This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Twisted amide electrophiles enable cyclic peptide sequencing

Serge Zaretsky, Vishal Rai, Gerald Gish, Matthew W. Forbes, Michael Kofler, Joy C. Y. Yu, Tony Pawson, and Andrei K. Yudin*

There is an ever-increasing interest in synthetic methods that not only enable peptide macrocyclization, but also facilitate downstream application of the synthesized molecules. We have found that aziridine amides are stereoelectronically attenuated in a macrocyclic environment such that non-specific interactions with biological nucleophiles are reduced or even shut down. The electrophilic reactivity, revealed at high pH, enables peptide sequencing by mass spectrometry, which will further broaden the utility of aziridine amide-containing libraries of macrocycles.

Macrocycles continue to attract attention on behalf of the chemistry community. Their unique structural features have found many uses in areas such as medicinal chemistry, coordination chemistry, and nanotechnology. Cyclic peptides, as a subset of macrocycles, are noteworthy for their biological applications, specifically in the field of protein-protein interactions. Our group has been interested in exploring the conformational consequences of placing strained amides within the structures of cyclic peptides. Herein we demonstrate the surprising compatibility of aziridine amides with biological nucleophiles, and also showcase the utility of preserving the latent electrophile in a cyclic peptide sequencing application.

Torsional strain provides an opportunity to stereoelectronically “challenge” one of nature’s most ubiquitous structural units – the amide bond. The electrophilic reactivity of the amide carbon is modulated through delocalization of the nitrogen lone pair of electrons, but in twisted amides this conjugation is significantly diminished. Seminal studies in the area of twisted amides have focused on using anti-Bredt tricyclic systems that enforce the nitrogen lone pair out of conjugation with the acyl substituent. Kirby and co-workers have shown that the nitrogen center in 1-aza-2-adaman-tanone displays nucleophilic amine reactivity, while the carbonyl group acts as a ketone-like electrophile. A study of twisted lactams by the Aubé group suggested that a minimum 50° distortion of the amide was required to observe nucleophilic amine reactivity. By resorting to the Schmidt-Aubé reaction, Stoltz and co-workers have prepared the long-sought 2-quinuclidone as a HBF₄ salt. While these examples highlight the unusual properties of twisted amides, they do not suggest obvious applications.

Figure 1. (a) Aziridine amide-bearing cyclic peptides made by peptide cyclization with an Azy residue (1) or by aziridine aldehyde-mediated macrocyclization (2); (b) Aziridine amide bis-electrophile reactivity by SN₂ or transacylation.

We have developed several methods for generating macrocycles with embedded aziridine amides (Figure 1a) and demonstrated downstream reactivity of these compounds with numerous nucleophiles. Aziridine amides are ambident electrophiles, meaning that they can partake in both SN₂ and transacylation reactions (Figure 1b). This dual electrophilicity is caused by the strained three-membered aziridine ring and the electrophilic carbonyl group, which exhibits diminished delocalization of the nitrogen lone pair of electrons compared to regular amides. As reactive intermediates, aziridine amides have been primarily used in SN₂ reactions with soft nucleophiles. Meanwhile, hard
nucleophiles are known to trigger the transacylation reaction, a process that has gained comparatively little attention.²⁵

We now report that stereoelectronic attenuation of the aziridine amide can prevent non-specific reactivity with biological nucleophiles. Additionally, we report that preserving the electrophilic carbonyl group of the aziridine amide enables cyclic peptide sequencing by chemoselective linearization and tandem mass spectrometry.

![Figure 2](image-url)

**Figure 2.** Comparison of ketone-like character in $^{13}$C of larger cycles (2) vs. 6-membered rings (3). 2a & 2b: $R = ^{1}Bu, R^1 = H, R^2 = CH_2CH(CH_3)2, R^3 = H, R^4 = CH_2C_6H_5, R^* = H$ (2a); $R^* = CH_2CH(CH_3)2 (2b)$.

In piperazinones (3), the nitrogen lone pair cannot effectively conjugate with the carbonyl π* orbital, which is evident in the $^{13}$C chemical shift of 183-187 ppm (CDCl$_3$). Within larger rings, we noted a slightly more upfield chemical shift of the aziridine amide carbonyl group (Figure 2). Within 18-membered rings, the aziridine carbonyl shift appears at 183-184 ppm (CDCl$_3$), which is closer to the chemical shift of aziridine amides in linear systems (181-183 ppm, CDCl$_3$).²⁶ While the differences are small, the data still suggests a tendency for 6-membered rings to possess more ketone-like character in the aziridine amide. Our hypothesis was that larger rings would engender aziridine amides with greater stability towards hydrolysis and reduced non-specific reactivity through the transacylation pathway.²⁷ At the same time, we were interested in preventing non-specific $S_{n}2$ reactivity, which would be promoted by the increased N-to-C conjugation.

As a testing ground, we decided to evaluate a series of probes 3a, 2a, and 2b (Figure 2) in a biological setting. 3a was made from Pro and cyclic peptides 2a and 2b were prepared from H-Pro-Gly-Leu-Gly-Phe-0tt and the respective aziridine aldehyde reagents using our previously reported methods.²⁸ When 2a and 2b were evaluated for their reactivity with glutathione (GSH), only the unsubstituted aziridine analog 2a was found to react (Figure 3a). The thiol attack at the aziridine was confirmed by collision-induced dissociation/tandem mass spectrometry (CID/MS$^2$). Owing to its unhindered aziridine and non-attenuated acyl reactivity, probe 3a reacted with GSH by both the $S_{n}2$ and transamidation pathways.

We next investigated any reactivity against a broader biological context, a protein. For our goals, we sought a protein that: a) contained a number of surface-exposed nucleophilic residues and b) was capable of enzymatic activity. The Fes SH2-kinase was chosen as a model.²¹ The catalytic SH2-kinase domain of the protein tyrosine kinase Fes is a complex nucleophile containing eight nucleophilic residues and b) protein in a cysteine-independent manner.

![Figure 3](image-url)

**Figure 3.** (a) Increasing amidicity of the strained aziridine amide bond in 2a prevents the transacylation pathway seen for 3a with GSH; (b) Exposing FES SH2-kinase to cyclic aziridine amides leads to no reactivity with 2a, while 3a reacts with the protein in a cysteine-independent manner.

With biological compatibility established for the 18-membered ring macrocycles, we sought to further explore the “hard” carbon center of the acyl aziridine. When we subjected 2b to basic aqueous media (0.8 M NaOH), clean and site-selective hydrolysis of the aziridine amide took place to yield compound 4 (Figure 4). This hydrolysis product was composed of the linear peptide segment attached to the cyclization-linker at the N-terminus. Site-selective linearization has been previously applied to cyclic peptides,²⁵ but for cyclized peptides the technique is limited to ring scission by cyanogen bromide,²⁶ which is plagued by a number of side reactions.²⁷
Figure 4. Base-mediated hydrolysis of 2b to linearized peptide 4. 4 is fragmented by CID/MS\(^2\) to prominent \(a\)- and \(b\)- ions. Underlined ions correspond to characteristic fragmentation ions of the linker.

Conversely, regiospecific linearization at the aziridine amide is enabling because it allows for sequencing of macrocyclic peptides using CID/MS\(^2\). Compound 4 fragmented readily with CID/MS\(^2\) to characteristic \(a\)- and \(b\)- ions (Figure 3).\(^{28}\) BioAnalyst\(^{29}\) was used to annotate the spectra for sequence information using the most common 20 natural amino acids and a modified proline residue incorporating the cyclization linker. The highest scoring sequence tags corresponded to partial internal sequences and the highest scoring full sequence was the correct peptide sequence. The ability to site-specifically cleave and sequence aziridine-containing macrocycles is significant in that it allows for post-assay identification of a peptide. Hydrolysis with \(^{18}\)O-labelled water may provide further assistance in deciphering the \(y\)-ions.\(^{30}\)

The selectively linearized macrocyclic peptides do not lead to scrambled sequences and the cyclization tag serves as a reference point for downstream peptide sequencing. In contrast, conventional cyclic peptides can fragment at multiple positions, complicating interpretation of the CID/MS\(^2\) spectra.\(^{31}\) The cyclic topology also rules out Edman degradation for sequence elucidation and instead, elaborate protocols must be engineered to overcome these limitations.\(^{32}\) Accordingly, the challenge of de novo peptide sequencing has limited biological screening techniques, such as affinity selection-mass spectrometry, to linear peptides and not their cyclic counterparts.\(^{33}\) The aziridine amide enabled linearization technique should provide the tools needed to screen and sequence libraries of biologically stable cyclic peptide probes.\(^{34}\)

In summary, we have evaluated the reactivity of aziridine amides embedded within peptidic rings and demonstrated that the 18-membered ring macrocycles are stereoelectronically attenuated such that non-specific reactivity with biological nucleophiles such as glutathione or Fes SH2-kinase is reduced, or altogether absent. While aziridines are typically known for irreversible reactivity with biomolecules, our study opens the door to novel applications of this class of molecules. The key benefit of retaining a strained amide is the ability selectively hydrolyze it under harsh basic conditions. The chemoselective linearization of aziridine amide-containing macrocycles enabled routine sequencing by CID/MS\(^2\) without the need for multistage mass spectrometry or complex algorithms.\(^{35}\) Because the undesired \(S_N2\) chemistry has been shut down, the macrocycles are amenable to site-specific modification (e.g. enzymes) followed by deconvolution with site-selective linearization and sequencing. These findings will lead to novel biological screening applications of cyclic peptides with embedded aziridine amides.

Funding for this work was provided by the Natural Sciences and Engineering Research Council (NSERC). S.Z. thanks the Ontario Graduate Scholarship (OGS) program for financial support.

Notes and references
\(^a\) Dr. S. Zaretsky, Dr. V. Rai, Dr. M. W. Forbes, Ms. J. C. Y. Yu, Prof. Dr. A. K. Yudin. Davenport Research Laboratories, Department of
Organic & Biomolecular Chemistry: Communication

5. (a) S. E. Northfield, C. K. Wang, C. I. Schroeder, T. Durek, M. W. Kan, J. Dr. G. Gish, Dr. M. Kofler, Prof. Dr. T. Pawson, Lu nenfeld Tanenbaum Research Institute, Mount Sinai Hospital, 600 University Ave., Toronto, ON, MSG 1X5, (Canada).