CHEMISTRY
A European Journal

A Journal of
ChemPubSoc Europe

Accepted Article

Title: Features of Auxiliaries that Enable Native Chemical Ligation beyond Glycine and Cleavage via Radical Fragmentation

Authors: Simon Ferdinand Loibl, Andre Dallmann, Kathleen Hennig, Carmen Juds, and Oliver Seitz

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *Chem. Eur. J.* 10.1002/chem.201705927

Link to VoR: http://dx.doi.org/10.1002/chem.201705927

Supported by
ACES

WILEY-VCH

10.1002/chem.201705927

WILEY-VCH

FULL PAPER

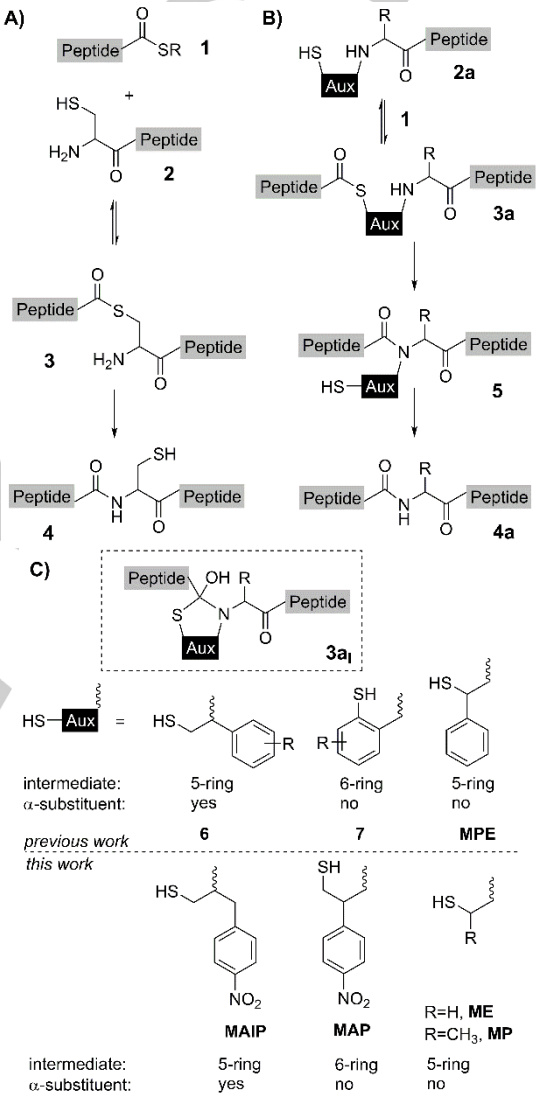
Features of Auxiliaries that Enable Native Chemical Ligation beyond Glycine and Cleavage via Radical Fragmentation

Simon F. Loibl, Andre Dallmann, Kathleen Hennig, Carmen Juds and Oliver Seitz*[a]

Abstract: The native chemical ligation (NCL) is an invaluable tool in the total chemical synthesis of proteins. Ligation auxiliaries overcome the requirement for cysteine. However, the reported auxiliaries remained limited to glycine-containing ligation sites and the acidic conditions applied for cleavage of the typically applied N-benzyl-type linkages promote side reactions. With the aim to improve upon both ligation and cleavage, we systematically investigated alternative ligation scaffolds that challenge the N-benzyl dogma. The study revealed that auxiliary-mediated peptide couplings are fastest when the ligation proceeds via 5-membered rather than 6-membered rings. Substituents in α-position of the amine shall be avoided. We observed, perhaps surprisingly, that additional β-substituents accelerated the ligation conferred by the β-mercaptoethyl scaffold. We also describe a potentially general means to remove ligation auxiliaries by treatment with an aqueous solution of triscarboxyethylphosphine (TCEP) and morpholine at pH 8.5. NMR analysis of a ^{13}C-labeled auxiliary showed that cleavage most likely proceeds through a radical-triggered oxidative fragmentation. High ligation rates provided by β-substituted 2-mercaptoethyl scaffolds, their facile introduction as well as the mildness of the cleavage reaction are attractive features for protein synthesis beyond cysteine and glycine ligation sites.

Introduction

Peptide ligation reactions are the key to the total chemical synthesis of proteins. At current, the native chemical ligation (NCL) [1] is providing the most reliable access to proteins up to a size of 304 amino acids.[2] The reaction involves two unprotected peptide segments i.e. a peptide thioester **1** and a cysteine peptide **2** (Scheme 1). At pH ≥ 6 thioesters undergo thiol exchange reactions, which eventually leads to a thioester intermediate **3**. This sets the stage for an intramolecular S→N acyl shift yielding the native peptide bond in the final ligation product **4**. NCL reactions proceed efficiently in aqueous solution at lower millimolar concentration of peptides. However, the requirement for N-terminal cysteine in the C-terminal peptide fragment limits the breadth of NCL chemistry. Target proteins may lack cysteine or contain cysteine at positions that are difficult to access by NCL chemistry such as Pro-Cys, Thr-Cys, Ile-Cys and Val-Cys junctions.[3]



Scheme 1: Concepts of A) NCL and B) auxiliary-mediated peptide ligation. C) Chemical structures of established (left) Nᵅ-auxiliaries and novel base-labile Nᵅ-auxiliaries (MAP, 3-mercapto-2-aryl-propyl; MAIP, 2-mercapto-2'-aryl-isopropyl; ME; 2-mercaptoethyl; MP, 2-mercaptopropyl: MPE, 2-mercapto-2-phenyl-ethyl).

Ligation auxiliaries have been developed with the aim to expand the applicability of NCL (Scheme 1B). This approach relies on the N-terminal attachment of a cleavable scaffold that offers a mercapto group in a position that facilitates the intramolecular S→N acyl shift in thioester intermediate **3a**.[4] The pioneering work of Dawson[5] and Kent[6] focused on N-benzyl type auxiliaries (**6** and **7** in Scheme 1C) which were designed to

[a] S. F. Loibl, Dr. A. Dallmann, K. Hennig, C. Juds, Prof. Dr. O. Seitz
Institut für Chemie
Humboldt-Universität zu Berlin
Brook-Taylor-Str. 2, 12489 Berlin (Germany)
E-mail: oliver.seitz@chemie.hu-berlin.de

Supporting information for this article is given via a link at the end of the document.

Accepted Manuscript

This article is protected by copyright. All rights reserved.

enable cleavage by acidolysis[6-7] or photolysis[8]. However, the rather low ligation rates at ligation junctions beyond glycine as well as the vulnerability of the formed peptide bond to acid-induced amide cleavage prevented a broad acceptance of this ligation technology.

In the ligation-desulfurization method,[9] the mercapto group is connected to the framework of the N-terminal amino acid. This approach provides for fast ligation reactions and the advent of metal-free radical desulfurization[10] sparked the development of a variety of thiolated amino acid building blocks.[11] However, only 3 out of the 15 thiolated amino acids used in NCL reactions are commercially available. As a result, the application of the ligation-desulfurization approach is limited to labs skilled in organic synthesis.

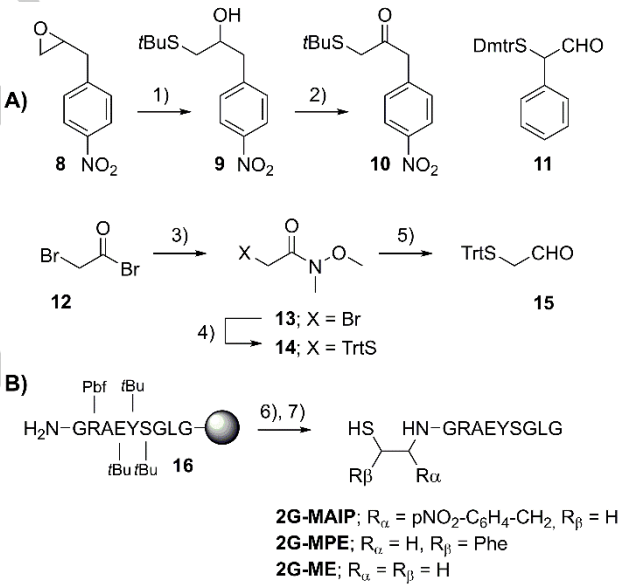
An ideal method would provide access to arbitrary ligation sites without the efforts required for the preparation of a panel of building blocks. Given the ease by which ligation auxiliaries are introduced upon reductive amination in the last step of automated solid-phase peptide synthesis, we reconsidered the design of auxiliary scaffolds. We sought to i) improve ligation rates by avoiding steric bulk at the ligation site and ii) prevent cleavage of the formed N-alkyl amides by getting around the acidic cleavage conditions that typically induce S→N acyl shift reactions. This analysis guided us to the development of new auxiliary scaffolds such as the 2-mercapto-2-phenethyl (MPE) auxiliary[12], which is the first auxiliary that escapes the need for glycine at the ligation junction and permits auxiliary removal under mild basic conditions when S→N acyl shifts are favored[13] over N→S acyl shifts. Previous reports stressed the importance of native chemical ligation-type reactions proceeding through 5-membered ring transition states.[4-5, 14] However, despite a 6-membered ring intermediate (**3a**$_{\mathrm{I}}$) the reactivity of mercaptobenzyl-type auxiliaries **7** seems comparable to the 2-mercapto-1-phenethyl auxiliaries **6**. The seemingly high reactivity was attributed to a template effect conferred by the mercaptobenzyl system.[5] On the other hand, steric bulk was identified as the main reason for the generally low ligation rate obtained with the benzylic auxiliaries **6** and **7**. We felt that the reasons as to why some aminothiol systems react faster in NCL-type chemistry than others seemed rather unclear.

To explore the requirements for a high reactivity in auxiliary-mediated ligations we embarked on a comparative study of different auxiliary scaffolds. Herein, we evaluate 4 different ligation auxiliaries. By comparing the 3-mercapto-2-arylpropyl (MAP) and 2-mercapto-2'-aryl-isopropyl (MAIP) auxiliaries with the recently reported 2-mercapto-2-phenethyl (MPE) auxiliary we shine a light on the role of i) the size of the ring formed in the NCL transition state and ii) the substituent in α position to the amino group (Scheme 1C). To facilitate reactions at sterically demanding ligation sites we sought options to increase scaffold flexibility at the N-attachment site. We therefore included the 2-mercaptoethyl (ME) auxiliary in this study. A comparison of the ME auxiliary with the 2-mercaptopropyl (MPE) auxiliary will expose the role of substituents in the β position to the amine. Of note, the ME scaffold was the first ligation auxiliary reported in the seminal paper by Canne et al.[4] However, a method for the removal of this simplest ligation auxiliary imaginable was still lacking. We show that the mercaptoethyl group as well as the other studied auxiliaries are susceptible to cleavage under mildly basic (pH 8.5), radical conditions. We also introduce a revised mechanism for cleavage of the 2-mercapto-2-phenethyl auxiliary.

Results and Discussion

Scaffold synthesis and introduction of auxiliaries in peptides

Native chemical ligation reactions involving the 2-mercaptoethyl auxiliaries MAIP, MPE and ME will proceed through 5-membered cyclic ligation intermediates. By contrast, the 3-mercaptopropyl auxiliary MAP will lead to a 6-membered ring. For the coupling of 2-mercaptoethyl-type auxiliaries with a C-terminal peptide fragment we chose a reductive amination route (Scheme 2A). Ketone **10** was envisioned as a suitable building block for the introduction of the MAIP auxiliary. We commenced the synthesis by allylation of 4-nitroaniline[15] and subsequent oxidation of the terminal alkene[16] to the epoxide **8**. Treatment with *tert*.-butylmercaptane in presence of tetrabutylammonium fluoride (TBAF) induced the regioselective opening of the epoxide. The β-hydroxy thioether **9** was subsequently converted to the desired ketone **10** by Dess-Martin oxidation.



Scheme 2: Synthesis of A) precursors for auxiliary introduction via B) reductive amination. 1) *t*BuSH, TBAF, 79%; 2) Dess-Martin periodinane, DCM/*t*BuOH, 58%; 3) N, O HNMeOMe, DIPEA, DCM; 4) TrtSH, DIPEA, DCM, 71% (over two steps); 5) $LiAlH_4$, THF, toluene, 87%; 6) **10**, $NaCNBH_3$, AcOH, $CH(OMe)_3$, MeOH, DMF or **11/15**, $NaCNBH_3$, AcOH, *i*PrOH, NMP; 7) **2G-MAIP**: TFA/TIS then TFA/TFMSA/anisole, 20%; **2G-MPE**: TFA/TIS, 25%; **2G-MPE**: TFA/TIS, 18%. The letter "G" in 2G represents N-terminal glycine.

For introduction of the 2-mercapto-2-phenethyl-auxiliary MPE via reductive amination we had recently reported the 6-step synthesis of aldehyde **11**. A similar building block (**15**) should allow the introduction of the unsubstituted 2-mercaptoethyl auxiliary (ME). Bromoacetylbromide was converted to the Weinreb amide **13** prior to treatment with tritylmercaptane. The

Accepted Manuscript

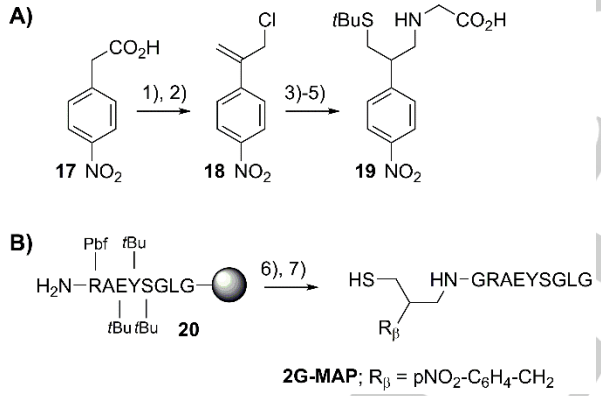
This article is protected by copyright. All rights reserved.

10.1002/chem.201705927

reaction of **14** with lithiumaluminumtetrahydride afforded aldehyde **15**.

In the next step, we equipped the resin-bound peptides **16** with the auxiliaries MAIP, MPE and ME by means of reductive amination with tBu-S-protected ketone **10** or aldehydes **11** and **15** (Scheme 2B). The aldehydes **11** and **15** were dissolved in a mixture of N-methylpyrolidone (NMP) and methanol and were allowed to react with peptide resin **16** in presence of $NaCNBH_3$ and acetic acid. The reductive amination of ketone **10** proceeded best when the solution of $NaCNBH_3$ in DMF/methanol also contained the condensing agent trimethylorthoformiate. After reductive amination, acidolytic cleavage released the desired auxiliary-peptides **2G-MAIP**, **2G-MPE** and **2G-ME** in 18-25% overall yield.

For the introduction of the 6-ring-forming MAP auxiliary, we prepared glycine building block **19** (Scheme 3A). First, 4-nitrophenylacetic acid (**17**) was treated with paraformaldehyde in a decarboxylative double Mannich-type reaction. The N-alkylated morpholine derivative was converted to the chloride **18** by reaction with isobutyl chloroformiate. Following N-alkylation of glycine *t*Bu-ester, a TBAF-promoted addition of *tert*.-butylmercaptane and ester acidolysis furnished the desired auxiliary-glycine conjugate **19**. The *t*Bu-S-protected monomer **19** was coupled to the resin bound peptide (Scheme 3B) and afforded the auxiliary peptide **2G-MAP** (22% yield) after TFA/TFMSA-treatment and HPLC-purification.



Scheme 3: Synthesis of A) precursors for auxiliary introduction via B) coupling of a preformed amino acid building block. 1) morpholine, $(CH_2O)_n$, toluene, 47%; 2) *i*BuOCOCl, toluene, 76%; 3) GlyO*t*Bu, DIPEA, DCM, 86%; 4) *t*BuSH, TBAF, 79%; 5) HCl, dioxane, 99%; 6) PyBOP, DIPEA, DMF; 7) TFA/TIS then TFA/TFMSA/anisole, 22%.

Peptide ligations

Next we compared the reactivity of the auxiliary scaffolds in native chemical ligation reactions involving the peptide thioesters **1G**, **1A** and **1L** (Fig. 1A, the letters G, A and L indicate the C-terminal amino acid). The flexible γ-mercaptopropyl (MAP) auxiliary in **2G-MAP** enabled nearly quantitative formation of the Gly-Gly junction in **5GG** within less than 1 hour ($t_{½}$ = 4 min, Fig. 1B, Table 1). The α-substituted β-mercaptoalkyl auxiliary in **2G-MAIP** required 4 hours ($t_{½}$ = 14 min) to deliver >90% ligation product (Fig. 1C). The reactivity difference became more pronounced in Ala-Gly-forming ligation reactions (Fig. 1D). Based on the reaction half-times, the MAP scaffold provided 6-fold higher reactivity than the MAIP scaffold (**5AG-MAIP**: $t_{½}$ = 417 min, **5AG-MAP**: $t_{½}$ = 66 min). Likewise, the MAP auxiliary afforded the sterically more demanding Leu-Gly junction in >90 % yield within 24 h. In this reaction, the MAIP auxiliary provided < 70% yield. We tested the two auxiliaries in ligations beyond glycine junctions. However, the reaction including alanine peptide thioester **1A** and N-terminal alanine auxiliary peptides **2A-MAP** and **2A-MAIP** furnished no ligation product (see Fig. S14).

At the first glance, it seems surprising that auxiliary-mediated ligation via 6-membered ring intermediates proceeded faster than reactions via 5-membered intermediates. However, we assumed that the results hint at the importance of the substituent in α-position to the secondary amine. Usually, ring closure via a 5-exo-trig process is preferred to a 6-exo-trig process. However, the 5-exo-trig reaction with the MAIP auxiliary is penalized by the α-substituent, which presumably hinders access to the secondary amine. We reasoned that the 2-mercapto-2-arylethyl auxiliary (MPE) should enable faster ligations than the 3-mercapto-2-arylpropyl auxiliary (MAP). While both auxiliaries lack the α-substituent the reaction on the MPE scaffold can proceed via the preferred 5-exo-trig rather than the 6-exo-trig process. Indeed, the 2-mercapto-2-phenyl ethylauxiliary (MPE) provided the highest reactivity and afforded Ala-Gly and Leu-Gly (Fig. 1E) junctions in 3-fold lower reaction half time than the 3-mercapto-2-arylpropyl (MAP) auxiliary (Fig. 1D, Table 1). Most importantly, the reactivity of the MPE auxiliary was sufficiently high to allow the formation of sterically demanding ligation sites that did not contain glycine (Fig. 1I). After 24 h reaction time the MPE auxiliary afforded the Ala-Asn and the Leu-Asn junction in **5AN-MPE** and **5LN-MPE** in 79 % and 63 % yield, respectively (Table 1).

As shown previously, the MPE scaffold is the first auxiliary to enable native chemical ligation beyond glycine.[12] We were interested to elucidate the role of the β-aryl substituent and examined ligation with the unsubstituted 2-mercaptoethyl auxiliary. Perhaps surprisingly, the unsubstituted ME auxiliary

Table 1. Yields[a] and half times[b] of auxiliary-induced NCL reactions (see Fig. 1A).

Junction	MAP	MAIP	MPE	ME
Gly-Gly	96 % (2 h) $t_{½}$ ≈ 4 min	95 % (4 h) $t_{½}$ ≈ 14 min	97 % (2 h) $t_{½}$ ≈ 5 min	/
Ala-Gly	98 % (24 h) $t_{½}$ ≈ 66 min	81 % (24 h) $t_{½}$ ≈ 417 min	96 % (24 h) $t_{½}$ ≈ 25 min	/
Leu-Gly	94 % (24 h) $t_{½}$ ≈ 190 min	66 % (24 h) $t_{½}$ ≈ 480 min	96 % (24 h) $t_{½}$ ≈ 55 min	70 % (24 h) $t_{½}$ ≈ 75 min
Ala-Asn	No product	No product	79 % (24 h) $t_{½}$ ≈ 290 min	75 % (24 h) $t_{½}$ ≈ 550 min
Leu-Asn	No product	No product	63 % (24 h) $t_{½}$ ≈ 960 min	34 % (24 h) $t_{½}$ ≈ 2200 min

[a] based on HPLC analysis after time indicated in parenthesis; [b] estimated from reaction time course.

Accepted Manuscript

This article is protected by copyright. All rights reserved.

10.1002/chem.201705927

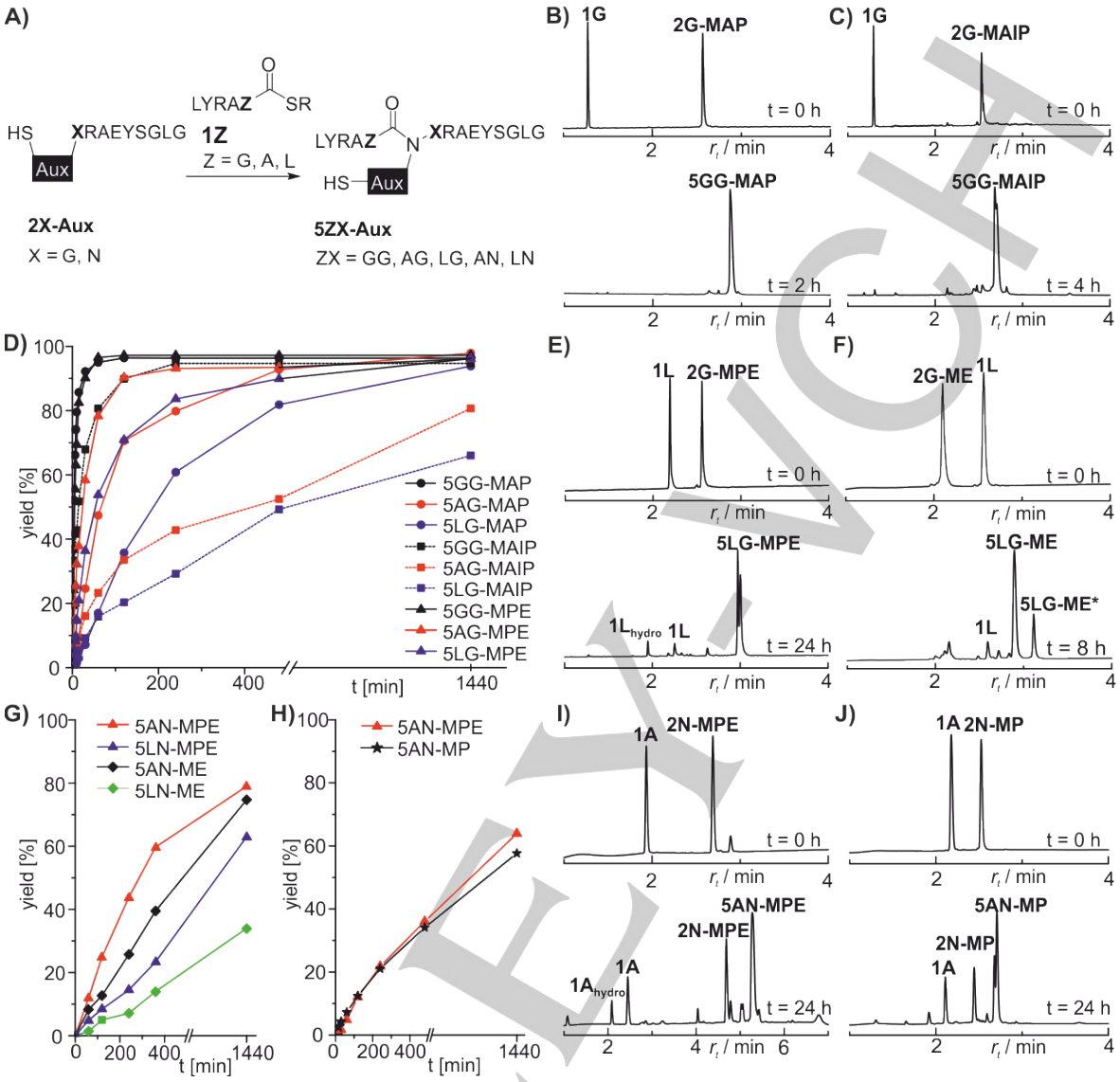


Figure 1: a) Auxiliary-mediated ligation reaction between peptide thioesters **1Z** and auxiliary modified peptides **2X-Aux** (the letters Z and X represent the C-terminal amino acid of the peptide thioester and the N-terminal amino acid of the auxiliary peptide, respectively). UPLC analysis of Gly-Gly ligation with the B) MAP and C) MAIP auxiliary. D) Time course of Gly-Gly, Ala- Gly and Leu-Gly ligation reactions using MAP, MAIP and MPE auxiliaries. UPLC analysis of Leu-Gly ligation on E) MPE and F) ME auxiliaries (**5LGMe***, double acylated product obtained upon reaction of **5LGMe** with thioester **1L**, see Fig S10A for structure). Time course of G) Ala-Asn and Leu-Asn ligations using MPE and ME auxiliaries. Comparison of the H) time course and UPLC analysis of Ala-Asn-ligation on the I) MPE and J) MP auxiliary. Conditions: 2-5 mM peptides, 20 mM TCEP, 100 mM Na_2HPO_4, 3 vol% PhSH, rt, pH 7.5.

proved less reactive than the MPE auxiliary. While there was little difference in native chemical ligation at the Leu-Gly junction (Fig. 1F), the unsubstituted ME scaffold required approx. double as much time as the MPE scaffold to afford the Ala-Asn and Leu-Asn bonds in **5AN-ME** and **5LN-ME**. With the MPE auxiliary the Leu-Asn junction was established in > 60 % yield after 24 h whereas the ME-mediated reaction proceeded in 34% yield only.

The inferior reactivity of the unsubstituted ME auxiliary seemed counter-intuitive. A closer inspection exposes a rather linear behavior of the reaction time course (Fig. 1G), which suggests that ligations at sterically demanding junctions bear resemblance to a zero order process. Such a case may result from a rate limiting intramolecular S→N acyl shift which proceeds from a small steady state concentration of the thioester intermediate **3a**; a notion which has been put forward already in previous reports on auxiliary-mediated ligations. [7b, 14, 17] Indeed, UPLC analyses show formation of the thioester intermediate, which precedes the S→N acyl shift (Fig. S15). Since the S→N acyl shift progresses via a cyclic intermediate we assumed that under these circumstances the native chemical ligation can be analyzed in terms of a cyclization reaction. According to the Thorpe-Ingold effect cyclizations are accelerated when a ring

Accepted Manuscript

This article is protected by copyright. All rights reserved.

10.1002/chem.201705927

hydrogen atom is substituted by more space filling groups.[18] To substantiate that the reactivity difference between the MPE and the ME auxiliary is due to the Thorpe-Ingold effect we evaluated the 2-mercaptopropyl (MP) auxiliary (see SI for details on synthesis). The time courses of the Ala-Asn ligation mediated by the MPE and the MP auxiliaries were nearly superimposable (Fig. 1H). We conclude that, perhaps surprisingly, β-substitution of β-mercaptoalkyl auxiliaries increases rather than decreases the rate of auxiliary-mediated native chemical ligation.

Auxiliary removal

The results of the ligation experiments exposed the MPE scaffold as the most reactive ligation auxiliary. In previous work, we had demonstrated that treatment of ligation products such as **5GG-MPE** with triscarboxyethylphosphine (TCEP) and morpholine at pH 8.5 in aqueous buffer resulted in the quantitative removal of the auxiliary (Fig. 2A,B). Based on the observation that 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) − a persistent radical used as radical scavenger – inhibited the cleavage reaction (Fig. S13),[12] we assumed a radical pathway for auxiliary cleavage. Initially, we had hypothesized that the benzylic radical obtained upon TCEP-mediated radical desulfurization in presence of oxygen would trigger a β-fragmentation which would lead to styrene and an amide radical. However, in presence of $Na_2S_2O_8$ as radical starter (Fig. 2D) we noticed the formation of an N-formylpeptide species **4GGf**. This is indicative of a C-C bond scission and questioned the postulated cleavage mechanism.

To elucidate the fate of the auxiliary we analyzed the reaction by NMR spectroscopy. To help identify auxiliary debris we incubated an α,β-^{13}C-labelled analogue of ligation product **5GG-MPE** (see SI for synthesis) with an aqueous solution (5% D_2O in water) of TCEP and morpholine and followed the reaction by heteronuclear single quantum coherence (HSQC)-spectroscopy. The reaction mixture was extracted with $CDCl_3$. The UPLC analysis showed that $CDCl_3$ extraction lead to the disappearance of peak **21** (compare Fig.2B with Fig. 2C). The 2D-HSQC spectra (see Fig. 2E and Fig. 2F) of the two phases exposed four distinct species in the lower field. With the help of reference materials the two major components were identified as benzaldehyde **21** and N-formyl-morpholine **22** and the two remaining species as traces of formic acid (**23**) and N-formylated peptide (<5% according to UPLC analysis in Fig. 2B). The formation of benzaldehyde was also observable in 1D-^{1}H-NMR measurements of the mixture obtained after cleavage of non-labeled MPE. The signals of the

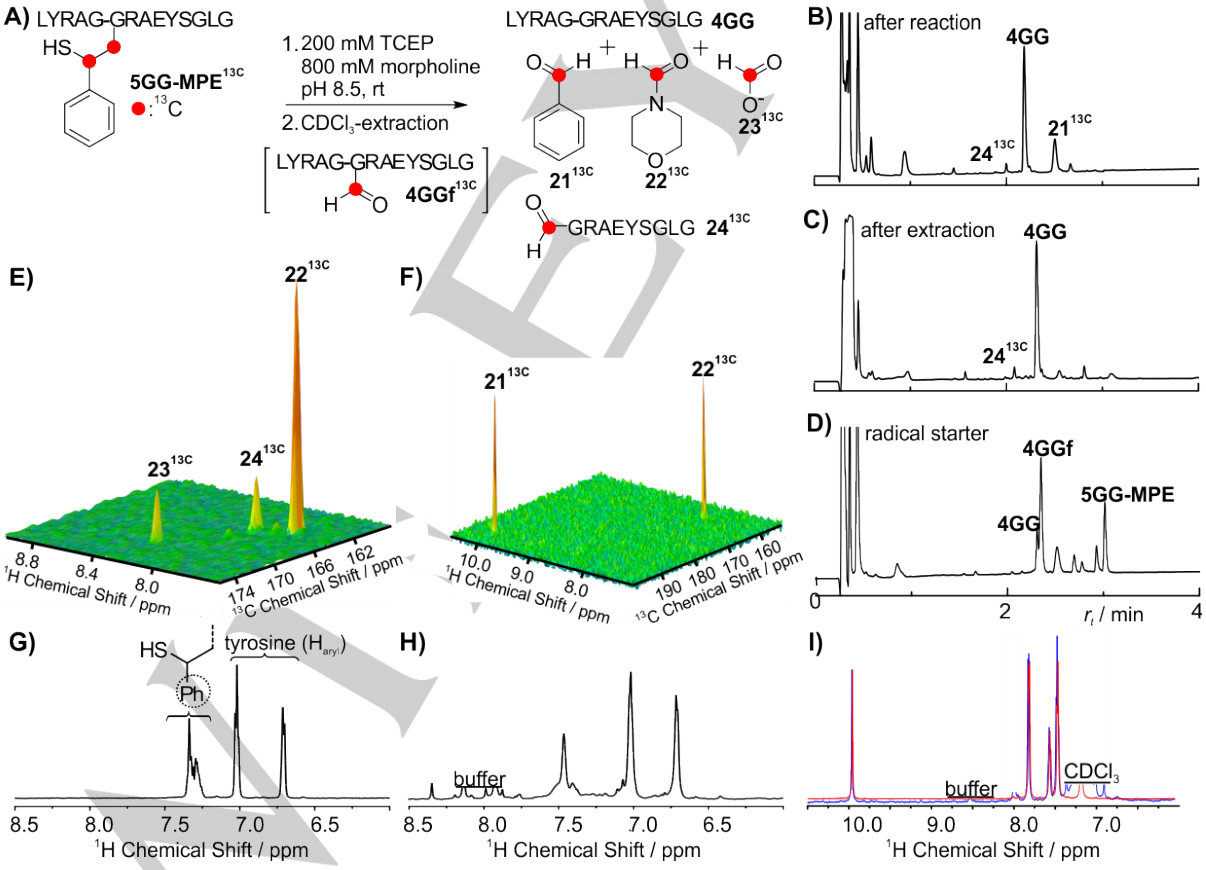


Figure 2: A) Removal of the ^{13}C-labeled 2-mercaptophenethyl auxiliary (MPE) from ligation product **5GG-MPE^{13C}**. UPLC-analysis of the reaction mixture obtained upon incubation of **5GG-MPE^{13C}** (B), C)) or **5GG-MPE** (D)) after B) 16 h, C) after extraction with $CDCl_3$ and D) after 1h, when cleavage was performed in presence of 100mM $Na_2S_2O_8$. HSQC-NMR spectra of the E) aqueous and F) organic phase after extraction of the reaction mixture with $CDCl_3$.^{1}H-NMR spectra of **5GG-MPE** G) before and H) after cleavage. I) Comparison of NMR spectra obtained from the $CDCl_3$ extract (red curve) and benzaldehyde in $CDCl_3$ (blue curve) as authentic reference.

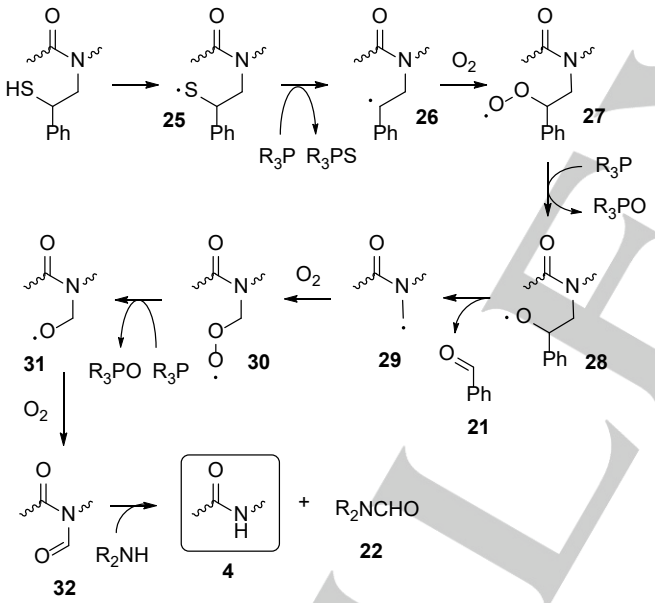
Accepted Manuscript

This article is protected by copyright. All rights reserved.

10.1002/chem.201705927

MPE auxiliary experienced a downfield shift upon cleavage (compare Fig. 2G and Fig. 2H). Inspection of the organic phase obtained after $CDCl_3$ extraction revealed a pattern characteristic of a phenyl ring system (Fig. 2I, red curve). Most notable was a signal at δ = 9.97 ppm, which hinted at benzaldehyde. The comparison of the ^{1}H-NMR spectra with authentic reference material (Fig. 2I, blue curve) confirmed the presence of benzaldehyde in the reaction mixture.

Based on these results, we revise the postulated mechanism for cleavage of the 2-mercapto-2-phenyl ethyl auxiliary (Scheme 4). Under basic conditions, the thiolate should be readily converted to the thiyl-radical by air-oxidation.[19] The high concentration of phosphine triggers the desulfurization[10] and furnishes the benzyl radical **26**. In the absence of a radical scavenger (such as the high concentration of mercaptanes used in radical desulfurization[10, 20]) benzyl radicals should rapidly react with molecular oxygen dissolved in the aqueous buffer. Such reactions are barrierless and proceed with very high speed.[21] As described by Buckler[22] and Howard[23] the resulting peroxyradical **27** should readily react with TCEP under formation of the alkoxy-radical intermediate **28**, which based on literature evidence[24] is expected to undergo β-fragmentation by release of benzaldehyde.[24b, 25]



Scheme 4. Proposed cleavage mechanism of the 2-mercaptophenethyl (MPE) auxiliary induced by TCEP and morpholine.

The remaining amidoalkyl radical **29** is a known species[26] which likely is processed into the α-alkoxy radical **31** by means of the reaction with molecular oxygen and subsequent deoxygenation by TCEP. Without a connected downhill process β-fragmentation of this radical (to an amide radical and formaldehyde) should be disfavored. Rather, the formation of the formyl species **22** and **23** indicates that the amidoalkoxy radical **31** undergoes oxidation to the N-formyl peptide **32** (such as **4GGf** in Fig. 2). This intermediate is observable when radical reactions are accelerated upon use of radical initiator (Fig. 2D). In the last step, morpholine (or hydroxide anions) attack the N-formylpeptide at the most accessible electrophilic position i.e. the formyl carbon leading to the native peptide **4** and N-formyl morpholine **22** (or formiate **23**).

The repeated involvement of oxygen raises the question as to how the reaction would proceed when oxygen is limiting. Auxiliary cleavage was performed at reduced concentration of oxygen (Fig. S22) and we noticed the formation of an N-methylated peptide by-product, which suggests that the amidoalkyl radical **29** can be scavenged by hydrogen atom abstraction. The postulated mechanism of auxiliary cleavage also provides an explanation for the occurrence of N-terminally formylated fragment **24**, which we observed when the auxiliary was removed from **5GG-MPE** (see Fig. 2E). In principle, any auxiliary-modified fragment **2G-MPE** that remains after the ligation reaction will be converted to the N-formyl species upon auxiliary cleavage. The N-terminal formamide is not reactive enough to undergo aminolysis at pH 8.5. However, despite our efforts to remove remaining auxiliary peptide **2G-MPE**, we still observed the formation of < 5% N-terminal formamide **24**. We therefore assume that the side-reaction points to a peculiarity of formylated Xxx-Gly amide bonds. Due to the absence of an α-substituent at glycine, the nuclophilic attack of the bisacyl amide may occur also at the backbone. For all other ligation junctions we expect that nucleophilic attack will almost exclusively occur at the less hindered formyl group.

The mechanism in Scheme 4 describes a potentially generic approach to remove ligation auxiliaries under mildly basic conditions. Indeed, both the MAIP and the MAP auxiliaries were removed by treatment of the ligation products **5GG-MAIP** and **5GG-MAP**, respectively, with the TCEP/morpholine mixture (Fig. 3A). The HPLC analysis suggested that the β-mercaptopropyl scaffold of the MAIP auxiliary (Fig. 3B) reacted more smoothly than the γ-mercaptopropyl scaffold of the MAP auxiliary (Fig. 3C). We assume that the cleavage of the MAP auxiliary proceeds through three alkoxy radical intermediates and scission of two carbon-carbon bonds. This rather lengthy path probably is prone to interference. Interestingly, the radical oxidative fragmentation outlined in Scheme 4 also offered a means to remove the simplest of all auxiliaries i.e. the mercaptoethyl (ME) scaffold. Cleavage of the ME group from the Gly-Gly ligation product proceeded as rapidly as cleavage of the MPE group. However, the noisy base-line in the UPLC trace (Fig. 2D) suggests that cleavage of the ME auxiliary results in more by-products than cleavage of the MPE auxiliary. A slightly higher purity of the crude material was obtained in cleavage reactions of the 2-mercapto-2-propyl (MP) auxiliary (Fig. 3E). Though not optimized, the removal of the ME but not the MPE auxiliary from Gln-Phe and Ala-Asn bonds proved problematic (Fig. S21). Given that the MPE auxiliary provides for the highest yields in both ligation and cleavage reactions we consider the MPE auxiliary as the current gold standard in auxiliary mediated native chemical ligation.

Conclusions

In this work we have extended the repertoire of auxiliary-mediated

Accepted Manuscript

This article is protected by copyright. All rights reserved.

10.1002/chem.201705927

peptide ligation by introducing four new auxiliaries, which can be

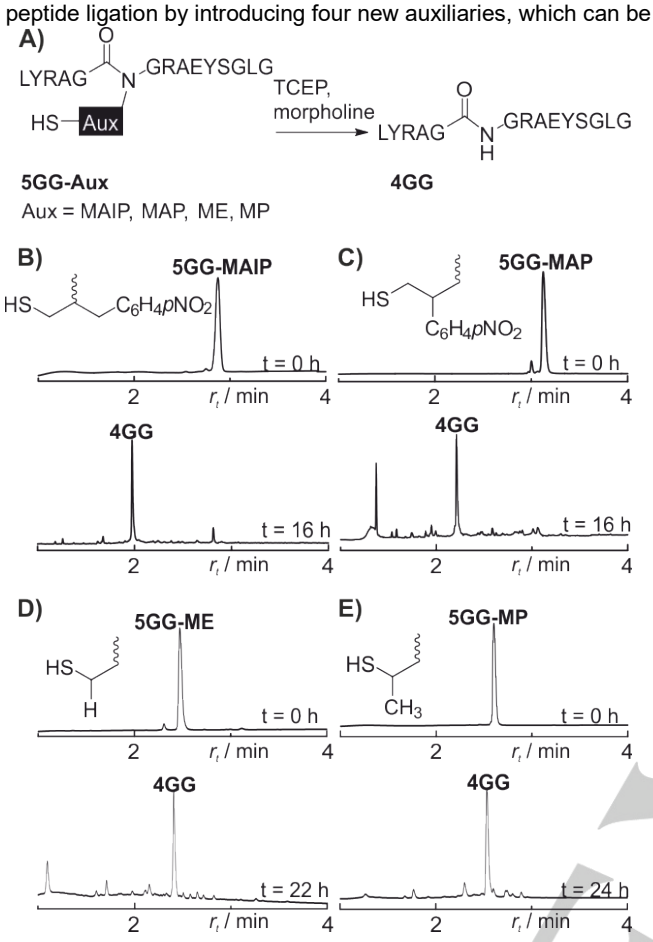


Figure 3. A) Treatment with TCEP/morpholine as a generic method for inducing cleavage of ligation auxiliaries. UPLC analysis of cleavage of the B) MAIP auxiliary; C) MAP auxiliary; D) ME auxiliary and E) MP auxiliary. Conditions: **5GG-MAIP** and **5GG-MAP**: 0.5 mM peptides, 20 mM TCEP, 140 mM morpholine pH 9-9.5, 40°C; **5GG-ME** and **5GG-MP**: 0.5 mM peptides, 100 mM TCEP, 400 mM morpholine pH 8.5, 40°C.

cleaved under mildly basic conditions in presence of TCEP. The comparison of 5 different scaffolds suggests that extended native chemical ligation is limited to glycine containing ligation sites as long as the reaction proceeds through 6-membered rather than 5-membered rings and auxiliaries contain substituents in α-position of the amine. The study included the reactivity assessment of three structurally similar ligation auxiliaries that were based on α-unsubstituted β-mercaptoethyl scaffolds. We observed, perhaps surprisingly, that additional β-substituents increased the ligation rates. For example, the 2-mercapto-2-phenethyl (MPE) and 2-mercaptopropyl (MP) auxiliaries formed an Ala-Asn junction at a 2-fold higher rate than the structurally simplest mercaptoethyl (ME) auxiliary. We attribute the increased reactivity to the Thorpe Ingold effect, which may gain importance when the S→N acyl shift becomes rate limiting.

A notable observation was: all of the studied ligation auxiliaries were susceptible to cleavage with an aqueous mixture of TCEP and morpholine at pH 8.5. The supression of cleavage in presence of a radical scavenger such as TEMPO suggests a radical pathway of auxiliary removal. To elucidate the apparently general mechanism we analyzed the cleavage of the MPE auxiliary, the synthetic utility of which had been proven in the total synthesis of the 48- and 126 amino acid long proteins opistoporin-2[12] and mucin-1[27]. NMR analysis of cleavage reactions performed with isotopically labeled auxiliary revealed benzaldehyde and formylmorpholine as the major components of auxiliary debris. This and the observation of amide-linked formyl peptide intermediates points to an oxidative fragmentation reaction which most likely involves the attack of dissolved oxygen on the alkyl radical formed upon radical desulfurization. Owing to the high concentration of TCEP the resulting alkyl peroxyradicals will rapidly be converted to alkoxyradicals. The cleavage mechanism probably proceeds through two types of alkoxyradicals with distinct reactivity. One type undergoes β-fragmentation and therefore provides an aldehyde along with a new alkylamido radical on the remaining auxiliary fragment shortened by one carbon unit. Upon reaction with oxygen and TCEP this radical gives rise to the second type of alkoxy radical, which in the absence of an energetically accessible β-fragmentation pathway will react with oxygen under formation of an amide-linked formyl group. The latter is subject to aminolysis (or hydrolysis) which eventually affords the auxiliary-free peptide.

The combination of radical oxygenation, phosphine induced cleavage of O-O bonds and subsequent β-fragmentation or oxidation of the resulting alkyloxy radicals suggests a potentially generic means for the cleavage of amide-linked mercaptoalkyl auxiliaries. According to that, the mercapto group serves as the entry point for the induction of the radical oxidation-fragmentation cascade which eventually results in the formation of a base-labile, amide-linked formyl peptide. The mechanism plausibly explains why such different auxiliaries as 3-mercapto-2-arylpropyl, 2-mercapto-2'-aryl-isopropyl, 2-mercapto-2-phenethyl, 2-mercaptopropyl and even 2-mercaptoethyl scaffolds are cleavable by exposure to aqueous TCEP and morpholine.

The high ligation rate provided by the β-substituted 2-mercaptoethyl scaffolds, their facile introduction by means of reductive amination as well as the mildness of the cleavage reaction, which seems particularly attractive for the synthesis of posttranslationally modified peptides, are appealing features for protein synthesis beyond cysteine and glycine ligation sites. However, we are aware of potential shortcomings. Auxiliary cleavage will affect ligation-unrelated cysteine side chains. We have previously demonstrated that methionine residues are tolerated and that side chain protecting groups such as the acetamido (ACM) group protect cysteine side-chains from side reactions.[12] A potential problem may arise from the requirement for oxygen in the auxiliary cleavage reaction. In our experiments, oxygen was made available by performing the experiments at small micromol scale in plastic vials. At large scales oxygen may be limiting and it may therefore prove necessary to explore alternative sources of oxygen.

Even though the usefulness of the MPE auxiliary has been demonstrated previously,[12] the two key observations i.e. i) β-substituents increase ligation rate and ii) TCEP/morpholine cleaves ligation auxiliaries through a radical oxidation-

Accepted Manuscript

This article is protected by copyright. All rights reserved.

10.1002/chem.201705927

fragmentation cascade suggest means for potentially improved designs of cleavable ligation auxiliaries. In ongoing work we deliberately avoid α-substituents and rather focus on variations of β-substituents, which on the one hand may enable faster ligation reactions and on the other hand allow tuning of the speed of radical formation and radical fragmentation. We expect that β-substituted β-mercaptoethyl scaffolds will become a highly useful tool for the rapid total synthesis of (posttranslationally modified) proteins by native chemical ligation beyond cysteine and glycine.

Experimental Section

Synthesis of auxiliary building blocks:

Alcohol 9: The epoxide **8** [16] (869 mg, 4.85 mmol) was dissolved in neat *tert*.-butylmercaptane (1.10 ml, 880 mg, 9.75 mmol) and a catalytical amount of TBAF*3H_2O (27 mg, 0.25 mmol) was added. The reaction mixture was stirred at 50°C for 18 h and subsequently concentrated *in vacuo*. The desired product (1.03 g, 3.81 mmol, 79%) was isolated as a yellow oil after purification by flash-chromatography on silica-gel using cyclohexane/ethylacetate (4/1) as a mobile phase. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.14-8.11 (m, J_1 = 9.0 Hz, , J_2 = 2.5 Hz, J_3 =2.0 Hz, 2H), 7.41-7.39 (m, J_1 = 9.0 Hz, , J_2 = 2.5 Hz, J_3 =2.0 Hz, 2H), 3.94-3.89 (m, 1H), 2.96- 2.93 (m, J_1 = 13.5 Hz, , J_2 = 4.5 Hz, 1H) 2.89-2.85 (m, J_1 = 14.0 Hz, , J_2 = 8.0 Hz, 1H), 2.76-2.72 (m, J_1 = 13.0 Hz, , J_2 = 4.5 Hz, 1H), 2.62 (s, 1H), 2.58-2.54 (m, J_1 = 13.0 Hz, , J_2 = 7.5 Hz, 1H), 1.29 (s, 9H). ^{13}C-NMR (126 MHz, $CDCl_3$): δ = 146.8, 146.3, 130.4, 123.6, 70.6, 42.8, 42.5, 35.9, 31.1. HRMS (ESI): m/z = 270.1160 ($C_{13}H_{20}NO_3S$ $(M+H)^+$, calc.: 270.1164); $C_{13}H_{19}NO_3S$: 269.35 g·mol^{-1}.

S-*t*Bu-protected ketone 10: To a solution of the alcohol **9** (969 mg, 3.60 mmol) in a mixture of DCM/tert.butanol (22 ml, 9/1; v/v) was added Dess-Martin periodinane (12.6 g, 4.43 mmol, 15 wt% solution in DCM). After 3.5 h the reaction was stopped by addition of a mixture of aqueous $NaHCO_3$ (10 wt%, 10 ml) and aqueous $NaHSO_3$ (saturated, 10 ml). The aqueous phase was separated and extracted with ethylacetate (2x 20 ml). The combined organic phases were dried over $MgSO_4$ and subsequently concentrated *in vacuo*. The desired product (558 mg, 2.09 mmol, 58%) was isolated as a yellow solid after purification by flash-chromatography on silica-gel using cyclohexane/ethylacetate (4/1) as a mobile phase. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.16-8.14 (m, 2H), 7.39-7.37 (m, 2H), 4.06 (s, 2H), 3.36 (s, 2H), 1.30-1.29 (m, 9H). ^{13}C-NMR (126 MHz, $CDCl_3$): δ = 203.6, 147.1, 141.7, 130.7, 123.8, 46.4, 43.9, 39.5, 30.8. HRMS (ESI): m/z = 266.0856 ($C_{13}H_{16}NO_3S$ $(M-H)^-$, calc.: 266.0845); $C_{13}H_{17}NO_3S$: 267.34 g·mol^{-1}.

Weinreb amide 14: To an ice-cold solution of *N*, *O*-dimethylhydroxylamine hydrochloride (0.50 g, 5.13 mmol) in DCM (50 ml) was added DIPEA (1.33 g, 1.79 ml, 10.3 mmol) and dropwise bromoacetylbromide **12** (1.04 g, 446 µL, 5.13 mmol). The mixture was allowed to warm to room temperature and stirred for 2.5 h. The organic phase was washed with sat. $NaHSO_4$ solution (2 x 15 ml), dried over $MgSO_4$ and concentrated *in vacuo*. The residue was dissolved in DCM (25 ml). To this solution was added triphenylmethyl mercaptan (1.70 g, 6.15 mmol) and DIPEA (0.66 g, 0.89 ml, 5.10 mmol). After 16 h the mixture was concentrated *in vacuo.* The desired product (1.37 g, 3.62 mmol, 71%; $C_{23}H_{23}NO_2S$; MW = 377.50 g·mol^{-1}) was isolated as white solid after purification by flash-chromatography on silica-gel using cyclohexane/ethyl-acetate/EDMA (2/1/0.01) as a mobile phase. ^{1}H-NMR (300 MHz, $CDCl_3$): δ = 7.50-7.46 (m, 6H), 7.35-7.21 (m, 9H), 3.48 (s, 3H), 3.13 (s, 3H), 3.10 (s, 2H).

Aldehyde 15: To a solution of weinreb amide **14** (250 mg, 662 µmol) in dry THF (5 ml) was added dropwise $LiAlH_4$ (207 µL, 725 µmol, 3.5 M in THF/toluene) at -78°C. After 30 min $NaHSO_4$ solution (25 ml, 5 wt%) was added and the mixture was allowed to warm to room temperature. The mixture was diluted by addition of DCM (50 ml), the organic phase was separated, washed with sat. $NaHSO_4$ solution (2 x 50 ml, 5 wt%), dried over $MgSO_4$ and concentrated *in vacuo*. The desired product (183 mg, 575 µmol, 87%) was isolated as brown oil after purification by flash-chromatography on silica-gel using cyclohexane → cyclohexane/ethylacetate/EDMA (2/1/0.01) as a mobile phase. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.77-8.76 (t, J = 2.5 Hz, 1H), 7.38-7.35 (m, 3H), 7.24-7.16 (m, 12H), 3.02-3.01 (d, J = 2.5 Hz, 2H). ^{13}C-NMR (126 MHz, $CDCl_3$): δ = 196.7, 144.2, 129.7, 128.3, 127.3, 67.6, 42.6.

Allyl chloride 18: To a suspension of 4-nitrophenylacetic acid **17** (10.0 g, 55.2 mmol) and paraformaldehyde (4.01 g, 132 mmol) in toluene (25 ml) was added morpholine (4.8 ml, 4.75 g, 55. 6 mmol). The reaction mixture was heated (90°C, 1 h→ 100°C, 4 h) and subsequently concentrated *in vacuo*. The residue was dissolved in DCM (50 ml), the organic phase washed with water (2 x 25 ml) and concentrated *in vacuo*. The desired morpholine derivative (6.34 g, 25.6 mmol, 47%) was isolated as a yellow solid after purification by flash-chromatography on silica-gel using cyclohexane/ethylacetate (3/1) as a mobile phase. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.18-8.16 (dd, J_1 = 9.0 Hz, J_2 = 0.5 Hz, 2H), 7.71-7.68 (m, , J_1 = 9.0 Hz, J_2 = 2.5 Hz, 2H) 5.64 (d, J = 1.0 Hz, 1H) , 5.41 (d, J = 1.0 Hz, 1H), 3.66-3.64 (t, J = 4.5 Hz, 4H), 3.35 (d, J = 0.5 Hz, 2H), 2.46-2.45 (t, J = 4.5 Hz, 4H). ^{13}C-NMR (126 MHz, $CDCl_3$): δ = 146.7, 142.4, 127.3, 123.6, 119.3, 67.1, 63.5, 53.5. HRMS (ESI): m/z = 249.1235 ($C_{19}H_{24}NO_5S_2$ $(M+H)^+$, calc.: 249.1239); $C_{13}H_{16}N_2O_3$: 248.28 g·mol^{-1}.To a solution of the morpholine derivative (2.85 g, 11.5 mmol) in toluene (5 ml) was added dropwise isobutyl chloroformate (1.8 ml, 1.89 g, 13.8 mmol). After 18 h the reaction mixture was concentrated *in vacuo*. The desired product (1.73 g, 8.77 mmol, 76%) was isolated as a yellow oil after purification by flash-chromatography on silica-gel using cyclohexane/ethylacetate (3/1) as a mobile phase. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.17-8.15 (dd, J_1 = 8.5 Hz, J_2 = 1.0 Hz, 2H), 7.62-7.60 (d, J = 8.5 Hz, 2H), 5.69 (s, 1H), 5.62 (s, 1H), 4.48-4.47(d, J = 0.5 Hz, 2H). ^{13}C-NMR (126 MHz, CDCl3): δ = 147.4, 144.0, 142.3, 127.0, 123.7, 120.3, 45.9. HRMS (ESI): m/z = 220.0138 ($C_9H_8ClNO_2Na$ $(M+Na)^+$, calc.: 220.0141); $C_9H_8ClNO_2$: 197.62 g·mol^{-1}.

MAP-auxiliary glycine conjugate 19: To a solution of the allyl chloride **18** (435 mg, 2.20 mmol) in DCM (10 ml) was added subsequently glycine tert-butyl ester hydrochloride (1.11 g, 6.60 mmol) and DIPEA (1.13 ml, 1.50 g, 11.7 mmol). After 16 h the reaction mixture was washed with (2 x 10 ml), dried over $MgSO_4$ and concentrated *in vacuo*. The desired glycine intermediate 1 (555 mg, 1.90 mmol, 86%) was isolated as a yellow oil after purification by flash-chromatography on silica-gel using cyclohexane/ethylacetate (4/1) as a mobile phase. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.19-8.16 (dt, J_1 = 9.0 Hz, J_2 = 2.5 Hz, 2H), 7.65-7.62 (dt, J_1 = 9.0 Hz, J_2 = 2.5 Hz, 2H), 5.57 (s, 1H), 5.44 (d, J = 1.0 Hz, 1H), 3.68 (d, J = 1.0 Hz, 2H), 3.30 (s, 2H), 1.77 (s, 1H), 1.46 (s, 9H). ^{13}C-NMR (126 MHz, $CDCl_3$): δ = 171.7, 147.3, 146.4, 144.4, 127.1, 123.8, 117.6, 81.5, 52.9, 50.8, 28.2. HRMS (ESI): m/z = 293.1496 ($C_{15}H_{21}N_2O_4$ $(M+H)^+$, calc.: 293.1501); $C_{15}H_{20}N_2O_4$: 292.33 g·mol^{-1}. The glycine intermediate1 (494 mg, 1.69 mmol) was dissolved in neat *tert*.-butylmercaptane (238 µL, 190 mg, 2.11 mmol) and a small amount of TBAF·3H_2O (27 mg, 0.09 mmol) was added. The reaction mixture was stirred at rt for 16 h and concentrated *in vacuo*. The desired intermediate 2 (508 mg, 1.34 mmol, 79%) was isolated as a yellow oil after purification by flash-chromatography on silica-gel using cyclohexane/ethylacetate (4/1) as a mobile phase. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.18-8.16 (d, J = 3.5 Hz, 2H), 7.41-7.39 (d, J = 3.5 Hz, 2H), 3.27-3.19 (dd, , J_1 = 17.0 Hz, J_2 = 6.5 Hz, 2H), 3.10-3.04 (m , 2H) 2.98-2.94 (m, 2H), 2.91-2.87 (m, 2H), 2.76-2.72 (m, 2H), 1.51 (s, 1H), 1.42 (s, 9H), 1.25 (s, 9H). ^{13}C-NMR (126 MHz, $CDCl_3$): δ = 171.5, 150.4, 147.0,

Accepted Manuscript

This article is protected by copyright. All rights reserved.

128.8, 123.9, 81.4, 54.0, 51.8, 46.9, 42.5, 32.3, 30.9, 28.2. HRMS (ESI): m/z = 383.1998 ($C_{19}H_{31}N_2O_4S$ $(M+H)^+$, calc.: 383.2005); $C_{19}H_{30}N_2O_4S$: 382.52 g·mol^{-1}. For removal of the *t*Bu ester intermediate 2 (2.40 g, 6.27 mmol) was dissolved in 4 M HCl in dioxane (5 ml). The reaction mixture was stirred at rt for 2 h and subsequently concentrated *in vacuo*. The product was co-evaporated with toluene (3 x 5 ml). The desired MAP-glycine conjugate (2.25 g, 6.20 mmol, 99%) was obtained as a white solid. ^{1}H-NMR (500 MHz, CD_3OD): δ = 8.28-8.26 (d, *J* = 8.5 Hz, 2H), 7.63-7.61 (d, *J* = 8.5 Hz, 2H), 3.88- 3.80 (m, J_1 = 17.0 Hz, , J_1 = 8.0 Hz, 2H), 3.63-3.50 (m, J_1 = 9.5 Hz, J_2 =5.0 Hz, 2H), 3.41-3.35 (m, *J* = 7.5 Hz, 1H), 3.01-2.86 (m, J_1 = 7.5 Hz, J_2 =5.0 Hz 2H), 1.29 (s, 9H). ^{13}C-NMR (126 MHz, CD_3OD): δ = 168.9, 149.2, 148.1, 130.6, 125.1, 52.7, 45.1, 43.6, 33.2, 31.2. HRMS (ESI): m/z = 327.1374 ($C_{15}H_{23}N_2O_4S$ $(M+H)^+$, calc.: 327.1379); $C_{15}H_{23}ClN_2O_4S$: 362.87 g·mol^{-1}.

Synthesis of auxiliary peptides

The peptide **16** was synthesized on a Rink-amide resin (Fmoc-strategy, see SI.). The N-terminal Fmoc-group was removed prior to reductive amination or coupling. After reaction times indicated the *t*Bu-protected auxiliary peptides were cleaved from the peptide-resin with a mixture of TFA/TIS (95/5, v/v), precipitated in ether and treated with a mixture of TFA/TFMSA/anisole (8/1/1, v/v/v) at 0°C to liberate the mercapto-groups (deprotection of *t*Bu-thioether, see Fig. S4 and S6)). The desired auxiliary peptides were purified by semi-preparative HPLC (gradients of solvent B (0.1% TFA, 1% H_2O, 98.9% acetonitrile) in solvent A (0.1% TFA, 1% acetonitrile, 98.9% H_2O) and isolated as white solids after lyophilisation.

2G-MAIP: For reductive amination, the peptidyl-resin (20 µmol) was incubated with a mixture of ketone **10** and $NaCNBH_3$ (15 eq each, c = 0.4 M) in DMF/MeOH/AcOH/TMOF (3/3/3/1) for 16 h. Semi-preparative HPLC: 3-45% B in 30 min; UPLC: t_R = 2.7 min (3-40% B in 4 min); ESI-MS: 1103.6 ($(M+H)^+$, calc.: 1103.5), 552.4 ($(M+2H)^{2+}$, calc.: 552.3); $C_{47}H_{70}N_{14}O_{15}S$: 1103.20 g·mol^{-1}; yield: 3.97 µmol (20%; A_{280} = 0.99, V= 4 ml).

2G-MPE and **2N-MPE** were synthesized as described recently.[12]

2G-ME and **2N-ME**: For reductive amination, the peptidyl-resins were incubated with a mixture of aldehyde **15** and $NaCNBH_3$ (15 eq each relative to resin-loading, c = 0.4 M) in NMP/*i*PrOH (3/1) with 5 vol% AcOH for 16 h. Semi-preparative HPLC: 3-40% B in 4 min. **2G-ME**: Synthesis scale:17.9 µmol; UPLC-MS: t_R = 2.2 min; m/z = 968.6 ($(M+H)^+$, calc.: 968.5), 484.9 ($(M+2H)^{2+}$, calc.: 484.7); $C_{40}H_{65}N_{13}O_{13}S$: 968.1 g·mol^{-1}; yield: 3.18 µmol (18%; A_{280} = 0.408, V = 1.0 ml). **2N-ME**: synthesis scale: 12.8 µmol; UPLC-MS: t_R = 2.2 min (3-40% B in 4 min); m/z = 1025.6 ($(M+H)^+$, calc.: 1025.5), 513.3 ($(M+2H)^{2+}$, calc.: 513.2); $C_{42}H_{68}N_{14}O_{14}S$: 1025.2 g·mol^{-1}; yield: 4.91 µmol (38%; A_{280} = 0.628, V = 1.0 ml).

2G-MAP: The peptide resin (6.3 µmol) was treated with a mixture of the MAP-auxiliary glycine conjugate **19** (6 eq, c = 0.4 M in DMF), PyBOP (6 eq) and DIPEA (18 eq) for 1 h. UPLC: t_R = 1.78 min (3-40% B in 2 min); ESI-MS: 1103.5 ($(M+H)^+$, calc.: 1103.5), 552.4 ($(M+2H)^{2+}$, calc.: 552.3); $C_{47}H_{70}N_{14}O_{15}S$: 1103.20 g·mol^{-1}; 1.4 µmol (22%; A_{280} = 0.69, V = 2 ml).

Peptide ligations

Auxiliary peptide **2X-Aux** (1 eq) and peptide thioester **1Z** (1-1.1 eq) were united (from stock solutions) and lyophilized. The lyophilized peptides were dissolved under argon atmosphere in degassed ligation buffer (20 mM TCEP, 100 mM Na_2HPO_4, pH = 7.5, 3 vol% thiophenol) to a final concentration of 2-5 mM (details see SI). The progress of the reaction was monitored by UPLC and HPLC-MS analysis. Peak areas were determined with due regard to the molar extinction coefficient of the peptides.

5GG-MAIP: UPLC: t_R = 2.7 min (3-40% B in 4 min); ESI-MS: 1663.6 ($(M+H)^+$, calc.: 1663.8), 832.5 ($(M+2H)^{2+}$, calc.: 832.4), 555.5 ($(M+3H)^{3+}$, calc.: 555.4); $C_{73}H_{110}N_{22}O_{21}S$: 1663.9 g·mol^{-1}. **5GG-MAP:** UPLC: t_R = 2.9 min (3-40% B in 4 min); ESI-MS: 1663.7 ($(M+H)^+$, calc.: 1663.8), 832.5 ($(M+2H)^{2+}$, calc.: 832.4), 555.5 ($(M+3H)^{3+}$, calc.: 555.4); $C_{73}H_{110}N_{22}O_{21}S$: 1663.9 g·mol^{-1}. **5AG-MAIP**: UPLC: t_R = 2.7 min (3-40% B in 4 min); ESI-MS: 1677.6 ($(M+H)^+$, calc.: 1677.8), 839.5 ($(M+2H)^{2+}$, calc.: 839.4); 560.2 ($(M+3H)^{3+}$, calc.: 559.9); $C_{74}H_{112}N_{22}O_{21}S$ 1677.9 g·mol^{-1}. **5AG-MAP**: UPLC: t_R = 3.0 min (3-40% B in 4 min); ESI-MS: 1677.5 ($(M+H)^+$, calc.: 1677.8), 839.5 ($(M+2H)^{2+}$, calc.: 839.4); 560.2 ($(M+3H)^{3+}$, calc.: 559.9); $C_{74}H_{112}N_{22}O_{21}S$ 1677.9 g·mol^{-1}. **5LG-MAIP**: UPLC: t_R = 7.1-7.3 min (3-40% B in 10 min); ESI-MS: 1720.8 ($(M+H)^+$, calc.: 1719.9), 860.5 ($(M+2H)^{2+}$, calc.: 860.4); 574.2 ($(M+3H)^{3+}$, calc.: 574.0); $C_{77}H_{118}N_{22}O_{21}S$ 1712.0 g·mol^{-1}. **5LG-MAP:** UPLC: t_R = 7.3-7.5 min (3-40% B in 10 min); ESI-MS: 1719.8 ($(M+H)^+$, calc.: 1719.9), 860.5 ($(M+2H)^{2+}$, calc.: 860.4); 574.2 ($(M+3H)^{3+}$, calc.: 574.0); $C_{77}H_{118}N_{22}O_{21}S$: 1712.0 g·mol^{-1}. **5LG-ME**: UPLC-MS: t_R = 2.9 min (3-40% B in 4 min); m/z = 793.4 ($(M+2H)^{2+}$, calc.: 792.9), 529.2 ($(M+3H)^{3+}$, calc.: 528.9); $C_{70}H_{113}N_{21}O_{19}S$: 1584.9 g·mol^{-1}. **5AN-MPE**: UPLC-MS: t_R = 3.8 min (3-30% B in 8 min); m/z = 800.5 ($(M+2H)^{2+}$, calc.: 800.4), 534.1 ($(M+3H)^{3+}$, calc.: 533.9); $C_{69}H_{110}N_{22}O_{20}S$: 1599.8 g·mol^{-1}. **5AN-ME**: UPLC-MS: t_R = 3.8 min (3-30% B in 6 min); m/z = 839.1 ($(M+2H)^{2+}$, calc.: 838.4), 559.7 ($(M+3H)^{3+}$, calc.: 559.2.6); $C_{75}H_{114}N_{22}O_{20}S$: 1675.9 g·mol^{-1}. **5LN-MPE:** UPLC-MS: t_R = 4.4 min (3-30% B in 8 min); m/z = 821.8 ($(M+2H)^{2+}$, calc.: 821.4), 548.2 ($(M+3H)^{3+}$, calc.: 548.0); $C_{72}H_{116}N_{22}O_{20}S$: 1641.9 g·mol^{-1}. **5LN-ME**: UPLC-MS: t_R = 3.8 min (3-30% B in 6 min); m/z = 860.2 ($(M+2H)^{2+}$, calc.: 859.4), 573.6 ($(M+3H)^{3+}$, calc.: 573.3); $C_{78}H_{120}N_{22}O_{20}S$: 1718.0 g·mol^{-1}. **5AN-MP** UPLC-MS: t_R = 2.6-2.8 min (3-40% B in 4 min); m/z = 807.8 ($(M+2H)^{2+}$, calc.: 807.4), 538.9 ($(M+3H)^{3+}$, calc.: 538.6); $C_{70}H_{112}N_{22}O_{20}S$: 1613.9 g·mol^{-1}.

Synthesis of ligation products

Auxiliary peptides and peptide thioesters were united from stock solutions in a 1:1 stoichiometry. The solution was lyophilized. The residue was dissolved under argon atmosphere in degassed buffer (20 mM TCEP, 100 mM Na_2HPO_4, pH = 7.5, 3 vol% thiophenol) to final concentrations as indicated. For isolation of ligation product, the mixture was separated by semi-preparative HPLC and the united product fractions were lyophilized.

5GG-MAIP: Synthesis scale: 2.0 µmol; concentration: 4 mM; reaction time: 2 h, semi-preparative HPLC: 3-45% B in 30 min; UPLC-MS: t_R = 2.87 min (3-40% B in 2 min); 832.7 ($(M+2H)^{2+}$, calc.: 832.4), 555.8 ($(M+3H)^{3+}$, calc.: 555.4); $C_{73}H_{110}N_{22}O_{21}S$: 1663.85 g·mol^{-1}; yield: 1.32 µmol (66%; A_{280} = 0.74, V = 2 ml).

5GG-MAP: Synthesis scale: 2.3 µmol; concentration: 3 mM; reaction time: 3 h; semi-preparative HPLC: 3-45% B in 30 min; UPLC: t_R = 3.12 min (3-40% B in 2 min); ESI-MS: 833.1 ($(M+2H)^{2+}$, calc.: 832.4), 555.5 ($(M+3H)^{3+}$, calc.: 555.4); $C_{73}H_{110}N_{22}O_{21}S$: 1663.85 g·mol^{-1}; yield: 1.12 µmol (56%; A_{280} = 0.23, V = 5.5 ml).

5GG-MPE: see reference[12]

5GG-ME: Synthesis scale: 1.0 µmol; concentration: 2 mM; reaction time: 1.5 h; semi-preparative HPLC: 3-40% B in 30 min; UPLC-MS: t_R = 2.5 min (3-40% B in 4 min); m/z = 765.1 ($(M+2H)^{2+}$, calc.: 764.9), 510.7 ($(M+3H)^{3+}$, calc.: 510.3); $C_{66}H_{105}N_{21}O_{19}S$: 1528.8 g·mol^{-1}; yield: 0.73 µmol (73%; A_{280} = 0.373, V = 0.5 ml).

5GG-MP: Synthesis scale: 1.0 µmol; concentration: 2 mM; reaction time: 4 h; semi-preparative HPLC: 3-40% B in 30 min; UPLC-MS: t_R = 2.6 min (3-40% B in 4 min); m/z = 772.3 ($(M+2H)^{2+}$, calc.: 771.9), 515.2 ($(M+3H)^{3+}$,

Accepted Manuscript

This article is protected by copyright. All rights reserved.

calc.: 514.9); $C_{67}H_{107}N_{21}O_{19}S$: 1542.8 g·mol^{-1}; yield: 0.63 µmol (63%; A_{280} = A_{280} = 0.323, V = 0.5 ml).

NMR-experiments

Cleavage from ^{13}C-labeled ligation product **5GG-MPE***: The lyophilized ^{13}C-labbeled ligation product **S31** (300 nmol, for synthesis see SI) was dissolved in of a solution of 200 mM TCEP and 800 mM morpholine in a mixture of (H_2O/D_2O, 95/5, v/v) to a final concentration of 0.5 mM. The mixture was agitated at rt. After 24 h the aqueous phase was extracted with deuterated chloroform (3 x 200 µL) and the organic phase was dried over $MgSO_4$. The aqueous and organic phase were analysed by UPLC-MS and NMR spectroscopy. Please consult the SI (Table S1) for relevant signals of the ^{1}H-^{13}C-HSQC analysis, detailed experimental setup and preparation of reference materials. Cleavage from unlabeled **5GG-MPE**: The lyophilized ligation product **5GG-MPE** (350 nmol, for synthesis see ref [12]) was dissolved in a solution of 200 mM TCEP and 800 mM morpholine in deuterated water to a final concentration of 0.5 mM. The mixture was agitated at 40 °C. After 16 h the aqueous phase was extracted with deuterated chloroform (3 x 200 µL) and the organic phase was dried over $MgSO_4$. The aqueous and organic phase were analysed by UPLC-MS and NMR spectroscopy (Fig. S20).

Auxiliary cleavage

5GG-MAIP→4GG: The lyophilized ligation product **55G-MAIP** (177 nmol) was dissolved in 354 µL of a solution of 20 mM TCEP and 140 mM morpholine in water. The mixture was agitated at 40 °C. After 22h, 93% of the mixture was submitted to purification by semi-preparative HPLC followed by lyophilization. **4GG**: UPLC: t_R = 1.15 min (3-40% B in 2 min); ESI-MS: 1469.6 ($(M+H)^+$, calc.: 1468.8), 735.0 ($(M+2H)^{2+}$, calc.: 734.9); 490.5 ($(M+3H)^{3+}$, calc.: 490.3); $C_{64}H_{101}N_{21}O_{19}$: 1468.61 g·mol^{-1}; yield: 82 nmol, 50%; A_{280} = 0.165, V_{Probe} = 500 µL.

5GG-MAP→4GG: The lyophilized ligation product **55G-MAP** (200 nmol) was dissolved in 400 µL of a solution of 20 mM TCEP and 140 mM morpholine in water. The mixture was agitated at 40 °C. After 16 h, 93% of the mixture was submitted to purification by semi-preparative HPLC followed by lyophilization. **4GG**: 83 nmol, 45% yield.

5GG-ME→4GG: The lyophilized ligation product **55G-ME** was dissolved in a solution of 100 mM TCEP and 400 mM morpholine in water to a final concentration of 0.5 mM. The mixture was agitated at 40 °C. The progress of the reaction was monitored by UPLC-MS analysis. **4GG**: UPLC-MS: t_R = 2.4 min (3-40% B in 4 min); m/z = 735.1 ($(M+2H)^{2+}$, calc.: 734.9), 490.5 ($(M+3H)^{3+}$, calc.: 490.3); $C_{64}H_{101}N_{21}O_{19}$: 1468.6 g·mol^{-1}.

5GG-MP→4GG: The lyophilized ligation product **S29GG** was dissolved in a solution of 100 mM TCEP and 400 mM morpholine in water to a final concentration of 0.5 mM. The mixture was agitated at 40 °C. The progress of the reaction was monitored by UPLC-MS analysis. **4GG**: UPLC-MS: t_R = 2.4 min (3-40% B in 4 min); m/z = 735.1 ($(M+2H)^{2+}$, calc.: 734.9), 490.5 ($(M+3H)^{3+}$, calc.: 490.3); $C_{64}H_{101}N_{21}O_{19}$: 1468.6 g·mol^{-1}.

Acknowledgements

The authors gratefully acknowledge the financial support by the Deutsche Forschungsgemeinschaft (Se819-15/1, SPP 1623).

Keywords: peptide ligation • radical reactions • protecting groups • peptide coupling • fragmentation

[1] a) P. Dawson, T. Muir, I. Clark-Lewis, S. Kent, *Science* **1994**, *266*, 776-779; b) S. Bondalapati, M. Jbara, A. Brik , *Nat. Commun.* **2016**, *8*, 407-418; c) S.B.H. Kent, *Bioorg. Med. Chem.* **2017**, *25*, 4926-4937.

[2] K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon, A. Brik, *Angew. Chem. Int. Ed.* **2011**, *50*, 6137-6141.

[3] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068-10073.

[4] L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891-5896.

[5] J. Offer, P. E. Dawson, *Org. Lett.* **2000**, *2*, 23-26.

[6] P. Botti, M. R. Carrasco, S. B. H. Kent, *Tetrahedron Lett.* **2001**, *42*, 1831-1833.

[7] a) T. Kawakami, K. Akaji, S. Aimoto, *Org. Lett.* **2001**, *3*, 1403-1405; b) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642-4646.

[8] a) T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2003**, *44*, 6059-6061; b) C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, *Biorg. Med. Chem.* **2004**, *12*, 2749-2757; c) C. Bello, S. Wang, L. Meng, K. W. Moremen, C. F. W. Becker, *Angew. Chem. Int. Ed.* **2015**, *54*, 7711-7715; d) C. Nadler, A. Nadler, C. Hansen, U. Diederichsen, *Eur. J. Org. Chem.* **2015**, *2015*, 3095-3102.

[9] L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526-533.

[10] Q. Wan, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-9252.

[11] a) D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064-10065; b) J. Chen, Q. Wan, Y. Yuan, J. Zhu, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2008**, *47*, 8521-8524; c) C. Haase, H. Rohde, O. Seitz, *Angew. Chem. Int. Ed.* **2008**, *47*, 6807-6810; d) K. S. Ajish Kumar, M. Haj-Yahya, D. Olschewski, H. A. Lashuel, A. Brik, *Angew. Chem. Int. Ed.* **2009**, *48*, 8090-8094; e) R. Yang, K. K. Pasunooti, F. Li, X.-W. Liu, C.-F. Liu, *J. Am. Chem. Soc.* **2009**, *131*, 13592-13593; f) J. Chen, P. Wang, J. Zhu, Q. Wan, S. J. Danishefsky, *Tetrahedron* **2010**, *66*, 2277-2283; g) Z. Harpaz, P. Siman, K. S. A. Kumar, A. Brik, *ChemBioChem* **2010**, *11*, 1232-1235; h) Z. Tan, S. Shang, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2010**, *49*, 9500-9503; i) S. Shang, Z. Tan, S. Dong, S. J. Danishefsky, *J. Am. Chem. Soc.* **2011**, *133*, 10784-10786; j) P. Siman, S. V. Karthikeyan, A. Brik, *Org. Lett.* **2012**, *14*, 1520-1523; k) L. R. Malins, K. M. Cergol, R. J. Payne, *ChemBioChem* **2013**, *14*, 559-563; l) R. E. Thompson, B. Chan, L. Radom, K. A. Jolliffe, R. J. Payne, *Angew. Chem. Int. Ed.* **2013**, *52*, 9723-9727; m) L. R. Malins, K. M. Cergol, R. J. Payne, *Chem. Sci.* **2014**, *5*, 260-266; n) L. R. Malins, R. J. Payne, *Aust. J. Chem.* **2015**, *68*, 521-537; o) J. Sayers, R. E. Thompson, K. J. Perry, L. R. Malins, R. J. Payne, *Org. Lett.* **2015**, *17*, 4902-4905.

[12] S. F. Loibl, Z. Harpaz, O. Seitz, *Angew. Chem. Int. Ed.* **2015**, *54*, 15055-15059.

[13] J. Vizzavona, F. Dick, T. Vorherr, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1963-1965.

[14] J. Offer, *Peptide Science* **2010**, *94*, 530-541.

[15] F. Ek, O. Axelsson, L.-G. Wistrand, T. Frejd, *J. Org. Chem.* **2002**, *67*, 6376-6381.

[16] E. L. A. COMPANY, *WO2005/54202 A1* **2005**, 37.

[17] C. Marinzi, S. J. Bark, J. Offer, P. E. Dawson, *Biorg. Med. Chem.* **2001**, *9*, 2323-2328.

[18] a) R. M. Beesley, C. K. Ingold, J. F. Thorpe, *Journal J. Chem. Soc. Trans.* **1915**, *107*, 1080-1106; b) M. E. Jung, G. Piizzi, *Chem. Rev.* **2005**, *105*, 1735-1766.

[19] a) M. V. Trivedi, J. S. Laurence, T. J. Siahaan, *Curr. Protein Pept. Sci.* **2009**, *10*, 614-625; b) G. Rajpal, P. Arvan, in *Handbook of Biologically Active Peptides (Second Edition)*, Academic Press, Boston, **2013**, pp. 1721-1729.

[20] C. Walling, R. Rabinowitz, *J. Am. Chem. Soc.* **1959**, *81*, 1243-1249.

Accepted Manuscript

This article is protected by copyright. All rights reserved.

[21] a) K. U. Ingold, *Acc. Chem. Res.* **1969**, *2*, 1-9; b) Y. Murakami, T. Oguchi, K. Hashimoto, Y. Nosaka, *J. Phys. Chem. A*, **2007**, *111*, 13200-13208; c) W. Sander, S. Roy, K. Bravo-Rodriguez, D. Grote, E. Sanchez-Garcia, *Chem. Eur. J.* **2014**, *20*, 12917.

[22] S. A. Buckler, *J. Am. Chem. Soc.* **1962**, *84*, 3093-3097.

[23] E. Furimsky, J. A. Howard, *J. Am. Chem. Soc.* **1973**, *95*, 369-374.

[24] a) P. Gray, A. Williams, *Chem. Rev.* **1959**, *59*, 239-328; b) J. K. Kochi, *J. Am. Chem. Soc.* **1962**, *84*, 1193-1197; c) F. D. Greene, M. L. Savitz, F. D. Osterholtz, H. H. Lau, W. N. Smith, P. M. Zanet, *J. Org. Chem.* **1963**, *28*, 55-64; d) C. Walling, A. Padwa, *J. Am. Chem. Soc.* **1963**, *85*, 1593-1597; e) J. D. Bacha, J. K. Kochi, *J. Org. Chem.* **1965**, *30*, 3272-3278.

[25] a) M. E. Snook, G. A. Hamilton, *J. Am. Chem. Soc.* **1974**, *96*, 860-869; b) H. A. Headlam, M. J. Davies, *Free Radical Biol. Med.* **2002**, *32*, 1171-1184.

[26] P. Renaud, L. Giraud, *Synthesis* **1996**, *1996*, 913-926.

[27] S. F. Loibl, Z. Harpaz, R. Zitterbart, O. Seitz, *Chem. Sci.* **2016**, *7*, 6753-6759.

WILEY-VCH

Accepted Manuscript

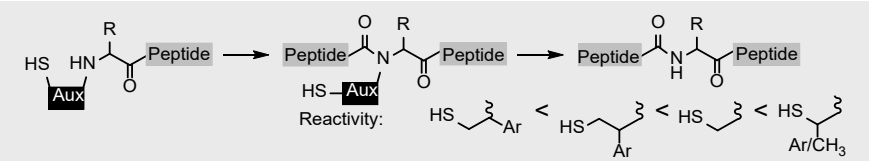
This article is protected by copyright. All rights reserved.

10.1002/chem.201705927

WILEY-VCH

Entry for the Table of Contents (Please choose one layout)

FULL PAPER



Efficient peptide ligation auxiliaries have a β-mercaptoethyl scaffold and avoid substituent in α–position to the amine. Surprisingly, β-substituents accelerate ligation reactions. All presented auxiliaries can be removed under mild-basic conditions. NMR analysis of a ^{13}C-labelled auxiliary indicated that auxiliary cleavage most likely proceeds through a radical oxidation-fragmentation cascade.

*Simon F. Loibl, Andre Dallmann, Kathleen Hennig, Carmen Juds and Oliver Seitz**

Page No. – Page No.

Features of Auxiliaries that Enable Native Chemical Ligation beyond Glycine and Cleavage via Radical Fragmentation

WILEY-VCH

Accepted Manuscript

This article is protected by copyright. All rights reserved.