Rapid Cross-Metathesis for Reversible Protein Modifications via Chemical Access to Se-Allyl-selenocysteine in Proteins

Yuya A. Lin, Omar Boutureira,† Lukas Lercher, Bhaskar Bhushan, Robert S. Paton, and Benjamin G. Davis*

Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Mansfield Road, Oxford OX1 3TA, U.K.

Supporting Information

ABSTRACT: Cross-metathesis (CM) has recently emerged as a viable strategy for protein modification. Here, efficient protein CM has been demonstrated through biomimetic chemical access to Se-allyl-selenocysteine (Seac), a metathesis-reactive amino acid substrate, via dehydroalanine. On-protein reaction kinetics reveal a rapid reaction with rate constants of Seac-mediated-CM comparable or superior to off-protein rates of many current bioconjugations. This use of Se-relayed Seac CM on proteins has now enabled reactions with substrates (allyl GlcNAc, N-allyl acetamide) that were previously not possible for the corresponding sulfur analogue. This CM strategy was applied to histone proteins to install a mimic of acetylated lysine (KAc, an epigenetic marker). The resulting synthetic H3 was successfully recognized by antibody that binds natural H3-K9Ac. Moreover, Cope-type selenoxide elimination allowed this putative marker (and function) to be chemically expunged, regenerating an H3 that can be rewritten to complete a chemically enabled "write (CM)−erase (ox)−rewrite (CM)" cycle.

Developing strategies for site-specific protein modifications is an ongoing challenge in chemical biology.1−3 For over a decade, allylic alcohols have been identified as reactive in olefin metathesis.4,5 However, only in recent years has a generalized allylic heteroatom effect begun to emerge in cross-metathesis (CM).5 A mechanism involving pre-coordination to the metal center of catalyst 1 has been proposed to explain the observed reactivity of privileged substrates, such as amino acid S-allylcysteine (Sac, Scheme 1).6 This "heteroatom-relay" has enabled CM to be applied as a bioconjugation technique in selective protein modifications,7−11 and so greatly increased the benchmark for complexity in the substrate scope. Such putative beneficial effects of allylic chalcogens suggest a likely increase in the reactivity of higher chalcogens as Lewis bases that may be more preferred, softer ligands for ruthenium complexes. Indeed, simple allyl selenides are reactive substrates in aqueous CM.11 We show here that installation into peptides and proteins of the simplest amino acid residue containing an allyl selenide moiety, the unnatural amino acid Se-allyl-selenocysteine (Seac), enables rapid and efficient CM.

Possible approaches to incorporating Seac into proteins were considered that relied on installation of the Se heteroatom either co- (either using Seac directly or from a precursor such as selenocysteine (Sec)) or post-translationally. Direct co-

© 2013 American Chemical Society

Published: July 26, 2013

dx.doi.org/10.1021/ja403191g | J. Am. Chem. Soc. 2013, 135, 12156−12159

Supporting Information

Scheme 1. Allyl Chalcogen-Assisted Cross-Metathesis

where X = S or Se and R = H for Sec or Seac
Scheme 2. Biomimetic Synthesis of 5-Se-Allyl-selenocysteine via Conjugate Addition of Allyl Selenolate to Dehydroalanine (Table S1)

![Scheme 2. Biomimetic Synthesis of 5-Se-Allyl-selenocysteine via Conjugate Addition of Allyl Selenolate to Dehydroalanine (Table S1)](image)

53% of the desired Seac product 4 (Table S1). Allyl selenocyanate (2c), a potentially more stable,24 alternative allyl selenolate source, was investigated next. Please, due to closer pre-reduction, the desired addition product 4 was obtained with an improved yield of 86% after 1 h at room temperature (Scheme 2a). Notably, although allyl selenolate is sensitive to acidic conditions (allylselenole quickly decomposes to 2b in the absence of radical inhibitor23), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

With promising results observed on amino acid models, we next explored installation of Seac via this biomimetic strategy in two differing protein structural motifs/folds: three-layer α/β-β/α-Rossman-fold protein subtilisin from Bacillus lentus (SBL) and all-β-helix protein 275–61Dha (SBL) and αβ-61Dha ( termed Npβ).25 respectively. Accordingly, Dha substrate proteins SBL-156Dha (5) and Npβ-61Dha (7) were synthesized from corresponding cysteine mutants following reported bis-alkylative elimination procedures.21 Use of selenolate derived from 2a or 2b failed on proteins and led to only a mixture of products. However, allyl selenolate solution generated from 2c allowed successful biomimetic conjugate addition (Scheme 2b); LC-MS analysis revealed >95% conversion of SBL-156Dha (5) to the desired product SBL-156Seac (6a) after 1 h at room temperature. Even addition to the more sterically demanding26 61 site in the absence of radical inhibitor23), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

With promising results observed on amino acid models, we next explored installation of Seac via this biomimetic strategy in two differing protein structural motifs/folds: three-layer α/β-β/α-Rossman-fold protein subtilisin from Bacillus lentus (SBL) and all-β-helix protein 275–61Dha (SBL) and αβ-61Dha ( termed Npβ).25 respectively. Accordingly, Dha substrate proteins SBL-156Dha (5) and Npβ-61Dha (7) were synthesized from corresponding cysteine mutants following reported bis-alkylative elimination procedures.21 Use of selenolate derived from 2a or 2b failed on proteins and led to only a mixture of products. However, allyl selenolate solution generated from 2c allowed successful biomimetic conjugate addition (Scheme 2b); LC-MS analysis revealed >95% conversion of SBL-156Dha (5) to the desired product SBL-156Seac (6a) after 1 h at room temperature. Even addition to the more sterically demanding26 61 site in the absence of radical inhibitor23), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

With promising results observed on amino acid models, we next explored installation of Seac via this biomimetic strategy in two differing protein structural motifs/folds: three-layer α/β-β/α-Rossman-fold protein subtilisin from Bacillus lentus (SBL) and all-β-helix protein 275–61Dha (SBL) and αβ-61Dha ( termed Npβ).25 respectively. Accordingly, Dha substrate proteins SBL-156Dha (5) and Npβ-61Dha (7) were synthesized from corresponding cysteine mutants following reported bis-alkylative elimination procedures.21 Use of selenolate derived from 2a or 2b failed on proteins and led to only a mixture of products. However, allyl selenolate solution generated from 2c allowed successful biomimetic conjugate addition (Scheme 2b); LC-MS analysis revealed >95% conversion of SBL-156Dha (5) to the desired product SBL-156Seac (6a) after 1 h at room temperature. Even addition to the more sterically demanding26 61 site in the absence of radical inhibitor23), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

With promising results observed on amino acid models, we next explored installation of Seac via this biomimetic strategy in two differing protein structural motifs/folds: three-layer α/β-β/α-Rossman-fold protein subtilisin from Bacillus lentus (SBL) and all-β-helix protein 275–61Dha (SBL) and αβ-61Dha ( termed Npβ).25 respectively. Accordingly, Dha substrate proteins SBL-156Dha (5) and Npβ-61Dha (7) were synthesized from corresponding cysteine mutants following reported bis-alkylative elimination procedures.21 Use of selenolate derived from 2a or 2b failed on proteins and led to only a mixture of products. However, allyl selenolate solution generated from 2c allowed successful biomimetic conjugate addition (Scheme 2b); LC-MS analysis revealed >95% conversion of SBL-156Dha (5) to the desired product SBL-156Seac (6a) after 1 h at room temperature. Even addition to the more sterically demanding26 61 site in the absence of radical inhibitor23), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

With promising results observed on amino acid models, we next explored installation of Seac via this biomimetic strategy in two differing protein structural motifs/folds: three-layer α/β-β/α-Rossman-fold protein subtilisin from Bacillus lentus (SBL) and all-β-helix protein 275–61Dha (SBL) and αβ-61Dha ( termed Npβ).25 respectively. Accordingly, Dha substrate proteins SBL-156Dha (5) and Npβ-61Dha (7) were synthesized from corresponding cysteine mutants following reported bis-alkylative elimination procedures.21 Use of selenolate derived from 2a or 2b failed on proteins and led to only a mixture of products. However, allyl selenolate solution generated from 2c allowed successful biomimetic conjugate addition (Scheme 2b); LC-MS analysis revealed >95% conversion of SBL-156Dha (5) to the desired product SBL-156Seac (6a) after 1 h at room temperature. Even addition to the more sterically demanding26 61 site in the absence of radical inhibitor23), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

With promising results observed on amino acid models, we next explored installation of Seac via this biomimetic strategy in two differing protein structural motifs/folds: three-layer α/β-β/α-Rossman-fold protein subtilisin from Bacillus lentus (SBL) and all-β-helix protein 275–61Dha (SBL) and αβ-61Dha ( termed Npβ).25 respectively. Accordingly, Dha substrate proteins SBL-156Dha (5) and Npβ-61Dha (7) were synthesized from corresponding cysteine mutants following reported bis-alkylative elimination procedures.21 Use of selenolate derived from 2a or 2b failed on proteins and led to only a mixture of products. However, allyl selenolate solution generated from 2c allowed successful biomimetic conjugate addition (Scheme 2b); LC-MS analysis revealed >95% conversion of SBL-156Dha (5) to the desired product SBL-156Seac (6a) after 1 h at room temperature. Even addition to the more sterically demanding26 61 site in the absence of radical inhibitor23), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

azide–alkyne cycloaddition,31,32 and inverse electron-demand Diels–Alder cycloaddition,33–36 which have typically been studied in small-molecule models (“off-protein”). Under pseudo-first-order conditions with respect to the protein, starting material and product were monitored by LC-MS at several time points during reactions (Figure 1a).37 Data analyzed using nonlinear, single-exponential regression gave pseudo-first-order rate constants (Figure 1b).

Determination of pseudo-first-order rate constants at various allyl alcohol concentrations and analysis of the results using linear regression revealed second-order rate constants for Seac-mediated CM. The on-protein second-order rate constant for the CM of allyl alcohol with Seac-protein was determined to be $0.01 \pm 0.004 \text{ M}^{-1} \text{s}^{-1}$ (Figure 1c,d), whereas the value for on-protein Sac-mediated CM in an identical protein context (site 156) was some 10-fold lower ($0.31 \pm 0.0015 \text{ M}^{-1} \text{s}^{-1}$) under the same conditions (Figures S21 and S22). Interestingly, Figure 1d indicated an initially linear increase in observed rate with respect to allyl alcohol concentration and gradually plateaued at concentrations $>20 \text{ mM}$; this suggests that the catalyst plays a critical role under these conditions, presumably becoming “saturated” at high concentrations of allyl alcohol, leading to a shift in rate-limiting step (see also SI section 2.10 for off-protein catalyst dependency). Preliminary computations (using methyl allyl chalcogenides, see SI section 3) suggest a reaction profile in which the Ru-carbene intermediate after “Se-relay” is unusually stable and enhanced in stability over “S-relay”; this intermediate is also accessed through a lowered transition state.

When comparing the second-order rate constants with those from widely used bioconjugation techniques in chemical biology,
the rate of on-protein Seac-assisted CM was found to be ~150-fold faster than that of off-protein Staudinger ligation (k = 0.002 M⁻¹ s⁻¹).⁴⁸ and comparable to faster variants of off-protein strain-promoted azide–alkyne cycloaddition (k = 0.002–0.9 M⁻¹ s⁻¹)⁴⁹–⁵¹ and certain tetrazine-based cycloadditions (k > 0.15 M⁻¹ s⁻¹).²⁷ Notably, however, recent select examples of catalyzed oxime ligation⁴² and tetrazine-based reactions are still more rapid (Figure 2).²⁷,²⁸,⁴³,⁴⁴ It seems likely that local and function-dependent activity may be enhanced, consistent with CM-assisted protein modiﬁcation at position 9 suggested a strategy for chemical recapitulation of enzyme-mediated reaction conditions could not be achieved, we extrapolate (see SI section 2.10) the rate of on-protein Seac-assisted CM to be at least comparable and may be enhanced, consistent with CM’s sterically hindered nature.

Next, having established usefully enhanced on-protein rates, we examined substrate breadth. Model Seac-tagged protein SBL-156Sec 6a was tested with challenging metathesis substrate GlcNAc 10; this olefin is essentially unreactive in sulfur-relayed CM with SBL-156Sac 9a). Plesantly, reaction reached 75% conversion after 1.5 h at 37 °C (Scheme 4). Another biologically relevant substrate, N-allyl acetamide (11), was used in CM with protein 6a. Acetamide 11 has been suggested to poison catalyst 1, and also fails with SBL-156Sac 9a).¹¹,₁⁴,₄⁶ Despite poisoning,⁴⁷ after 20 min at 37 °C, LC-MS revealed conversion to desired 6d (Scheme 4), albeit in 30% yield.

After demonstrating CM on SBL-Seac, we tested Se-relayed CM on more functionally relevant protein systems. Histone proteins are key chromatin components; post-translational modifications (e.g., K-Ac) are critical regulators of structure and function.⁵⁸,⁹⁷ First, Seac was installed into key epigenetic site 9 of histone H3. After generation of H3-9Dha,²⁷ subsequent allyl selenolate addition to Dha 12 proceeded successfully to yield Seac protein 13 after 1 h at room temperature (Scheme 5).

With H3-9Sec (13) in hand, we carried out model CM with allyl alcohol. Coloration implying metal coordination to 13 was observed; such coordination can necessitate metal scavenging prior to MS analysis. Use of 3-mercaptopropionic acid 51,⁵０ allowed ready MS monitoring of H3 CM reactions. Pleasingly, this revealed CM of H3-9Sec and allyl alcohol with nearly full conversion to CM-modiﬁed H3 14 (Scheme 6a). This ability to modify site 9 suggested a strategy for chemical recapitulation (under benign conditions) of a “write–read–erase” cycle that is observed (enzyme-mediated) in histone epigenetics.⁴⁸,⁴⁹ As a “write” step we used the more challenging N-allyl acetamide oleﬁn in CM; complete conversion was observed to 15 (Scheme 6a). As a “read” step, CM-modiﬁed H3 15 was analyzed by an antibody raised to natural epigenetic marker N-acetyl lysine at position 9 of H3. We were pleased to find that this antibody successfully cross-reacted and recognized the CM-installed modiﬁcation at position 9 as a K9Ac PTM mimic (Scheme 6b). Next, as an “erase” step the CM-installed K9Ac mimic “mark” was removed with full conversion using Cope-type elimination through the generation of labile selenoxide using mild peroxide oxidation to regenerate Dha. Met-containing H3 may be prone to side reactions;⁵² this was tested and conﬁrmed through the use of a Met-free H3 variant in which the mark was cleanly expunged without side reaction (see SI sections 2.11–2.14). Finally, a full chemical “write–read–erase–rewrite–read” cycle was performed (SI section 2.12) to demonstrate iterative histone “switching”.

In conclusion, we have demonstrated efﬁcient chemical incorporation of Seac into proteins. This has facilitated rapid cross-metathesis for protein modiﬁcations; determination of on-protein rates shows that these outstrip or are comparable to many of the so-called “click” reactions in chemical biology. Using such allyl selenide (Seac) proteins, we were able to access a broader substrate scope in Se-relayed CM than was possible with
allyl sulfides. Direct access to Seac in proteins creates opportunities for potential PTM mimicry via CM as demonstrated here: a K9Ac mimic was successfully installed, recognized, and removed (write–read–erase) in histone H3. Although, to the best of our knowledge, such a chemical “write–read–erase” approach has not been explored before, one could envisage a incorporation of allyl chalcogenide amino acids into proteins to Rovira i Virgili, C/Marcel

We note too that this work provides access to a protected (allylated) Sec in proteins (that could be revealed using, e.g., Pd(0)) and also further motivates developments in genetic incorporation of allyl chalcogenide amino acids into proteins to access more generally the wider biological applications of cross-metathesis.

**ASSOCIATED CONTENT**

Supporting Information

Full procedures and protein ESI-MS. This material is available free of charge via the Internet at http://pubs.acs.org.

**AFFECTED CONTENT**

Corresponding Author

ben.davis@chem.ox.ac.uk

Present Address

O.B.: Departamento de Química Analítica i Orgànica, Universitat Rovira i Virgili, C/Marcel·lí Domingo s/n, Tarragona, Spain.

Notes

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We thank EPSRC (Y.A.L., L.L.), CRUK (L.L.), Rhodes Trust (B.B.), and EU FP7 Marie Curie programme (O.B.) for funding. B.G.D. is a Royal Society Wolfson Merit Research Award Recipient.

**REFERENCES**

(1) Sletten, E. M.; Bertozzi, C. R. Angew. Chem., Int. Ed. 2009, 48, 6974.
(2) Chalker, J. M.; Bernardes, G. J. L.; Davis, B. G. Acc. Chem. Res. 2011, 44, 730.
(3) Takaoka, Y.; Ojida, A.; Hamachi, I. Angew. Chem., Int. Ed. 2013, 52, 4088.
(4) Hoye, T. R.; Zhao, H. Org. Lett. 1999, 1, 1123.
(5) Lin, Y. A.; Davis, B. G. Beilstein J. Org. Chem. 2010, 6, 1219.
(6) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. J. Am. Chem. Soc. 2008, 130, 9642.
(7) Brik, A. Adv. Synth. Catal. 2008, 350, 1661.
(8) Kirshenbaum, K.; Arora, P. S. Nat. Chem. Biol. 2008, 4, 527.
(9) Lin, Y. A.; Chalker, J. M.; Davis, B. G. ChemBioChem 2009, 10, 959.
(10) Ai, H.-W.; Shen, W.; Brustad, E.; Schultz, P. G. Angew. Chem., Int. Ed. 2010, 49, 935.
(11) Lin, Y. A.; Chalker, J. M.; Davis, B. G. J. Am. Chem. Soc. 2010, 132, 16805.
(12) Foster, S. J.; Kraus, R. J.; Garner, H. E. Arch. Biochem. Biophys. 1986, 251, 77.
(13) Ganther, H. E. Carcinogenesis 1990, 20, 1657.
(14) Stols, J. F.; Basu, P.; Santini, J. M.; Orenland, R. S. Annu. Rev. Microbiol. 2006, 60, 107.
(15) Amér, E. S. J.; Sarioglu, H.; Lottspeich, F.; Holmgren, A.; Böck, A. J. Mol. Biol. 1999, 292, 1003.
(16) Aldag, C.; Bröcker, M. J.; Hohn, M. J.; Prat, L. H.; G.; Plummer, A.; Soll, D. Angew. Chem., Int. Ed. 2013, 52, 1441.