DOI: 10.1002/cbic.200900002

Olefin Metathesis for Site-Selective Protein Modification

Yuya A. Lin, Justin M. Chalker, and Benjamin G. Davis*[a]

For a reaction to be generally useful for protein modification, it must be site-selective and efficient under conditions compatible with proteins: aqueous media, low to ambient temperature, and at or near neutral pH. To engineer a reaction that satisfies these conditions is not a simple task. Olefin metathesis is one of most useful reactions for carbon–carbon bond formation, but does it fit these requirements? This minireview is an account of the development of olefin metathesis for protein modification. Highlighted below are examples of olefin metathesis in peptidic systems and in aqueous media that laid the groundwork for successful metathesis on protein substrates. Also discussed are the opportunities in protein engineering for the genetic introduction of amino acids suitable for metathesis and the related challenges in chemistry and biology.

Introduction

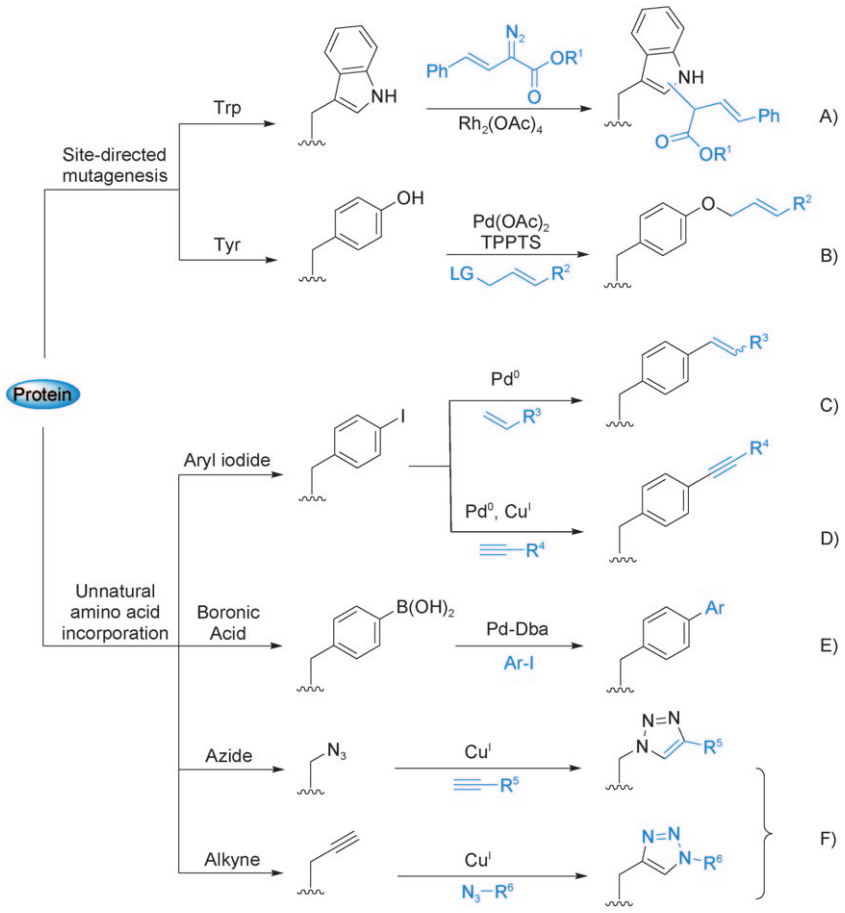
Chemical modification of proteins is a powerful method for study and modulation of macromolecular function. Covalent attachment of fluorescent probes, for example, allows proteins to be tracked both in vitro and in vivo.[1,2] Tethering antigens to protein carriers is a current route to synthetic vaccine candidates for HIV,[3] cancer,[4] malaria,[5] and pathogenic bacteria.[6] Additionally, the study of post-translational modifications (PTMs) is facilitated by chemical access to modified proteins that are otherwise difficult to isolate from natural sources in pure form.[7,8] Finally, protein function may be enhanced, altered, or imparted entirely by a selective modification.[9] Given the range of opportunities in chemistry, biology, and medicine, it is not surprising that new strategies for the selective modification of proteins have continued to develop at a rapid pace.

Most strategies for chemical protein modification rely on the nucleophilic side chains of amino acids such as lysine, cysteine, or aspartic and glutamic acids.[10] Although these modifications are widespread, the reaction space they occupy is relatively small and new strategies are needed for evermore precise and sophisticated endeavors. For a reaction to be generally useful for protein modification, it must be site-selective and efficient under conditions compatible with proteins: aqueous media, low to ambient temperature, and at or near neutral pH. Moreover, the reaction must tolerate salts and surfactants commonly used to stabilize proteins.[11] To engineer a reaction that satisfies these conditions is not a simple task.

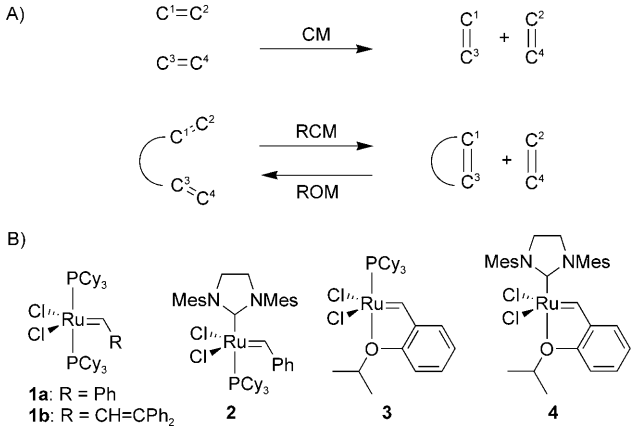
Transition-metal-catalyzed reactions are attractive candidates for site-selective protein modification. These transformations are ubiquitous in organic synthesis and mediate the formation of bonds that are otherwise difficult or impossible to create.[12] These reactions are flexible and can often be tuned by judicious selection of ligands and additives. Moreover, the repertoire of transition-metal-mediated reactions in water has expanded in recent years.[13,14] While there is still a disparity between the widespread use of transition metals in small-molecule synthesis and their modest deployment in protein modification, key advances have been made (Scheme 1). The Francis group has pioneered several modifications of natural residues using transition metals.[15] Tryptophan alkylation with rhodium carbenoids[16] and allylation of tyrosine by palladium π–allyl complexes[17] are among their accomplishments. For unnatural residues, Heck and Sonogashira reactions at *p*-iodophenylalanine have been described.[18–20] Progress in Suzuki cross-coupling at this residue[21] and *p*-boronophenylalanine[22] has also been reported. Finally, the copper-catalyzed azide–alkyne [3+2] cycloaddition has been employed widely in protein modification.[23–26]

A glaring omission from the examples in Scheme 1 is olefin metathesis. It is indisputable that olefin metathesis is one of most useful reactions for carbon–carbon bond formation (Scheme 2).[27–30] In part, the broad utility of olefin metathesis is a consequence of the exquisite selectivity and functional group tolerance of ruthenium-based metathesis catalysts. Using olefin metathesis to modify proteins is a stringent test of this chemoselectivity and functional group compatibility. Moreover, the capacity to form carbon–carbon bonds on protein surfaces creates many new and exciting opportunities in biology.[31] This minireview is an account of the development of olefin metathesis for protein modification. Highlighted below are examples of olefin metathesis in peptidic systems and in aqueous media that laid the groundwork for successful metathesis on protein substrates.[32] Also discussed are the opportunities in protein engineering for the genetic introduction of amino acids suitable for metathesis. Finally, we outline prospective opportunities and challenges in chemistry and biology that stem from the use of olefin metathesis on proteins.

[a] *Y. A. Lin, J. M. Chalker, Prof. B. G. Davis*
Department of Chemistry, University of Oxford,
Chemistry Research Laboratory
12 Mansfield Road, Oxford OX1 3TA (UK)
Fax: (+44) 44-1865-285002
E-mail: ben.davis@chem.ox.ac.uk



Scheme 1. Examples of transition-metal-catalyzed protein modification. A) Trp alkylation with rhodium carbenoids. B) Modification of Tyr with π-allylpalladium electrophiles. TPPTS = sodium triphenylphosphine trisulfonate. C) Heck reaction. D) Sonogashira coupling. E) Suzuki coupling. F) Copper-catalyzed azide–alkyne [3+2] cycloaddition.



Scheme 2. A) Selected types of olefin metathesis: CM = cross metathesis, RCM = ring-closing metathesis, ROM = ring-opening metathesis. B) Conventional ruthenium-based metathesis catalysts

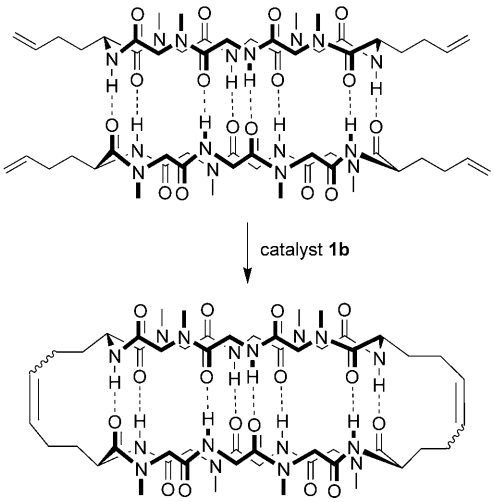
The Challenge of Olefin Metathesis on Proteins

A number of challenges must be addressed for efficient olefin metathesis on a protein surface. Firstly, the protein must contain an alkene. Tirrell and co-workers have reported the genetic incorporation of homoallylglycine (Hag) as a methionine surrogate in proteins through the use of methionine auxotrophic bacterial hosts.[33,34] The authors clearly recognized the opportunity for modification by olefin metathesis, but metathesis at Hag is at present limited to peptide models in organic solvent. Thus, a second challenge is that the reaction must work efficiently in water at or below 37 °C. Olefin metathesis reactions often involve long reaction times in organic solvent at elevated temperatures—not inviting conditions for a protein. Thirdly, the reaction must tolerate hundreds of side chains that might chelate the metal center and sequester or poison the catalyst. Fourthly, the secondary and tertiary structure of the protein must not compromise access to the desired site of modification. Finally, because proteins are often only available at low concentrations, the reaction must be rapid to achieve full conversion. These criteria are daunting and have thwarted efforts in metathesis on proteins for more than a decade since Tirrell's first incorporation of Hag into proteins. We now turn to the fundamental studies that motivated and guided the development of olefin metathesis on proteins.

Olefin Metathesis on Peptides

The overriding motivation for using olefin metathesis to modify peptides is to install nonlabile carbon–carbon bonds.[35] These modifications are typically of two varieties. The first type is a crosslink that might stabilize peptide secondary structures, with the goal of imparting better metabolic stability and higher binding affinity towards biological targets. The second type of modification is post-synthetic labeling with functional tags such as carbohydrates, thereby providing nonlabile analogues of natural PTMs. We discuss key examples of these modifications in turn.

One of the earliest examples of olefin metathesis on peptides was reported by Clark and Ghadiri.[36] They demonstrated sequential cross metathesis and ring-closing metathesis (RCM) between two cyclic peptides containing Hag residues in the presence of catalyst **1b** (Scheme 3). The resulting structure is a β-sheet-like cylindrical structure. This example is remarkable in that it is not only one of the earliest examples of olefin meta-

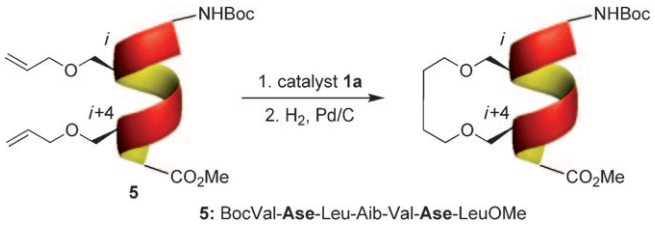


Scheme 3. Peptide cylinder by CM-RCM assisted by intermolecular hydrogen bonding.

thesis in peptide synthesis, but an example of cross metathesis templated by intermolecular hydrogen bonding. This early example highlights the influence of peptide preorganization on the outcome of the metathesis and the ability to "covalently capture" self-assembled secondary structure.

RCM, probably the most developed metathesis reaction, has proven useful in the synthesis of unnatural amino acids and peptidomimetics. Exploratory work in this area delineated the scope of ring sizes and sidechains that are useful in the synthesis of cyclic amino acids.[37–39] This chemistry provides access to several β-turn analogues that mimic the natural role of β-turns in stabilizing short peptides.[40,41] β-Turns are common structural features and comprise about 25% of all amino acids in proteins. Moreover, they often serve as recognition elements on the protein surface; this makes them attractive targets for RCM.[42]

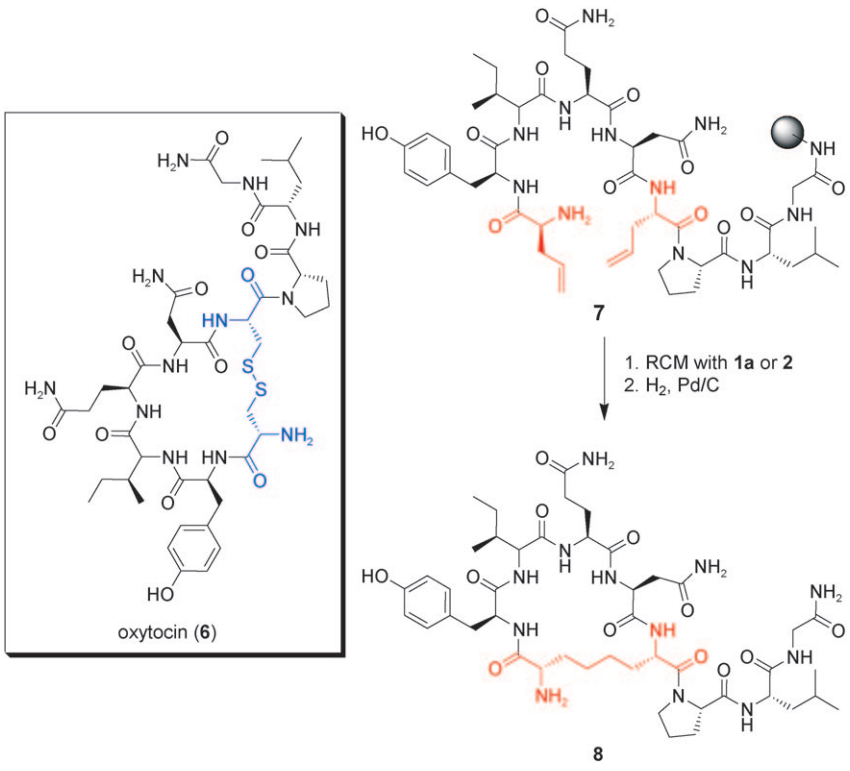
RCM has also been used to crosslink α-helices to induce structural rigidity.[43] In one of the first examples, by Grubbs, an α-helix consisting of seven amino acids was modified to contain *O*-allyl ethers at its *i* and *i*+4 positions, which were approximately aligned on the turn of the helix.[44,45] RCM with catalyst **1a** followed by hydrogenation provided the crosslinked α-helix (Scheme 4). A systematic study of the crosslink position and helix stability was later undertaken by Verdine.[46]



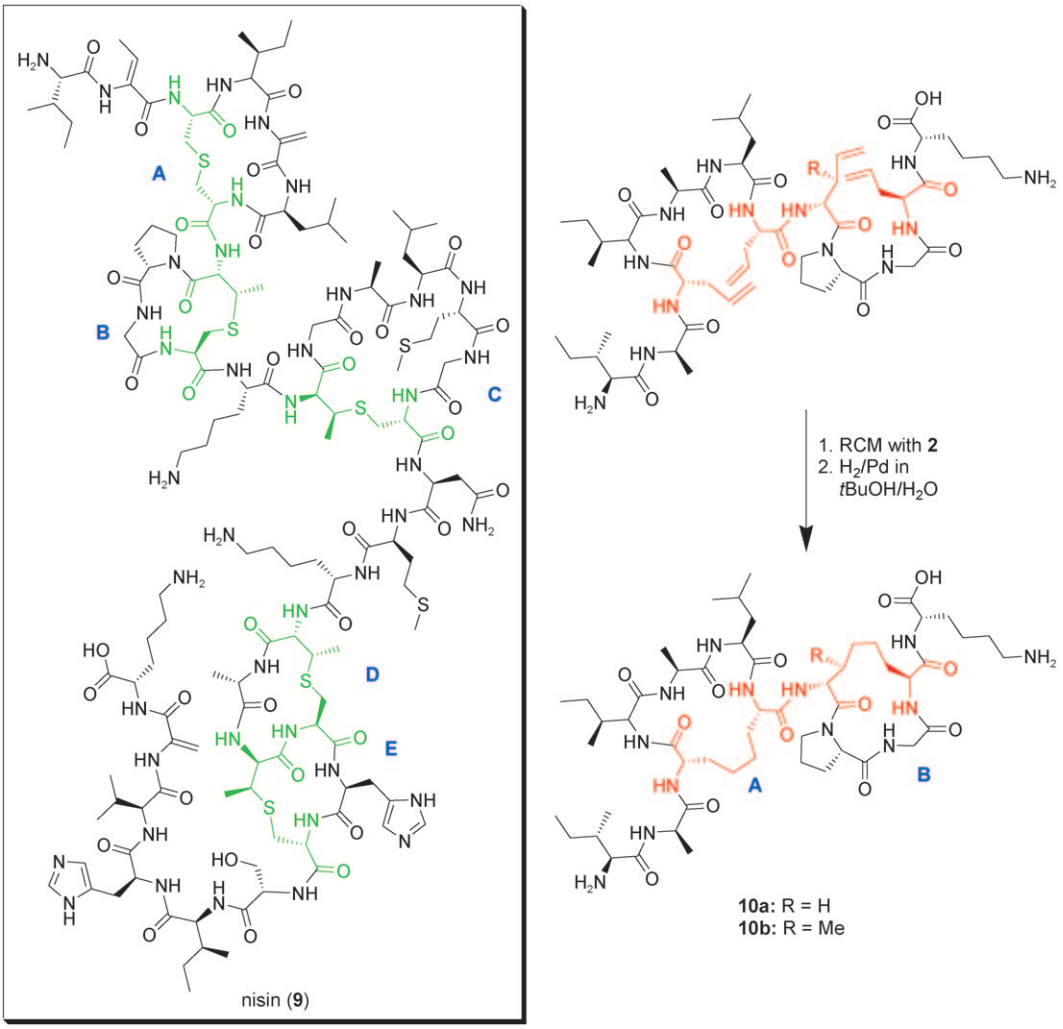
Scheme 4. Crosslinked helices by RCM.

Disulfide bridges play important roles in stabilizing the secondary and tertiary structures of peptides and proteins. However, they are labile under reducing conditions. As a result, many workers have turned to RCM to create the carbon equivalents of disulfide bridges, the so-called "dicarba analogues". Building on Grubbs' early use of RCM to replace disulfides by all-carbon analogues,[37,38] Vederas and co-workers synthesized the dicarba analogue of the hormone oxytocin (**6**, Scheme 5).[47,48] The linear peptide **7**, in which the two cysteine residues were replaced with allylglycine, was constructed on a solid support. RCM was then carried out on-resin, followed by cleavage from the solid support. Hydrogenation afforded the cyclic oxytocin analogue **8** (Scheme 5), which was shown to have a longer half-life in vivo than the natural counterpart.

Similar strategies guided efforts in the synthesis of analogues of the lantibiotic nisin (**9**, Scheme 6) in which thioether linkages were replaced with all-carbon linkages.[49] It was also found that the alkane-bridged mimics of the nisin A and B rings had activity comparable to that of the natural lantibiotic.



Scheme 5. Synthesis of an oxytocin dicarba analogue by RCM.



Scheme 6. The lantibiotic nisin (**9**) and its alkane-bridged AB ring mimics **10a** and **10b**.

Apart from Ghadiri's cross metathesis of preorganized peptides, the examples discussed so far are essentially restricted to RCM. To modify proteins post-translationally by cross metathesis, it is instructive to consider peptide models and the relative reactivities of prospective side chains. Studies by Gibson on cross metathesis of unsaturated amino acids revealed that reactivity in metathesis increased with the length of the side chain when catalyst **1a** was used (Scheme 7A).[50,51] It is noteworthy that Schrock-type molybdenum catalysts promote cross metathesis of unreactive partners such as vinylglycine.[52] However, the air- and moisture-sensitivity of this catalyst does not, at present, bode well for application to metathesis on deprotected peptides in aqueous media.

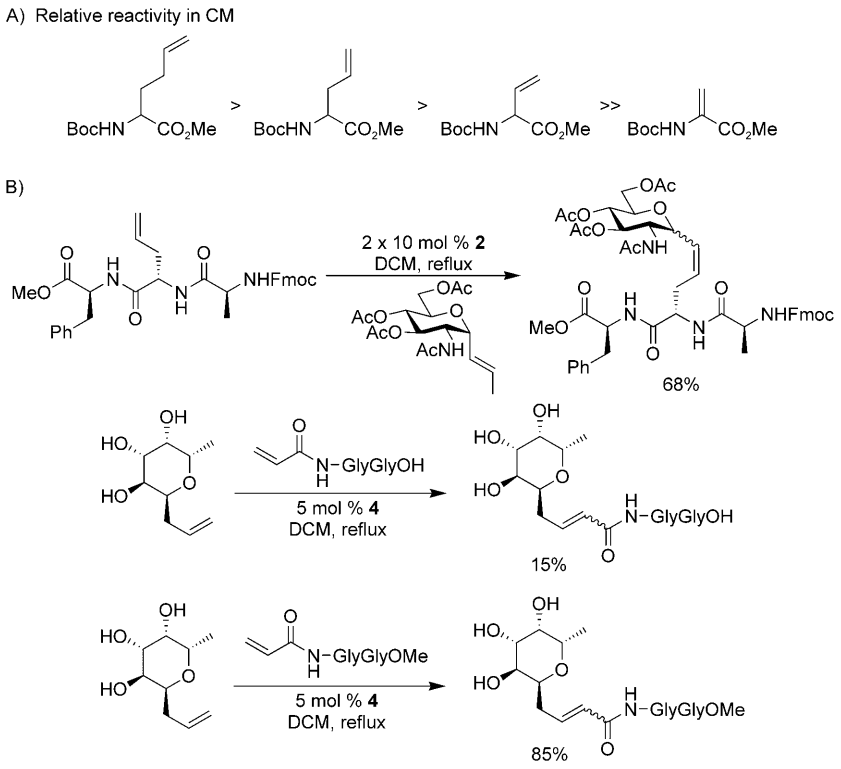
Biologically important glycopeptides have also been synthesized by cross metathesis between sugars and peptides. Early examples by the Roy laboratory demonstrated the feasibility of glycopeptide synthesis by cross metathesis.[53] McGarvey later refined considerations of protecting groups and reaction conditions to favor cross metathesis over self metathesis, with the most promising results being obtained with catalyst **2**.[54] Blechert recently revisited glycopeptide synthesis by cross metathesis and detailed the compatibility of metathesis catalysts with certain amino acid residues.[55] Histidine and tryptophan were problematic, whereas methionine was notably tolerated in metathesis reactions when the Hoveyda–Grubbs second-generation catalyst (**4**) was used. In that report, deprotected carbohydrates were also demonstrated as useful substrates in metathesis reactions. Deprotected peptides were also used, but a free carboxy terminus was sometimes detrimental (Scheme 7B). Collectively, these examples motivate the use of cross metathesis as a route to nonlabile *C*-glycopeptides and glycoprotein analogues of the natural *O*-linked and *N*-linked glycopeptides and glycoproteins.

Although these examples were all carried out in organic solvents, the results demonstrated promising functional group compatibility and substrate scope for olefin metathesis in peptidic systems. The relative reactivities of various unsaturated amino acids in metathesis can also guide amino acid selection for incorporation into proteins. Successful results with Hag were particularly promising, because it can be genetically incorporated into proteins. Finally, the general strategy of using metathesis to stabilize peptide secondary structures or to

tether carbohydrates covalently can, in principle, be extended to proteins. In order for these strategies to be applicable to proteins, developments in aqueous olefin metathesis were required.

Aqueous Olefin Metathesis

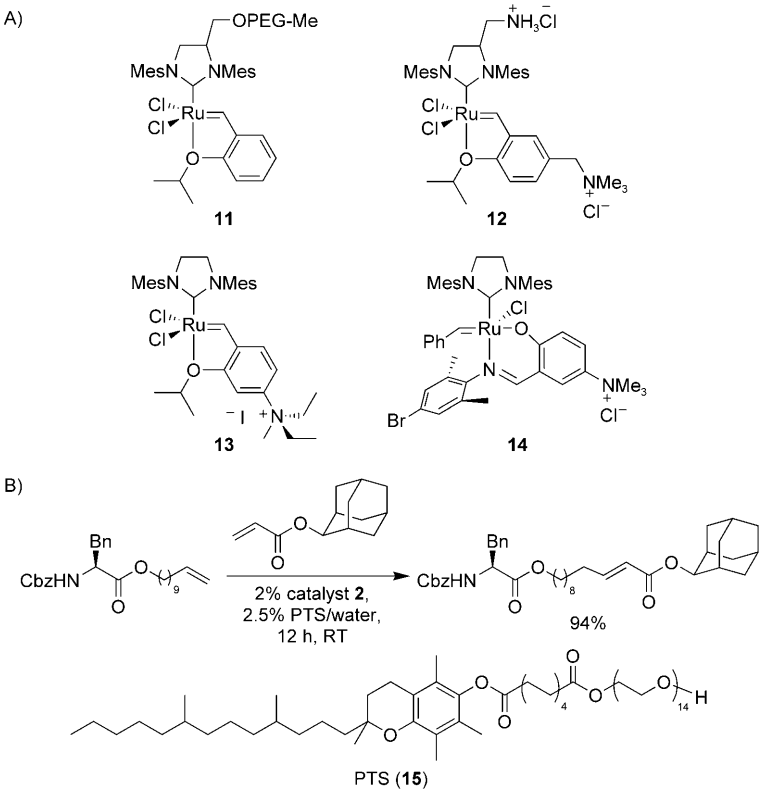
Aqueous olefin metathesis is a rapidly developing area and has been recently reviewed in detail.[56] Although this research is primarily driven by the ultimate goal of sustainable chemistry, progress in this field is immediately relevant to olefin metathesis on biomolecules that require aqueous media. Selected catalysts useful in homogenous olefin metathesis are discussed below, along with other notable strategies for aqueous olefin metathesis that might prove useful in bioconjugation.



Scheme 7. A) Relative CM reactivities of a selection of unsaturated amino acids. B) Glycopeptide syntheses by CM.

Several water-soluble (pre)catalysts have been developed for olefin metathesis. The solubility of these catalysts in water is imparted by ligands containing hydrophilic poly(ethylene glycol) (PEG) or quaternary ammonium groups. A selection is given in Scheme 8 A. The PEGylated catalyst **11** was the first to promote a range of RCM in neat water.[57] The self metathesis of allyl alcohol was also the first reported cross metathesis in water. Another catalyst in which the *N*-heterocyclic carbene (NHC) ligand was well-defined and soluble in water by virtue of pendant ammonium groups was later reported.[58] Grela and coworkers have reported the installation of ammonium groups on the alkylidene portion of the catalyst. The ammonium group imparts water solubility to the precatalyst and weakens the O–Ru coordination, resulting in fast initiation. This catalyst is active in RCM, ene–yne, and cross metathesis of simple substrates in protic solvents, including water.[59,60] Raines has also reported an addition to water-soluble metathesis catalysts containing a salicylaldimine ligand with a pendant ammonium group. This catalyst was active in RCM of a range of substrates in aqueous solvent.[61]

Other studies have focused on the use of water-insoluble catalysts in aqueous media with the use of a cosolvent. Blechert reported good RCM with catalysts **2** or **4** in aqueous DMF or methanol.[62] Modest activity in CM was observed for simple olefins.



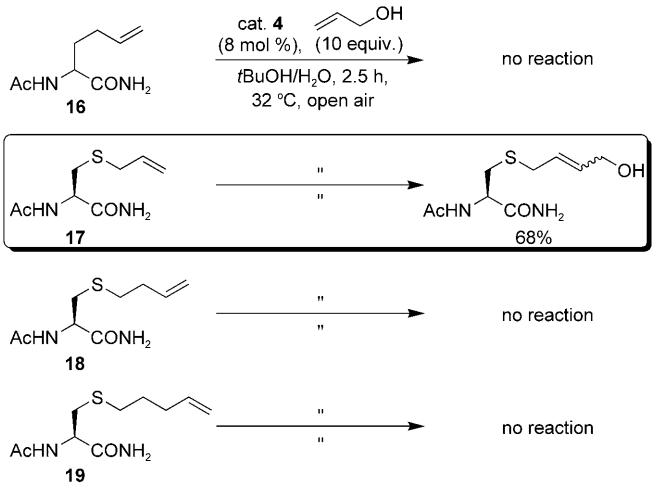
Scheme 8. A) Examples of water-soluble metathesis catalysts. B) Cross metathesis through the use of a non-ionic amphiphile.

Raines reported the use of dimethoxyethane or acetone in water to aid aqueous RCM of several dienes and self metathesis of allyl alcohol in the presence of commercially available catalyst **4**.[63] Despite these developments, aqueous CM is largely limited to simple, reactive olefins such as allyl alcohol. The substrate scope in aqueous CM can be expanded to more complex substrates with emulsion chemistry. Lipshutz has widened the scope of aqueous olefin metathesis with the use of non-ionic amphiphiles such as **15**.[64] The efficiency of this method is likely due to increased effective concentrations of reactants and catalyst in the micelles formed by the surfactant. It is not clear at this point whether nanometer micelles can both solubilize the metathesis catalyst and still allow reaction on a protein surface. Nevertheless, the advance in aqueous metathesis is notable.

While the development of water-soluble catalysts for olefin metathesis is a milestone in aqueous chemistry, there is not yet a catalyst that is completely resistant to decomposition in water. A substituted olefin can slow decomposition because alkylidenes are more stable in water than methylidenes, but the rate of metathesis might also be compromised.[65] Metathesis catalyst decomposition in protic solvents and water has been investigated, but is still not completely understood.[66] In most cases, decomposition results in ruthenium hydrides, which are inactive in metathesis but prone to mediate carbon–carbon double bond isomerization and migration.[67] These side reactions point to the need for a comprehensive investigation of the coordination sphere and its ability to attenuate or prevent catalyst decomposition. An immediate compromise to catalyst stability in water—particularly for the purpose of protein modification—is to use a substrate that can react at a much greater rate than catalyst decomposition.

Allyl Sulfides: Privileged Substrates in Olefin Metathesis

In exploratory work in aqueous metathesis, our lab focused on cross metathesis of unsaturated amino acids with allyl alcohol in the presence of the Hoveyda–Grubbs second-generation catalyst (**4**).[32] Unnatural amino acid derivatives such as Hag and *S*-alkenyl cysteines were screened because they could potentially be incorporated into proteins either genetically or chemically if they proved reactive in metathesis. At the outset, no effort was made to exclude oxygen and *t*BuOH was used to solubilize the catalyst. Hag was screened first because its incorporation into proteins was well-established[33,34] and metathesis in organic solvent with this amino acid had already been demonstrated.[50,51] Disappointingly, no cross metathesis was observed with Hag in aqueous media. Our fortunes changed with *S*-allylcysteine (Sac). Sac derivative **17** gave a 68% yield of the desired CM product after only 2.5 h of reaction time at 32 °C. Hag and *S*-butenyl- and *S*-pentenylcysteine underwent no productive CM under the same reaction conditions (Scheme 9). Allyl amines and allyl ethers were also examined, but allyl sulfides remained the most reactive substrates in cross metathesis in aqueous media. It must be noted that catalyst degradation does occur in such aqueous systems, but the reactions with allyl sulfides were sufficiently high in turnover frequency to outcompete catalyst decomposition.



Scheme 9. Efficient cross metathesis of *S*-allylcysteine.

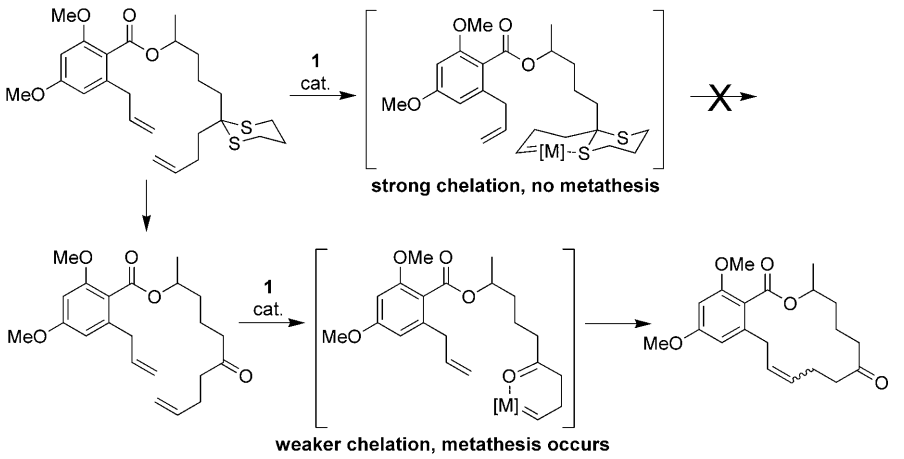
The efficient metathesis of allyl sulfides in water was striking for two reasons. Firstly, before this result the most complex cross metathesis in homogenous aqueous media had been the self metathesis of simple, reactive substrates such as allyl alcohol. Secondly, the efficient metathesis was observed with sulfur-containing substrates. Sulfur is often problematic in transition-metal-catalyzed reactions because its high affinity to the soft metal center can poison the catalyst. Indeed, there have been several cases of olefin metathesis in which sulfides were detrimental.[68–71] In Fürstner's synthesis of the macrocycle Zeranol, for instance, the key step involved RCM of a molecule containing a 1,3-dithiane unit.[71] Cyclization was not observed in the presence of the first-generation Grubbs catalyst. The outcome was explained by nonproductive chelation of the proximal sulfur atom to the ruthenium. Indeed, deprotection to the corresponding ketone enabled efficient cyclization (Scheme 10).

In the light of this result and other similar complications with sulfur-containing substrates, Fürstner suggested "a quite general incompatibility of the ruthenium based metathesis catalysts with substrates containing sulfur(II) donor sites."[71] Indeed, this is essentially correct; allyl sulfides are an exception to the rule. Moreover, for allyl sulfides sulfur is not simply tolerated by the catalyst, it *enhances* the substrate reactivity. Although metathesis reactions of allyl sulfides in organic solvents have been reported, the enhanced reactivity went unnoticed or was not apparent under the conditions employed.[72,73] The basis for this enhanced reactivity of allyl sulfides in metathesis is discussed next.

Sulfur-assisted cross metathesis

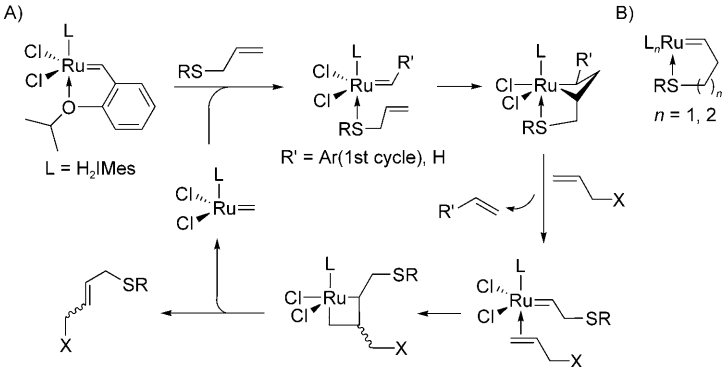
The efficient reactivity of allyl sulfides in CM was explained with a mechanism invoking sulfur pre-coordination to ruthenium (Scheme 11 A). Because of the soft natures of second row transition metals, sulfur is preferred over oxygen as a donor to

the second-generation precatalyst. This leads to rapid formation of the reactive alkylidene species, which initiates subsequent metathesis events. The decreased reactivities of butenyl and pentenyl sulfides may be attributed to the unproductive five- or six-membered chelates depicted in Scheme 11 B and are consistent with Fürstner's observations with the dithiane in Scheme 10. Apparently, the chelate formed with allyl sulfides is too strained to sequester the catalyst and rapid turnover is observed.



Scheme 10. Sulfide chelation can inhibit RCM.

The sulfur relay mechanism may be related to other cases in which remote functionality enhances the rate of olefin metathesis. Hoye noted the influence of both the steric and electronic characters of the allylic substituents in RCM in linalool and related dienes.[74] A free allylic hydroxyl group greatly enhanced RCM relative to the corresponding methyl ether or unsubstituted starting material (Scheme 12 A). Fürstner has also shown that in certain cases, macrocycle synthesis by RCM can be favored over oligomerization when strategically placed carbonyls can coordinate to ruthenium (Scheme 12 B, C).[75] In these examples, as in the case of the allyl sulfide, the increased rate of metathesis appears to derive from an intermediate chelate that brings the reactive centers into close proximity. The chelate cannot be too stable or turnover will stop and the remote functionality will hinder metathesis.



Scheme 11. A) Sulfur-assisted cross metathesis. B) Unproductive chelates.

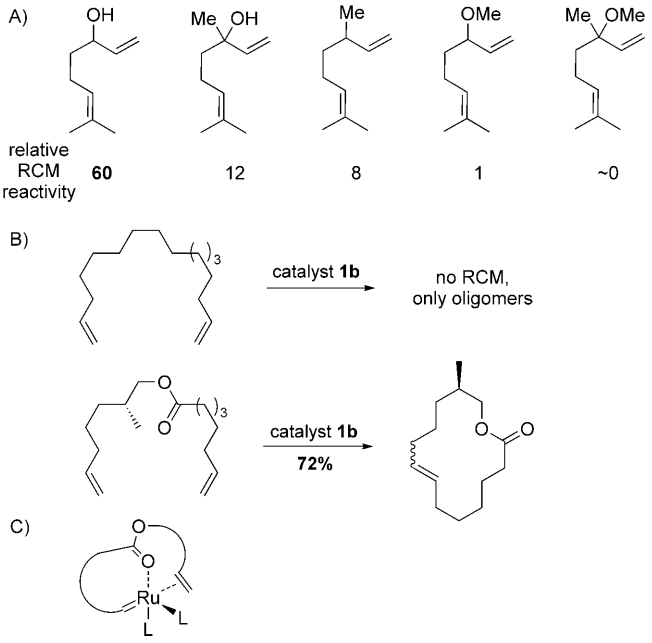
Importantly, the amino acid *S*-allylcysteine was found to react efficiently in cross metathesis, even in aqueous media. This breakthrough motivated its incorporation into proteins so that the ultimate goal of metathesis on protein surfaces could be achieved. We now turn to both chemical and genetic strategies for incorporation of allyl sulfides into proteins.

Chemical Access to Allyl Sulfides into Proteins

Several chemical routes to *S*-allylcysteine can be envisioned. Addition to dehydroalanine (Dha) by allyl thiol, direct allylation, and desulfurization of the allyl disulfide are potentially practical for proteins. The basis for each transformation, and the scope and limitations, are discussed next.

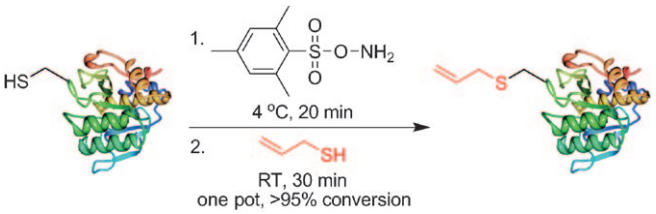
Conjugate addition to Dha

An efficient way to access *S*-allylcysteine in proteins is by conjugate addition of allyl thiol to dehydroalanine. Dha is an effective Michael acceptor for thiol nucleophiles. Multiple routes to Dha-containing proteins have been reported. An early example was the elimination of the phenylmethylsulfonyl fluoride



Scheme 12. A) Allylic hydroxyl activation of RCM. B) Macrocycle synthesis by carbonyl-relayed RCM. C) Putative chelate in carbonyl-relayed RCM.

(PMSF) adduct of the catalytic serine of serine proteases.[76,77] Another route is the elimination of dialkylated cysteine residues.[78] Both of these methods require either high pH or elevated temperature and so may not be useful for any given protein. Milder methods have recently emerged. The Schultz group has demonstrated efficient incorporation of the unnatural residue phenylselenocysteine and its subsequent oxidative elimination to Dha by treatment with hydrogen peroxide.[79] Our own lab has reported a novel oxidative elimination directly from cysteine by use of *O*-mesitylenesulfonylhydroxylamine (MSH).[80] Once Dha is obtained, *S*-allylcysteine in proteins is accessed readily by conjugate addition of allyl thiol. We have successfully demonstrated this one-pot reaction sequence on a single cysteine mutant of subtilisin *Bacillus lentus* (Scheme 13).[32] The advantage of this route is that it is fast, easy, and chemoselective. Unfortunately, the diastereoselectivity of thiol addition to Dha in peptides is typically low, resulting in epimeric mixtures.[81,82] Some sequences may promote selectivity[83] and on the surface of a protein the outcome will be highly dependent on the local stereochemical environment. For the purpose of evaluating the metathesis reaction, the diastereoselectivity was not an immediate concern. However, it is useful to have complimentary allylation methods that avoid diastereomeric product mixtures. Prospective solutions are discussed next.



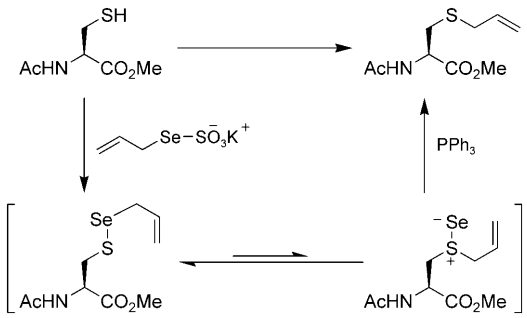
Scheme 13. Synthesis of Sac-containing proteins by conjugate addition to Dha

Direct allylation

One way to avoid epimerization at cysteine is by direct alkylation. There is good precedence for semi-selective alkylation of cysteine on proteins.[10] Nonselective alkylation at other nucleophilic residues such as lysine and histidine is the major drawback, but this can sometimes be avoided by judicious electrophile selection, together with careful control of pH, stoichiometry, reaction time, and temperature.[84] Alkylations with α-halocarbonyls, maleimides, and other alkyl halides have been used extensively.[10] Although direct allylation of cysteine with allyl chloride or similar electrophiles on the surface of a protein has not been reported, the successful alkylations mentioned above bode well for freely accessible cysteines. This alkylation is currently under investigation in our laboratories.

Desulfurization

Disulfide formation followed by desulfurization is another route to *S*-allylcysteine. The desulfurative rearrangement of allylic selenosulfides and disulfides has been explored by Crich in the synthesis of a range of *S*-allylcysteine derivatives (Scheme 14).[85–87] The allylic selenosulfides and disulfides were prepared by treatment of the free thiol with allylic selenosulfates and thiosulfates, respectively. The dechalcogenative rearrangement was then mediated by treatment with triphenylphosphine (Scheme 14). This method has been shown to work on unprotected amino acids in water[87] and is likely to be explored in protein modification. Moreover, this method resolves the issue of epimer formation by the Dha route and is likely more selective than direct allylation, provided that the protein modified does not have natural disulfides susceptible to reduction.



Scheme 14. *S*-Allylcysteine through dechalcogenative allylic selenosulfide rearrangement.

Genetic Incorporation of Allyl Sulfides into Proteins

Perhaps the most general route to proteins containing unnatural amino acids is by translation in *E. coli* hosts.[88–90] The residue may be incorporated as a surrogate for another residue of similar size and electronic character. Alternatively, a unique codon of an expanded genetic code may be used in the translation of these residues. The prospective use of each method for the incorporation of *S*-allylcysteine or other allyl sulfides into proteins is discussed next.

S-Allylcysteine as a methionine surrogate

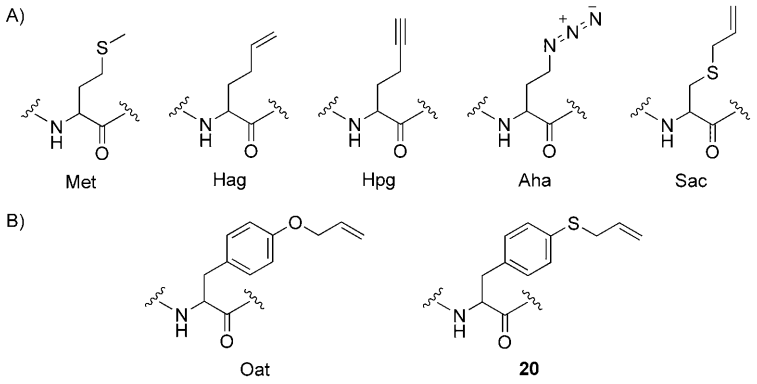
Genetic incorporation of unsaturated amino acids as methionine surrogates in methionine auxotrophic *E. coli* has been well-documented.[33,34] These amino acids are incorporated because their side chains have a size and electronic character similar to methionine. As mentioned in earlier sections, Hag is a potential handle for olefin metathesis. Other methionine surrogates useful for protein modification are homopropargylglycine (Hpg)[34] and azidohomoalanine (Aha).[91] Both have been used extensively in conjugation chemistry. In our own efforts, we considered that *S*-allylcysteine might be a suitable methionine surrogate because its side chain is not too different in size and polarity from the methionine side chain. We expressed a single Met mutant of a model glycosidase in a methionine auxotrophic *E. coli* in Met-depleted media.[32] The incorporation of *S*-allylcysteine, however, was quite low. Nonetheless, unam-

biguous incorporation was verified by mass spectrometry and peptide mapping. This modest, but promising, starting point has motivated further efforts to incorporate related amino acids as analogues for methionine and other low-incidence residues.

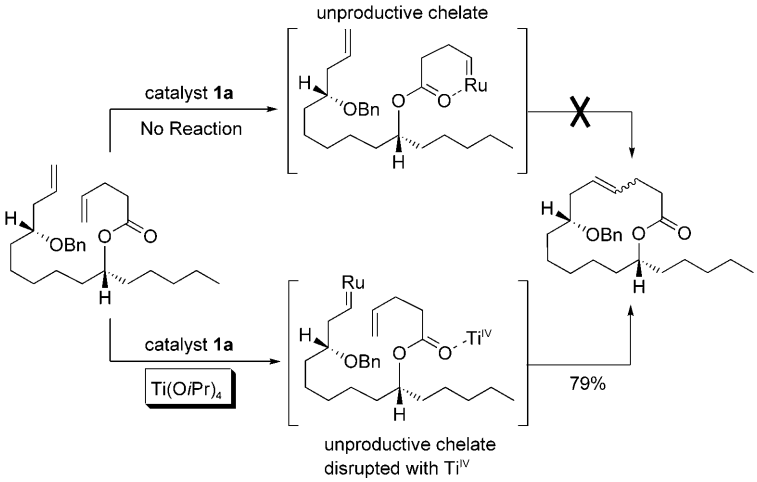
Codon suppression

Reassignment of the amber stop codon is a powerful strategy for the introduction of unnatural amino acids into proteins, and in principle a large set of allyl sulfide metathesis handles could be incorporated.[92] This method, developed extensively by the Schultz laboratory, uses the amber nonsense codon (UAG) as a unique codon for a particular unnatural amino acid.[93] Many examples of unnatural amino acids have been incorporated, including residues suitable for metathesis such as *O*-allyltyrosine (Oat). It seems reasonable that allyl sulfide analogue **20** could be incorporated in a similar manner, provided that the residue is metabolically stable (Scheme 15 B).

As such residues become accessible in standard translation systems, opportunities for olefin metathesis on proteins are likely to become more common. At present, the only reported method for accessing proteins with allyl sulfides is by addition of allyl thiol to Dha.[32] While there are clearly many alternative approaches, our ultimate goal of using olefin metathesis on proteins was in sight. We now turn to the key modification.



Scheme 15. A) Methionine and surrogates for incorporation into proteins by methionine auxotrophic *E. coli*. B) *O*-Allyltyrosine and sulfur analogue **20**.



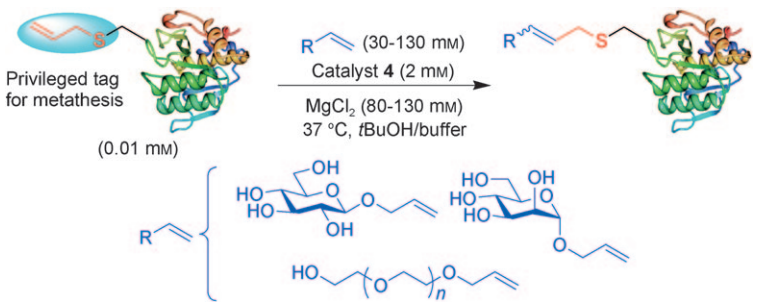
Scheme 16. Use of a Lewis acid to disrupt unproductive chelation.

Protein Modification by Cross Metathesis

With Sac-containing protein in hand, we sought to take advantage of the unique reactivity of allyl sulfides in cross metathesis. Initial attempts at cross metathesis with allyl alcohol in the presence of catalyst **4** and *tert*-butanol as a cosolvent returned only unmodified protein. However, careful analysis of the reaction mixture formed from only protein and catalyst revealed a compound with a mass corresponding to a possible metalloprotein species, a putative intermediate that was unreactive in metathesis. We speculated that chelation from nearby amino acid residues might sequester the catalyst from further metathesis events. Fürstner has shown that similar non-productive chelation in metathesis reaction can be disrupted by the use of a hard Lewis acid such as $Ti(OiPr)_4$ (Scheme 16).[94] Other reports of similar unproductive chelation and Lewis acid rescue in olefin metathesis have also emerged.[95,96]

Magnesium(II) as an additive in CM

For biocompatibility, $MgCl_2$ was chosen as a mild Lewis acid to prevent non-productive chelation. Remarkably, once magnesium(II) was included in the buffer, efficient cross metathesis was immediately observed on the protein with, initially, allyl alcohol as a partner olefin. This breakthrough allowed us to explore more biologically relevant modifications (Scheme 17). Glycosylation was achieved by cross metathesis of the Sac-containing protein with monosaccharide allyl glycosides. Glycoproteins are involved in cell signaling during inflammation and immune response. Surface-bound carbohydrates are also critical markers for protein folding.[97,98] Access to homogenous natural samples is difficult and cross metathesis opens new access to well-defined glycoprotein constructs.[99] Protein PEGylation, an important modification for increasing the metabolic half-lives and shelf lives of



Scheme 17. Cross metathesis on SBL-Sac156.

therapeutic proteins,[100] was also accomplished on the protein in good yield (Scheme 17).

Our immediate goal of adapting olefin metathesis to protein modification was met with successful cross metathesis at *S*-allyl cysteine. This is a new addition to a rapidly expanding interest in metal-mediated protein modification[15] and a new standard in substrate complexity in olefin metathesis. The opportunities and challenges uncovered during these efforts will be welcomed as our lab and others build from these initial discoveries. We conclude with an outlook over these possibilities that necessitate focused effort in both chemistry and biology.

Outlook

The enhanced reactivities of allyl sulfides in aqueous cross metathesis enable olefin metathesis on protein surfaces. Whereas we have used *S*-allylcysteine, other allyl sulfides will likely find use in metathesis. Accordingly, there are many opportunities in unnatural amino acid incorporation into proteins. Other privileged handles that assist metathesis in the same way as allyl sulfides may also be discovered. Many challenges in olefin metathesis on proteins are worthy of discussion, and several insightful points have already been raised by Kirshenbaum and Arora.[101] Firstly, the necessary accessibility of amino acid metathesis partners has not been fully assessed. Is it possible to modify a hindered active site or helix? This is yet to be determined. Secondly, *E*/*Z* mixtures may result after metathesis, but at present determination of the stereoselectivity on proteins is a challenge. Hydrogenation of these side chains may be necessary for complete homogeneity. Thirdly, the use of cosolvents to solubilize commercially available metathesis (pre)catalysts also warrants attention. The use of water-soluble metathesis (pre)catalysts could avoid the use of organic cosolvents that might not be compatible with certain protein samples. Moreover, new water-compatible, reactive catalysts may eventually allow metathesis at residues such as Hag that are at present not sufficiently reactive in water. Fourthly, a relatively limited set of metathesis partners was screened in our preliminary report. The full scope and functional group tolerance is determined not by the protein alone, but also by the metathesis partner and its unique steric and electronic character. Ultimately, with translation capacity of unsaturated amino acid residues, efficient water-soluble catalysts, and a more detailed understanding of optimal metathesis partners, the opportunity for in vivo protein modification by olefin metathesis may become reality. With access to these modified proteins, their use in deciphering and influencing a number of biochemical processes becomes possible. These questions are driving our current efforts to assess fully the scope of olefin metathesis in protein modification and its use in biology.

Acknowledgements

We gratefully acknowledge the EPSRC (Y.A.L.), and the Rhodes Trust (J.M.C.) for financial support.

Keywords: allyl sulfides · aqueous chemistry · olefin metathesis · protein modifications · unnatural amino acids

[1] I. S. Carrico, *Chem. Soc. Rev.* **2008**, *37*, 1423–1431.
[2] T. L. Foley, M. D. Burkart, *Curr. Opin. Chem. Biol.* **2007**, *11*, 12–19.
[3] R. D. Astronomo, H.-K. Lee, C. N. Scanlan, R. Pantophlet, C.-Y. Huang, I. A. Wilson, O. Blixt, R. A. Dwek, C.-H. Wong, D. R. Burton, *J. Virol.* **2008**, *82*, 6359–6368.
[4] L. M. Krug, G. Ragupathi, K. K. Ng, C. Hood, H. J. Jennings, Z. Guo, M. G. Kris, V. Miller, B. Pizzo, L. Tyson, V. Baez, P. O. Livingston, *Clin. Cancer Res.* **2004**, *10*, 916–923.
[5] P. H. Seeberger, D. B. Werz, *Nature* **2007**, *446*, 1046–1051.
[6] V. Verez-Bencomo, V. Fernández-Santana, E. Hardy, M. E. Toledo, M. C. Rodríguez, L. Heynngnezz, A. Rodriguez, A. Baly, L. Herrera, M. Izquierdo, A. Villar, Y. Valdés, K. Cosme, M. L. Deler, M. Montane, E. Garcia, A. Ramos, A. Aguilar, E. Medina, G. Toraño, I. Sosa, I. Hernandez, R. Martínez, A. Muzachio, A. Carmenates, L. Costa, F. Cardoso, C. Campa, M. Diaz, R. Roy, *Science* **2004**, *305*, 522–525.
[7] B. G. Davis, *Science* **2004**, *303*, 480–482.
[8] D. P. Gamblin, S. I. van Kasteren, J. M. Chalker, B. G. Davis, *FEBS J.* **2008**, *275*, 1949–1959.
[9] D. Qi, C.-M. Tann, D. Haring, M. D. Distefano, *Chem. Rev.* **2001**, *101*, 3081–3111.
[10] G. T. Hermanson, *Bioconjugation Techniques*, 2nd ed., Academic Press, San Diego, **2008**.
[11] E. Y. Chi, S. Krishnan, T. W. Randolph, J. F. Carpenter, *Pharm. Res.* **2003**, *20*, 1325–1336.
[12] L. S. Hegedus, *Transition Metals in the Synthesis of Complex Organic Molecules*, 2nd ed., University Science Books, Sausalito, **1999**.
[13] C. I. Herrerías, X. Yao, Z. Li, C.-J. Li, *Chem. Rev.* **2007**, *107*, 2546–2562.
[14] C.-J. Li, *Chem. Rev.* **2005**, *105*, 3095–3166.
[15] J. M. Antos, M. B. Francis, *Curr. Opin. Chem. Biol.* **2006**, *10*, 253–262.
[16] J. M. Antos, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 10256–10257.
[17] S. D. Tilley, M. B. Francis, *J. Am. Chem. Soc.* **2006**, *128*, 1080–1081.
[18] I. S. Carrico, *Ph.D. Thesis*, etd-09082003–09110526, California Institute of Technology (USA), **2004**.
[19] K. Kodama, S. Fukuzawa, H. Nakayama, T. Kigawa, K. Sakamoto, T. Yabuki, N. Matsuda, M. Shirouzu, K. Takio, K. Tachibana, S. Yokoyama, *ChemBioChem* **2006**, *7*, 134–139.
[20] K. Kodama, S. Fukuzawa, H. Nakayama, K. Sakamoto, T. Kigawa, T. Yabuki, N. Matsuda, M. Shirouzu, K. Takio, S. Yokoyama, K. Tachibana, *ChemBioChem* **2007**, *8*, 232–238.
[21] A. Ojida, H. Tsutsumi, N. Kasagi, I. Hamachi, *Tetrahedron Lett.* **2005**, *46*, 3301–3305.
[22] E. Brustad, M. L. Bushey, J. W. Lee, D. Groff, W. Liu, P. G. Schultz, *Angew. Chem.* **2008**, *120*, 8344–8347; *Angew. Chem. Int. Ed.* **2008**, *47*, 8220–8223.
[23] Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J. Am. Chem. Soc.* **2003**, *125*, 3192–3193.
[24] S. Sen Gupta, J. Kuzelka, P. Singh, W. G. Lewis, M. Manchester, M. G. Finn, *Bioconjugate Chem.* **2005**, *16*, 1572–1579.
[25] P.-C. Lin, S.-H. Ueng, M.-C. Tseng, J.-L. Ko, K.-T. Huang, S.-C. Yu, A. K. Adak, Y.-J. Chen, C.-C. Lin, *Angew. Chem.* **2006**, *118*, 4392–4396; *Angew. Chem. Int. Ed.* **2006**, *45*, 4286–4290.
[26] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105–1109.
[27] R. H. Grubbs, *Handbook of Metathesis*, Wiley-VCH, Weinheim, **2003**.
[28] A. H. Hoveyda, A. R. Zhugralin, *Nature* **2007**, *450*, 243–251.
[29] A. Fürstner, *Angew. Chem.* **2000**, *112*, 3140–3172; *Angew. Chem. Int. Ed.* **2000**, *39*, 3012–3043.
[30] S. J. Connon, S. Blechert, *Angew. Chem.* **2003**, *115*, 1944–1968; *Angew. Chem. Int. Ed.* **2003**, *42*, 1900–1923.
[31] For a recent review on olefin metathesis in chemical biology: J. B. Binder, R. T. Raines, *Curr. Opin. Chem. Biol.* **2008**, *12*, 767–773.
[32] Y. A. Lin, J. M. Chalker, N. Floyd, G. J. L. Bernardes, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 9642–9643.
[33] J. C. M. van Hest, D. A. Tirrell, *FEBS Lett.* **1998**, *428*, 68–70.

[34] J. C. M. Van Hest, K. L. Kiick, D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, *122*, 1282–1288.
[35] A. Brik, *Adv. Synth. Catal.* **2008**, *350*, 1661–1675.
[36] T. D. Clark, M. R. Ghadiri, *J. Am. Chem. Soc.* **1995**, *117*, 12364–12365.
[37] S. J. Miller, R. H. Grubbs, *J. Am. Chem. Soc.* **1995**, *117*, 5855–5856.
[38] S. J. Miller, H. E. Blackwell, R. H. Grubbs, *J. Am. Chem. Soc.* **1996**, *118*, 9606–9614.
[39] F. P. J. T. Rutjes, H. E. Schoemaker, *Tetrahedron Lett.* **1997**, *38*, 677–680.
[40] J. Pernerstorfer, M. Schuster, S. Blechert, *Chem. Commun.* **1997**, 1949–1950.
[41] B. E. Fink, P. R. Kym, J. A. Katzenellenbogen, *J. Am. Chem. Soc.* **1998**, *120*, 4334–4344.
[42] G. D. Rose, L. M. Glerasch, J. A. Smith, *Adv. Protein Chem.* **1985**, *37*, 1–109.
[43] L. K. Henchey, A. L. Jochim, P. S. Arora, *Curr. Opin. Chem. Biol.* **2008**, *12*, 692–697.
[44] H. E. Blackwell, R. H. Grubbs, *Angew. Chem.* **1998**, *110*, 3469–3472; *Angew. Chem. Int. Ed.* **1998**, *37*, 3281–3284.
[45] H. E. Blackwell, J. D. Sadowsky, R. J. Howard, J. N. Sampson, J. A. Chao, W. E. Steinmetz, D. J. O'Leary, R. H. Grubbs, *J. Org. Chem.* **2001**, *66*, 5291–5302.
[46] C. E. Schafmeister, J. Po, G. L. Verdine, *J. Am. Chem. Soc.* **2000**, *122*, 5891–5892.
[47] J. L. Stymiest, B. F. Mitchell, S. Wong, J. C. Vederas, *Org. Lett.* **2003**, *5*, 47–49.
[48] J. L. Stymiest, B. F. Mitchell, S. Wong, J. C. Vederas, *J. Org. Chem.* **2005**, *70*, 7799–7809.
[49] N. Ghalit, J. F. Reichwein, H. W. Hilbers, E. Breukink, D. T. S. Rijkers, R. M. J. Liskamp, *ChemBioChem* **2007**, *8*, 1540–1554.
[50] S. E. Gibson, V. C. Gibson, S. P. Keen, *Chem. Commun.* **1997**, 1107–1108.
[51] S. C. G. Biagini, S. E. Gibson, S. P. Keen, *J. Chem. Soc. Perkin Trans. 1* **1998**, 2485–2499.
[52] O. Brümmer, A. Rückert, S. Blechert, *Chem. Eur. J.* **1997**, *3*, 441–446.
[53] Y.-J. Hu, R. Roy, *Tetrahedron Lett.* **1999**, *40*, 3305–3308.
[54] G. J. McGarvey, T. E. Benedum, F. W. Schmidtmann, *Org. Lett.* **2002**, *4*, 3591–3594.
[55] P. Brüchner, D. Koch, U. Voigtmann, S. Blechert, *Synth. Commun.* **2007**, *37*, 2757–2769.
[56] D. Burtscher, K. Grela, *Angew. Chem.* **2009**, *121*, 450–462; *Angew. Chem. Int. Ed.* **2009**, *48*, 442–454.
[57] S. H. Hong, R. H. Grubbs, *J. Am. Chem. Soc.* **2006**, *128*, 3508–3509.
[58] J. P. Jordan, R. H. Grubbs, *Angew. Chem.* **2007**, *119*, 5244–5247; *Angew. Chem. Int. Ed.* **2007**, *46*, 5152–5155.
[59] A. Michrowska, Ł. Gułajski, Z. Kaczmarska, K. Mennecke, A. Kirschning, K. Grela, *Green Chem.* **2006**, *8*, 685–688.
[60] Ł. Gułajski, A. Michrowska, J. Narożnik, Z. Kaczmarska, L. Rupnicki, K. Grela, *ChemSusChem* **2008**, *1*, 103–109.
[61] J. B. Binder, I. A. Guzei, R. T. Raines, *Adv. Synth. Catal.* **2007**, *349*, 395–404.
[62] S. J. Connon, M. Rivard, M. Zaja, S. Blechert, *Adv. Synth. Catal.* **2003**, *345*, 572–575.
[63] J. B. Binder, J. J. Blank, R. T. Raines, *Org. Lett.* **2007**, *9*, 4885–4888.
[64] B. H. Lipshutz, G. T. Aguinaldo, S. Ghorai, K. Voigtritter, *Org. Lett.* **2008**, *10*, 1325–1328.
[65] D. M. Lynn, B. Mohr, R. H. Grubbs, L. M. Henling, M. W. Day, *J. Am. Chem. Soc.* **2000**, *122*, 6601–6609.
[66] M. B. Dinger, J. C. Mol, *Organometallics* **2003**, *22*, 1089–1095.
[67] M. B. Dinger, J. C. Mol, *Eur. J. Inorg. Chem.* **2003**, 2827–2833.
[68] S. K. Armstrong, B. A. Christie, *Tetrahedron Lett.* **1996**, *37*, 9373–9376.
[69] Y.-S. Shon, T. R. Lee, *Tetrahedron Lett.* **1997**, *38*, 1283–1286.
[70] J. L. Mascareñas, A. Rumbo, L. Castedo, *J. Org. Chem.* **1997**, *62*, 8620–8621.
[71] A. Fürstner, G. Seidel, N. Kindler, *Tetrahedron* **1999**, *55*, 8215–8230.
[72] F. D. Toste, A. K. Chatterjee, R. H. Grubbs, *Pure Appl. Chem.* **2002**, *74*, 7–10.
[73] G. Spagnol, M.-P. Heck, S. P. Nolan, C. Mioskowski, *Org. Lett.* **2002**, *4*, 1767–1770.
[74] T. R. Hoye, H. Zhao, *Org. Lett.* **1999**, *1*, 1123–1125.
[75] A. Fürstner, K. Langemann, *Synthesis* **1997**, 792–803.
[76] D. H. Strumeyer, W. N. White, D. E. Koshland, Jr., *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 931–935.
[77] H. Weiner, W. N. White, D. G. Hoare, D. E. Koshland, Jr., *J. Am. Chem. Soc.* **1966**, *88*, 3851–3859.
[78] T. J. Holmes, R. G. Lawton, *J. Am. Chem. Soc.* **1977**, *99*, 1984–1986.
[79] J. Wang, S. M. Schiller, P. G. Schultz, *Angew. Chem.* **2007**, *119*, 6973–6975; *Angew. Chem. Int. Ed.* **2007**, *46*, 6849–6851.
[80] G. J. L. Bernardes, J. M. Chalker, J. C. Errey, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 5052–5053.
[81] Y. Zhu, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1189–1192.
[82] D. P. Galonić, W. A. van der Donk, D. Y. Gin, *Chem. Eur. J.* **2003**, *9*, 5997–6006.
[83] U. Schmidt, E. Öhler, *Angew. Chem.* **1976**, *88*, 54; *Angew. Chem. Int. Ed. Engl.* **1976**, *15*, 42.
[84] M. W. Crankshaw, G. A. Grant, *Modification at Cysteine*, Wiley, Hoboken, **1996**, Unit 15.1.
[85] D. Crich, V. Krishnamurthy, T. K. Hutton, *J. Am. Chem. Soc.* **2006**, *128*, 2544–2545.
[86] D. Crich, F. Brebion, V. Krishnamurthy, *Org. Lett.* **2006**, *8*, 3593–3596.
[87] D. Crich, V. Krishnamurthy, F. Brebion, M. Karatholuvhu, V. Subramanian, T. K. Hutton, *J. Am. Chem. Soc.* **2007**, *129*, 10282–10294.
[88] L. Wang, P. G. Schultz, *Angew. Chem.* **2005**, *117*, 34–68; *Angew. Chem. Int. Ed.* **2005**, *44*, 34–66.
[89] A. J. Link, M. L. Mock, D. A. Tirrell, *Curr. Opin. Biotechnol.* **2003**, *14*, 603–609.
[90] T. Hohsaka, M. Sisido, *Curr. Opin. Chem. Biol.* **2002**, *6*, 809–815.
[91] K. L. Kiick, E. Saxon, D. A. Tirrell, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 19–24.
[92] J. Xie, P. G. Schultz, *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775–782.
[93] C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, P. G. Schultz, *Science* **1989**, *244*, 182–188.
[94] A. Fürstner, K. Langemann, *J. Am. Chem. Soc.* **1997**, *119*, 9130–9136.
[95] P. Wipf, W. S. Weiner, *J. Org. Chem.* **1999**, *64*, 5321–5324.
[96] Q. Yang, W.-J. Xiao, Z. Yu, *Org. Lett.* **2005**, *7*, 871–874.
[97] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.
[98] K. Ohtsubo, J. D. Marth, *Cell* **2006**, *126*, 855–867.
[99] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579–601.
[100] J. M. Harris, R. B. Chess, *Nat. Rev. Drug Discovery* **2003**, *2*, 214–221.
[101] K. Kirshenbaum, P. S. Arora, *Nat. Chem. Biol.* **2008**, *4*, 527–528.

Received: January 3, 2009

