that it was more than 90% complete after only 5 min. Finally, the hydrodisulfide **67.4** was readily reduced by treatment with dithiothreitol, to liberate the desired coupled peptide **67.5**.

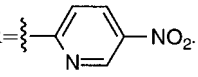
6.5. Ligation by amide formation

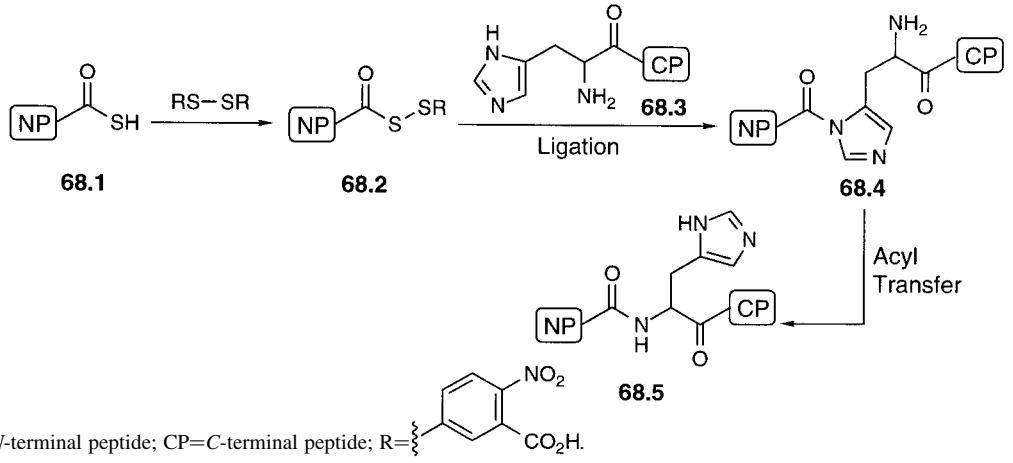
One final approach to peptide segment coupling that has been reported by Tam, makes use of an *N*-terminal histidine



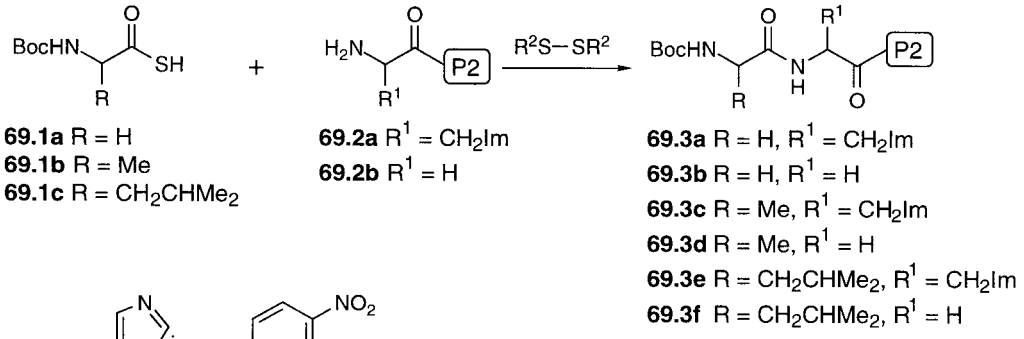
Scheme 66. NP=*N*-terminal peptide; CP=*C*-terminal peptide; R=  .



Scheme 67. P1=HSGYVGARC(Acm)EHADLLA; P2=FFGTARFLVQEDKP; Acm=$CH_2NHCOMe$; R=  .



Scheme 68. NP=*N*-terminal peptide; CP=*C*-terminal peptide; R=



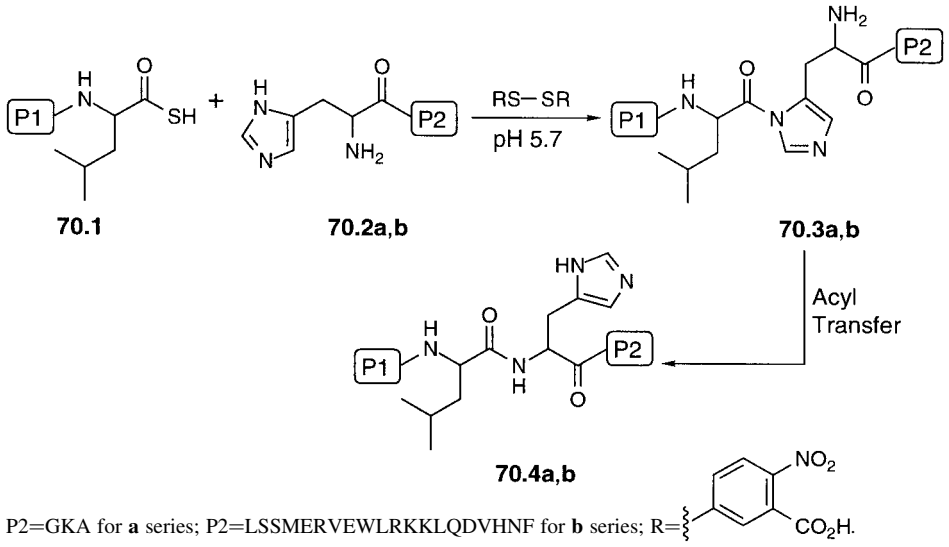
Scheme 69. P2=GKA; Im= ; R²=

to aid in ligation and coupling.[91] This approach (Scheme 68) provides the third example of a practical coupling technique that does not involve, either directly or indirectly, a cysteine residue. The histidine in question is positioned at the *N*-terminus of the *C*-terminal segment (**68.3**) and functions in much the same way as the terminal cysteine in the thioester exchange approaches of Kent and Tam, discussed above. In the present case, however, it is, of course, the imidazole rather than a thiol that acts as the nucleophilic ligation component. To effect ligation, the imidazole nitrogen attacks an activated thioacid moiety at the carboxyl terminus of the *N*-terminal peptide (**68.1**), and generates the transient amide **68.4**. The activated form of the thioacid is the disulfide **68.2**, which is generated in situ by treatment of the thioacid with 5,5′-dithiobis(2-nitrobenzoic acid). As mentioned earlier (see Section 5.1.1), acyl disulfides (cf. **68.2**) can react further to give a thioester; however, the rate of this process appears[91] to be sensitive to the reaction conditions as well as to the nature of the thioacid and, in the present case, it is believed[91] that the disulfide is actually the species involved in the ligation. Evidently, the aryl disulfide is an adequate leaving group. Once the transient amide **68.4** has formed, spontaneous intramolecular acyl transfer occurs via a six-membered transition state, and gives the coupled product **68.5**.

Model studies (Scheme 69) on the histidine-based coupling strategy involved the *C*-terminal peptides **69.2a** and **69.2b**, and the thioacids derived from *N*-Boc-protected glycine (**69.1a**), alanine (**69.1b**), and leucine (**69.1c**). Peptide **69.2b** provided a control that could be used for comparison to the *N*-terminal histidine peptide, in order to determine if the ligation product indeed formed prior to acyl transfer. Each of the *C*-terminal peptides used contained an unprotected lysine residue, and so the problem of selectivity of acylation between the α-amino and the ϵ-amino groups was also examined. At pH 5, 5.7 and 6.5 each of the thioacids **69.1a–c**, once activated [with 5,5′-dithiobis(2-nitrobenzoic acid)], reacted preferentially with the α-amino group instead of the ϵ-amino group of the tetrapeptides **69.2a** and **69.2b**. At pH 5.7 and 6.5 both a higher selectivity of acylation in favor of the α-amino group, as well as a higher reaction rate, were observed for the peptide with the *N*-terminal histidine (**69.2a**) as compared to the one without (**69.2b**). The relative product distribution for coupling between each of the thioacids and each of the *C*-terminal tetrapeptides at pH 5.7 is shown in Table 15. As further confirmation of the role of the *N*-terminal histidine in the coupling process, a competition experiment was carried out in which an equimolar mixture of the two *C*-terminal peptides (**69.2a**,**b**) was treated with thioacid **69.1a** that had been activated by 5,5′-dithiobis(2-

Table 15. Product distribution for reaction of **69.1a–c** with **69.2a**,**b**[91]

Thiocarboxylic acid	*C*-Terminal peptide			
	Coupling yield (%)		Acylation selectivity (α/ϵ)	
	69.2a	**69.2b**	**69.2a**	**69.2b**
69.1a	85	40	16	7
69.1b	82	50	9	5
69.1c	75	20	4	1



Scheme 70. P1=SKA; P2=GKA for **a** series; P2=LSSMERVEWLRKKLQDVHNF for **b** series; R=

nitrobenzoic acid). In this case, the product corresponding to coupling with the histidyl peptide **69.3a** was obtained in greater than 90% yield, whereas less than 10% of the other product (**69.3b**) was formed.

The amide ligation-based coupling procedure was further tested[91] by applying it to the synthesis of an 8-residue peptide (**70.4a**, Scheme 70) and a 25-residue peptide (**70.4b**). For each synthesis, tetrapeptide **70.1** was used as the *N*-terminal thioacid-derivatized peptide. This peptide was combined with either the 4-residue *C*-terminal peptide **70.2a**, or the 21-residue *C*-terminal peptide **70.2b**, in the presence of 5,5′-dithiobis(2-nitrobenzoic acid) at pH 5.7, to give the presumed ligation product (**70.1**+**70.2a**,**b**→**70.3a**,**b**). Each acyl transfer reaction proceeded as expected under these conditions, and gave the 8-residue coupled product in 75% yield, and the 25-residue product in 60% yield.

7. Conclusion

The development of methods for peptide segment coupling based on prior ligation and proximity-induced intramolecular acyl transfer has evolved considerably since the early reports and speculations made by Brenner, Wieland, and others. In this respect, the Kemp group has provided a substantial amount of kinetic evidence in an effort to define the limits to which its own methods are applicable, and has demonstrated proof of principle by synthesis of a few peptides of moderate size. The Kent group, on the other hand, has demonstrated their method in many cases, and has provided some insight into the mechanistic aspects of their approach. The Tam group has also established proof of principle and, additionally, has explored several variations of the ligation step. With the recent application of Kent's Native Chemical Ligation Strategy to the solid phase, as well as its use with recombinantly derived peptide segments, a new era of research in segment coupling appears to be unfolding. It remains to be seen, however, what the practical limitations of these recent developments will be, but the techniques appear extremely promising in terms of offering rapid access to peptides of sizes that are currently inaccessible by synthetic methods.

8. List of Abbreviations

Acm=acetamidomethyl (–CH_2NHCOMe); Dnp=2,4-dinitrophenylsulfenyl; Dns=5-(dimethylamino)-1-naphthylsulfonyl; HFIP=hexafluoroisopropyl alcohol; Scm=sulfenylcarbomethoxy (-SCO_2Me).

Acknowledgements

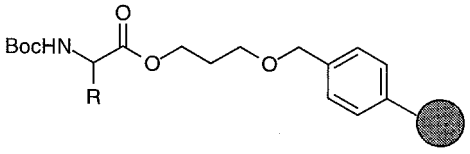
The author gratefully acknowledges Professor D. L. J. Clive for his advice and encouragement and for many insightful discussions during the preparation of this manuscript. Several of the investigators, whose work is reviewed, very kindly clarified some minor points in their publications, and particular thanks are due to Professor D. S. Kemp for his immensely helpful review of the manuscript. The author also thanks both the Alberta Heritage Foundation for Medical Research and the Killam Foundation for financial support.

References

1. Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int J. Pep. Protein Res.* **1992**, *40*, 180–193.
2. (a) Zoller, M. J.; Smith, M. *Methods Enzymol.* **1983**, *100*, 468–500. (b) Hutchison III, C. A.; Phillips, S.; Edgell, M. H.; Gillam, S.; Jahnke, P.; Smith, M. *J. Biol. Chem.* **1978**, *253*, 6551–6560. (c) Knowles, J. R. *Science* **1987**, *236*, 1252–1258.
3. (a) Cornish, W. V.; Mendel, D.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 621–633. (b) Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R. *J. Am. Chem. Soc.* **1989**, *111*, 8013–8014.
4. Curtius, T. *J. Prakt. Chem.* **1881**, *24*, 239–240.
5. Fischer, E. *Chem. Ber.* **1902**, *35*, 1095–1106.
6. Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
7. (a) Bayer, E. *Angew. Chem. Int., Ed. Engl.* **1991**, *30*, 113–129. (b) Lloyd-Williams, P.; Albericio, F.; Giralt, E. *Tetrahedron* **1993**, *48*, 11065–11133.
8. Kent, S. B. H. *Annu. Rev. Biochem.* **1988**, *57*, 957–989.

9. See also, for example, Ref. 1.
10. *Peptide Chemistry at the Medical Research Council of Molecular Biology, Cambridge*; Compiled by Sheppard, R. C.; Printed privately: Cambridge, 1997.
11. It is possible, although not routine, to go beyond 50 amino acid residues using solid phase peptide synthesis: See Ref. 1 and (a) Li, C. H.; Yamashiro, D. *J. Am. Chem. Soc.* **1970**, *92*, 7608–7609. (b) Sharp, J. J.; Robinson, A. B.; Kamen, M. D. *J. Am. Chem. Soc.* **1973**, *95*, 6097–6108. (c) Gutte, B.; Merrifield, R. B. *J. Am. Chem. Soc.* **1969**, *91*, 501–502.
12. See, for example: (a) Hirschmann, R.; Nutt, R. F.; Veber, D. F.; Vitali, R. A.; Varga, S. L.; Jacob, T. A.; Holly, F. W.; Denkewalter, R. G. *J. Am. Chem. Soc.* **1969**, *91*, 507–508, and earlier papers in the series. (b) Kaiser, E. T. *Acc. Chem. Res.* **1989**, *22*, 47–54. (c) Aimoto, S.; Mizoguchi, N.; Hojo, H.; Yoshimura, S. *Bull. Chem. Soc. Jpn.* **1989**, *62*, 524–531. (d) Romovacek, H.; Dowd, S. R.; Kawasaki, K.; Nishi, N.; Hofmann, K. *J. Am. Chem. Soc.* **1979**, *101*, 6081–6091. (e) Meienhofer, J. *Biopolymers* **1981**, *20*, 1761–1784. (e) Yajima, H.; Fujii, N. *J. Am. Chem. Soc.* **1981**, *103*, 5867–5871.
13. (a) Capon, B. *Q. Rev., Chem. Soc.* **1964**, *18*, 45–111. (b) Page, M.; Jencks, W. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1678–1683. (c) Kirby, A. *Adv. Phys. Org. Chem.* **1980**, *17*, 183–278.
14. Menger, F. M. *Acc. Chem. Res.* **1985**, *18*, 128–134.
15. Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. *Justus Liebigs Ann. Chem.* **1953**, *583*, 129–149.
16. (a) Brenner, M.; Zimmermann, J. P.; Wehrmüller, J.; Quitt, P.; Photaki, I. *Experientia* **1955**, *11*, 397–399. (b) Brenner, M.; Zimmermann, J. P.; Wehrmüller, J.; Quitt, P.; Hardtmann, A.; Schneider, W.; Beglinger, U. *Helv. Chim. Acta* **1957**, *40*, 1497–1517.
17. Brenner, M. In *Peptides: Proceedings of the Eighth European Peptide Symposium*; Beyerman, H. C.; van den Linde, A.; van den Brink, W. M., Eds.; North-Holland: Amsterdam, 1967, pp 1–7.
18. Paulus, H. *Chem. Soc. Rev.* **1998**, *27*, 375–386.
19. (a) Proudfoot, A. E. I.; Rose, K.; Wallace, C. J. A. *J. Biol. Chem.* **1989**, *264*, 8764–8770. (b) Wallace, C. J. A.; Corthésy, B. E. *Protein Eng.* **1986**, *1*, 23–27. (c) Homandberg, G. A.; Chaiken, I. M. *J. Biol. Chem.* **1980**, *255*, 4903–4909. (d) Dyckes, D. F.; Creighton, T.; Sheppard, R. C. *Nature* **1974**, *247*, 202–204. (e) Cf. Homandberg, G. A.; Mattis, J. A.; Laskowski, M., Jr. *Biochemistry* **1978**, *17*, 5220–5227. (f) Offord, R. E. *Semisynthetic Proteins*; Wiley: New York, 1980.
20. (a) Mills, K. V.; Lew, B. M.; Jiang, S.-Q.; Paulus, H. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3543–3548. (b) Lew, B. M.; Mills, K. V.; Paulus, H. *J. Biol. Chem.* **1998**, *273*, 15 887–15 890.
21. Kemp, D. S. *Biopolymers* **1981**, *20*, 1793–1804.
22. Fotouhi, N.; Galakatos, N. G.; Kemp, D. S. *J. Org. Chem.* **1989**, *54*, 2803–2817.
23. Liu, C.-F.; Tam, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 4149–4153.
24. Kemp, D. S.; Roberts, D.; Hoyng, C.; Grattan, J.; Vellaccio, F.; Reczek, J. In *Peptides: Chemistry, Structure, and Biology*; Walter, R., Ed.; Ann Arbor Science: Ann Arbor, Michigan, 1975, pp 295–305.
25. Kemp, D. S.; Vellaccio, Jr. F. *J. Org. Chem.* **1975**, *40*, 3003–3004.
26. A subsequent publication (see Ref. 21) indicated that products resulting from *O*⟶*O* acyl transfer are also observed.
27. Kemp, D. S.; Gratten, J. A.; Reczek, J. *J. Org. Chem.* **1975**, *40*, 3465–3466.
28. Imine formation from salicylaldehydes occurs with unusually large rate and equilibrium constants. See: Green, R. W.; Sleet, R. J. *Aust. J. Chem.* **1969**, *22*, 917–919 and references cited therein.
29. Although it is not clear from the published work (see Ref. 27), the acyl transfer presumably proceeds by initial *O*⟶*O* transfer between the phenolic hydroxyls, followed by *O*⟶*N* transfer.
30. Kemp, D. S.; Wang, S. W.; Busby III, S.; Hugel, G. *J. Am. Chem. Soc.* **1970**, *92*, 1043–1055.
31. Kemp, D. S.; Vellaccio, Jr. F. *J. Org. Chem.* **1975**, *40*, 3464–3465.
32. Kemp, D. S.; Kerkman, D. J.; Leung, S.-L.; Hanson, G. *J. Org. Chem.* **1981**, *46*, 490–498.
33. Worrall, D. E. *J. Am. Chem. Soc.* **1927**, *49*, 1598–1605.
34. Kemp, D. S.; Choong, S.-L. H.; Pekaar, J. *J. Org. Chem.* **1974**, *39*, 3841–3847.
35. Kemp, D. S.; Leung, S.-L.; Kerkman, D. J. *Tetrahedron Lett.* **1981**, *22*, 181–184.
36. (a) Kemp, D. S.; Kerkman, D. J. *Tetrahedron Lett.* **1981**, *22*, 185–186. (b) Kemp. D. S.; Buckler, D. R.; Galakatos, N. G.; Kirkman, D. *J. Org. Chem.* **1989**, *54*, 3853–3858.
37. Simpson, R. B. *J. Am. Chem. Soc.* **1961**, *83*, 4711–4717.
38. Kemp, D. S.; Galakatos, N. G. *J. Org. Chem.* **1986**, *51*, 1821–1829.
39. Kemp, D. S.; Galakatos, N. G.; Bolin, D. In *Peptides: Synthesis, Structure, Function*; Rich, D., Gross, E., Eds.; Pierce Chemical Company: Rockford, IL, 1981, pp 73–79.
40. Kemp, D. S.; Galakatos, N. G.; Bowen, B.; Tan, K. *J. Org. Chem.* **1986**, *51*, 1829–1838.
41. March, J. *Advanced Organic Chemistry: Reactions Mechanisms, and Structure*, 4th ed.; Wiley: New York, 1992, p 211.
42. Kemp, D. S.; Carey, R. I.; Dewan, J. C.; Galakatos, N. G.; Kerkman, D.; Leung, S.-L. *J. Org. Chem.* **1989**, *54*, 1589–1603.
43. (a) Menger, F. M.; Smith, J. H. *J. Am. Chem. Soc.* **1972**, *94*, 3824–3829. (b) Johnson, S. L. *Adv. Phys. Org. Chem.* **1967**, *5*, 237–330.
44. Although the data in Tables 4 and 5 suggest that DMSO would be a kinetically unsuitable solvent, it was, in fact, the only solvent that allowed acyl transfer at an acceptable rate, presumably because the reaction requires participation of a strong hydrogen-bonding acceptor (Kemp, D. S. In *Peptides: Proceedings of the Eleventh European Peptide Symposium*; Nesvadba, H., Ed.; North-Holland: Amsterdam, 1972, pp 1–19).
45. In earlier publications (see Refs. 35,40), an EM of between 0.1 and 0.5 M was reported for this system. Presumably the more recent value of 3–14 M is a more accurate measurement.
46. Hanson, G.; Kemp, D. S. *J. Org. Chem.* **1981**, *46*, 5441–5443.
47. Kemp, D. S.; Buckler, D. R. *Tetrahedron Lett.* **1991**, *32*, 3009–3012.
48. Although it is not stated in the original publication, I assume that only extended conformations of the transition state are accessible with this template.
49. Kemp, D. S.; Buckler, D. R. *Tetrahedron Lett.* **1991**, *32*, 3013–3016.
50. Kemp, D. S.; Galakatos, N. G.; Dranginis, S.; Ashton, C.; Fotouhi, N.; Curran, T. P. *J. Org. Chem.* **1986**, *51*, 3320–3324.
51. Cf. (a) Khurgin, Yu. I.; Dmitrieva, M. G. *Tetrahedron* **1965**, *21*, 2305–2312. (b) Pless, J.; Boissonnas, R. A. *Helv. Chim. Acta* **1963**, *48*, 1609–1625.
52. McBride, B. J.; Kemp, D. S. *Tetrahedron Lett.* **1987**, *28*, 3435–3438.
53. (a) Brois, S. J.; Pilot, J.; Barnum, H. *J. Am. Chem. Soc.* **1970**, *92*, 7629–7631. (b) Hiskey, R. G.; Li, C.; Vunnam, R. *J. Org.*

Chem. **1975**, *40*, 950–953. (c) Kamber, B. *Helv. Chim. Acta* **1973**, *56*, 1370–1381.
54. Kemp, D. S.; Fotouhi, N. *Tetrahedron Lett.* **1987**, *28*, 4637–4640.
55. Prepared (see Ref. 22) by reaction of **29.5** (R=OAc) with the cyclohexylamine salt of Cbz-L-Cys(Scm).
56. Kemp, D. S.; Carey, R. I. *J. Org. Chem.* **1989**, *54*, 3640–3646.
57. In other applications (see Ref. 22) of the revised Prior Thiol Capture Strategy (see Scheme 36), non-cysteine side-chains, as well as the α-amino group of the template-linked *N*-terminal peptide segment (cf. **36.1**) were also protected (see Ref. 58).
58. Kemp, D. S.; Carey, R. I. *J. Org. Chem.* **1993**, *58*, 2216–2222.
59. Kemp has explored (see Ref. 56) the use of a novel thiazolidine protecting group as a substitute for Acm. This thiazolidine is potentially advantageous in that it can be hydrolyzed selectively under conditions that are not detrimental to the other thiol protecting groups in the peptide, thus releasing a free thiol group for reaction with ScmCl.
60. The problem of converting the SAcm into the SScm moiety in the presence of the unprotected nucleophilic side-chains of lysine (see Refs. 22,58) and histidine (see Ref. 61) has been studied. Histidine functionalization by ScmCl is suppressed by using formic acid–acetic acid–DMF for conversion of a terminal Cys(Acm) into Cys(Scm).
61. Kemp, D. S.; Carey, R. I. *Tetrahedron Lett.* **1991**, *32*, 2845–2848.
62. Chalfoun, D. J. Ph.D Dissertation, Massachusetts Institute of Technology, Massachusetts, 1996.
63. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.
64. Bruice, T. C.; Benkovic, S. J. *Bioorganic Mechanisms;* Benjamin: New York, 1966; Vol. 1, pp 259–297.
65. Leu-enkephalin is: YGGFL. Voet, D.; Voet, J. G. *Biochemistry*, 2nd ed.; Wiley: New York, 1995, p 1274.
66. Muir, T. W.; Dawson, P. E.; Kent, S. B. H. *Methods Enzymol.* **1997**, *289*, 266–298.
67. Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* **1997**, *119*, 4325–4329.
68. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.
69. The thioesters in this case were in the form shown in Table 8, where the thiol component is a portion of the linker used for solid phase synthesis of the peptide segment in question. Cleavage from the solid support gave the thioesters LYRAX-SR directly, thereby avoiding the additional step of thioacid derivatization.
70. Lu, W.; Qasim, M. A.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 8518–8523.
71. Canne, L. E.; Walker, S. M.; Kent, S. B. H. *Tetrahedron Lett.* **1995**, *36*, 1217–1220.
72. Cf. Canne, L. E.; Ferré-D'Amaré, A. R.; Burley, S. K.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 2998–3007.
73. Kent has developed an efficient procedure for the preparation of this compound. See Ref. 71.
74. Several other peptides besides those discussed in this Review have been made (see Refs. 66,75), but no details are given in the original publications.
75. Canne, L. E.; Bark, S. J.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896.
76. Cotton, G. J.; Ayers, B.; Xu, R.; Muir, T. W. *J. Am. Chem. Soc.* **1999**, *121*, 1100–1101.
77. Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. A.; Kent, S. B. H. *J. Am. Chem. Soc.* **1999**, *121*, 8720–8727.
78. Canne, L. E.; Winston, R. L.; Kent, S. B. H. *Tetrahedron Lett.* **1997**, *38*, 3361–3364.
79. Canne, L. E.; Ferré-D'Amaré, A. R.; Burley, S. K.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 2998–3007.
80. The exact form of the *C*-terminus at this point is not clear from the original publication. The *C*-terminal peptide is either released from the resin as a thioester, in which case it is presumably used as such, or it is released in the form of a thioacid, in which case it is presumably derivatized as a thioester, and then elaborated.
81. The published results (see Ref. 75) do not indicate whether any attempt was made to separate the two peptides; hence, it is not clear if the *C*-terminal peptides used in the model ligations 4 and 5 (see Table 9) were pure stereoisomers.
82. Liu, C.-F.; Tam, J. P. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6584–6588.
83. Tam, J. P.; Rao, C.; Liu, C.-F.; Shao, J. *Int J. Pep. Protein Res.* **1995**, *45*, 209–216.
84. The following species has also been used as a solid support for these purposes. See Ref. 23.



85. Liu, C.-F.; Rao, C.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 307–312.
86. For proline surrogates, see: Tam, J. P.; Miao, Z. *J. Am. Chem. Soc.* **1999**, *121*, 9013–9022 and references cited therein.
87. Tam, J. P.; Lu, Y.-A.; Liu, C.-F.; Shao, J. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12 485–12 489.
88. Tam, J. P.; Yu, Q. *Biopolymers* **1998**, *46*, 319–327.
89. Heinrikson, R. L. *J. Biol. Chem.* **1971**, *246*, 4090–4096.
90. Liu, C.-F.; Rao, C.; Tam, J. P. *Tetrahedron Lett.* **1996**, *37*, 933–936.
91. Zhang, L.; Tam, J. P. *Tetrahedron Lett.* **1997**, *38*, 3–6.

Biographical Sketch

Don Coltart obtained a Bachelor of Science (Honours) degree in biochemistry from the University of Manitoba, Canada, and then a Master of Science degree in organic chemistry from the same institution. His MSc work was in the area of natural product synthesis based on asymmetric Diels–Alder reactions. Currently, as a doctoral candidate at the University of Alberta, he is working under the supervision of Professor D. L. J. Clive in the area of natural product synthesis.