Selenocysteine Derivatives for Chemoselective Ligations

Matt D. Gieselman, Yantao Zhu, Hao Zhou, Danica Galonic, and Wilfred A. van der Donk*[a]

KEYWORDS:
enzymes · glycopeptides · ligations · protein engineering · selenium

1. Introduction

Selenocysteine (Sec or U) is often referred to as the 21st natural amino acid because it is found in a number of naturally occurring enzymes in archaea, eubacteria, and eukarya.[1] Sec is inserted cotranslationally into these proteins and has its own $tRNA^{Sec}$ that recognizes the opal stop codon UGA.[2] Both cis (intramolecular) elements and trans (intermolecular) acting factors are instrumental in decoding UGA for Sec insertion rather than aborting translation.[3] In bacteria, a stem-loop structure in the mRNA immediately downstream of the UGA codon (the selenocysteine insertion sequence, SECIS) is recognized by a specialized selenocysteine elongation factor (SelB, Figure 1 A). The quaternary complex of SelB, Sec – $tRNA^{Sec}$, guanosine triphosphate (GTP), and the SECIS is then directed to the A site of the ribosome for Sec incorporation opposite the UGA codon.

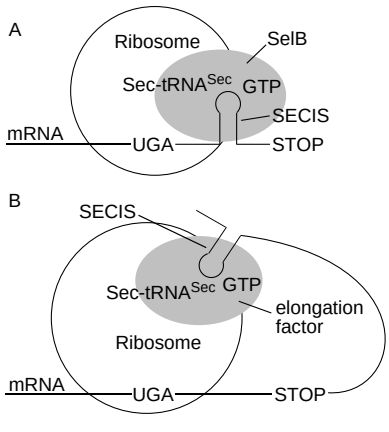


Figure 1. *Selenocysteine incorporation in eubacteria (A) and eukarya (B). SelB (grey oval) binds Sec-$tRNA^{Sec}$, GTP, and a stem loop structure (SECIS) in the mRNA in bacteria. The quaternary complex directs the $tRNA^{Sec}$ to the ribosome (white circle). In eubacteria, the SECIS is located just downstream of the UGA codon, while in eukarya it is in the 3′ untranslated region. The SelB analogue in eukaryotes has not yet been identified.*

In eukarya, the SECIS is in the 3′ untranslated region (Figure 1 B), whereas in bacteria it is located in the translated region of the mRNA.[4] Therefore, at present Sec cannot be incorporated site-specifically into proteins at nonnative positions by using heterologous bacterial expression systems unless the selenocysteine is to be located at one of the last ≈three amino acids at the C terminus.[5] This is unfortunate since the site-specific incorporation of Sec can be a very useful tool to understand enzyme catalysis, to modulate the properties of metal centers in numerous metalloproteins, and to introduce a unique site for subsequent chemoselective chemical modification.

Wu and Hilvert prepared the first artificial selenoprotein by selective activation of the active site serine of subtilisin with the protease inhibitor phenylmethanesulfonyl fluoride.[6] Subsequent treatment of the sulfonylated enzyme with high concentrations of hydrogen selenide (0.5 M) resulted in nucleophilic displacement producing selenosubtilisin. An alternative approach features in vivo charging of $tRNA^{Cys}$ with selenocysteine under cysteine-deprived growth conditions.[7, 8] Whereas this approach extends potential applications beyond proteins that contain a single reactive serine residue, disadvantages include substitution of Cys by Sec at multiple positions as well as a Sec incorporation efficiency of ≈70 – 80% giving rise to a heterogeneous protein mixture. In the past year, alternative means for high-level incorporation of Sec into peptides and proteins have been developed.[9–11] We describe here the use of selenocysteine derivatives for chemoselective ligations.

2.1. Properties of Selenocysteine

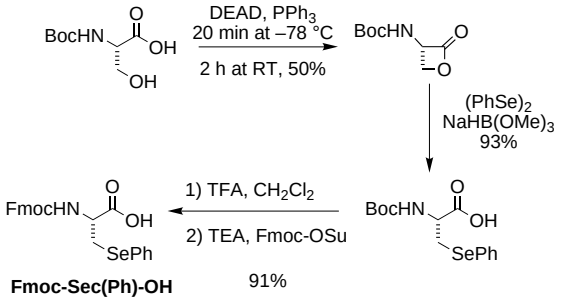
A number of unique properties exhibited by Sec and Sec derivatives have attracted interest in incorporating these amino acids into peptides and proteins. First, mutation of Cys to Sec represents the most conservative substitution that can be used to address a specific role assigned to the cysteine residue. The selenol in Sec is significantly more acidic than the thiol in Cys (pKa values of 5.2 – 5.7 and 8.5, respectively),[12, 13] and a selenolate is a much better nucleophile than a thiolate.[14, 15] These characteristics can be used to investigate proteins that utilize a cysteine nucleophile for catalysis, or they alternatively may allow site-specific chemical modification of a protein at pH 7. Furthermore, diselenides and selenosulfides have much

[a] *Prof. W. A. van der Donk, M. D. Gieselman, Y. Zhu, H. Zhou, D. Galonic*
University of Illinois at Urbana/Champaign
Department of Chemistry
600 S. Mathews Avenue
Urbana, Il 61801 (USA)
Fax: (+1) 217-244-8068
E-mail: vddonk@scs.uiuc.edu

lower reduction potentials than the corresponding disulfides;[7, 14, 16] a fact that can be exploited in protein folding or to investigate proteins involved in dithiol/disulfide interchange. A third interesting property of selenides involves their fast and mild oxidative elimination to give olefins in high yields.[17] In the case of selenocysteine derivatives this would generate dehydroalanines that are found in many natural products and can be employed as electrophilic 'handles' for further chemoselective elaboration.

2.2. Synthesis of Selenocysteine Derivatives

Walter and co-workers reported one of the first syntheses of optically pure selenocysteine derivatives.[18, 19] A key step in the preparation involved the nucleophilic displacement of an *O*-tosylated L-serine derivative. Comparison of optical rotations of the products with resolved, enantiomerically pure Sec derivatives revealed that the stereochemical integrity at the α-carbon atom was preserved. A later alternative approach was reported by Shirahama and co-workers who reduced diphenyldiselenide with sodium metal and reacted the resulting selenolate with *tert*-butoxycarbonyl (Boc) protected serine β-lactone.[20] More recently, this approach has been modified by performing an in situ reduction of diphenyldiselenide with sodium trimethoxyborohydride ($NaBH(OMe)_3$; Scheme 1).[21] In this latter study, the Boc group was removed, followed by 9-fluorenylmethoxycarbonyl (Fmoc) protection in a one-pot procedure to yield Fmoc-Sec(Ph)-OH, which was used for automated solid-phase peptide synthesis (SPPS).

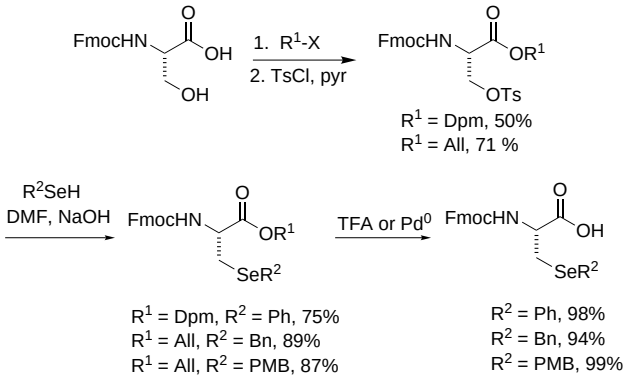


Scheme 1. *Synthesis of Fmoc-Sec(Ph)-OH via L-serine β-lactone. DEAD = diethylazodicarboxylate, TFA = trifluoroacetic acid, TEA = triethylamine, Fmoc-OSu = 9-fluorenylmethoxycarbonyl succinate.*

The routes discussed so far are not readily adaptable to prepare peptides containing unprotected selenocysteine. The currently most useful derivative to accomplish this goal contains a *p*-methoxybenzyl (PMB) protecting group on the selenol moiety.[22, 23] Preparation of Fmoc-Sec(PMB)-OH has been accomplished by treating a β-halogenated alanine derivative with a diselenide anion, generated by reduction of selenium powder with either lithium triethylborohydride or sodium borohydride.[9, 22, 24, 25] The resulting selenocystine derivative was subsequently reduced and reacted with PMBCl resulting in the protected selenol. Fujii and co-workers have used the protected amino acid in model studies and to assemble Cys-to-Sec mutants of α-atrial natriuretic peptides, 28-residue peptides containing an intramolecular disulfide bridge important for in vivo function.[22, 26] In these studies, the PMB group could be removed in an oxidative manner with either I_2 in acetic acid or dimethylsulfoxide in TFA to give the symmetric diselenide dimer. Alternatively, the presence of a strong Lewis acid such as trimethylsilyl trifluoromethanesulfonate (TMSOTf) in TFA yielded the free selenol. Moroder and co-workers have made extensive use of the PMB-protected selenocysteine in their studies of the oxidative folding of disulfide containing peptides in which one or both of the Cys residues were replaced with Sec.[27–30] An octapeptide with two Cys residues corresponding to a portion of glutaredoxin was used to determine the redox potential of diselenides in unconstrained peptides.[16] By using CD spectroscopy it was demonstrated that the reduction potential of the diselenide was -381 mV, whereas the mixed selenosulfide and disulfide were reduced at -326 mV and -180 mV, respectively. This difference of ≈ 200 mV between disulfide and diselenide is significant and should facilitate the study of catalytic dithiol/disulfide interchange processes as it may be feasible to "trap" intermediates that are otherwise not kinetically accessible.

Although the PMB protecting group has proven very important for the assembly of Sec-containing peptides, it also presents several drawbacks. Fmoc-Sec(PMB)-OH is prone to racemization during activation and coupling steps. In addition, once incorporated into peptides, PMB-protected Sec tends to deselenate to some extent during iterative piperidine deprotection steps, yielding piperidine adducts to the initially formed dehydroalanines.[23] Moreover, in our hands, iodine-mediated deprotection of Sec(PMB)-containing peptides requires optimization for each peptide, and the presence of other Cys residues in the sequence can complicate matters.[9] Although these impediments require special precautions, the difficulties are greatly diminished when selenocysteine is only present at the N terminus (see below) and, therefore, is not exposed to multiple rounds of piperidine treatment during Fmoc-based SPPS.[31]

The syntheses of Fmoc-Sec(PMB)-OH described above are either lengthy or not easily scalable, presenting limitations when gram quantities of protected Sec are desired for SPPS. We recently reported a synthesis of Fmoc-Sec derivatives that can be performed on large scale.[9] Fmoc-L-serine was protected either as a diphenylmethyl (Dpm) or allyl (All) ester (Scheme 2), and the

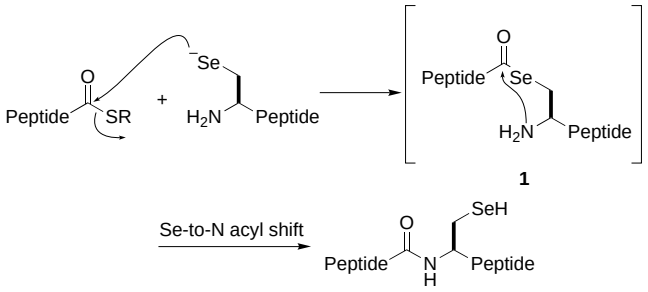


Scheme 2. *Improved scalable synthesis of Fmoc-Sec derivatives. pyr = pyridine, Dpm = diphenylmethyl, All = allyl, DMF = N,N-dimethylformamide, Bn = benzyl.*

hydroxy group was activated with *p*-toluenesulfonyl chloride (TsCl) in pyridine. Deprotonation of aryl or alkyl selenols with aqueous NaOH and reaction with the activated serines yielded the protected amino acids. Subsequent deprotection of the carboxylate with TFA or catalytic palladium provided optically pure Fmoc-Sec derivatives as demonstrated by both optical rotations and HPLC with a chiral stationary phase. The advantages of this route lie in its brevity and scalability. Notably, Fmoc-Sec(Ph)-OH and Fmoc-Sec(PMB)-OH were synthesized on greater than 10 g scale, and all purifications in the synthesis of Fmoc-Sec(Ph)-OH were accomplished by recrystallization of the crude reaction mixtures.[9]

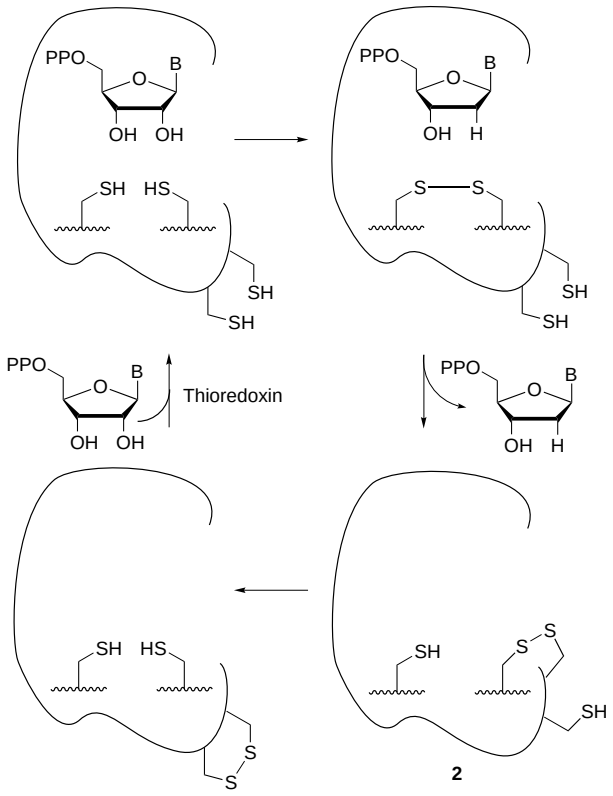
3. Native Chemical Ligation with Selenocysteine

Synthetic introduction of Sec into peptides is an attractive alternative to recombinant methods due to the complications associated with decoding the UGA stop codon described in the introduction. A number of synthetic peptides and proteins have been prepared by SPPS that incorporated selenocysteine by using Boc- or Fmoc-Sec(PMB)-OH.[22, 26, 28–30, 32, 33] By using the native chemical ligation technique developed by Kent and co-workers,[34] the ability to synthetically introduce a free selenol (or diselenide) into a peptide permits entry into larger peptides/proteins. In the selenocysteine version of native chemical ligation, two ligation partners, one a C-terminal peptide thioester and the other a peptide containing an unprotected Sec/selenocystine at its N terminus, are mixed together along with a reducing agent. After an initial trans selenoesterification, a selenoester **1** is formed (Scheme 3). This intermediate rearranges



Scheme 3. *Selenocysteine-mediated native chemical ligation.*

through an Se-to-N acyl shift to give the thermodynamically more stable native peptide bond. Several recent model studies have shown the feasibility of this process. In our laboratory, a number of Sec containing analogues of the C terminus of ribonucleotide reductase (RNR) were synthesized.[9] RNR contains four redox active cysteines. Two cysteines in the active site provide the two electrons required to reduce ribonucleotides to deoxyribonucleotides.[35] The disulfide generated in the active site is reduced back to the active form of the protein by dithiol/disulfide interchange with two Cys residues near the C terminus (Scheme 4). These C terminal cysteines shuttle reducing equivalents from thioredoxin into the active site of RNR. The proposed

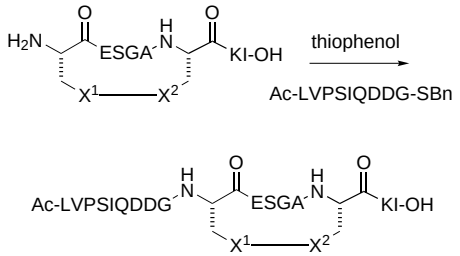


Scheme 4. *Reduction of nucleotides to deoxynucleotides by ribonucleotide reductase (RNR). The two C-terminal cysteines shuttle reducing equivalents from thioredoxin into the active site. C-Terminal sequence of RNR from* E. coli = *LVPSIQDDG**C**ESGA**C**KI. B = base, PP = diphosphate.*

disulfide intermediate **2** between one of the active site thiols and one of the C terminal thiols has never been observed, presumably because it is kinetically invisible.[36] Given the significant difference in redox potentials between disulfides and selenosulfides discussed above, replacement of one of the C-terminal cysteines with selenocysteine may allow the detection of the selenosulfide analogue of **2**. Consequently, we focused our efforts on developing selenocysteine-mediated native chemical ligations with applications in RNR in mind.

Three permutations of the nucleophilic coupling partner in which one or both of the Cys residues were replaced by Sec were investigated (Table 1). The purified Sec-containing peptides were prepared in high yields by Fmoc-based SPPS with the Sec(PMB) monomer, and the thioester coupling partner was prepared by SPPS by using 2-chlorotriphenylmethyl resin.[37] At a concentration of 3 mM, the reaction of the Sec-containing peptides was slower than most literature reports for ligations with cysteine even though the latter are typically performed at lower concentrations.[38] These findings are consistent with the previous observation that diselenides and selenosulfides are thermodynamically more stable than disulfides. Hence, with only the weakly reducing thiophenol present to reduce the diselenide/selenosulfides, only a small equilibrium concentration of a reactive species is produced. Raines and co-workers showed that once a free selenocysteine is generated, the native chemical ligation itself actually proceeds much faster than with cysteine

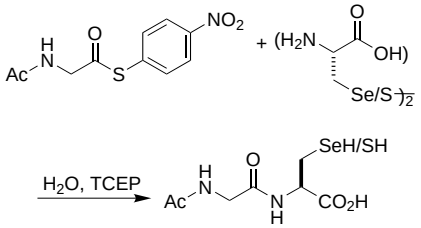
Table 1. *Selenocysteine-mediated native chemical ligation.*



X^1	X^2	Conc. [mM]	time [h]	yield [%][a]
Se	Se	2.6	27.5	48
S	Se	2.9	16	76
Se	S	3.2	20	60

[a] Yields after HPLC purification.

(Scheme 5).[10] Glycine p-nitrophenylthioester was treated with either cystine or selenocystine in the presence of TCEP, a stronger and irreversible reductant. The observed rate of *p*-nitrophenylthiolate formation was higher with selenocysteine than with cysteine over a pH range of 5–8. At pH 5.0, for example, the selenocysteine reaction was 10^3 times faster than the ligation with cysteine, a fact consistent with the lower pK_a values of selenols.



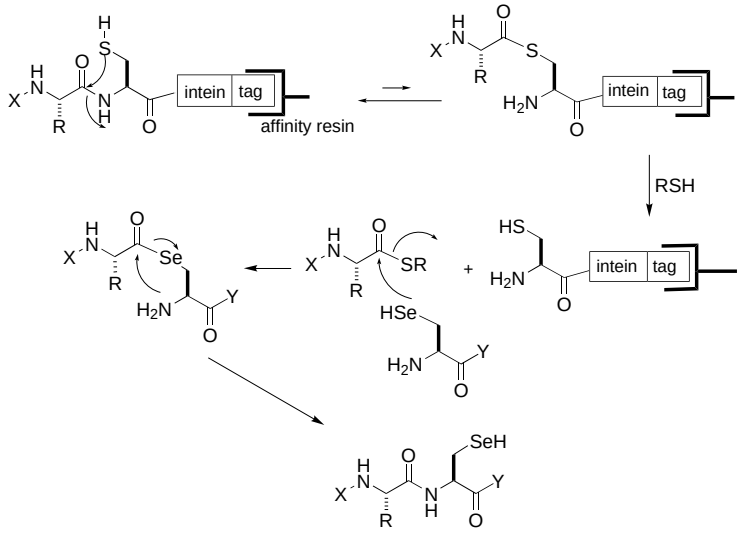
Scheme 5. *Reactions used to compare the rates of native chemical ligation with cystine/selenocystine in the presence of triscarboxyethylphosphine (TCEP).*

Hilvert and co-workers have extended Sec-mediated native chemical ligation to the synthesis of bovine pancreatic trypsin inhibitor (BPTI),[11] containing six Cys residues and three disulfide bonds. Synthetic $BPTI^{1-37}$ and C38U-$BPTI^{38-58}$ fragments were ligated in the presence of TCEP to yield C38U-BPTI, which was subsequently folded to its native structure. The mutant exhibited properties that were very similar to the wild-type protein including its stoichiometric inhibition of trypsin.

4. Incorporating Selenocysteine into Proteins through Expressed Protein Ligation

Expressed protein ligation (EPL), a powerful extension of native chemical ligation, was first reported in 1998[39] and has become increasingly popular in recent years for protein engineering.[40–42] With the success of Sec-mediated native chemical ligation, it became clear that Sec could also be used at the point of ligation in expressed protein ligation. The process of selenocysteine-mediated EPL is briefly outlined in Scheme 6. A target protein truncated at its C terminus is overexpressed in *Escherichia coli* as a fusion to an intein domain and affinity tag (either a hexa-His tag or chitin binding domain, CBD). The cell lysate is passed through the affinity column to bind the target protein and separate it from other cellular constituents. Inteins self-catalyze a rearrangement reaction from an amide to a thioester at a cysteine residue within their sequence. When an external thiol is added to the resin, a trans-thioesterification takes place that cleaves the truncated target protein from the resin while leaving the intein attached. Addition of an unprotected synthetic peptide with Sec at its N terminus initiates another trans-thioesterification, followed by an Se-to-N acyl shift to yield a fully deprotected mutant protein with a Sec incorporated at a specific site within the polypeptide. Raines and co-workers have recently demonstrated the feasibility of Sec-mediated EPL with the semisynthesis of a C110U mutant of ribonuclease A (RNase A),[10, 43] which in its wild-type form contains eight cysteines involved in four disulfide bonds. RNase A^{1-109} was overexpressed in *E. coli* as an intein–CBD fusion protein, and C110U-RNase $A^{110-124}$ was chemically synthesized. After ligation, C110U RNase A was shown to have essentially identical activity to the wild-type protein, thereby validating the methodology.

In collaboration with Lu and Berry, we prepared an engineered variant of azurin by selenocysteine-mediated EPL.[44] Azurin is an electron transport protein in bacteria and a member of the type 1 family of blue copper proteins. The protein has a number of characteristic properties including its intense blue color ($\varepsilon_{625nm} = 5000\,M^{-1}\,cm^{-1}$), a small parallel EPR hyperfine splitting (53 G), fast electron transfer rates, and a redox potential of $\approx$318 mV against the normal hydrogen electrode.[45, 46] The crystal structure of azurin (Figure 2) revealed a short Cu–S bond



Scheme 6. *Selenocysteine-mediated expressed protein ligation. X = recombinant target protein truncated at its C terminus. Y = synthetic peptide corresponding to the truncated C terminus.*

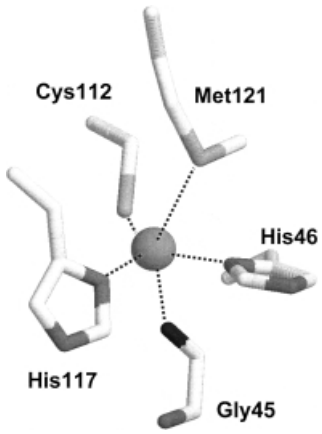


Figure 2. *The active site of azurin from Pseudomonas aeruginosa (PDB file: 4AZU).*[47]

to Cys112.[47, 48] Experimental and theoretical studies have suggested that the high covalency of this interaction is one of the key features responsible for azurin's properties.[49–51] Site-directed mutagenesis studies designed to further investigate this covalency have been unsuccessful as they led to loss of the type 1 copper site.[52–55] By using expressed protein ligation with selenocysteine, we have been able to replace the key cysteine residue with Sec. Synthetic C112U azurin$^{112-128}$ was ligated with expressed azurin$^{1-111}$ – intein – CBD providing C112U azurin, as determined by ESI-MS. The mutant bound copper stoichiometrically to yield a blue protein, but the spectroscopic features of C112U azurin were significantly different from wild-type.[44] The UV/Vis spectrum of C112U azurin displayed a red-shifted charge-transfer band with a somewhat lower intensity ($\varepsilon = 4000\,\text{M}^{-1}\,\text{cm}^{-1}$). Also, EPR spectroscopy of the mutant revealed a much larger parallel hyperfine splitting (101 G). These data are consistent with a less covalent Cu–Se interaction while maintaining some of the characteristics of blue copper proteins. Like selenosubtilisin,[6, 56] this is another example of a protein with significantly altered properties upon substitution of cysteine by selenocysteine.

5. Synthesis of Dehydropeptides and Their Use for Chemoselective Ligations

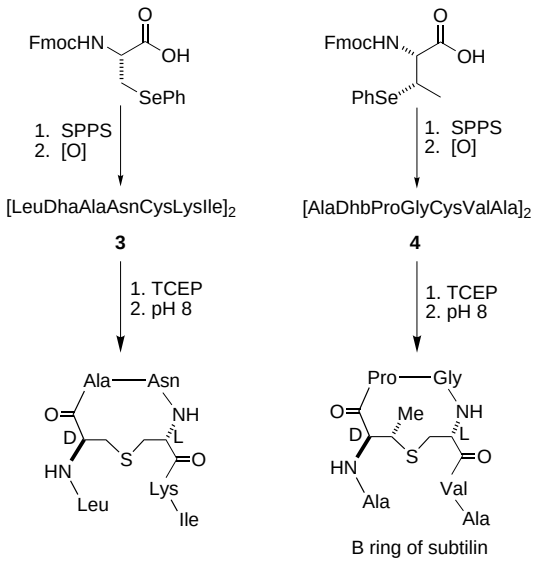
Dehydroalanine (Dha) is found in a number of natural products and can impart interesting biological activities on polypeptides.[57, 58] Furthermore, dehydroamino acids can serve as very useful precursors to peptide conjugates. Dehydroalanine-containing peptides have been prepared by a number of methods. The most common approach is the activation and elimination of serine residues or Hofmann elimination of 2,3-diaminopropionic acid.[59–61] These methods, however, preclude the presence of other unprotected serine or threonine residues in the former case or lysine residues in the latter case. A more versatile approach reported recently is the pyrolytic or basic elimination of cysteine sulfoxides.[62] The drawback of this methodology lies in its reported incompatibility with protected Cys residues.[63]

An alternative facile, site-specific and chemoselective method for introducing dehydroalanine into globally deprotected peptides employs *Se*-phenylselenocysteine. This amino acid can be conveniently incorporated into peptides by using standard SPPS (see Section 2.2.), and after cleavage from the resin the Sec derivative can be converted into Dha through a mild, chemoselective oxidation with either hydrogen peroxide or sodium periodate.[21] Table 2 demonstrates the compatibility with unprotected amino acids and with Cys residues that are protected either with *tert*-butyl-S (*t*BuS) groups to form disulfides or with triphenylmethyl (Trt) groups (entries 5 – 7). Methionine and tryptophan residues are also compatible with the method provided that the conditions (temperature, stoichiometry) are chosen such that their oxidation is minimized (entries 8 and 9).

Table 2. *Synthesis of dehydropeptides.*

Entry	Peptide	Dehydropeptide	Oxidant	Yield [%]
1	Fmoc-GLPU(Ph)VIA	Fmoc-GLPDhaVIA	$NaIO_4$	72
2	Fmoc-ISVU(Ph)RSTS	Fmoc-ISVDhaRSTS	$NaIO_4$	67
3	Ac-GLPU(Ph)VIA	Ac-GLPDhaVIA	H_2O_2	82
4	Ac-ISVU(Ph)RSTS	Ac-ISVDhaRSTS	$NaIO_4$	82
5	Ac-GGC(*t*BuS)PU(Ph)VIA	Ac-GGC(*t*BuS)PDhaVIA	$NaIO_4$	84
6	LU(Ph)PGC(Trt)VG	LDhaPGC(Trt)VG	$NaIO_4$	80
7	RIAU(Ph)IALC(*t*BuS)K	RIADhaIALC(*t*BuS)K	$NaIO_4$	72
8	AMU(Ph)A	AMDhaA	$NaIO_4$	62
9	Fmoc-WU(Ph)-ODpm	Fmoc-WDha-ODpm	$NaIO_4$	64

The mild oxidative elimination introduces an electrophilic handle into unprotected peptides (for example, **3** in Scheme 7), which has been exploited to synthesize a number of cyclic thioether peptides called lanthionines.[21] These lanthionines are the signature structural motif found in lantibiotics, a class of

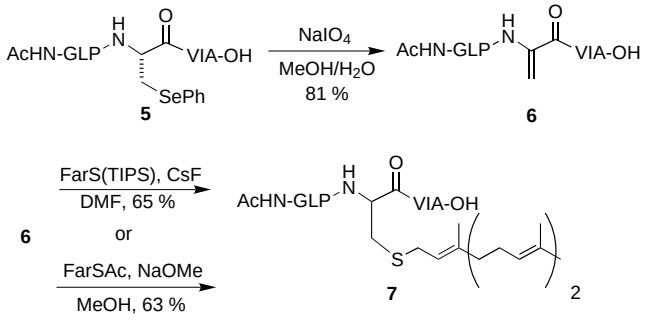


Scheme 7. *Biomimetic formation of* meso*-lanthionines and methyllanthionines by stereoselective intramolecular cyclization.*

posttranslationally modified peptide antibiotics.[64] The intramolecular biomimetic Michael additions provided single diastereomers by stereoselective protonation of the enolate intermediates. Bradley and co-workers reported similar stereoselective cyclizations and assigned the stereochemistry of the products as the natural *meso*-lanthionines by using NMR spectroscopy methods.[62] More recently, we have investigated biomimetic cyclizations with dehydrobutyrine-containing peptides obtained by oxidative elimination of (2*R*,3*S*)-3-methyl-*Se*-phenylselenocysteine residues (for example, **4** in Scheme 7). Of the four possible stereoisomers that could be formed in the Michael addition to form the B ring of the lantibiotic subtilin, only the natural (2*S*,3*S*,6*R*)-methyllanthionine was obtained; this demonstrates that, in the context of a constrained peptide, the local chiral environment can bias the stereoselectivity of these Michael additions.[65]

In addition to their use for intramolecular Michael additions, dehydroalanines can also be versatile electrophiles for the introduction of external nucleophiles.[66] Peptide conjugates such as glycopeptides, prenylated peptides, and lipopeptides are essential to a number of cellular processes, and unnatural peptide conjugates are powerful tools for biophysical studies and drug delivery. Therefore, rapid and high yielding synthetic strategies to assemble peptide conjugates have been the focus of numerous research efforts. One often adopted methodology calls for the synthesis of conjugated amino acid building blocks followed by their introduction into peptides.[67, 68] This building block approach, however, is linear and as such labor intensive if a library is desired that only differs in the pendant conjugates. In addition, the presence of the appendage may not be compatible with existing optimized SPPS protection and coupling schemes. A conceptually attractive alternative involves chemoselective ligations between tailor-made unprotected peptides and conjugates in aqueous solution. This concept of coupling two mutually and uniquely reactive functional groups was initially developed to extend the size limit of chemically synthesized peptides and proteins;[69, 70] it has more recently been applied to the synthesis of conjugates.[71–73] Such convergent approaches mandate an orthogonal reactive moiety on the peptide and a number of different electrophilic handles have been introduced for this purpose. Most often used are aldehydes installed on side chains or at the N terminus of the peptide,[74, 75] by periodate treatment of an appropriate precursor.[76–79] Alternatively, specialized amino acid derivatives have been incorporated site-specifically into the peptide during SPPS[80] that are later unmasked in some cases.[81] Condensation of the carbonyl moiety with functionalities on the conjugate such as aminooxy groups,[75, 82–84] hydrazides,[77, 84, 85] or 1,2-aminothiols[69] has been used to prepare a plethora of macromolecules.

Almost all these ligation strategies produce nonnative linkages between the peptide and the conjugate, and in some cases (for example, oximes) a mixture of isomers may be formed. We reasoned that the product of Michael reactions between dehydroalanine-containing peptides and thioglycosides or prenylthiols would provide conjugates through linkages that are either identical to (prenyl) or very close mimics of (*S*-linked glycopeptides) the native structures. Furthermore, an expanded arsenal of available orthogonal ligation strategies is potentially useful for tandem ligations.[75, 86] To illustrate the use of dehydroalanines, a heptapeptide (**5**) corresponding to the C terminus of N-Ras was synthesized with Sec(Ph) installed at the site of a Cys residue (Scheme 8). Oxidative elimination yielded **6**,



Scheme 8. *Convergent chemoselective ligation strategy for the synthesis of the C terminus of human N-Ras. Far = farnesyl, TIPS = triisopropylsilyl.*

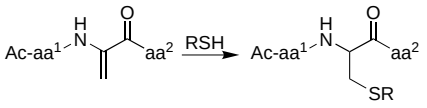
and Michael additions with unmasked farnesyl thiolates provided **7** in good overall yield. The convergent nature of this approach is illustrated in Table 3 since the same dehydroalanine precursor was used to prepare a variety of analogues. This method can also be applied to the synthesis of glycopeptides. These structures are particularly suited for convergent ligation strategies since they are not readily prepared by either traditional synthetic or recombinant techniques. In addition, the well-known heterogeneity caused by the existence of various isoforms of glycoproteins hampers structure – reactivity studies. Compared with the various nonnative linkages that have been used to prepare *O*-linked glycopeptide mimics through chemoselective ligations,[71, 84, 87] the isosteric *S*-linked conjugates are perhaps the closest analogues. Another enticing feature of these structures is their reported higher chemical stability compared with their *O*-linked counterparts.[88] The potential of Michael additions to dehydroalanines for the assembly of *S*-linked glycopeptides was evaluated with a series of tripeptides and 1-thiosugars (Table 4).[66] Addition of either protected or unprotected β-1-thiosugars resulted in the formation of exclusively β-linked glycopeptides in good yields (entries 1 – 5). Use of 1-thio-*N*-acetyl-α-D-galactose afforded the desired α-linked glycopeptide, which is an important analogue of the core structure of *O*-linked glycoproteins (entry 6). Importantly, no special protecting-group strategy beyond that used for Fmoc-based SPPS is required for either the generation of the dehydroalanine handle

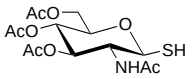
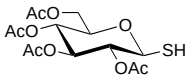
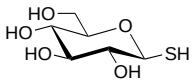
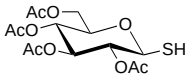
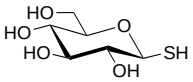
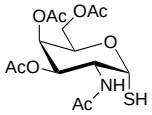
Table 3. *Convergent synthesis of peptide conjugates.*[a]

Electrophile	Nucleophile	Yield [%]
6	GerSAc	72
6	GerGerS(TIPS)	76
6	Dimethylallylthioacetate	69
6	AcSK	67

[a] Ger = geranyl, TIPS = triisopropylsilyl.

Table 4. *Synthesis of glycopeptides by Michael ligation.*

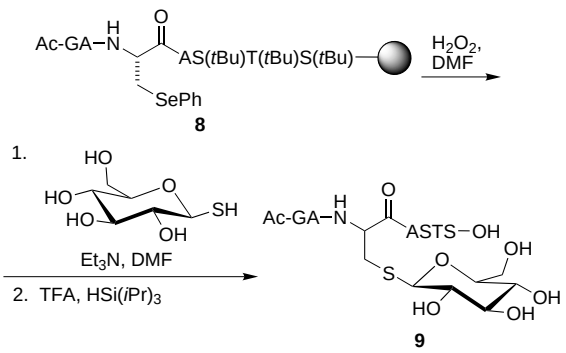


Entry	aa[1]	aa[2]	RSH	Yield [%]
1	Ala	Ala-O*t*Bu		62
2	Ala	Ala-O*t*Bu		62
3	Ala	Ala-O*t*Bu		74
4	Pro	Val-O*t*Bu		88
5	Pro	Val-O*t*Bu		91
6	Gly	Gly-NHMe		79[a]

[a] D. Galonic, D. Y. Gin, W. A. van der Donk, unpublished results.

or the Michael addition. Moreover, all steps used to generate and derivatize dehydroalanines are compatible with solid-phase chemistry. For instance, resin-bound peptide **8** was prepared by Fmoc-based SPPS, followed by oxidative elimination, and Michael addition with unprotected thioglucose (Scheme 9). Subsequent cleavage from the resin, and HPLC purification provided glycoconjugate **9** in 45% yield based on the loading of the Wang resin.[66]

The major advantage of this methodology involves the defined stereochemical integrity at the anomeric center and the accessibility of both α and β anomers. However, an obvious current drawback is the lack of diastereoselectivity at the peptidic α-carbon atom that provides two diastereomers, even though these are readily separated by HPLC. The lack of substrate- or reagent-controlled stereoselectivity is undoubtedly due to the inherently fast rate of protonation of the enolate intermediate in protic solvents. Consequently, to overcome this impediment, the barrier for the stereodetermining step must be raised. One potential avenue to accomplish this would be to exploit radical conjugate additions. Owing to the high bond dissociation energy of O–H bonds, protic solvents are poor hydrogen-atom donors whereas the thio sugars are much better reductants. The energies of the diastereomeric transition states for hydrogen-atom transfer generating either the L- or D-configuration at the α-carbon atom may therefore be sufficiently different that selectivity can be achieved. In fact, Kessler and co-workers have shown the utility of dehydroalanines for addition of anomeric glycosyl radicals providing *C*-linked glycopeptides.[89] Furthermore, the feasibility of stereoselective radical additions to dehydroalanines has been demonstrated in a nonpeptidic context.[90–92] Efforts are currently underway in our laboratory to combine these approaches for the development of a chemo- and stereoselective radical ligation. An alternative approach that would assure formation of only one isomer at the α position features the ligation of 1-thiosugars to cyclic sulfamidates derived from L-serine. The feasibility of this strategy has been shown at the amino acid level[93] and more recently with small peptides.[94]



Scheme 9. *Solid-phase synthesis of glycopeptides by using chemoselective Michael addition to dehydroalanine.*

6. Conclusion

Synthetic methodology has emerged in recent years to introduce selenocysteine into peptides. Combined with selenocysteine-mediated native and expressed protein ligation, these accomplishments set the stage for the application of selenocysteine as a tool to investigate protein structure and function. Given the characteristic physicochemical properties of Sec, including its low pK_a, low redox potential and high nucleophilicity, site-specific introduction of Sec can either address questions regarding enzymatic reaction mechanisms or provide a chemoselective handle to introduce biophysical probes. The $S = 1/2$ ^{77}Se nucleus may also find applications in NMR and EPR spectroscopy and the presence of Se can aid in X-ray studies. In addition, the utility of Sec-containing peptides as precursors to dehydropeptides has been demonstrated by the chemical synthesis of a number of peptide conjugates. The sequence of SPPS, mild oxidative elimination, and Michael ligation allows rapid entry into a variety of conjugates including prenylated, glycosylated, and lipidated peptides.

Our work on selenocysteine derivatives has been supported by the National Institutes of Health (Grant: GM58822), the Burroughs Welcome Fund (Grant: APP 1920), and the Beckman Foundation (a BYI award). W.A.V. is a Cottrell Scholar of the Research Corporation and an Alfred P. Sloan Fellow.

[1] T. C. Stadtman, *Annu. Rev. Biochem.* **1996**, *65*, 83.
[2] W. Leinfelder, E. Zehelein, M. A. Mandrand-Berthelot, A. Böck, *Nature* **1988**, *331*, 723.
[3] S. Commans, A. Böck, *FEMS Microbiol. Rev.* **1999**, *23*, 335.
[4] S. C. Low, M. J. Berry, *Trends Biochem. Sci.* **1996**, 203.

[5] E. S. J. Arnér, H. Sarioglu, F. Lottspeich, A. Holmgren, A. Böck, *J. Mol. Biol.* **1999**, *292*, 1003.
[6] Z. P. Wu, D. Hilvert, *J. Am. Chem. Soc.* **1989**, *111*, 4513.
[7] S. Müller, H. Senn, B. Gsell, W. Vetter, C. Baron, A. Böck, *Biochemistry* **1994**, *33*, 3404.
[8] S. Boschi-Muller, S. Muller, A. Van Dorsselaer, A. Böck, G. Branlant, *FEBS Lett.* **1998**, *439*, 241.
[9] M. D. Gieselman, L. Xie, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1331.
[10] R. J. Hondal, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2001**, *123*, 5140.
[11] R. Quaderer, A. Sewing, D. Hilvert, *Helv. Chim. Acta* **2001**, *84*, 1197.
[12] B. Nygard, *Ark. Kemi* **1967**, *27*, 341.
[13] A. P. Arnold, K. S. Tan, D. L. Rabenstein, *Inorg. Chem.* **1986**, *25*, 2433.
[14] R. E. Huber, R. S. Criddle, *Arch. Biochem. Biophys.* **1967**, *122*, 164.
[15] R. G. Pearson, H. R. Sobel, J. Songstad, *J. Am. Chem. Soc.* **1968**, *90*, 319.
[16] D. Besse, F. Siedler, T. Diercks, H. Kessler, L. Moroder, *Angew. Chem.* **1997**, *109*, 915; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 883.
[17] T. Wirth, *Angew. Chem.* **2000**, *112*, 3890; Angew. *Chem. Int. Ed.* **2000**, *39*, 3741.
[18] D. Theodoropoulos, I. L. Schwartz, R. Walter, *Biochemistry* **1967**, *6*, 3927.
[19] J. Roy, W. Gordon, I. L. Schwartz, R. Walter, *J. Org. Chem.* **1970**, *35*, 510.
[20] M. Sakai, K. Hashimoto, H. Shirahama, *Heterocycles* **1997**, *44*, 319.
[21] N. M. Okeley, Y. Zhu, W. A. van der Donk, *Org. Lett.* **2000**, *2*, 3603.
[22] T. Koide, H. Itoh, A. Otaka, H. Yasui, M. Kuroda, N. Esaki, K. Soda, N. Fujii, *Chem. Pharm. Bull.* **1993**, *41*, 502.
[23] D. Besse, L. Moroder, *J. Pept. Sci.* **1997**, *3*, 442.
[24] H. Tanaka, K. Soda, *Methods Enzymol.* **1987**, *143*, 240.
[25] E. M. Stocking, J. N. Schwarz, H. Senn, M. Salzmann, L. A. Silks, *J. Chem. Soc. Perkin Trans. 1* **1997**, 2443.
[26] T. Koide, H. Itoh, A. Otaka, M. Furuya, Y. Kitajima, N. Fujii, *Chem. Pharm. Bull.* **1993**, *41*, 1596.
[27] L. Moroder, D. Besse, H.-J. Musiol, S. Rudolph-Bohner, F. Siedler, *Biopolymers* **1996**, *40*, 207.
[28] S. Fiori, S. Pegoraro, S. Rudolph-Bohner, J. Cramer, L. Moroder, *Biopolymers* **2000**, *53*, 550.
[29] S. Pegoraro, S. Fiori, S. Rudolph-Bohner, T. X. Watanabe, L. Moroder, *J. Mol. Biol.* **1998**, *284*, 779.
[30] S. Pegoraro, S. Fiori, J. Cramer, S. Rudolph-Bohner, L. Moroder, *Protein Sci.* **1999**, *8*, 1605.
[31] E. Atherton, R. C. Sheppard, *Solid Phase Peptide Synthesis, A Practical Approach*, Pierce Chemical Company, Rockford, IL, **1989**.
[32] T. Oikawa, N. Esaki, H. Tanaka, K. Soda, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 3057.
[33] K. Rajarathnam, B. D. Sykes, B. Dewald, M. Baggiolini, I. Clark-Lewis, *Biochemistry* **1999**, *38*, 7653.
[34] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776.
[35] J. Stubbe, W. A. van der Donk, *Chem. Rev.* **1998**, *98*, 705.
[36] S. Licht, J. Stubbe in *Comprehensive Natural Products Chemistry, Vol. 5* (Eds.: D. H. R. Barton, K. Nakanishi, O. Meth-Cohn), Pergamon Press, New York, **1999**, p. 163.
[37] S. Futaki, K. Sogawa, J. Maruyama, T. Asahara, M. Niwa, *Tetrahedron Lett.* **1997**, *38*, 6237.
[38] P. E. Dawson, M. J. Churchill, M. R. Ghadiri, S. B. H. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325.
[39] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705.
[40] T. C. Evans, M. Q. Xu, *Biopolymers* **1999**, *51*, 333.
[41] B. Ayers, U. K. Blaschke, J. A. Camarero, G. J. Cotton, M. Holford, T. W. Muir, *Biopolymers* **1999**, *51*, 343.
[42] C. J. Noren, J. M. Wang, F. B. Perler, *Angew. Chem.* **2000**, *112*, 458; *Angew. Chem. Int. Ed.* **2000**, *39*, 451.
[43] R. J. Hondal, R. T. Raines, *Methods Enzymol.* **2002**, *347*, 70.
[44] S. Berry, M. Gieselman, M. J. Nilges, W. A. van der Donk, Y. Lu, *J. Am. Chem. Soc.* **2002**, *124*, 2084.
[45] H. B. Gray, B. G. Malmstrom, R. J. P. Williams, *J. Biol. Inorg. Chem.* **2000**, *5*, 551.
[46] E. I. Solomon, M. J. Baldwin, M. D. Lowery, *Chem. Rev.* **1992**, *92*, 521.
[47] H. Nar, A. Messerschmidt, R. Huber, M. van de Kamp, G. W. Canters, *J. Mol. Biol.* **1991**, *221*, 765.
[48] E. T. Adman, *Adv. Protein Chem.* **1991**, *42*, 145.
[49] J. A. Guckert, M. D. Lowery, E. I. Solomon, *J. Am. Chem. Soc.* **1995**, *117*, 2817.
[50] E. I. Solomon, K. W. Penfield, A. A. Gewirth, M. D. Lowery, S. E. Shadle, J. A. Guckert, L. B. LaCroix, *Inorg. Chim. Acta* **1996**, *243*, 67.
[51] U. Ryde, M. H. M. Olsson, B. O. Roos, J. O. A. De Kerpel, K. Pierloot, *J. Biol. Inorg. Chem.* **2000**, *5*, 565.
[52] T. J. Mizoguchi, A. J. Di Bilio, H. B. Gray, J. H. Richards, *J. Am. Chem. Soc.* **1992**, *114*, 10076.
[53] M. Piccioli, C. Luchinat, T. J. Mizoguchi, B. E. Ramirez, H. B. Gray, J. H. Richards, *Inorg. Chem.* **1995**, *34*, 737.
[54] S. Faham, T. J. Mizoguchi, E. T. Adman, H. B. Gray, J. H. Richards, D. C. Rees, *J. Biol. Inorg. Chem.* **1997**, *2*, 464.
[55] S. DeBeer, C. N. Kiser, G. A. Mines, J. H. Richards, H. B. Gray, E. I. Solomon, B. Hedman, K. O. Hodgson, *Inorg. Chem.* **1999**, *38*, 433.
[56] Z. P. Wu, D. Hilvert, *J. Am. Chem. Soc.* **1990**, *112*, 5647.
[57] H.-G. Sahl, R. W. Jack, G. Bierbaum, *Eur. J. Biochem.* **1995**, *230*, 827.
[58] R. C. Lau, K. L. Rinehart, *J. Antibiot.* **1994**, *47*, 1466.
[59] U. Schmidt, A. Lieberknecht, J. Wild, *Synthesis* **1988**, 159.
[60] D. Ranganathan, K. Shah, N. Vaish, *J. Chem. Soc. Chem. Commun.* **1992**, 1145.
[61] C. Blettner, M. Bradley, *Tetrahedron Lett.* **1994**, *35*, 467.
[62] S. Burrage, T. Raynham, G. Williams, J. W. Essex, C. Allen, M. Cardno, V. Swali, M. Bradley, *Chem. Eur. J.* **2000**, *6*, 1455.
[63] S. A. Burrage, T. Raynham, M. Bradley, *Tetrahedron Lett.* **1998**, *39*, 2831.
[64] H. G. Sahl, G. Bierbaum, *Annu. Rev. Microbiol.* **1998**, *52*, 41.
[65] H. Zhou, W. A. van der Donk, *Org. Lett.* **2002**, *4*, 1335.
[66] Y. Zhu, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1189.
[67] H. Herzner, T. Reipen, M. Schultz, H. Kunz, *Chem. Rev.* **2000**, *100*, 4495.
[68] O. Seitz, *ChemBioChem* **2000**, *1*, 215.
[69] J. P. Tam, Q. T. Yu, Z. W. Miao, *Biopolymers* **1999**, *51*, 311.
[70] D. M. Coltart, *Tetrahedron* **2000**, *56*, 3449.
[71] S. E. Cervigni, P. Dumy, M. Mutter, *Angew. Chem.* **1996**, *108*, 1325; Angew. *Chem. Int. Ed. Engl.* **1996**, *35*, 1230.
[72] G. A. Lemieux, C. R. Bertozzi, *Trends Biotechnol.* **1998**, *16*, 506.
[73] S. Peluso, B. Imperiali, *Tetrahedron Lett.* **2001**, *42*, 2085.
[74] O. Melnyk, J. A. Fehrentz, J. Martinez, H. Gras-Masse, *Biopolymers* **2000**, *55*, 165.
[75] L. E. Canne, A. R. Ferre-D'Amare, S. K. Burley, S. B. H. Kent, *J. Am. Chem. Soc.* **1995**, *117*, 2998.
[76] K. F. Geoghegan, J. G. Stroh, *Bioconjugate Chem.* **1992**, *3*, 138.
[77] H. F. Gaertner, K. Rose, R. Cotton, D. Timms, R. Camble, R. E. Offord, *Bioconjugate Chem.* **1992**, *3*, 262.
[78] O. Melnyk, J. S. Fruchart, C. Grandjean, H. Gras-Masse, *J. Org. Chem.* **2001**, *66*, 4153.
[79] J. A. Fehrentz, M. Paris, A. Heitz, J. Velek, F. Winternitz, J. Martinez, *J. Org. Chem.* **1997**, *62*, 6792.
[80] L. A. Marcaurelle, C. R. Bertozzi, *Tetrahedron Lett.* **1998**, *39*, 7279.
[81] J. C. Spetzler, T. Hoeg-Jensen, *J. Pept. Sci.* **2001**, *7*, 537.
[82] G. Tuchscherer, *Tetrahedron Lett.* **1993**, *34*, 8419.
[83] K. Rose, *J. Am. Chem. Soc.* **1994**, *116*, 30.
[84] E. C. Rodriguez, L. A. Marcaurelle, C. R. Bertozzi, *J. Org. Chem.* **1998**, *63*, 7134.
[85] K. Rose, W. G. Zeng, P. O. Regamey, I. V. Chernushevich, K. G. Standing, H. F. Gaertner, *Bioconjugate Chem.* **1996**, *7*, 552.
[86] C. Grandjean, H. Gras-Masse, O. Melnyk, *Chem. Eur. J.* **2001**, *7*, 230.
[87] Y. M. Zhao, S. B. H. Kent, B. T. Chait, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 1629.
[88] D. Horton, J. D. Wander, in *Carbohydrates: Chemistry and Biochemistry, Vol. 4B* (Eds.: W. W. Pigman, D. Horton), Academic Press, New York, **1990**, p. 799.
[89] H. Kessler, V. Wittmann, M. Koeck, M. Kottenhahn, *Angew. Chem.* **1992**, *104*, 874; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 902.
[90] J. R. Axon, A. L. J. Beckwith, *J. Chem. Soc. Chem. Commun.* **1995**, 549.
[91] R. G. Gasanov, L. V. Ilinskaya, M. A. Misharin, V. I. Maleev, N. I. Raevski, N. S. Ikonnikov, S. A. Orlova, N. A. Kuzmina, Y. N. Belokon, *J. Chem. Soc. Perkin Trans. 1* **1994**, 3343.
[92] M. P. Sibi, Y. Asano, J. B. Sausker, *Angew. Chem.* **2001**, *113*, 1333; *Angew. Chem. Int. Ed.* **2001**, *40*, 1293.
[93] S. B. Cohen, R. L. Halcomb, *Org. Lett.* **2001**, *3*, 405.
[94] S. B. Cohen, R. L. Halcomb, *J. Am. Chem. Soc.* **2002**, *124*, 2534.

Received: February 8, 2001 [M361]