Organic LETTERS

Letter

pubs.acs.org/OrgLett

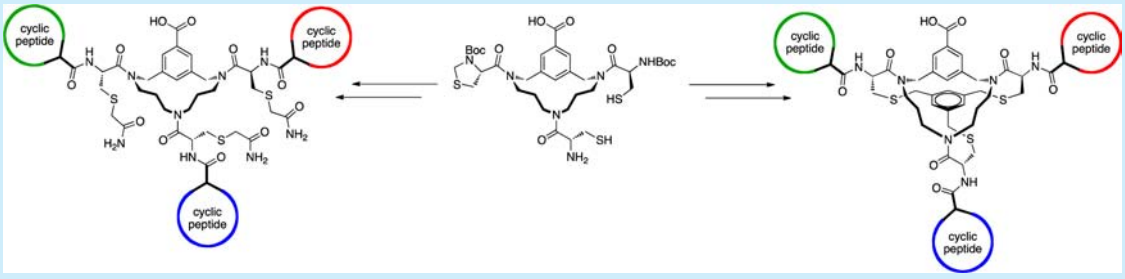
Efficient Synthesis of Protein Mimics by Sequential Native Chemical Ligation

Helmus van de Langemheen,[†] H. (Linda) C. Quarles van Ufford,[†] John A. W. Kruijtzer,[†] and Rob M. J. Liskamp*[,†,‡]

[†]Medicinal Chemistry & Chemical Biology, Utrecht Institute for Pharmaceutical Sciences, Department of Pharmaceutical Sciences, Faculty of Sciences, Utrecht University, P.O. Box 80082, 3508 TB, Utrecht, The Netherlands

[‡]School of Chemistry, Joseph Black Building, University of Glasgow, University Avenue, Glasgow, G12 8QQ, United Kingdom

Supporting Information



ABSTRACT: Synthetic mimics of protein surfaces have the potential to become inhibitors of protein−protein interactions or even synthetic vaccines. However, the synthesis of these complicated molecular constructs is still difficult. Here we describe an efficient and versatile synthesis of protein mimics containing up to three different cyclic peptides. Using a sequential native chemical ligation strategy, peptide loops containing a thioester handle were introduced onto a triazacyclophane scaffold bearing orthogonal protected cysteine residues.

The enormous diversity of molecular character of protein surfaces offers fantastic challenges for developing synthetic approaches for their mimicry by significantly smaller molecular constructs than the original proteins. This diversity is evident from greatly varying molecular patches or areas on their surface and responsible for a whole range of protein−protein or protein−ligand interactions. In the latter interactions the ligand can be virtually any small to large molecule. The most striking examples of protein−ligand interactions are probably antibody−antigen interactions, in which molecularly different areas of an antibody, reminiscent to different peptide segments present in the antigen binding loops, the so-called complementary determining regions, are usually capable of high affinity and highly selective binding of a peptide, protein, carbohydrate, nucleic acid or even small molecule antigen. Increasingly the different areas or sites of a protein, which together are essential for binding of a ligand, are indicated by the term "discontinuous epitope".

As part of a program to address the synthetic challenges for construction of protein mimics encompassing different peptide segments corresponding to different sites on the surface of a protein, we have been involved in (1) the development of syntheses of different scaffolds for attachment of (cyclic) peptides,[1] (2) synthetic approaches for attachment of different (cyclic) peptides to scaffolds,[2] and (3) the generation of libraries of the resulting protein mimics.[3]

In developing synthetic approaches for the construction of these relatively complex biomolecular constructs of "intermediate size",[4] increasingly accompanying challenges are found in the inclusion of practicality and efficiency of the synthesis in terms of yields and purities but also with respect to rapid expansion to collections or libraries.

To our knowledge, the availability of convergent methods for introducing both different and several peptide loops onto suitable scaffolds is still very limited. The most seminal contribution in this area was probably the template-assembled synthetic protein (TASP) originally developed by Mutter and Vuilleumier[6] and further developed by, among others, Dumy et al. to the so-called regioselectively addressable functionalized template (RAFT).[7] Another elegant example by Beyermann et al.[8] described the introduction of three different cyclic peptides, albeit without lysine residues, onto a linear peptide scaffold.[8]

Here we describe a very efficient and versatile synthesis of defined protein mimics containing up to three different cyclic peptide loops. We think that this *convergent* synthetic approach will open up a plethora of possibilities for construction of all kinds of protein mimics varying from, for example, the preparation of synthetic vaccines to perhaps even synthetic antibodies.

Received: February 25, 2014
Published: April 7, 2014

dx.doi.org/10.1021/ol500604h | *Org. Lett.* 2014, 16, 2138−2141

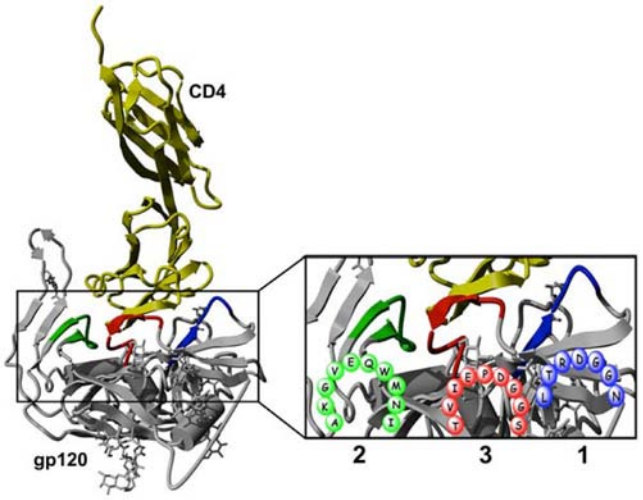


Figure 1. Conserved discontinuous epitope of the CD4 binding site on gp120.[2c,10]

As an illustration toward the former application, we have incorporated here by native chemical ligation (NCL)[9] different cyclic peptides in the synthesized protein mimics, which correspond to the sequences of the peptide segments of loops in gp120 together forming the discontinuous epitope interacting with CD4 (Figure 1).[2c,10]

We started from our earlier developed semiorthogonally protected triazacyclophane (TAC) scaffold (**1**, Figure 2),[1b] which we have also used for the nonstop solid phase synthesis of three different cyclic peptides on the TAC-scaffold corresponding to loops of the discontinuous epitope of HIV-gp120.2c. However, this approach has the inherent disadvantage of utilizing a linear continuous peptide synthesis strategies and requires, even with high yields per coupling step, extensive purification yielding relatively small amounts of the desired protein mimic. This scaffold (**1**) was now converted to scaffold **6**, which is amenable to a *convergent* synthetic approach in which by native chemical ligation three cyclic peptides can be introduced subsequently.

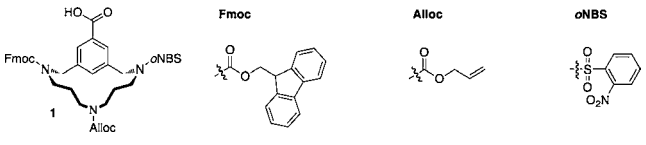
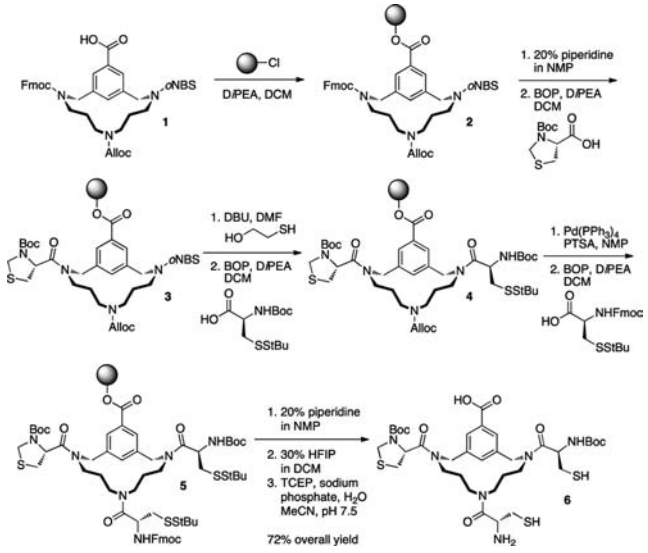


Figure 2. Structure of the triazacyclophane (TAC) scaffold.[1b]

To be able to introduce three different cyclic peptides on the TAC-scaffold by NCL, three selectively addressable cysteine residues were required. This was achieved by having one cysteine residue unprotected, one cysteine residue with its amino group protected with the acid labile Boc-group, and both the amino group and thiol group of the third residue protected as a thiazolidine (Thz)[11] moiety. The unprotected cysteine was of course immediately capable of undergoing NCL; the Boc-protected cysteine residue had a free thiol, but needed the free amino group for the S to N-acyl shift for completion of the NCL reaction, which became only available after treatment with TFA. The third cysteine residue underwent NCL after treatment with methoxy amine hydrochloride liberating both amino and thiol group.

The synthesis (Scheme 1) of the required tricysteine containing scaffold for stepwise NCL started by attachment of the TAC-template (**1**)[1b] to a 2-chlorotrityl resin, using D*i*PEA as the base to yield resin-bound TAC-scaffold **2**. Next, the Fmoc protective group was cleaved, followed by the coupling of the thiazolidine residue affording resin-bound scaffold **3**. Next, the *o*NBS-group of **3** was removed by treatment with 2-mercaptoethanol and DBU, followed by coupling with Boc-Cys(StBu)-OH to yield resin-bound scaffold **4**. Palladium-assisted cleavage of the Alloc protective group allowed the introduction of the last cysteine, still completely protected, residue to yield resin-bound protected scaffold **5**. This last cysteine residue will be the residue for the first NCL-step. Its Fmoc-group was removed by treatment with piperidine, followed by cleavage of the scaffold from the resin using 30% hexafluoroisopropanol (HFIP) in DCM. Finally, the two StBu sulfhydryl protective groups were removed by treatment with tris(2-carboxyethyl)-phosphine hydrochloride (TCEP) and the resulting scaffold **6**, which was now ready for conducting stepwise NCL, was purified by preparative HPLC and obtained in an excellent overall yield of 72% (Scheme 1).

Scheme 1. Synthesis of TAC-Scaffold 6



In this synthesis strategy, it was not possible to introduce the thiazolidine residue with an Fmoc protective group on the N-terminus instead of a Boc-group, since the next reaction step, the *o*NBS cleavage, would also lead to cleavage of the Fmoc group. Because our designed sequential NCL strategy already included a Boc deprotection step, it was decided to couple a Boc protected thiazolidine to the scaffold. By doing so, both Boc groups, from the cysteine and thiazolidine residues, will be cleaved simultaneously.

Scaffold **6** was now ready for successive NCL with different cyclic peptide thioesters for which an efficient solid phase synthesis method was recently described.[12] Using this approach cyclic peptide thioesters having amino acid sequences LTRDGGN (**7a**, loop 1), INMWQEVGKA (**7b**, loop 2), and SGGDPEIVT (**7c**, loop 3) were obtained (Figure 3).

Since the "left" and "right" positions are interchangeable in the scaffold because of the possibility of ring-flipping of the cyclophane ring, the protein mimic construct [1−2−3] is identical to the molecular construct [3−2−1], in which the

dx.doi.org/10.1021/ol500604h | *Org. Lett.* 2014, 16, 2138−2141

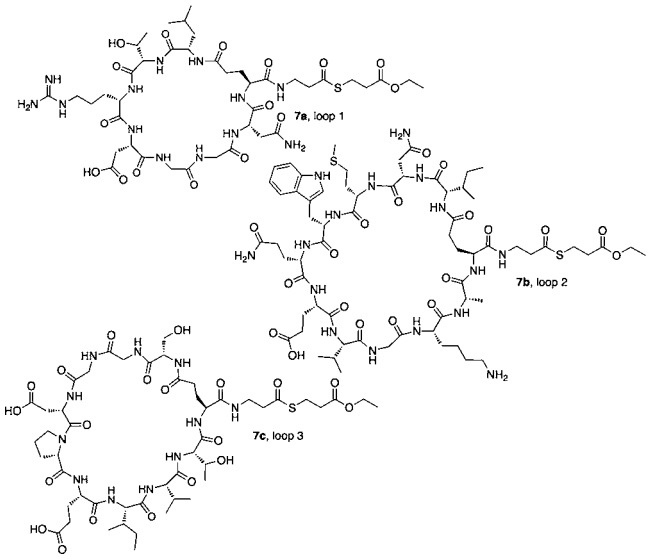


Figure 3. Structures of the LTRDGGN (**7a**, loop 1), INMWQEVGKA (**7b**, loop 2), and SGGDPEIVT (**7c**, loop 3) gp120 peptide thioester loops.[12]

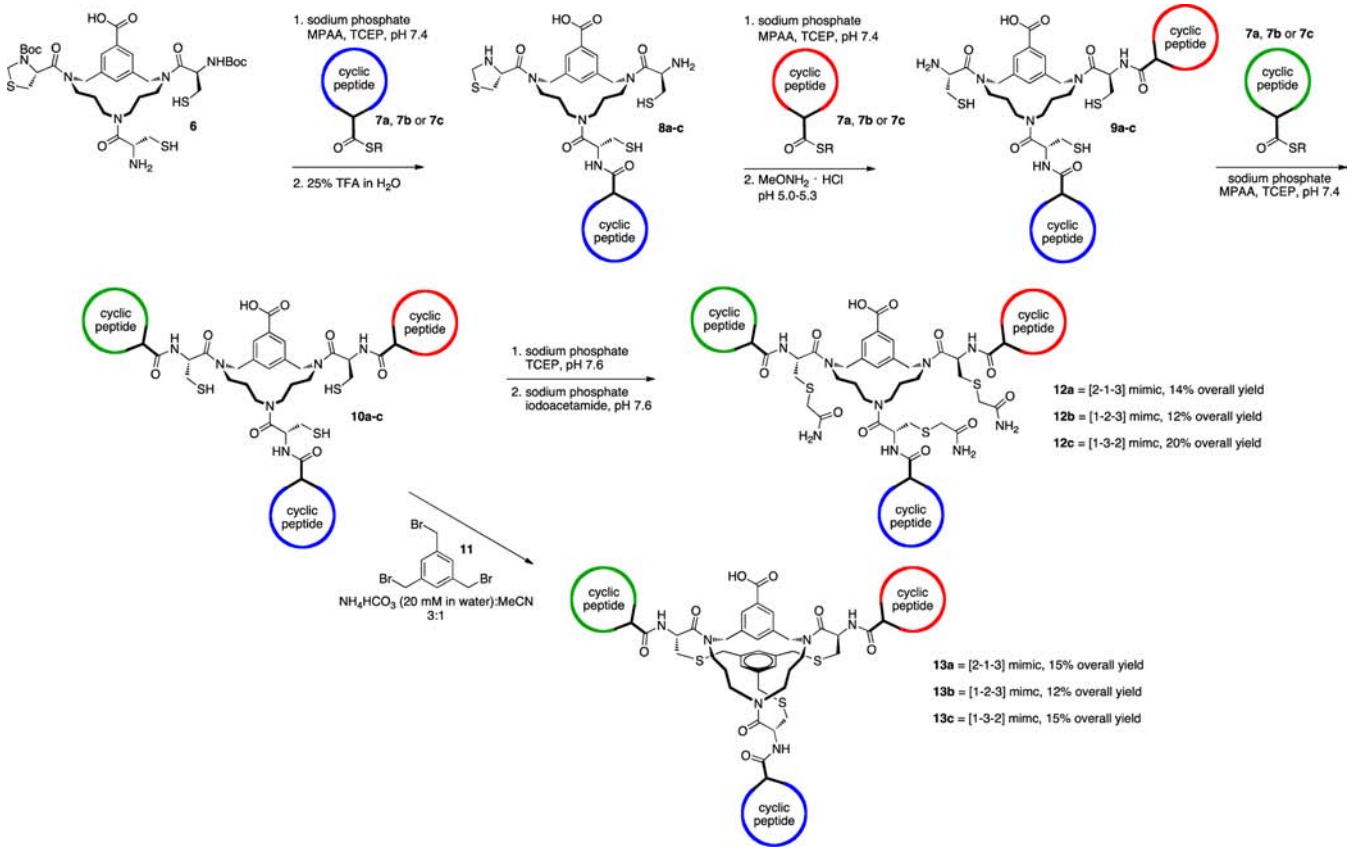
number indicates the peptide loop number (Figure 3). Thus, the following protein mimics were synthesized by successive NCL: [2−1−3], [1−2−3], and [1−3−2] as will be described below.

The first NCL reaction with the unprotected cysteine in TAC-scaffold **6** (3 μmol, 1 mM) was carried out with cyclic peptide thioester **7a**, **7b**, or **7c** (3 μmol, 1 mM) in ligation buffer (200 mM sodium phosphate, 200 mM 4-mercaptophenylacetic acid (MPAA), 40 mM TCEP, and pH 7.4) to give the single ligated product. Next, both Boc protective groups were removed using 25% TFA in water to yield the deprotected single ligated products **8a−c**. Then the second cyclic peptide was attached to the scaffold by NCL, followed by the thiazolidine ring-opening-deprotection of the third cysteine residue by methoxyamine hydrochloride to yield **9a−c**. Finally, NCL incorporation of the third peptide loop afforded the triple ligated protein gp120 mimics **10a−c** (Scheme 2).

For biological evaluation purposes, it is sensible to alkylate ("cap") the resulting thiols of the cysteine residues present in these gp120 protein mimics. This was achieved using iodoacetamide or 1,3,5-tris(bromomethyl)benzene (**11**). The latter reagent was used earlier by Timmerman et al.[13] for the simultaneous alkylation of three cysteine residues in a peptide, which we wanted to achieve too with our protein mimics containing three cysteine residues. Iodoacetamide alkylation gave the iodoacetamide alkylated gp120 mimics [2−1−3] (**12a**), [1−2−3] (**12b**), and [1−3−2] (**12c**) in overall yields of 14, 12, and 20%, respectively, corresponding to good average yields per step of 75, 73, and 79% (Scheme 2).

Fortuitously, even treatment with excess with tris(bromomethyl)benzene **11** led to monoalkylated protein gp120 mimics [2−1−3] (**13a**), [1−2−3] (**13b**), and [1−3−2] (**13c**) in overall yields of 15, 12, and 15%, respectively (corresponding to an average yield per step of 76, 73, and 76%), in which all three cysteine thiols had reacted (Scheme 2). It was expected that this simultaneous alkylation will constrain the TAC-scaffold considerably. Preliminary molecular modeling results showed

Scheme 2. General Strategy for Performing Sequential NCL on TAC-Scaffold 6 to Obtain gp120 Mimics 12a−c and 13a−c



dx.doi.org/10.1021/ol500604h | *Org. Lett.* 2014, 16, 2138−2141

that this simultaneous alkylation led to a flattening of protein mimic construct.

The ability of these triple ligated molecular constructs to act as protein mimics of gp120 was evaluated in a gp120-capture ELISA assay. The binding of the gp120 discontinuous epitope mimics to CD4 was addressed in a gp120-capture ELISA experiment (Figure 4). The individual loops, TAC-scaffold as well as an equimolar mixture of loops gave hardly any inhibition of the gp120−CD4 interaction. These results indicated that an assembly of at least two cyclic peptides was required for biological acitivity The data obtained from the ELISA experiments of the gp120 protein mimics **12a−12c** (capped with iodoacetamide), showed a moderate inhibitory activity and, although statistically not significant, a trend of reduction of inhibitory activity may be observed going from **12a** to **12c**, which seemed to be less the case in the tris(bromomethyl)benzene alkylated protein gp120 mimics **13a−13c**. In the latter protein mimics, preliminary molecular experiments indicated a more flattened structure, possibly leading to more similar inhibitory activities.

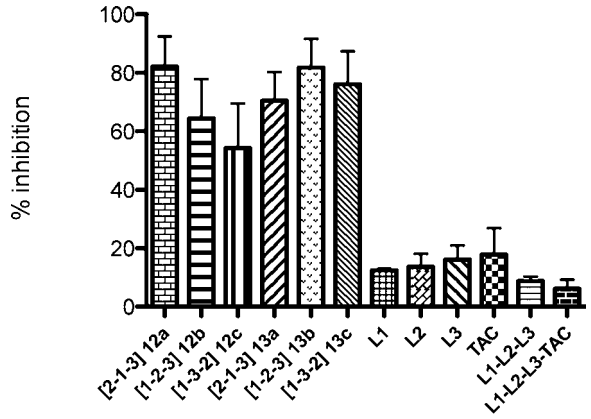


Figure 4. Results of the competitive gp120(IIIB)-CD4 ELISA for all gp120 mimics, different cyclic peptide thioesters loops, TAC-scaffold, equimolar mixture of the cyclic peptide thioesters and equimolar mixture of the cyclic peptide thioesters as well as the TAC-scaffold. The concentration of the mimics **12a−c**, the thioester loops (L1, L2 and L3), and TAC-scaffold was 31 μM. The concentration of the mimics **13a−c** was 32 μM. In the mixtures each individual component (L1, L2, L3, and TAC-scaffold) was present in a concentration of 31 μM.

In conclusion, we have designed an efficient synthetic strategy for the sequential introduction of cyclic peptides to a newly designed molecular scaffold. In this method, up to three different cyclic peptide thioesters can be introduced sequentially to this new three-cysteine derivative containing a modified TAC-scaffold, using the very versatile NCL reaction. However, it should be emphasized that this procedure is not limited to attachment of *peptide* thioesters, and extending the scope to other small to large thioesters is under present investigation. The required TAC-scaffold (**6**) was efficiently synthesized on the solid support and obtained in a high yield. Moreover, all synthesized gp120 protein mimics were obtained in good overall yields. Fortuitously, alkylating the sulfhydryl groups of protein mimics **10a−c** with 1,3,5-tris(bromomethyl)benzene gave only monoalkylated products. In fact this led to a novel constrained multicyclic scaffold, which might be more flat, reflecting the observed similar inhibitory activity of the resulting protein mimics. The sequential NCL strategy used for successful introduction of different cyclic peptides described in this paper may indicate its potential for the efficient chemical construction of other protein mimics of a variety of discontinuous epitope containing proteins and may ultimately lead even to synthetic antibodies by furnishing suitable scaffolds with the appropriate CDR-peptide loops.

■ ASSOCIATED CONTENT

Supporting Information

Experimental procedures, HPLC analytical data for **6**, **12a−c**, and **13a−c**, LC−MS analytical data for **8a−c**, **9a−c**, **10a−c**, **12a−c**, and **13a−c**, and MALDI-TOF analysis for **12a−c** and **13a−c**. Molecular modeling of the tris(bromomethyl)benzene alkylated TAC-scaffold. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: r.m.j.liskamp@uu.nl, robert.liskamp@glasgow.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was (partly) financed by Chemical Sciences of The Netherlands Organization for Scientific Research (NWO). We gratefully acknowledge Javier Sastre Toraño for the HRMS measurements.

■ REFERENCES

(1) (a) Rump, E. T.; Rijkers, D. T. S.; Hilbers, H. W.; De Groot, P. G.; Liskamp, R. M. J. *Chem.—Eur. J.* **2002**, *8*, 4613. (b) Opatz, T.; Liskamp, R. M. J. *Org. Lett.* **2001**, *3*, 3499.

(2) (a) Chamorro, C.; Liskamp, R. M. J. *J. Comb. Chem.* **2003**, *5*, 794. (b) Hijnen, M.; van Zoelen, D. J.; Chamorro, C.; Gageldonk, P.; Mooi, F. R.; Berbers, G.; Liskamp, R. M. J. *Vaccine* **2007**, *25*, 6807. (c) Chamorro, C.; Kruijtzer, J. A. W.; Farsaraki, M.; Balzarini, J.; Liskamp, R. M. J. *Chem. Commun.* **2009**, 821.

(3) (a) Mulder, G. E.; Kruijtzer, J. A. W.; Liskamp, R. M. J. *Chem. Commun.* **2012**, *48*, 10007. (b) Mulder, G. E.; Quarles van Ufford, H. C.; Ameijde, J. van; Brouwer, A. J.; Kruijtzer, J. A. W.; Liskamp, R. M. J. *Org. Biomol. Chem.* **2013**, *11*, 2676.

(4) This term is increasingly used by us and others[5] to distinguish these molecules from the so-called "small molecules" with a MW around 500 Da and the large "biologics" (nanobodies etc.) with MW around 15 000 Da and beyond.

(5) Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D. *Chem. Biol. Drug Des.* **2013**, *81*, 136.

(6) Mutter, M.; Vuilleumier, S. *Angew. Chem., Int. Ed.* **1989**, *28*, 535.

(7) (a) Boturyn, D.; Coll, J.-L.; Garanger, E.; Favrot, M.-C.; Dumy, P. *J. Am. Chem. Soc.* **2004**, *126*, 5730. (b) Singh, Y.; Dolphin, G. T.; Razkin, J.; Dumy, P. *ChemBioChem* **2006**, *7*, 1298.

(8) Pritz, S.; Kreatke, O.; Klose, A.; Klose, J.; Rothemund, S.; Fechner, K.; Bienert, M.; Beyermann, M. *Angew. Chem., Int. Ed.* **2008**, *47*, 3642.

(9) (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776. (b) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10068.

(10) (a) Franke, R.; Hirsch, T.; Overwin, H.; Eichler, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 1253. (b) Kwong, P. D.; Wyatt, R.; Robinson, J.; Sweet, R. W.; Sodroski, J.; Hendrickson, W. A. *Nature* **1998**, *393*, 648.

(11) (a) Villain, M.; Vizzavona, J.; Rose, K. *Chem. Biol.* **2001**, *8*, 673. (b) Bang, D.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534.

(12) Van de Langemheen, H.; Brouwer, A. J.; Kemmink, J.; Kruijtzer, J. A. W.; Liskamp, R. M. J. *J. Org. Chem.* **2012**, *77*, 10058.

(13) Timmerman, P.; Beld, J.; Puijk, W. C.; Meloen, R. H. *ChemBioChem* **2005**, *6*, 821.