



Doctoral Thesis

## Chemical protein synthesis by $\alpha$ -ketoacid-hydroxylamine ligation

**Author(s):**

Wucherpfennig, Thomas G.

**Publication Date:**

2016

**Permanent Link:**

<https://doi.org/10.3929/ethz-a-010666498> →

**Rights / License:**

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

DISS. ETH Nr. 23525

**CHEMICAL PROTEIN SYNTHESIS BY  
 $\alpha$ -KETOACID-HYDROXYLAMINE LIGATION**

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH ZURICH)

Presented by

**Thomas Georg Wucherpfennig**

Master of Science in Chemistry, ETH Zurich,  
Born on 07.12.1986  
Citizen of Germany

Accepted on the recommendation of  
Prof. Dr. Jeffrey W. Bode, examiner  
Prof. Dr. Peter Kast, co-examiner  
Prof. Dr. Pablo Rivera-Fuentes, co-examiner

2016

## Abstract

Proteins are ubiquitous biomacromolecules in nature and fulfill diverse, important functions. Innumerable studies in life sciences require homogenous peptides and proteins of defined structure and composition. In many cases, the chemical synthesis of peptides and proteins offers unique advantages and possibilities compared to isolation from natural sources or recombinant protein production.

Direct linear chemical synthesis techniques such as solid-phase peptide synthesis routinely provide access to peptides and smaller proteins consisting of about 50–60 amino acid residues. However, most bioactive proteins are significantly larger, and alternative, convergent approaches such as the well-established native chemical ligation (NCL) are required. Despite its broad application, this method has several disadvantages, including the complicated synthesis of peptide thioesters and the requirement for a cysteine residue at the ligation site. Recently, several improvements have partially overcome these limitations, but NCL cannot yet provide a general solution to all protein targets and alternative approaches are desirable.

The Bode group developed the  $\alpha$ -ketoacid-hydroxylamine (KAHA) ligation, a reaction using conditions and functional groups complementary to NCL. The key element of the KAHA ligation is the decarboxylative condensation of an  $\alpha$ -ketoacid with a hydroxylamine to form an amide bond. Since its initial discovery in 2006, this reaction has undergone constant innovation, such as the development of 5-oxaproline, the best hydroxylamine derivative to date in terms of stability and reactivity for peptide and protein synthesis. These advances render the KAHA ligation a very flexible and powerful synthetic method as demonstrated in the synthesis of several smaller proteins.

This dissertation focuses on both technical improvements to the KAHA ligation and its application in the chemical synthesis of larger biologically interesting proteins. During this work, it was unexpectedly found that the primary products of the KAHA ligation with 5-oxaproline are esters rather than the anticipated amides. Mild conditions to rearrange these depsipeptides and depsiproteins to the natural amides were developed. The formation of esters provided insights into a possible mechanism of the KAHA ligation with 5-oxaproline.

A major limitation hampering the general applicability of the KAHA ligation – the cumbersome and limited synthesis of peptide  $\alpha$ -ketoacids – was addressed. A novel protecting group for enantiopure  $\alpha$ -ketoacids was developed and implemented in monomers for Fmoc-SPPS. This protocol delivers C-terminal peptide  $\alpha$ -ketoacids directly upon resin cleavage without additional manipulation steps, and allows the inclusion of all canonical amino acids including cysteine, methionine and tryptophan. Further studies led to an orthogonal protecting group for  $\alpha$ -ketoacids, which is stable to the acidic resin cleavage

conditions and can be selectively removed on demand. Both technologies provide a general, streamlined and scalable synthesis of peptide  $\alpha$ -ketoacids and enable more flexible protein synthesis strategies.

KAHA ligations and the newly developed  $\alpha$ -ketoacid protecting groups were utilized to synthesize milligram quantities of pure SUMO3, an important ubiquitin-like modifier protein involved in the regulation of numerous cellular processes. Despite containing two homoserine mutations, the synthetic SUMO3 protein retained full bioactivity and was readily conjugated to target proteins by the enzymatic SUMOylation machinery.

Derived from this synthetic protein, a chemical probe for cellular protein SUMOylation was developed. The probe overcomes the highly dynamic and reversible nature of protein SUMOylation by incorporating a C-terminal mutation retarding the deconjugation by endogenous SUMO-specific proteases. A biotin-tag and a fluorescence label on the probe allowed facile detection and affinity purification of SUMO–protein conjugates. The probe was applied in *in vitro* SUMOylation reactions using crude cell lysates as substrates and SUMOylated proteins were readily enriched and identified using MS-based proteomic experiments. Projecting forward, this technology could provide a versatile platform for identifying and profiling cellular protein SUMOylation under well-defined conditions such as cellular stress or pathogen infections to obtain further insights into the effects of SUMOylation on protein function.

The KAHA ligation technology was utilized in the synthesis of irisin, a myokine that is speculated to play an important role in modulating human fat metabolism. Milligram quantities of unlabeled and fluorescence labeled irisin were synthesized and used in cell binding studies, indicating the presence of a specific irisin-binding receptor on certain adipocytes. In future work, chemically synthesized and further modified irisin could be a valuable tool for the identification of the irisin receptor, shining light on the role of this controversial hormone in human physiology.

## Zusammenfassung

Proteine sind in der Natur allgegenwärtige Biomakromoleküle und erfüllen unterschiedliche, wichtige Funktionen. Zahlreiche Studien in den Lebenswissenschaften benötigen homogene Peptide und Proteine mit definierter Struktur und Zusammensetzung. In vielen Fällen bietet die chemische Synthese der Peptide und Proteine einzigartige Vorteile und Möglichkeiten verglichen mit der Isolierung aus natürlichen Quellen oder rekombinanter Proteinproduktion.

Direkte, lineare chemische Synthesetechniken wie etwa die Festphasenpeptidsynthese (engl. *solid phase peptide synthesis*, SPPS) ermöglichen den Zugang zu Peptiden und kleineren Proteinen mit einer Größe von etwa 50–60 Aminosäuren. Die meisten bioaktiven Proteine sind allerdings deutlich größer und alternative, konvergente Ansätze wie etwa die gängige native chemische Ligation (NCL) werden benötigt. Trotz ihrer breiten Anwendung hat diese Methode einige Nachteile wie etwa die komplizierte Synthese der benötigten Peptidthioester und die Erfordernis eines Cysteins an der Ligationsstelle. In jüngster Zeit haben viele Verbesserungen diese Limitierungen zum großen Teil überwunden, trotzdem kann die NCL jedoch noch keine allgemeine Lösung für alle Zielproteine bieten und alternative Methoden sind erstrebenswert.

Die Forschungsgruppe Bode entwickelte die  $\alpha$ -Ketosäure-Hydroxylamin (KAHA)-Ligation, eine Reaktion, die bezüglich der Reaktionsbedingungen und der genutzten funktionellen Gruppen komplementär zur NCL ist. Das Schlüsselement der KAHA-Ligation ist die decarboxylative Kondensation einer  $\alpha$ -Ketosäure mit einem Hydroxylamin unter Knüpfung einer Amidbindung. Seit ihrer Entdeckung 2006 wurde die Reaktion kontinuierlich verbessert, wie etwa die Entwicklung von 5-Oxaprolin zeigt, dem bisher besten Hydroxylaminderivat bezüglich Stabilität und Reaktivität in der Peptid- und Proteinsynthese. Diese Fortschritte machen die KAHA-Ligation zu einer sehr flexiblen und mächtigen Synthesemethode, wie die Synthesen einiger kleinerer Proteine zeigten.

Diese Dissertation beschreibt sowohl technische Verbesserungen der KAHA-Ligation, als auch ihre Anwendung in der Synthese grösserer, biologisch interessanter Proteine. Im Verlauf dieser Arbeit wurde gefunden, dass die Primärprodukte der KAHA-Ligation mit 5-Oxaprolin nicht die erwarteten Amide, sondern Ester sind. Es wurden milde Bedingungen entwickelt, um diese Depsipeptide und Depsiproteine zu den natürlichen Amidinen umzulagern. Die Esterbildung gab Einblicke in einen möglichen Mechanismus der KAHA-Ligation mit 5-Oxaprolin.

Eine Hauptlimitierung, welche die breite Anwendbarkeit der KAHA-Ligation beeinträchtigt – die umständliche und eingeschränkte Synthese der benötigten Peptid- $\alpha$ -ketosäuren – wurde behoben. Eine neue Schutzgruppe für enantiomerenreine

$\alpha$ -Ketosäuren wurde entwickelt und in Monomere für die Fmoc-Festphasenpeptidsynthese implementiert. Diese Methode ergibt C-terminale Peptid- $\alpha$ -Ketosäuren direkt nach der Abspaltung von der Festphase ohne zusätzliche Manipulationsschritte und ermöglicht den Einbau aller kanonischen Aminosäuren, inklusive Cystein, Methionin und Tryptophan. Weitere Studien führten zur Entwicklung einer orthogonalen Schutzgruppe für  $\alpha$ -Ketosäuren, die unter den sauren Abspaltungsbedingungen stabil ist und selektiv in Lösung entfernt werden kann. Beide Technolgien ermöglichen eine generelle, vereinfachte und skalierbare Synthese von Peptid- $\alpha$ -Ketosäuren und ermöglichen flexiblere Strategien zur Proteinsynthese.

Die KAHA-Ligation und die neuentwickelten  $\alpha$ -Ketosäure-Schutzgruppen wurden verwendet, um reines SUMO3 – ein wichtiges ubiquitin-artiges Modifikationsprotein, dass an der Regulation zahlreicher zellulärer Prozesse beteiligt ist – auf Milligrammskala zu synthetisieren. Obwohl das synthetische SUMO3-Protein zwei Homoserinmutationen enthielt, bewahrte es seine volle biologische Aktivität und wurde durch die enzymatische SUMOylierungsmaschinerie an Zielproteine konjugiert.

Aufbauend auf diesem synthetischen Protein wurde eine chemische Sonde zur Untersuchung der SUMOylierung von zellulären Proteinen entwickelt. Diese Sonde überwindet den hochdynamischen und reversiblen Charakter der Protein-SUMOylierung durch eine Mutation, welche die Dekonjugation durch endogene, SUMO-spezifische Proteasen verhindert. Ein Biotin-Tag und ein Fluoreszenzfarbstoff ermöglichen die einfache Detektion und Anreicherung von SUMO-Protein-Konjugaten. Die Sonde wurde in *in vitro*-Reaktionen mit Zelllysaten als Substraten verwendet und die dabei erhaltenen SUMOylierten Proteine konnten leicht angereichert und mittels MS-basierten Proteomics-Experimenten identifiziert werden. Für zukünftige Projekte könnte diese Technologie eine vielfältige Plattform zur Identifikation und Profilierung der SUMOylierung zellulären Proteine unter definierten Bedingungen wie etwa zellulärem Stress oder Infektionen bieten, um dadurch den Einfluss der SUMOylierung auf die Funktion der Proteine besser verstehen zu können.

Die KAHA-Ligation wurde weiterhin zur Synthese von Irisin verwendet, einem Myokin, das vermutlich eine wichtige Rolle in der Regulation des menschlichen Fettmetabolismus spielt. Milligram-Mengen von ungelabeltem und Fluoreszenz-gelabeltem Irisin wurden synthetisiert und in einer Zellbindungsstudie verwendet, dessen Ergebnisse auf die Präsenz von spezifischen Irisin-Rezeptoren in bestimmten Adipozyten hindeuten. Für zukünftige Arbeiten könnte chemisch synthetisiertes und weiter modifiziertes Irisin ein nützliches Werkzeug zur Identifizierung des Irisinrezeptors sein, um dadurch die Rolle dieses kontroversen Hormons in der menschlichen Physiologie besser zu verstehen.