

Review

## Native Chemical Ligation: A Boon to Peptide Chemistry

Parashar Thapa, Rui-Yang Zhang, Vinay Menon and Jon-Paul Bingham \*

Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, HI 96822, USA

\* Author to whom correspondence should be addressed; E-Mail: jbingham@hawaii.edu; Tel.: +1-808-956-4864; Fax: +1-808-956-3542.

Received: 31 July 2014; in revised form: 2 September 2014 / Accepted: 2 September 2014 /

Published: 12 September 2014

---

**Abstract:** The use of chemical ligation within the realm of peptide chemistry has opened various opportunities to expand the applications of peptides/proteins in biological sciences. Expansion and refinement of ligation chemistry has made it possible for the entry of peptides into the world of viable oral therapeutic drugs through peptide backbone cyclization. This progression has been a journey of chemical exploration and transition, leading to the dominance of native chemical ligation in the present advances of peptide/protein applications. Here we illustrate and explore the historical and current nature of peptide ligation, providing a clear indication to the possibilities and use of these novel methods to take peptides outside their typically defined boundaries.

**Keywords:** native chemical ligation; cyclotides; conotoxin; convergent; sequential; auxiliary; thioester

---

### 1. Introduction

Peptides have been at the forefront of scientific research for the past several decades. Their structural diversity and high level of target selectivity position them as prime candidates for therapeutic development. To this end, it is desirable to be able to synthesize peptides under controlled laboratory conditions using highly efficient chemical reactions. The field of synthetic peptide chemistry was first made feasible by the development of Solid Phase Peptide Synthesis (SPPS) in the 1960s by Bruce Merrifield [1], for which he was awarded the 1984 Noble Prize in Chemistry. In addition to making peptide synthesis practical, this technique allows for the introduction of non-native

amino acids and post-translational modifications (PTMs) into the peptide sequence in a stepwise and regulated fashion. In spite of its many advantages one major disadvantage of SPPS is its length limitation. SPPS can only synthesize peptides restricted to ~50 amino acid residues in length [2]. This length limitation of SPPS can be attributed to the formation of intermolecular aggregates, addition/deletion of amino acids from the growing peptide chain, and a higher number of side reactions as the peptide chain increases in length [3,4]. Native Chemical Ligation (NCL), developed by the Kent laboratory [5], allows for the construction of longer peptide chains while retaining the native peptide backbone, which for many peptides is crucial for maintaining bioactivity. The basis of this chemical approach allows for the coupling of two individual fully deprotected peptide fragments in an aqueous solution to form a site-directed native peptide bond at the ligation site. NCL in conjugation with SPPS can be used to make functional proteins that are ~250 amino acids in length [6]. Peptides and proteins play critical roles within biological and physiological processes in all living organisms. The chemical synthesis of peptides has been developed for over 100 years [6]. During this period, NCL was one of the most important breakthroughs in peptide chemistry. In this review we discuss the evolution, advantages and potential applications of this technique.

## 2. History of Ligation Strategies

### 2.1. Hydrazone Ligation

One of the earliest forms of ligation [7] utilized a hydrazone linkage, an amine derivative of an imine. This bond is typically created by reacting a hydrazine moiety with an aldehyde resulting in a nucleophilic attack on the carbonyl carbon and the formation of the hydrazone bond through the loss of water. This highly chemo-selective reaction negates the need for side-chain protection groups, allowing molecules to remain small and water-soluble. However, this method is slow, typically taking 1–3 days to go to completion but can be accelerated through the use of polar, aprotic co-solvents like dimethyl sulfoxide (DMSO). This technique has been used by Cabezas *et al.*, to produce a synthetic HIV-1 vaccine [8] as well as peptide dendrimers [9] (refer to Table 1).

### 2.2. Oxime Ligation

An oxime bond can be created by reacting an oxyamine with a free carbonyl under acidic conditions. This reaction proceeds at a modest rate under acidic conditions. However, improved kinetics can be achieved either by using millimolar concentrations of each reactant or a large excess of one of the reagents. The reaction can be further accelerated by adding a nucleophilic catalyst such as aniline. This effectively changes the electrophile from a relatively unpopulated carbonyl, to the highly populated Schiff base of the aniline. This addition results in up to a 400-fold increase in the rate of oxime formation in mildly acidic aqueous conditions [10]. Oxime formation is also heavily influenced by temperature and co-solvents. Reactions conducted under high heat (37 °C) and/or using DMSO or dimethylformamide (DMF) as co-solvents saw a marked increase in oxime bond formation, as well as a substantial decrease in the number of observed side reactions in the case of the DMF [11]. Also, unlike other imines such as hydrazones, oximes are stable at physiological pH [10]. The reaction proceeds very close to completion (~90%), and the final polyoximes are easily purified and soluble in

both water and phosphate buffered saline (PBS) resulting in relatively good yields [9]. Oxime ligation has been used to successfully synthesize peptide dendrimers [11] and famously by Rose to synthesize homogenous, branched polypeptides with 195 amino acid residues [9] (refer to Table 1).

### 2.3. Thiazolidine Ligation

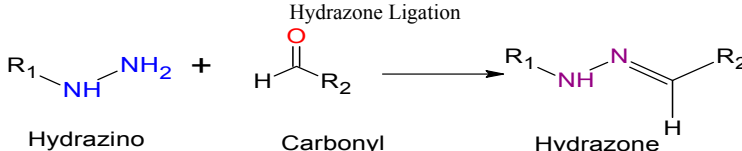
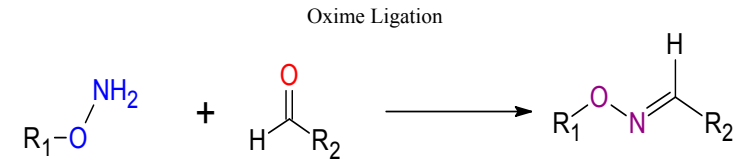
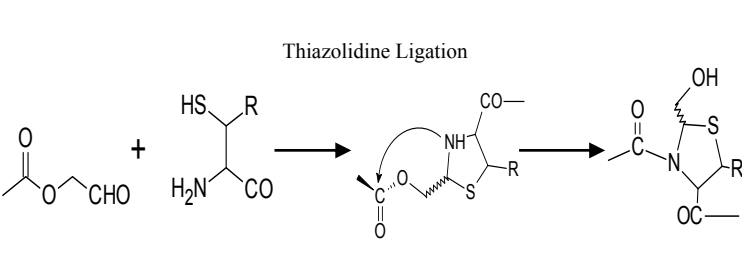
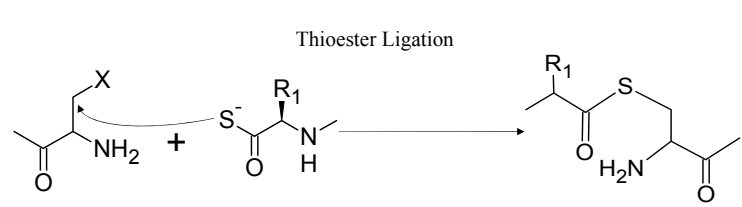
In order to achieve the goal of overcoming the kinetic barrier of ligating two large peptides, a highly specific reaction was necessary. This was first accomplished in 1994 by Liu and Tam by reacting the alkyl aldehyde glycolaldehyde with the mercaptamine moiety of a cysteine residue that functions as a weak base. These reagents coupled with acidic conditions, which effectively excludes stronger bases, like amines and guanidino groups, from reacting. The high specificity of this reaction makes side-chain protection groups unnecessary. This arrangement enables the formation of a stable 5-membered thiazolidine ring via a dehydration reaction. Ring formation is fast and complete, typically taking ~15 min at pH 5–6 and less than 5 min at neutral or basic pH. The final step in this reaction is the *O,N*-acyl rearrangement. This reaction, whereby the secondary amine of the newly formed thiazolidine ring performs a nucleophilic attack on the carbonyl carbon, constitutes the rate-determining step of this process. Since the amine is a weak base, the reaction itself proceeds best under acidic conditions and is also dependent on the steric and electronic hindrances of both the carbonyl and amino functionalities.

Liu and Tam originally hypothesized that the dependence on acidic pH is likely due to the electron-withdrawing capability of the carboxamide group, which leads to a decrease in the basicity of thiazolidine amine [12]. However thiazolidine ligation is made less desirable by its reaction conditions. In order to achieve the acidity necessary for the *O,N*-acyl rearrangement, trifluoroacetic acid (TFA) is typically used. Furthermore, a heavy metal catalyst, like  $\text{Ag}^+$ , is also needed [13]. This methodology pioneered by Liu and Tam [12] was subsequently used by Liu *et al.*, in order to synthesize analogs of HIV-1 protease [12], as well as for the creation of peptide dendrimers [11] (refer to Table 1).

### 2.4. Thioester Ligation

This technique, developed by Schnölzer and Kent, involves a thioester fragment performing a nucleophilic attack on an acyl halide in a  $\text{S}_{\text{n}}2$ -type mechanism, resulting in a ligated product that is indefinitely stable under mildly acidic conditions (pH 4.3). The reaction is rapid and goes almost to completion after 3 h, with extensive product formation detected via reverse phase-high performance liquid chromatography (RP-HPLC) after 45 min. However, due to the fact that the thioester bond is labile under basic conditions (pH > 7.5), this technique must be performed using Boc-chemistry which necessitates the use of HF (when using solution-based synthesis), or TFA (during solid phase peptide synthesis) for the deprotection step. This was used to construct HIV-1 protease from two unprotected peptide fragments [14,15]. All of the previously mentioned techniques rely on joining peptide fragments through the use of non-native bonds. However, criticism [15] of how this would affect the biological activity of the resulting peptide led to the development of Native Chemical Ligation (NCL) by the Kent lab, a technique which retains the native peptide bond between the ligated fragments, preserving the biological activity of the resulting peptide (refer to Table 1).

**Table 1.** Highlighting the advantages, disadvantages, and underlying mechanism of hydrazone, oxime, thiazolidine and thioester ligation.

Technique	Advantages	Disadvantages	Ref.
<p style="text-align: center;">Hydrazone Ligation</p>  <p>Hydrazino + Carbonyl (aldehyde) → Hydrazone</p>	<ul style="list-style-type: none"> <li>• Greater hydrolytic stability</li> <li>• Good water solubility</li> <li>• Highly chemo-selective</li> <li>• DMSO co-solvent accelerates reaction rate by 20×</li> <li>• Bond is stable at neutral pH (5–7)</li> </ul>	<ul style="list-style-type: none"> <li>• Reaction rate not affected by in pH</li> <li>• Rate only mildly affected by temperature (37 °C)</li> </ul>	[7–9,16]
<p style="text-align: center;">Oxime Ligation</p>  <p>Oxyamine + Carbonyl → Oxime</p>	<ul style="list-style-type: none"> <li>• Stable at room temp at pH 2–7</li> <li>• Fragments self-assemble under relatively mild conditions</li> <li>• Reaction goes close to completion (~90%)</li> <li>• Good yield</li> <li>• Poly-oximes are easily purified and water soluble</li> </ul>	<ul style="list-style-type: none"> <li>• Long reaction time (16–18 h)</li> <li>• Basic pH leads to more side reactions</li> </ul>	[9–11,16]
<p style="text-align: center;">Thiazolidine Ligation</p> 	<ul style="list-style-type: none"> <li>• Can create longer polypeptides (&gt;50 residues)</li> <li>• Good yield</li> <li>• Retains <math>K_d</math> similar to native</li> <li>• Acidic conditions prevents side reactions of aldehydes</li> <li>• Ring formation is quick and complete (5–15 min)</li> <li>• No protection groups</li> <li>• Highly selective reaction</li> </ul>	<ul style="list-style-type: none"> <li>• Heavy metal catalyst needed (<math>Ag^+</math>)</li> <li>• Highly acidic conditions (TFA)</li> <li>• Possible reduction in <math>V_{max}</math></li> <li>• <i>O,N</i>-Acyl rearrangement is slow and dependent on steric hindrance</li> </ul>	[11–13]
<p style="text-align: center;">Thioester Ligation</p> 	<ul style="list-style-type: none"> <li>• Ligation of N and C-terminal fragments occurs rapidly via <math>S_N2</math> reaction</li> <li>• Ligation product was indefinitely stable at pH 4.3 (ligation conditions)</li> <li>• Easily purified by RP-HPLC</li> <li>• Ligation site can be replaced with any amino acid</li> </ul>	<ul style="list-style-type: none"> <li>• HF is necessary for Boc deprotection</li> <li>• Thioester bond is labile at higher pH, with hydrolysis occurring at ~7.5, leading to a half-life of about 2 h</li> <li>• Must account for inversion of configuration that occurs with <math>S_N2</math> reactions</li> </ul>	[13–15]

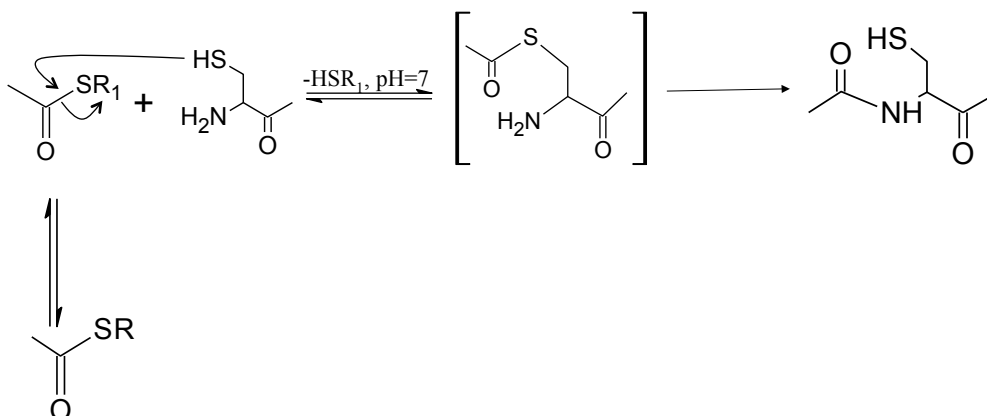
### 3. Native Chemical Ligation

#### 3.1. Native Chemical Ligation at Cysteine

Native chemical ligation has played a significant role in the advancement of peptide chemistry. NCL allows two peptide fragments to be joined in a covalent peptide bond at the site of ligation [5]. NCL helps overcome the shortcomings of Solid Phase Peptide Synthesis (SPPS) by allowing researchers to synthesize peptides that are greater than 50 amino acids in length [2,17]. The covalent nature of the bond at the site of ligation has been demonstrated through chemistry [5], X-ray crystallography [18] and NMR [19]. Another important feature of NCL is the absence of racemization at the site of ligation. Studies conducted by Lu *et al.*, demonstrated that the ligated product had no racemization within a limit of less than 1% D-amino acids [20]. The success of NCL can be attributed to the regioselective and chemoselective nature of the reaction. Ligation only occurs at the C-terminal thiol-ester and N-terminal cysteine thereby making the reaction highly specific and efficient. The presence of internal cysteine residues does not hinder this reaction and no additional protection groups are required for any of the amino acids found in the protein thereby making NCL an invaluable regioselective reaction [15,21,22]. This highly selective reaction is performed in an aqueous environment at neutral pH in the presence of denaturing agents. The presence of denaturing agents like 6 M guanidine hydrochloride allows the reaction to be performed at high concentration without causing aggregation of the reactants [2,5,6].

An NCL reaction involves three major steps. The first step involves the thiol-thioester exchange between the thioester containing peptide fragment and externally added alkyl or aryl thiol to create a new thioester containing peptide. In the second step a transthioesterification reaction occurs between the new thioester peptide and N-terminal cysteine to produce a thioester-linked intermediate. In the third step the thioester linked intermediate undergoes a rapid intramolecular S to N acyl transfer rearrangement to form a native peptide bond at the ligation site. The first two steps in NCL involving the exogenous thio-thiol ester exchange and transthioesterification are reversible, whereas the third step, which involves the rearrangement of intermediate, is irreversible. The irreversibility of the third step and the reversible nature of the first two steps cause the freely equilibrating intermediates to be diminished over time leading to the formation of ligated product [5,23] (Scheme 1).

**Scheme 1.** The mechanism of Native Chemical Ligation. R<sub>1</sub> can be an Alkyl or an Aryl group.

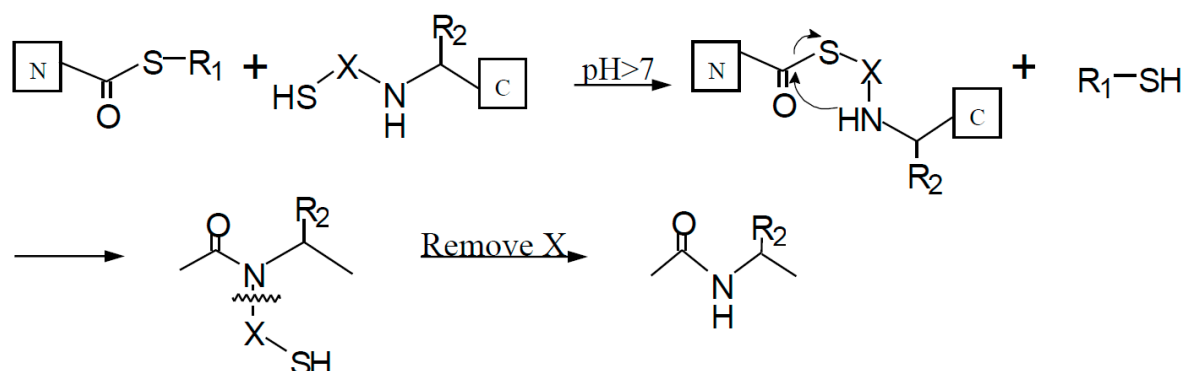


During the first thiol-thioester exchange the exogenous thiol could be an alkyl or an aryl thiol. However aryl thiols are preferred over alkyl thiols because aryl thiols have a lower  $pK_a$  and are better leaving groups. NCL reactions involving aryl thioesters can take days to go to completion, whereas reactions with alkyl thioesters happen relatively quickly [24]. Johnson *et al.* [25] examined fourteen different thiol catalysts chosen on the basis of their  $pK_a$  and found that aryl thiols were better catalysts than alkyl thiols. Furthermore they demonstrate that 4-mercaptophenylacetic acid (MPAA) is a non-malodorous water-soluble thiol catalyst that is highly efficient at optimizing NCL. The Kent laboratory has used MPAA in large syntheses such as HIV-1 protease [26] and RNase A [27] further validating the viability of MPAA as an efficient catalyst in NCL. Ligation rate is also dependent on the C-terminal thioester. Ligation occurs faster at less hindered amino acids and slowly at more hindered amino acids. The work of Dawson *et al.* [23,24] demonstrated that of the 20 amino acids at least 17 amino acids could be used as C-terminal thioester containing amino acids. Valine, isoleucine and proline are not suitable as C-terminal amino acids due to their slow ligation rates. NCL has undoubtedly been a boon to peptide chemistry and some of the applications are briefly discussed in this review.

### 3.2. Native Chemical Ligation at Non-Cysteine Sites: Auxiliary Mediated Ligation

One of the major limitations with NCL is the requirement of cysteine for ligation. In cases where cysteine is not present or is located at a position that is not amenable for ligation, NCL loses its effectiveness [28]. Auxiliary mediated ligation (AML) was designed to circumvent this problem and this technology has been relatively successful. AML involves an auxiliary group at the N-terminus that contains a free thiol which functions to replace the thiol in cysteine. The thiol-containing auxiliary reacts with the C-terminal thioester (Scheme 2) and upon removal of the auxiliary we get the desired product. Electron-rich  $N^\alpha$ -benzyl type systems such as the  $N^\alpha$ -(1-phenyl-2-mercaptoethyl) or the 4,5-dimethoxy-2-mercaptobenzyl (Dmb)-auxiliary have been the most successful auxiliary systems [29–33]. Auxiliary mediated ligation has been successfully implemented in the synthesis of the electron transfer protein cytochrome b562, the 62 amino acid SH3 domain from  $\alpha$ -spectrin, cyclic peptides, and glycopeptides [28,34,35]. A requirement of AML is the presence of glycine at one of the ligation sites due to the bulkiness of the auxiliary group.

**Scheme 2.** Mechanism depicting the steps of Auxiliary Mediated Ligation (AML), a technique used to join two peptide fragments resulting in large peptide molecules [36].



### 3.3. Ligation at Non Cysteine Amino Acids

NCL followed by desulfurization has been used to convert a cysteine residue to an alanine residue after ligation. Since alanine is more abundant than cysteine, ligation in conjugation with desulfurization can be used to expand the amino acids amenable to the ligation process [28]. This technology was reported by Yan and Dawson who utilized this strategy to synthesize microcin J25 [37], which is a 21 amino acid cyclic peptide that has one alanine residue and no cysteine residue. A cysteine residue was utilized to perform intramolecular ligation to give a cyclic peptide. The cyclic peptide was then subjected to a desulfurization reaction to give a native sequence.

The desulfurization reaction is performed in the presence of metal reagents. Several metal reagents such as Pd/Al<sub>2</sub>O<sub>3</sub>, Pd/C, Pd/BaSO<sub>4</sub> and Raney nickel have been used for the desulfurization reaction [28,37]. Cysteine residues not participating in desulfurization need to be protected during desulfurization. Kent *et al.* proposed the use of acetamidomethyl (Acm) as a protecting group for cysteine. Acm is stable to both Boc and Fmoc chemistry and can be removed using iodine [38]. One of the major problems encountered with metal-based desulfurization technology is the low yield of product, which results from peptide aggregation and adsorption on large metal surfaces. These desulfurization reactions are also prone to undesirable side reactions such as hydrogenation of tryptophan and demethylthiolation of methionine which further reduce the yield of final product [37]. Methods have been developed to eliminate the metal reagents used during desulfurization. To overcome some of the drawbacks associated with metal based desulfurization Wan and Danishefsky developed a technique for desulfurization of mercaptans by using trialkylphosphites under thermal and photochemical conditions [39].

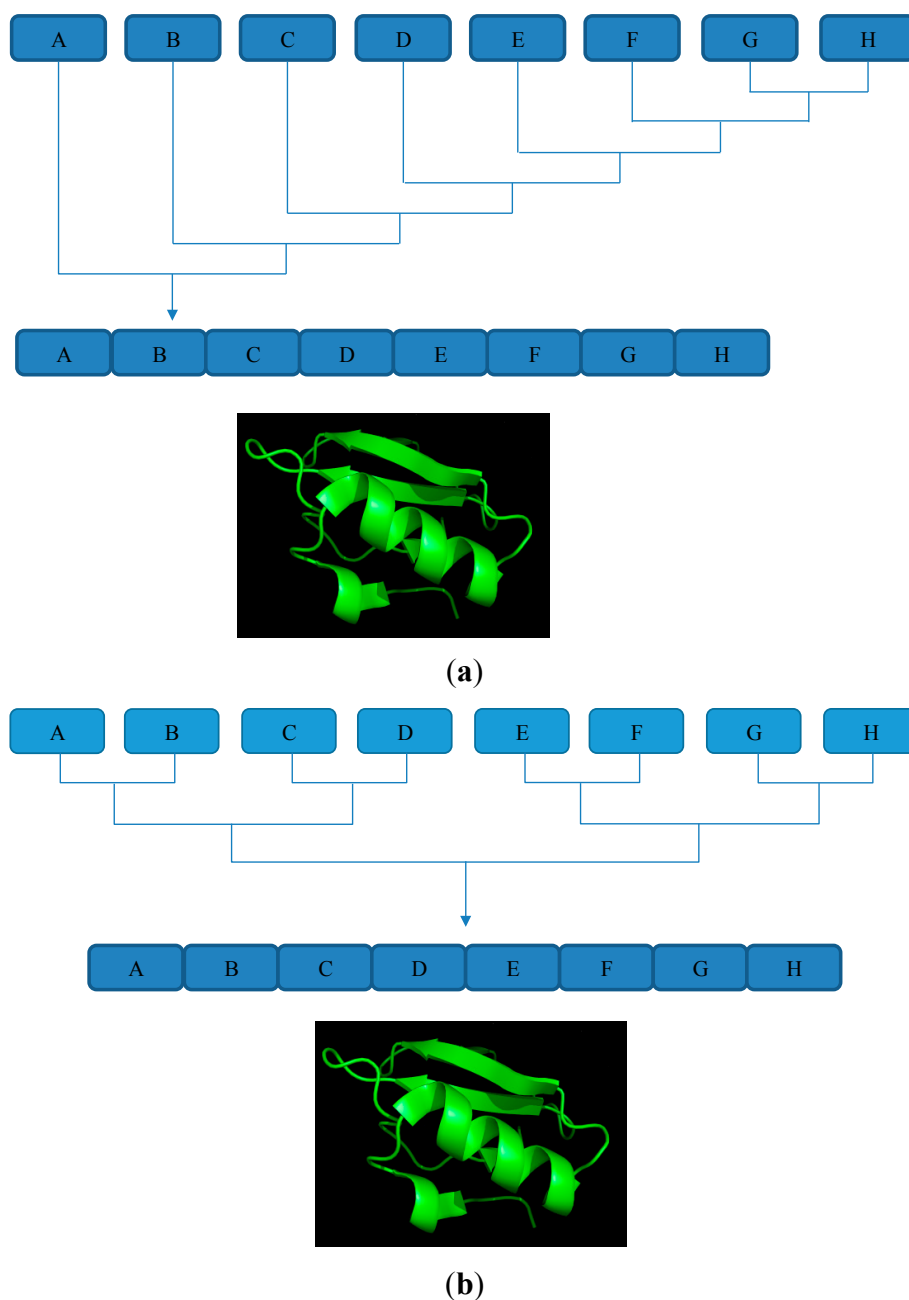
Recent methodologies have been developed in which amino acids other than alanine can be generated through desulfurization. The technology developed by Crich *et al.* [40] and Botti *et al.* [41] uses a modified phenylalanine in which the  $\beta$ -carbon has mercaptan group. This modified amino acid can be used in NCL and upon desulfurization is converted to phenylalanine. The major challenge in this technique is that  $\beta$ -mercaptophenylalanine is not commercially available and was developed by Crich *et al.* in a five step procedure using *threo*- $\beta$ -hydroxy-L-phenylalanine derivative as the starting material [40]. Modified valine has also been used in NCL. Seitz *et al.* [42] reported using  $\beta,\beta$ -dimethylcysteine as a valine alternative, whereas Yan *et al.* [37] have proposed the usage of  $\gamma$ -thiolated valine for NCL-desulfurization procedures.  $\beta,\beta$ -Dimethylcysteine is commercially available as penicillamine and was used to make two 22mer spanning segments of transmembrane spanning proteins STAT-1 and Syk-kinase. In both cases radical-mediated desulfurization reactions were performed to give valine [28]. As mentioned above, various attempts have been undertaken to perform NCL at non-cysteine residues and several successful strategies have been developed to overcome the cysteine requirement for NCL. All of the above mentioned technologies have expanded the repertoire of amino acids that can be utilized in native chemical ligation.

## 4. Sequential and Convergent Ligation

Native chemical ligation can be applied in the synthesis of functional proteins [15,43]. Sequential ligation and convergent ligation can be used to join multiple small fragments to give a large protein

(Figure 1). Careful synthetic design and efficient planning to minimize product loss and byproduct accumulation are extremely important when undertaking such a synthetic endeavor. Small high quality fragments of up to 30 amino acids residues are more desirable than longer fragments of lower quality. Since HPLC purification and lyophilization lead to handling losses, one-pot ligation or solid phase ligation can be undertaken to avoid HPLC purification of intermediates [36,44,45].

**Figure 1.** (a) Scheme showing sequential ligation strategy for joining peptide fragments together in order to produce a large peptide molecule [45]; (b) Scheme showing a strategy for ligating smaller fragments together in a convergent fashion in order to produce larger peptide molecules [45].



Some of the major strategies available for sequential ligation are one-pot ligation, His<sub>6</sub>-tag assisted ligation and solid phase ligation. A one-pot ligation strategy was developed to eliminate purification



process thereby reducing the material loss through handling. Sequential ligation of peptide fragments is undertaken in one reaction vessel to minimize handling losses. Using a 1,3-thiazolidine-4-carboxy (Thz) protecting group for the N-terminal cysteine of the peptide thioester fragment Bang *et al.* [44,46] were able to perform one pot ligation by adding methoxyamine hydrochloride and changing the pH of the reaction solution. Another strategy employed for one pot ligation is using a photolabile moiety to protect the N-terminal cysteine of peptide thioester fragment. Upon photoradiation the protecting group is removed and ligation is continued [47]. One pot sequential ligation strategy was adopted in the synthesis of crambin and ubiquitin [45]. The major drawback in one pot ligation is that byproduct and unreacted peptide fragments accumulate with each sequential reaction.

His<sub>6</sub>-tag ligation allows rapid purification and high recovery of target material after sequential ligation of the peptide fragment [48]. In this methodology the C-terminal peptide fragment has a His<sub>6</sub> tag that binds to a Ni-nitrilotriacetic acid (NTA) agarose column. The NCL reaction can be driven to near completion by using the peptide thioester fragment in excess. An advantage of this method is that since the C-terminal fragment is attached to a resin support, a simple washing procedure is used to eliminate the unreacted thioester fragment and non-tagged fragments. Another advantage of the His<sub>6</sub>-tag sequential ligation technique is that it allows the intermediates in the reactions to be eluted thereby providing analytical control at various points in the ligation process. A tetratrico peptide repeated protein and a crambin protein have been successfully made through His<sub>6</sub>-tag ligation strategy [45,48].

Solid phase sequential ligation adopts principle from solid phase peptide synthesis. Like peptide synthesis, the growing peptide fragments are attached to a resin support (Figure 1a). Similar to peptide synthesis excess reactants and undesired byproducts can be removed through filtration. Comparable to SPPS the target polypeptide can be generated by cleaving the polypeptide from the resin support [49,50]. An acid labile safety catch linker [49] and selectively cleavable ester linker [51] are used to grow the polypeptide on a resin support. Solid phase sequential ligation strategy has been successfully employed in the synthesis of C5a (74 amino acids), MIF (115 amino acids), and human group V secretory Phospholipase A2 (118 amino acids, six disulfide bonds) [45].

Another method for joining multiple peptide fragments is through convergent ligation. Unlike sequential ligation, in convergent ligation two halves of a target sequence are ligated separately and condensed to give the target polypeptide (Figure 1b). Convergent ligation is more efficient than sequential ligation because there is a lesser chance for byproduct accumulation and purification is easier in this approach [52,53]. Initially, convergent ligation led to the formation of a non-native peptide bond at the site of ligation [54]. Researchers utilized various functional groups to generate non-native bond at ligation site. A thioester bond forming ligation was developed that used thioacid and bromoacetyl functional groups. Additionally, an oxime bond forming ligation was developed through (aminoxy) acetyl and ketone groups [45]. Baca *et al.* reported synthesizing a HIV-1 protease analogue through convergent peptide ligation. This synthetic analogue showed similar catalytic activity to the native protein but possessed a non-native bond at various ligation sites [55]. With the advent of kinetically controlled ligation researchers were able to perform convergent ligations that generated native peptide bonds at the site of ligation [52]. Kinetically controlled ligation works on the principle that peptide- $\alpha$  thiophenylester reacts with a Cys-peptide faster than a peptide- $\alpha$  thioalkylester [56]. Kent *et al.*, proposed that, in the absence of external thiophenol under competitive reaction conditions, the large difference in the reaction rates between the  $\alpha$ -thiophenylester and

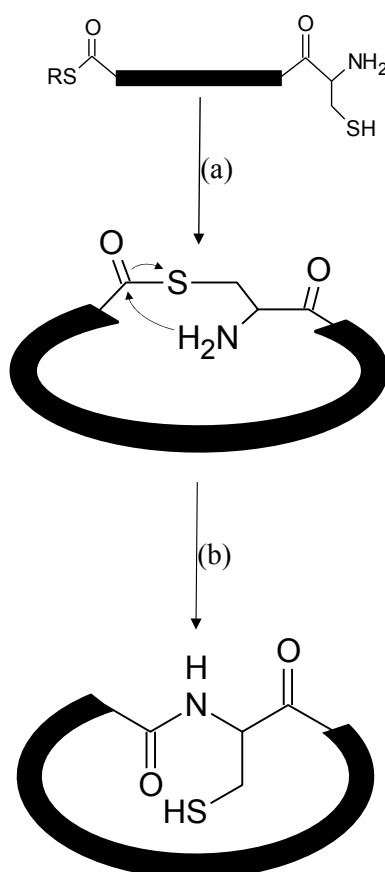
$\alpha$ -thioalkylester would make the  $\alpha$ -thioalkylester unreactive [45,52]. Kinetically controlled ligation has been successfully implemented in the synthesis of human lysozyme (130 amino acids) [57] and a covalent dimer of HIV-1 protease (203 amino acids) [58].

## 5. Application of Native Chemical Ligation

### 5.1. Cyclotides and Conotoxins

NCL technology can be implemented to make peptides in which the amino and carboxy termini are attached in a covalent bond to form a cyclic molecule. When the thioester and free cysteine functionalities are present on the same peptide fragment, an intramolecular NCL reaction can be used to generate a cyclic molecule (Scheme 3). Recently, naturally occurring cyclic molecules have been discovered in bacteria, plants and animals [59,60]. NCL reactions can be utilized to synthesize these circular proteins to find potential new drug leads and design scaffolds that aid in drug delivery.

**Scheme 3.** Mechanism showing the cyclization of a linear peptide. (a) Intramolecular nucleophilic attack results in a cyclic peptide joined by a thioester bond (b) S-N Acyl Shift produces the final cyclic molecule now joined by a native peptide bond.



Cyclotides are disulfide rich compounds naturally occurring in plants. They generally have about 30 amino acids and are predominantly found in the Rubiaceae (coffee) and Violaceae (violet) families [61–63]. Cyclotides fall into two major subfamilies, the bracelet family that contains two-thirds of cyclotides, and the mobius family that contains the remaining one-third. There is a third subfamily

called trypsin inhibitor family comprised of eight members [64,65]. Cyclotides are characterized by head to tail cyclization and contain six conserved cysteine residues arranged in a knotted topology. The cyclic cystine knot morphology provides cyclotides with exceptional stability thereby making them prime candidates for grafting studies and as templates for drug design [66,67]. These cyclic molecules have shown a vast array of activities ranging from antimicrobial, cytotoxic and anti-HIV. In nature cyclotides function as host defense agents and possess potent insecticidal activity [62,68–71].

Craik *et al.*, have used NCL to successfully conduct grafting studies on cyclotides. They combined stable cyclotide fragments with bioactive peptide epitopes to produce chimeric molecules with enhanced biopharmaceutical properties. As a proof of concept they grafted a highly polar region to a cyclotide and showed that the molecule still maintained its structural integrity [72]. Furthermore Gunasekera *et al.*, incorporated an antiangiogenic epitope (RRKRRR) referred to as the poly R sequence into kalata B1. The chimeric molecule possessed antiangiogenic activity and was more stable than the naked epitope [73]. Similarly Thongyoo *et al.*, have proposed that the active site in cyclotides loop 1 could be replaced with a sequence to inhibit a protease from the Foot and Mouth Disease virus [74]. Reengineering of cyclotides to produce chimeric molecules has shown promise in therapeutic development, and with further advancement this technology might produce future drugs.

Like cyclotides, conotoxins are natural products that are rich in cysteine residues thereby providing ample opportunities for NCL. Conotoxins are found in the venom of conesnails and act as potent antagonists of a range of neuronal receptors and ion channels [75].  $\omega$ -conotoxin MVIIA, a potent N-type calcium channel blocker extracted from *Conus magus*, was developed into ziconotide. This “conopeptide” is the first non-opioid IT treatment for the management of chronic refractory pain and is known by its brand name Prialt<sup>®</sup> [76–78]. One of the major hurdles in developing conopeptides as therapeutics is overcoming their low oral bioavailability [65]. Peptides and proteins have low oral bioavailability due to degradation caused by digestive enzymes present in the intestinal lumen [79–81]. Secondly, the inability of peptides to permeate through membranes owing to size and polarity also results in low bioavailability [79,82]. Consequently any method that improves the stability and bioavailability of peptides and proteins is highly desirable. Work done by Clark *et al.*, on cyclic  $\alpha$ -conotoxin MII [83], Clark *et al.*, on cyclic  $\alpha$ -conotoxin Vc1.1 [84] and Lovelace *et al.*, on cyclic  $\chi$ -conotoxin MrIA [85] has demonstrated that NCL can be used to cyclize conotoxins. Their work shows that when provided with a linker of right size the cyclic molecules can maintain their biological activity and are more stable than native peptides. Cyclization helps with stability by eliminating the free termini of a peptide and makes it less labile to exopeptidases. Additionally cyclization enhances stability by introducing rigidity and reducing the conformational energy of the unfolded state [86]. The added rigidity makes the peptide more resistant to proteases that cleave from the middle of the chain [87]. Conotoxins are a rich source of biological leads that might provide therapies for pain management, seizures, and Parkinson’s [84,88]. Through cyclization we can improve the therapeutic potential of conotoxins and enhance the number of potential leads in our pharmacological repertoire.

## 5.2. NCL in Overcoming Length Limitation Associated with SPPS

SPPS is a great tool for making peptides with non-natural amino acids or incorporating post-translational modification(s). However the size limitation imposed by SPPS puts constraints on its

usefulness. SPPS in conjugation with NCL can be used to overcome this hurdle. Initial ligation techniques introduced a non-native bond at the site of ligation (Section 1) but with the development of NCL ligations, the introduction of native bonds became possible.

One of the first total syntheses of a protein by chemical ligation was human immunodeficiency virus-1 protease (HIV-1 PR) [14]. The 22.5 kDa HIV-1 PR protein is a homodimer of two 99 amino acid polypeptide chains. The monomer was prepared by chemical ligation of a 51-residue peptide containing a C-terminal thioacid and a 48-residue peptide having an N-terminal alkyl bromide. However, the major disadvantage of this was producing a non-native bond to link the two unprotected peptide fragments.

With the advent of NCL the problem of non-native ligation was resolved. Human interleukin 8 (IL-8), which contains 72 amino acids was the first peptide synthesized using NCL [5]. SPPS was used separately to synthesize residues 1–33 of IL-8 with C-terminal thioester and residues 35–72 of IL-8 containing N-terminal free cysteine residue. Using NCL, a native bond was introduced at the point of ligation and the 72 amino acid peptide was generated. Human type II secretory phospholipase A<sub>2</sub> (sPLA<sub>2</sub>), an enzyme made up of 124 amino acids that is found in the  $\alpha$  granules of platelets and at inflammation sites, was also synthesized using NCL. sPLA<sub>2</sub> was prepared by ligating two fragments, the first fragment contained residues 1 to 58 and the second fragment contained residues 59 to 124 [22]. The following example describes the chemical synthesis of a multidomain anticoagulant protein, “microprotein S”, through modular assembly of synthetic peptide segments by native chemical ligation. In this case, the native peptide bond was produced between Cys-TSR (47–79)-thioester and Cys-EGF1 (80–116) to form the first segment. After that, the Gla-rich fragment (1–46)-thioester was ligated to the segment through the creation of another native peptide bond that completed the synthesis of full-length microprotein S [89]. NCL has allowed peptide chemists to synthesize functional proteins and is undoubtedly a great tool in synthetic protein chemistry (Table 2).

**Table 2.** Examples of the application of NCL to circumvent peptide size limitation. Cysteine residues used as ligation sites are in bold and larger font.  $\gamma$  = Gamma carboxyglutamic acid.

Toxin/Peptide	Application	Sequence	Ref.
Human interleukin 8 (IL-8)	Chemokine, chemoattractants for leukocytes	SAKELRCQCIKTYSKPFHPKFIKELRVIESGPHCANTEII VKLSDGRELCCLDPKEWVQRVVEKFLKRAENS	[5,90]
Human group II secretory phospholipase A <sub>2</sub> (sPLA <sub>2</sub> )	Hydrolysis of the fatty acid side chain ester bond	NLVNFHRMIKLLTTGKEAALSYGFGCHCGVGGRGSPK DATDRCCVTHDCCYKRLEKRGCGTKFLSYKFSNSGSRI TCAKQDSCRSQLCECDKAAATCFARNKTTYNKKYQY YSNKHCRGSTPRC	[22]
Microprotein S	Anticoagulant cofactor activity	NSLL $\gamma$ $\gamma$ TKQGNL $\gamma$ R $\gamma$ Cl $\gamma$ $\gamma$ LCNK $\gamma$ $\gamma$ AR $\gamma$ V $\gamma$ F $\gamma$ NDP $\gamma$ TDYFYP KYLGLCLRSFQTGLFTAARQSTNAYPDLRSCVNAIPDQC SPLPCNEDGYMSCKDGKASFTCTCKPGWQG EKCEFD	[89]
Barnase (Lys <sup>49</sup> -Cys <sup>49</sup> )	Microbial ribonuclease	AQVINTFDGVADYLQTYHKLPNDYITKSEAQALGWVA SKGNLADVAPGCSIGGDIFSNREGKLPKSGSRTWREA DINYTSGFRNSDRILYSSDWLIYKTTDHYQFTFKIR	[24]

### 5.3. Synthesis of Challenging Sequences

In addition to being a great tool to circumvent some of the length limitations associated with SPPS, NCL is also highly useful in synthesizing challenging proteins. Membrane proteins are difficult to synthesize because of their large hydrophobic peptide segments, these segments lead to difficulties with amino acid coupling, formation of aggregates, and poor peptide purity and solubility [91]. In order to avoid these problems, NCL is applied, joining multiple segments to obtain a full length and functional peptide chain. A 136-mer mechanosensitive ion channel from *Escherichia coli* (Ec-MscL) was also synthesized using NCL [92]. This chemically synthesized membrane protein had cysteine residues which were not present in the native sequence [93,94]. The introduction of cysteine residues aids in ligation but results in the presence of non-native sulfhydryl groups in the final protein. In this case, ligation of three peptide fragments through NCL resulted in a synthetic, non-native Ec-MscL protein possessing similar ion-channel function to the native protein. Another scenario where NCL has shown its usefulness is during the synthesis of snow flea antifreeze protein (sfAFP). This cysteine-rich protein has 81 amino acids in its sequence and contains Cys residues at positions 1, 13, 28, and 43 [95]. sfAFP has unusual properties such as thermal instability, therefore NCL was the desired route for synthesis. The antifreeze protein was synthesized by ligating four unprotected peptide fragments that contained three thioesters: Thz1–12-thioester; Thz13–27-thioester; Thz28–42-thioester; Cys43–81 [96,97]. Using convergent and sequential ligation many syntheses of long and difficult proteins have been accomplished. With further refinement and advancement this technology is bound to help further advance biomedical research (Table 3).

**Table 3.** Examples of the application of NCL to synthesize challenging peptide sequences. Cysteine residues used as ligation sites are in bold and larger font.

Toxin/Peptide	Application	Sequence	Ref.
Ec-MscL	Mechanosensitive ion channel	MSIIKEFREFAMRGNVVDLAVGVIIIGAAFGKIVSSLVADIIMPPLG LLIGGIDFKCFVTLRDAQGDIPAVVMHYGVFIQNVDFDLIVAF IFMAIKLINKLCRKKEEPAAPAPTKEEVLLTEIRDLLKEQNNRS	[92]
IbTx V16A/ D19-Cys-4-MeOBzl	Binding to BK Ca <sup>2+</sup> -activated K <sup>+</sup> channel (K <sub>Ca</sub> 1.1)	ZFTDVDCSVSKECWSACKX <sub>2</sub> LFGVDRGKCMGKKRCRCYQ (X <sub>2</sub> = Cys-4-MeOBzl)	[98]
Snow Flea Antifreeze Protein (sfAFP)	Inhibition of Ice Crystal Formation	CLGADGAHGVNGCPGTAGAAAGSVGGPGCDGGHGGNGGNGNP GCAGGVGGAGGASGGTGVGGRRGGKGGSGTPKGADGAPGAP	[96,97]
KcsA (Pro <sup>2</sup> -Ala <sup>2</sup> , Gln <sup>58</sup> -Ala <sup>58</sup> , Thr <sup>61</sup> -Ser <sup>61</sup> , Arg <sup>64</sup> -Asp <sup>64</sup> )	Potassium Channel	MAPMLSGLLARLVKLLLGRHGSALHWRAAGAATVLLVIVLLA GSYLAVLAERGAPGAALISYPDALWWSVETACTVGYGDLYPVT LWGRLVAVVMVAGITSFGLVTAALATWVFGREQERRGK	[99]

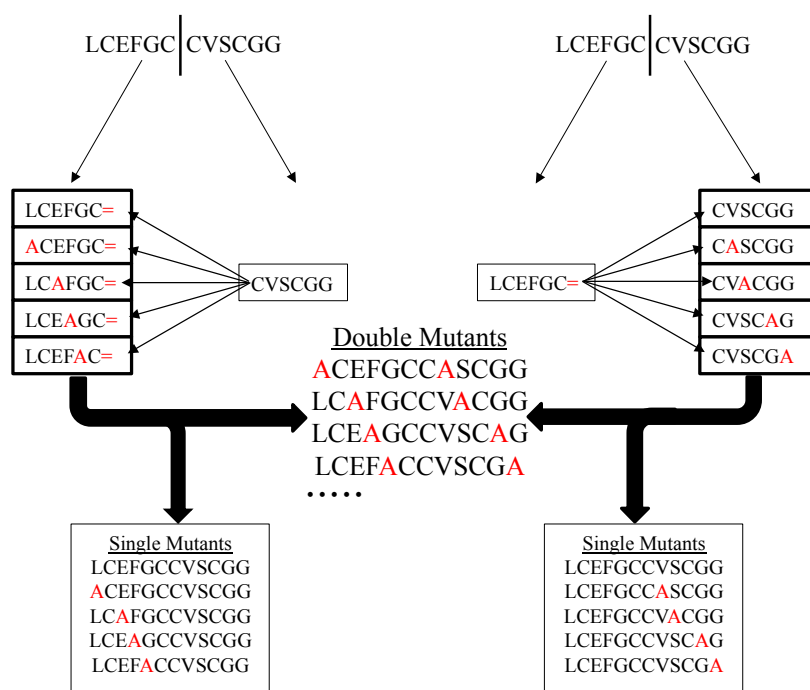
### 5.4. Use in Doing Alanine Scans to Study Structure Activity Relationships

Alanine scanning is a tool used in molecular biology and proteomics to determine the contribution of specific amino acid residues towards stability, functionality, and structure [100]. Furthermore, using alanine scanning it is possible to determine whether the specific residues contribute to bioactivity and protein-protein interactions. In this strategy, alanine is used to substitute each individual amino acid in the peptide sequence to assess the value of the substitution to the overall protein [101]. Since alanine is

non-bulky, chemically inert and can mimic the secondary structure of many other amino acids with its methyl functional group, alanine is the ideal model to replace other amino acids. Additionally, substitution with alanine removes all side chain atoms past the  $\beta$ -carbon, thus this method can be used to infer the roles of individual side chains in the final peptide. Site-directed mutagenesis is a powerful tool for researching protein structure-activity relationships and alanine-scanning mutagenesis has been particularly successful in systematically mapping the protein-binding interface [102,103]. Traditional alanine-scanning mutagenesis is a laborious endeavor because each mutant needs to be individually produced, purified and separately assessed in structural and functional assays. Binomial mutagenesis was developed as an alternative to the traditional method. This “next generation” technology was based on *in vivo* library-based mutagenesis and was considerably more attractive because a single library simultaneously provides information about many different residues. However it is still limited by the issues of heterogeneity of proteins, protein expression and protein folding [104]. In order to circumvent the need for protein purification and biophysical analysis, shotgun scanning was developed for mapping the functional epitopes of proteins. This methodology includes the concepts of conventional alanine mutagenesis and binomial mutagenesis along with phage display technology [101].

Using SPPS, researchers can synthesize different alanine substituents and one core peptide. In the next step, NCL can be used to ligate the core peptide to different fragments to generate one or two amino acid variants of the target peptide (Figure 2). NCL can be a powerful tool that can improve the efficiency of alanine scanning by: reducing the amount of required reagents, the number of syntheses, and time involved in conducting the process.

**Figure 2.** The figure demonstrates the application of Native Chemical Ligation in conducting an Alanine Scanning to study structure-activity relationships. The equal sign (=) designates a thioester functionality.



### 5.5. Conjugation with Recombinant Technology to Make Glycopeptides and Peptidomimetics

Glycoproteins have carbohydrates that are covalently attached to certain amino acid residues. A  $\beta$ -*N*-glycosidic linkage of *N*-linked glycoproteins connects *N*-acetylglucosamine to the amide residue of asparagine in Asn-Xaa-Ser/Thr motifs [105]. Alternatively an  $\alpha$ -*O*-glycosidic linkage between the glycan and hydroxyl groups of serine, threonine or tyrosine is called *O*-linked glycosylation [106]. Mannose, xylose, fructose and *N*-acetylgalactosamine are all found to be ligated to proteins through *O*-linked glycosylation [106]. Since many therapeutics are glycoproteins, several semi-synthetic strategies for glycoprotein preparation have been explored [107]. Moreover, recent developments in engineered cell culture employing yeast suggest that therapeutic proteins with a human-like glycosylation profile can be achieved [108].

However, human glycoproteins are heterogeneous and this level of glycosylation complexity cannot be wholly replicated in yeast recombinant systems, therefore, NCL could be an excellent platform for glycopeptide synthesis because it allows residues to be stoichiometrically and site-specifically glycosylated, while retaining native linkages. Therefore, NCL is the most widely used non-recombinant approach for the synthesis of glycoproteins. Diptericin  $\epsilon$ , an antibacterial glycopeptide, containing 82 amino acids and two *O*-linked glycosylation sites was the first major glycopeptide synthesized using NCL. In this case SPPS was used to make two peptide fragments: the first fragment consisted of a 24-residue segment with an *N*-terminal glycopeptide thioester and the second fragment consisted of a 58-residue glycopeptide containing an *N*-terminal cysteine residue. Both fragments were covalently bonded to *N*-acetylgalactosamine ( $\alpha$ -GalNAc, also known as the T<sub>N</sub> antigen) at the two glycosylation sites [109].

Traditionally, the formation of thioester is based on Boc chemistry; however, the acid deprotection condition of Boc chemistry is incompatible with glycopeptide thioesters, because low pH may lead to the breakage of glycosidic linkage [105]. Therefore, a “safety-catch” linker was employed for the synthesis of the *N*-terminal glycopeptide thioester fragment to circumvent this problem [110,111]. The safety-catch linker is normally unreactive with nucleophiles in basic and acidic conditions but, after completion of the synthesis, selective sulfonamide activation renders the linker susceptible to nucleophilic attack and allows the release of the final product. Lymphotoxin (Lptn), a glycoprotein with multiple *O*-glycosylation sites called “mucin type” [112], was also synthesized using NCL. Lptn is a 93-residue chemokine, which functions as a potent chemoattractant for T cells and natural killer cells [112,113]. The C-terminus of Lptn contains a mucin domain with up to eight *O*-glycosylation sites. During the synthesis of Lptn, the Lptn(1–48)-thioester fragment and the glycopeptide Lptn(49–93) fragment containing an *N*-terminal cysteine and eight  $\alpha$ -GalNAc were ligated to produce a 93 amino acid glycopeptide. Synthetic Lptn was subsequently assessed for its ability to bind its cognate chemokine receptor (XCR1). The synthetic glycopeptide activated the signal transduction cascade providing an increase in intracellular calcium concentrations [105,114], thereby suggesting that the synthetic analogue was similar in function to the native glycopeptide.

The first example of using NCL to construct an *N*-linked glycopeptide, containing a complex-type glycan, was RNase B [115]. In this case, Fmoc SPPS was used for the synthesis of glycodecapeptide thioester, two linkers were used on the resin before assembling the glycopeptide. The first of these linkers was the Ellman-sulfonamide safety-catch linker, which protected the glycosidic linkage [110].



The second was a Rink amide linker, which allowed for peptides to be released during assembly in order to assess the efficiency of the synthesis by HPLC and LC-MS. The glycopeptide containing the Cys-RNase(41–68) fragment was reacted with the C-terminal thioester of the other peptide fragment to produce RNase B (Refer to Table 4).

**Table 4.** Examples of the applications of NCL to synthesize glycopeptides and glycoproteins. Cysteine residues used as ligation sites are in bold and larger font. Glycosylation sites are shown with underlining residues.

Toxin/Peptide	Application	Sequence	Ref.
Diptericin $\epsilon$ (Cys <sup>25</sup> ,Glu <sup>29</sup> ,Glu <sup>45</sup> )	Antibacterial glycopeptide	DEKPKLILPT <b>P</b> APPNLPQLVGGGGCNRKEGFGVSVDAH QKVWTSSENGRHSIGV <b>T</b> PGYSQHLGGPYGNSRPDYRIG AGYSYNF	[109]
Lymphotactin (Lptn)	Chemoattractant (for T- cell and natural killer cell)	VGSEVSDKRTC <b>V</b> SLTTQRLPVSRIKTYTITTEGSLRAVIFI TKRGLKVCADPQATWVRDVVRSMDRKS <b>N</b> TRNNMIQ <b>T</b> K <b>P</b> T <b>G</b> I <b>Q</b> Q <b>S</b> T <b>N</b> T <b>A</b> V <b>L</b> L <b>I</b> G	[112]
RNase B Fragment	Cleavage of <i>N</i> -linked carbohydrates	MKSR <b>N</b> LTKDRCKPVNTFVHESLADVQAVCSQKNVAC KNG	[115]
CCL7 (MCP-3)	Monocyte specific chemotactic protein-3	Q <b>P</b> V <b>G</b> I <b>N</b> T <b>S</b> T <b>T</b> C <b>C</b> Y <b>R</b> F <b>I</b> N <b>K</b> K <b>I</b> P <b>K</b> Q <b>R</b> L <b>E</b> S <b>Y</b> R <b>R</b> T <b>S</b> S <b>H</b> C <b>P</b> R <b>E</b> A <b>V</b> I <b>F</b> K <b>T</b> K <b>L</b> D <b>K</b> E <b>I</b> C <b>A</b> D <b>P</b> T <b>Q</b> K <b>W</b> V <b>Q</b> D <b>F</b> M <b>K</b> H <b>L</b> D <b>K</b> K <b>T</b> Q <b>T</b> P <b>K</b> L	[116]

## 6. Conclusions

NCL developed by the Kent lab has been an indispensable tool for peptide chemists worldwide. By overcoming the size limitation imposed by SPPS, NCL has allowed the synthesis of many proteins such as HIV-1 protease, RNase B, and Ec-MscL. Researchers have applied the NCL technology to synthesize membrane proteins and cyclic peptides. Recently there has been an increase in the number of biologics entering the therapeutic market. By applying NCL there is potential to find new protein/peptide drug leads that were previously unattainable due to their size limitation or difficulty in synthesis. NCL is one of the major achievements of recent times and has made the field of synthetic peptide chemistry more efficient and diverse.

## Acknowledgments

We wish to acknowledge the continued financial support from the American Heart Association (Scientist Development Award 0530204N to J-P.B.), USDA TSTAR (#2009-34135-20067) and HATCH (HAW00595-R; J-P.B.) that have helped expand our horizons in peptide toxin bioengineering.

## Author Contributions

The authors contributed equally in formulation of this paper and its editing.

## Conflicts of Interest

The authors declare no conflict of interest, and have maintained professional and ethical standards in the work undertaken.



## References

1. Merrifield, R.B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
2. Dawson, P.E.; Kent, S.B. Synthesis of native proteins by chemical ligation. *Ann. Rev. Biochem.* **2000**, *69*, 923–960.
3. Grant, G.A. *Synthetic Peptides: A User's Guide*, 2nd ed.; Oxford University Press: Oxford, UK, 2002; p. 390.
4. Chan, W.C.; White, P.D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Oxford University Press: Oxford, UK, 2000; p. 346.
5. Dawson, P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776–779.
6. Chandrudu, S.; Simerska, P.; Toth, I. Chemical methods for peptide and protein production. *Molecules* **2013**, *18*, 4373–4388.
7. Day, A.C.; Whiting, M.C. Acetone hydrazone. *Org. Synth.* **1988**, *50*, 10–12.
8. Cabezas, E.; Wang, M.; Parren, P.W.; Stanfield, R.L.; Satterthwait, A.C. A structure-based approach to a synthetic vaccine for HIV-1. *Biochemistry* **2000**, *39*, 14377–14391.
9. Rose, K. Facile synthesis of homogeneous artificial proteins. *J. Am. Chem. Soc.* **1994**, *116*, 30–33.
10. Dirksen, A.; Hackeng, T.M.; Dawson, P.E. Nucleophilic catalysis of oxime ligation. *Angew. Chem. Int. Ed.* **2006**, *45*, 7581–7584.
11. Shao, J.; Tam, J.P. Unprotected peptides as building-blocks for the synthesis of peptide dendrimers with oxime, hydrazone, and thiazolidine linkages. *J. Am. Chem. Soc.* **1995**, *117*, 3893–3899.
12. Liu, C.F.; Tam, J.P. Chemical ligation approach to form a peptide-bond between unprotected peptide segments—Concept and model study. *J. Am. Chem. Soc.* **1994**, *116*, 4149–4153.
13. Liu, C.F.; Rao, C.; Tam, J.P. Orthogonal ligation of unprotected peptide segments through pseudoproline formation for the synthesis of HIV-1 protease. *J. Am. Chem. Soc.* **1996**, *118*, 307–312.
14. Schnolzer, M.; Kent, S.B. Constructing proteins by dovetailing unprotected synthetic peptides: Backbone-Engineered HIV protease. *Science* **1992**, *256*, 221–225.
15. Kent, S.B. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338–351.
16. Algar, W.R.; Prasuhn, D.E.; Stewart, M.H.; Jennings, T.L.; Blanco-Canosa, J.B.; Dawson, P.E.; Medintz, I.L. The controlled display of biomolecules on nanoparticles: a challenge suited to bioorthogonal chemistry. *Bioconjug. Chem.* **2011**, *22*, 825–858.
17. Muir, T.W.; Dawson, P.E.; Kent, S.B. Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol.* **1997**, *289*, 266–298.
18. Lu, W.; Randal, M.; Kossiakoff, A.; Kent, S.B. Probing intermolecular backbone H-bonding in serine proteinase-protein inhibitor complexes. *Chem. Biol.* **1999**, *6*, 419–427.
19. Xu, R.; Ayers, B.; Cowburn, D.; Muir, T.W. Chemical ligation of folded recombinant proteins: Segmental isotopic labeling of domains for NMR studies. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 388–393.
20. Lu, W.; Qasim, M.A.; Kent, S.B.H. Comparative total syntheses of turkey ovomucoid third domain by both stepwise solid phase peptide synthesis and native chemical ligation. *J. Am. Chem. Soc.* **1996**, *118*, 8518–8523.

21. Kent, S.; Sohma, Y.; Liu, S.; Bang, D.; Pentelute, B.; Mandal, K. Through the looking glass—A new world of proteins enabled by chemical synthesis. *J. Pept. Sci. off. Publ. Eur. Pept. Soc.* **2012**, *18*, 428–436.
22. Hackeng, T.M.; Mounier, C.M.; Bon, C.; Dawson, P.E.; Griffin, J.H.; Kent, S.B. Total chemical synthesis of enzymatically active human type II secretory phospholipase A2. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 7845–7850.
23. Hackeng, T.M.; Griffin, J.H.; Dawson, P.E. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.
24. Dawson, P.E.; Churchill, M.J.; Ghadiri, M.R.; Kent, S.B.H. Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives. *J. Am. Chem. Soc.* **1997**, *119*, 4325–4329.
25. Johnson, E.C.; Kent, S.B. Studies on the insolubility of a transmembrane peptide from signal peptide peptidase. *J. Am. Chem. Soc.* **2006**, *128*, 7140–7141.
26. Torbeev, V.Y.; Kent, S.B.H. Convergent Chemical Synthesis and Crystal Structure of a 203 Amino Acid “Covalent Dimer” HIV-1 Protease Enzyme Molecule. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 1667–1670.
27. Boerema, D.J.; Tereshko, V.A.; Kent, S.B.H. Total synthesis by modern chemical ligation methods and high resolution (1.1 Å) X-ray structure of ribonuclease A. *Pept. Sci.* **2008**, *90*, 278–286.
28. Rohde, H.; Seitz, O. Ligation-desulfurization: A powerful combination in the synthesis of peptides and glycopeptides. *Biopolymers* **2010**, *94*, 551–559.
29. Kawakami, T.; Akaji, K.; Aimoto, S. Peptide bond formation mediated by 4,5-dimethoxy-2-mercaptobenzylamine after periodate oxidation of the N-terminal serine residue. *Org. Lett.* **2001**, *3*, 1403–1405.
30. Low, D.W.; Hill, M.G.; Carrasco, M.R.; Kent, S.B.; Botti, P. Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6554–6559.
31. Offer, J.; Boddy, C.N.; Dawson, P.E. Extending synthetic access to proteins with a removable acyl transfer auxiliary. *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646.
32. Cardona, V.M.; Hartley, O.; Botti, P. Synthesis of cyclic peptides from unprotected precursors using removable N alpha-(1-(4-methoxyphenyl)-2-mercaptoethyl) auxiliary. *J. Pept. Res.* **2003**, *61*, 152–157.
33. Wu, B.; Chen, J.; Warren, J.D.; Chen, G.; Hua, Z.; Danishefsky, S.J. Building complex glycopeptides: Development of a cysteine-free native chemical ligation protocol. *Angew. Chem. Int. Ed.* **2006**, *45*, 4116–4125.
34. Payne, R.J.; Ficht, S.; Tang, S.; Brik, A.; Yang, Y.Y.; Case, D.A.; Wong, C.H. Extended sugar-assisted glycopeptide ligations: Development, scope, and applications. *J. Am. Chem. Soc.* **2007**, *129*, 13527–13536.
35. Tchertchian, S.; Hartley, O.; Botti, P. Synthesis of N alpha-(1-phenyl-2-mercaptoethyl) amino acids, new building blocks for ligation and cyclization at non-cysteine sites: Scope and limitations in peptide synthesis. *J. Org. Chem.* **2004**, *69*, 9208–9214.
36. Tam, J.P.; Xu, J.; Eom, K.D. Methods and strategies of peptide ligation. *Biopolymers* **2001**, *60*, 194–205.

37. Yan, L.Z.; Dawson, P.E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526–533.
38. Pentelute, B.L.; Kent, S.B. Selective desulfurization of cysteine in the presence of Cys(Acm) in polypeptides obtained by native chemical ligation. *Org. Lett.* **2007**, *9*, 687–690.
39. Wan, Q.; Danishefsky, S.J. Free-radical-based, specific desulfurization of cysteine: A powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252.
40. Crich, D.; Banerjee, A. Native chemical ligation at phenylalanine. *J. Am. Chem. Soc.* **2007**, *129*, 10064–10065.
41. Botti, P.; Tchertchian, S. Side-Chain Extended Ligation. WO2006133962 A1, December 21, 2006.
42. Haase, C.; Rohde, H.; Seitz, O. Native chemical ligation at valine. *Angew. Chem. Int. Ed.* **2008**, *47*, 6807–6810.
43. Kent, S. Origin of the chemical ligation concept for the total synthesis of enzymes (proteins). *Biopolymers* **2010**, *94*, iv–ix.
44. Bang, D.; Kent, S.B. A one-pot total synthesis of crambin. *Angew. Chem. Int. Ed.* **2004**, *43*, 2534–2538.
45. Lee, J.Y.; Bang, D. Challenges in the chemical synthesis of average sized proteins: Sequential vs. convergent ligation of multiple peptide fragments. *Biopolymers* **2010**, *94*, 441–447.
46. Bang, D.; Chopra, N.; Kent, S.B. Total chemical synthesis of crambin. *J. Am. Chem. Soc.* **2004**, *126*, 1377–1383.
47. Ueda, S.; Fujita, M.; Tamamura, H.; Fujii, N.; Otaka, A. Photolabile Protection for One-Pot Sequential Native Chemical Ligation. *Chem. Biochem.* **2005**, *6*, 1983–1986.
48. Bang, D.; Kent, S.B. His6 tag-assisted chemical protein synthesis. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5014–5019.
49. Brik, A.; Keinan, E.; Dawson, P.E. Protein synthesis by solid-phase chemical ligation using a safety catch linker. *J. Org. Chem.* **2000**, *65*, 3829–3835.
50. Johnson, E.C.; Durek, T.; Kent, S.B. Total chemical synthesis, folding, and assay of a small protein on a water-compatible solid support. *Angew. Chem. Int. Ed.* **2006**, *45*, 3283–3287.
51. Canne, L.E.; Botti, P.; Simon, R.J.; Chen, Y.; Dennis, E.A.; Kent, S.B. Chemical Protein Synthesis by Solid Phase Ligation of Unprotected Peptide Segments. *J. Am. Chem. Soc.* **1999**, *121*, 8720–8727.
52. Bang, D.; Pentelute, B.L.; Kent, S.B. Kinetically controlled ligation for the convergent chemical synthesis of proteins. *Angew. Chem. Int. Ed.* **2006**, *45*, 3985–3988.
53. Albericio, F. Developments in peptide and amide synthesis. *Curr. Opin. Chem. Biol.* **2004**, *8*, 211–221.
54. Canne, L.E.; Ferre-D'Amare, A.R.; Burley, S.K.; Kent, S.B. Total Chemical Synthesis of a Unique Transcription Factor-Related Protein: cMyc-Max. *J. Am. Chem. Soc.* **1995**, *117*, 2998–3007.
55. Baca, M.; Muir, T.W.; Schnolzer, M.; Kent, S.B. Chemical Ligation of Cysteine-Containing Peptides: Synthesis of a 22 kDa Tethered Dimer of HIV-1 Protease. *J. Am. Chem. Soc.* **1995**, *117*, 1881–1887.
56. Zhang, L.; Tam, J.P. Synthesis and Application of Unprotected Cyclic Peptides as Building Blocks for Peptide Dendrimers. *J. Am. Chem. Soc.* **1997**, *119*, 2363–2370.

57. Durek, T.; Torbeev, V.Y.; Kent, S.B. Convergent chemical synthesis and high-resolution x-ray structure of human lysozyme. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4846–4851.
58. Torbeev, V.Y.; Kent, S.B. Convergent chemical synthesis and crystal structure of a 203 amino acid “covalent dimer” HIV-1 protease enzyme molecule. *Angew. Chem. Int. Ed.* **2007**, *46*, 1667–1670.
59. Craik, D.J. Seamless proteins tie up their loose ends. *Science* **2006**, *311*, 1563–1564.
60. Trabi, M.; Craik, D.J. Circular proteins-no end in sight. *Trends Biochem. Sci.* **2002**, *27*, 132–138.
61. Craik, D.J.; Daly, N.L.; Bond, T.; Waine, C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.* **1999**, *294*, 1327–1336.
62. Svargard, E.; Burman, R.; Gunasekera, S.; Lovborg, H.; Gullbo, J.; Goransson, U. Mechanism of action of cytotoxic cyclotides: Cycloviolacin O2 disrupts lipid membranes. *J. Nat. Prod.* **2007**, *70*, 643–647.
63. Craik, D.J.; Daly, N.L.; Waine, C. The cystine knot motif in toxins and implications for drug design. *Toxicon* **2001**, *39*, 43–60.
64. Mylne, J.S.; Chan, L.Y.; Chanson, A.H.; Daly, N.L.; Schaefer, H.; Bailey, T.L.; Nguyencong, P.; Cascales, L.; Craik, D.J. Cyclic peptides arising by evolutionary parallelism via asparaginyl-endopeptidase-mediated biosynthesis. *Plant Cell* **2012**, *24*, 2765–2778.
65. Clark, R.J.; Craik, D.J. Native chemical ligation applied to the synthesis and bioengineering of circular peptides and proteins. *Biopolymers* **2010**, *94*, 414–422.
66. Daly, N.L.; Rosengren, K.J.; Craik, D.J. Discovery, structure and biological activities of cyclotides. *Adv. Drug Deliv. Rev.* **2009**, *61*, 918–930.
67. Craik, D.J.; Swedberg, J.E.; Mylne, J.S.; Cemazar, M. Cyclotides as a basis for drug design. *Expert. Opin. Drug Discov.* **2012**, *7*, 179–194.
68. Tam, J.P.; Lu, Y.A.; Yang, J.L.; Chiu, K.W. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8913–8918.
69. Gustafson, K.R.; McKee, T.C.; Bokesch, H.R. Anti-HIV cyclotides. *Curr. Protein Pept. Sci.* **2004**, *5*, 331–340.
70. Goransson, U.; Svargard, E.; Claeson, P.; Bohlin, L. Novel strategies for isolation and characterization of cyclotides: The discovery of bioactive macrocyclic plant polypeptides in the Violaceae. *Curr. Protein Pept. Sci.* **2004**, *5*, 317–329.
71. Barbeta, B.L.; Marshall, A.T.; Gillon, A.D.; Craik, D.J.; Anderson, M.A. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1221–1225.
72. Clark, R.J.; Daly, N.L.; Craik, D.J. Structural plasticity of the cyclic-cystine-knot framework: implications for biological activity and drug design. *Biochem. J.* **2006**, *394*, 85–93.
73. Gunasekera, S.; Foley, F.M.; Clark, R.J.; Sando, L.; Fabri, L.J.; Craik, D.J.; Daly, N.L. Engineering stabilized vascular endothelial growth factor-A antagonists: Synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides. *J. Med. Chem.* **2008**, *51*, 7697–7704.
74. Thongyoo, P.; Roque-Rosell, N.; Leatherbarrow, R.J.; Tate, E.W. Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org. Biomol. Chem.* **2008**, *6*, 1462–1470.

75. Terlau, H.; Olivera, B.M. Conus venoms: A rich source of novel ion channel-targeted peptides. *Physiol. Rev.* **2004**, *84*, 41–68.
76. Miljanich, G.P. Ziconotide: Neuronal calcium channel blocker for treating severe chronic pain. *Curr. Med. Chem.* **2004**, *11*, 3029–3040.
77. Olivera, B.M.; Cruz, L.J.; de Santos, V.; LeCheminant, G.W.; Griffin, D.; Zeikus, R.; McIntosh, J.M.; Galyean, R.; Varga, J.; Gray, W.R.; *et al.* Neuronal calcium channel antagonists. Discrimination between calcium channel subtypes using omega-conotoxin from *Conus magus* venom. *Biochemistry* **1987**, *26*, 2086–2090.
78. Lee, S.; Kim, Y.; Back, S.K.; Choi, H.W.; Lee, J.Y.; Jung, H.H.; Ryu, J.H.; Suh, H.W.; Na, H.S.; Kim, H.J.; *et al.* Analgesic effect of highly reversible omega-conotoxin FVIA on N type Ca<sup>2+</sup> channels. *Mol. Pain* **2010**, *6*, 97.
79. Lin, J.H. Pharmacokinetics of biotech drugs: Peptides, proteins and monoclonal antibodies. *Curr. Drug Metab.* **2009**, *10*, 661–691.
80. Haines, D.J.; Swan, C.H.; Green, J.R.; Woodley, J.F. Mucosal peptide hydrolase and brush-border marker enzyme activities in three regions of the small intestine of rats with experimental uraemia. *Clin. Sci.* **1990**, *79*, 663–668.
81. Woodley, J.F. Enzymatic barriers for GI peptide and protein delivery. *Crit. Rev. Ther. Drug Carr. Syst.* **1994**, *11*, 61–95.
82. Reichert, J.M. Monoclonal Antibodies as Innovative Therapeutics. *Curr. Pharm. Biotechnol.* **2008**, *9*, 423–430.
83. Clark, R.J.; Fischer, H.; Dempster, L.; Daly, N.L.; Rosengren, K.J.; Nevin, S.T.; Meunier, F.A.; Adams, D.J.; Craik, D.J. Engineering stable peptide toxins by means of backbone cyclization: Stabilization of the alpha-conotoxin MII. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13767–13772.
84. Clark, R.J.; Jensen, J.; Nevin, S.T.; Callaghan, B.P.; Adams, D.J.; Craik, D.J. The engineering of an orally active conotoxin for the treatment of neuropathic pain. *Angew. Chem. Int. Ed.* **2010**, *49*, 6545–6548.
85. Lovelace, E.S.; Armishaw, C.J.; Colgrave, M.L.; Wahlstrom, M.E.; Alewood, P.F.; Daly, N.L.; Craik, D.J. Cyclic MrIA: A stable and potent cyclic conotoxin with a novel topological fold that targets the norepinephrine transporter. *J. Med. Chem.* **2006**, *49*, 6561–6568.
86. Deechongkit, S.; Kelly, J.W. The effect of backbone cyclization on the thermodynamics of beta-sheet unfolding: Stability optimization of the PIN WW domain. *J. Am. Chem. Soc.* **2002**, *124*, 4980–4986.
87. Hubbard, S.J. The structural aspects of limited proteolysis of native proteins. *Biochim. Biophys. Acta* **1998**, *1382*, 191–206.
88. Quik, M.; Polonskaya, Y.; Kulak, J.M.; McIntosh, J.M. Vulnerability of 125I-alpha-conotoxin MII binding sites to nigrostriatal damage in monkey. *J. Neurosci.* **2001**, *21*, 5494–5500.
89. Hackeng, T.M.; Fernandez, J.A.; Dawson, P.E.; Kent, S.B.; Griffin, J.H. Chemical synthesis and spontaneous folding of a multidomain protein: anticoagulant microprotein S. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14074–14078.
90. Clark-Lewis, I.; Dewald, B.; Loetscher, M.; Moser, B.; Baggiolini, M. Structural requirements for interleukin-8 function identified by design of analogs and CXC chemokine hybrids. *J. Biol. Chem.* **1994**, *269*, 16075–16081.

91. Kimmerlin, T.; Seebach, D. “100 years of peptide synthesis”: Ligation methods for peptide and protein synthesis with applications to beta-peptide assemblies. *J. Pept. Res.* **2005**, *65*, 229–260.
92. Clayton, D.; Shapovalov, G.; Maurer, J.A.; Dougherty, D.A.; Lester, H.A.; Kochendoerfer, G.G. Total chemical synthesis and electrophysiological characterization of mechanosensitive channels from *Escherichia coli* and *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4764–4769.
93. Sukharev, S.I.; Blount, P.; Martinac, B.; Blattner, F.R.; Kung, C. A large-conductance mechanosensitive channel in *E. coli* encoded by *mscL* alone. *Nature* **1994**, *368*, 265–268.
94. Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry, C.E.; *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **1998**, *393*, 537–544.
95. Lin, F.H.; Graham, L.A.; Campbell, R.L.; Davies, P.L. Structural modeling of snow flea antifreeze protein. *Biophys. J.* **2007**, *92*, 1717–1723.
96. Pentelute, B.L.; Gates, Z.P.; Tereshko, V.; Dashnau, J.L.; Vanderkooi, J.M.; Kossiakoff, A.A.; Kent, S.B. X-ray structure of snow flea antifreeze protein determined by racemic crystallization of synthetic protein enantiomers. *J. Am. Chem. Soc.* **2008**, *130*, 9695–9701.
97. Pentelute, B.L.; Gates, Z.P.; Dashnau, J.L.; Vanderkooi, J.M.; Kent, S.B. Mirror image forms of snow flea antifreeze protein prepared by total chemical synthesis have identical antifreeze activities. *J. Am. Chem. Soc.* **2008**, *130*, 9702–9707.
98. Bingham, J.P.; Chun, J.B.; Ruzicka, M.R.; Li, Q.X.; Tan, Z.Y.; Kaulin, Y.A.; Englebretsen, D.R.; Moczydlowski, E.G. Synthesis of an iberiotoxin derivative by chemical ligation: a method for improved yields of cysteine-rich scorpion toxin peptides. *Peptides* **2009**, *30*, 1049–1057.
99. Valiyaveetil, F.I.; MacKinnon, R.; Muir, T.W. Semisynthesis and Folding of the Potassium Channel KcsA. *J. Am. Chem. Soc.* **2002**, *124*, 9113–9120.
100. Morrison, K.L.; Weiss, G.A. Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* **2001**, *5*, 302–307.
101. Weiss, G.A.; Watanabe, C.K.; Zhong, A.; Goddard, A.; Sidhu, S.S. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Nat. Acad. Sci. USA* **2000**, *97*, 8950–8954.
102. Cunningham, B.C.; Wells, J.A. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **1989**, *244*, 1081–1085.
103. Matthews, B.W. Structural and genetic analysis of the folding and function of T4 lysozyme. *FASEB J.* **1996**, *10*, 35–41.
104. Gregoret, L.M.; Sauer, R.T. Additivity of mutant effects assessed by binomial mutagenesis. *Proc. Nat. Acad. Sci. USA* **1993**, *90*, 4246–4250.
105. Payne, R.J.; Wong, C.H. Advances in chemical ligation strategies for the synthesis of glycopeptides and glycoproteins. *Chem. Commun.* **2010**, *46*, 21–43.
106. Masania, J.; Li, J.; Smerdon, S.J.; Macmillan, D. Access to phosphoproteins and glycoproteins through semi-synthesis, Native Chemical Ligation and N→S acyl transfer. *Org. Biomol. Chem.* **2010**, *8*, 5113–5119.
107. Davis, B.G. Synthesis of glycoproteins. *Chem. Rev.* **2002**, *102*, 579–602.

108. Pratt, M.R.; Bertozzi, C.R. Synthetic glycopeptides and glycoproteins as tools for biology. *Chem. Soc. Rev.* **2005**, *34*, 58–68.
109. Shin, Y.; Winans, K.A.; Backes, B.J.; Kent, S.B.H.; Ellman, J.A.; Bertozzi, C.R. Fmoc-Based Synthesis of Peptide- $\alpha$ Thioesters: Application to the Total Chemical Synthesis of a Glycoprotein by Native Chemical Ligation. *J. Am. Chem. Soc.* **1999**, *121*, 11684–11689.
110. Backes, B.J.; Ellman, J.A. An Alkanesulfonamide “Safety-Catch” Linker for Solid-Phase Synthesis. *J. Org. Chem.* **1999**, *64*, 2322–2330.
111. Backes, B.J.; Virgilio, A.A.; Ellman, J.A. Activation Method to Prepare a Highly Reactive Acylsulfonamide “Safety-Catch” Linker for Solid-Phase Synthesis. *J. Am. Chem. Soc.* **1996**, *118*, 3055–3056.
112. Marcaurelle, L.A.; Mizoue, L.S.; Wilken, J.; Oldham, L.; Kent, S.B.; Handel, T.M.; Bertozzi, C.R. Chemical synthesis of lymphotactin: a glycosylated chemokine with a C-terminal mucin-like domain. *Chemistry* **2001**, *7*, 1129–1132.
113. Hedrick, J.A.; Saylor, V.; Figueroa, D.; Mizoue, L.; Xu, Y.; Menon, S.; Abrams, J.; Handel, T.; Zlotnik, A. Lymphotactin is produced by NK cells and attracts both NK cells and T cells *in vivo*. *J. Immunol.* **1997**, *158*, 1533–1540.
114. Davis, B.G. Biochemistry. Mimicking posttranslational modifications of proteins. *Science* **2004**, *303*, 480–482.
115. Mezzato, S.; Schaffrath, M.; Unverzagt, C. An orthogonal double-linker resin facilitates the efficient solid-phase synthesis of complex-type N-glycopeptide thioesters suitable for native chemical ligation. *Angew. Chem. Int. Ed.* **2005**, *44*, 1650–1654.
116. Yamamoto, N.; Tanabe, Y.; Okamoto, R.; Dawson, P.E.; Kajihara, Y. Chemical Synthesis of a Glycoprotein Having an Intact Human Complex-Type Sialyloligosaccharide under the Boc and Fmoc Synthetic Strategies. *J. Am. Chem. Soc.* **2007**, *130*, 501–510.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).