Synthetic Enzymes

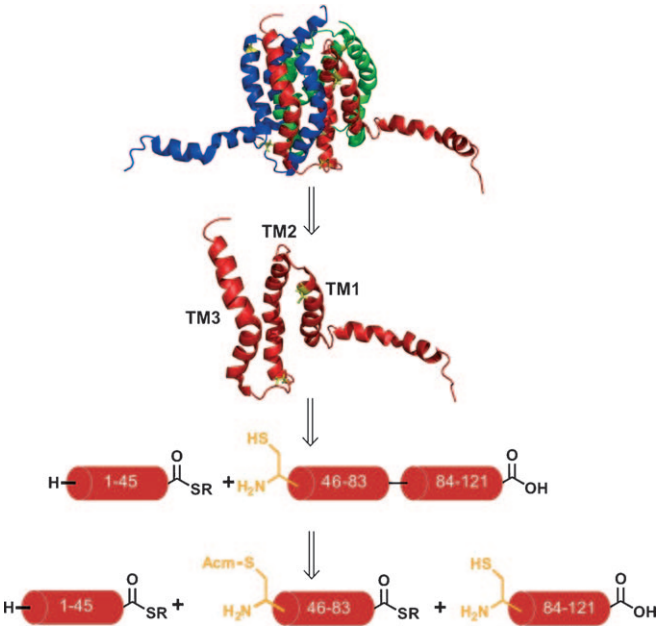
DOI: 10.1002/ange.201006686

Total Chemical Synthesis of an Integral Membrane Enzyme: Diacylglycerol Kinase from *Escherichia coli***

*Sunanda Lahiri, Manuel Brehs, Diana Olschewski, and Christian F. W. Becker**

Recent progress in chemical protein synthesis has provided access to many small to medium-sized proteins. However, the highly important class of membrane proteins comprising multimembrane-spanning receptors and ion channels as well as integral membrane enzymes remains elusive with regard to chemical synthesis. Only certain moderately sized membrane proteins have been generated by chemical protein synthesis or semisynthesis.[1–4] Synthesis of the constituent hydrophobic membrane-spanning peptide segments remains challenging owing to incomplete amino acid coupling steps and subsequent purification problems. Herein we report the chemical synthesis of *Escherichia coli* diacylglycerol kinase (DAGK), an integral membrane enzyme consisting of monomers with 121 amino acids each that form a homotrimer in a membrane environment as well as in detergent micelles.[5] These homotrimers catalyze the conversion of diacylglycerol into phosphatidic acid[6] and play a vital role in the lipid metabolism of Gram-negative bacteria, especially under conditions of environmental stress.[7–9]

DAGK is the smallest kinase described to date and is not related to other members of the large family of phosphotransferases. DAGK has served as a model system for the study of enzymatic processes in membranes, for stability testing, and for membrane-protein folding.[10–13] A high-resolution structure determined recently by NMR spectroscopy is in good agreement with previous predictions of DAGK topology with respect to its three transmembrane domains (TMs) and the N-terminal cytoplasmic domain.[14–16] This structure also revealed some specific features of the DAGK trimer, such as the high susceptibility of this integral membrane enzyme to point mutations that either affect its catalytic activity or lead to misfolding of the protein.[17] The DAGK homotrimer (molar mass of ≈40 kDa) contains three active sites that are created by swapping TM3 between subunits (Scheme 1). Even though much is known about the



Scheme 1. Synthetic strategy for *E. coli* diacylglycerol kinase (DAGK, PDB: 2kdc[14]). The DAGK monomer, which contains 121 amino acids, has been divided into three segments: DAGK1–45, DAGK46–83, and DAGK84–121, each of which is well within the reach of solid-phase peptide synthesis. Upon assembly of the three segments to the full-length DAGK1–121 by native chemical ligation (cysteine residues at ligation sites are indicated in yellow), the protein was transferred into detergent micelles, where assembly into the functional trimeric state occurs spontaneously.

folding and the overall three-dimensional structure of DAGK, many facets of the enzyme mechanism remain unclear. A robust synthetic strategy for DAGK would enable ready chemical manipulation of specific amino acids beyond the power of biological approaches, whereby the function of essential residues for substrate binding and phosphate transfer could be studied.

Owing to the lack of a high-resolution structure at the time, the synthesis described herein was planned solely on the basis of low-resolution structural information about DAGK topology and mutational analyses. We used these data to initially devise a synthetic route based on three segments that

[*] M. Brehs, Prof. Dr. C. F. W. Becker
Center of Integrated Protein Science and
Technische Universität München, Protein Chemistry
Lichtenbergstrasse 4, 85747 Garching (Germany)
Fax: (+49) 89-289-13345
E-mail: christian.becker@ch.tum.de
Homepage: http://www.ch.tum.de/proteinchemie
Dr. S. Lahiri,[+] Dr. D. Olschewski[#]
Max Planck Institute of Molecular Physiology
Otto-Hahn-Strasse 11, 44227 Dortmund (Germany)
[+] Current address:
Universität Freiburg, Institute of Physiology 2
Engesserstrasse 4, 79108 Freiburg (Germany)
[#] Current address: Lonza Ltd., 3930 Visp (Switzerland)

[**] We thank Katja Bäuml and Sascha Gentz for excellent technical assistance with peptide synthesis. We also thank Francis Lau for technical advice on establishing the DAGK activity assay. Financial support from the DFG is gratefully acknowledged.
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201006686.

are linked consecutively through native-chemical-ligation (NCL) reactions.

The retrosynthetic route depicted in Scheme 1 divides the 121 amino acid long DAGK monomer into three segments. Each of these segments comprises one transmembrane domain, as we wished to avoid the solid-phase synthesis of peptides with two membrane-spanning domains. Fortunately, at position 46, a native cysteine residue is located in an ideal position for a native-chemical-ligation reaction. However, no native cysteine residue exists between TM2 and TM3 that would lend itself as a second ligation site. Therefore, a cysteine residue was used in place of serine residue 84. Such a conservative mutation is well-tolerated in many proteins and often does not lead to changes in structure and function. To the best of our knowledge, no cysteine mutant of serine 84 has been described previously. On the basis of our results, this mutation has no detectable effect on the folding and function of DAGK (see below). The C-terminal amino acids were an alanine residue in position 45 of DAGK segment 1 and a glycine residue in position 83 of segment 2. Both residues are well-suited as thioester-carrying C-terminal amino acids owing to their fast ligation kinetics and low steric hindrance.[18] Such effects are crucial when native-chemical-ligation reactions are carried out with hydrophobic peptides that are difficult to solubilize. Favorable reaction conditions are of paramount importance for high reaction yields to be attained before precipitation and hydrolysis remove reactive components from the ligation mixture.

The preparation of milligram amounts of sufficiently pure peptide segments is a major concern in the synthesis of membrane proteins. We used solid-phase peptide synthesis based on *N*-butoxycarbonyl (Boc) protection with in situ neutralization and activation with *O*-(benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HBTU) for all three segments.[19] Optimized synthetic procedures were required, including the use of double coupling reactions for selected residues (see the Supporting Information). Milligram amounts of DAGK46–83 (the N-terminal cysteine residue carried an acetamidomethyl (Acm) thiol protecting group) and DAGK84–121 were obtained in good purity. We added trifluoroethanol (TFE) to acetonitrile–water mixtures as a solubilizing strategy for both crude peptide segments. Reversed-phase (RP) HPLC purification at 60 °C with buffers containing isopropanol gave the two peptides in final yields of 15 and 18 % (see Figures S1 and S2 in the Supporting Information).

The two peptide segments underwent native chemical ligation, each at a concentration of 0.5 M, in 8 M urea buffered at pH 7.8 with 300 mM sodium phosphate in the presence of 50 mM dodecylphosphocholine (DPC) or 20 mM octylglycoside (OG) as a solubilizing detergent.[4,20] Thiophenol (1 %) was used as a ligation mediator in combination with ethanethiol (0.5 %). The ligation product DAGK46–121 was obtained in 50 % yield after a reaction time of 36 h and purification by RP HPLC (Figure 1 A). Acm was used as the thiol-protecting group because we were not able to quantitatively remove the more convenient thiazolidine protecting group after its successful incorporation at the N terminus of DAGK46–83.[21] The Acm group was removed quantitatively by treatment with mercury(II) acetate in acetonitrile/water mixtures containing 3 % acetic acid.[22]

We also tested solubilizing strategies for DAGK45–83 to improve the handling properties of this segment. Increased solubility could translate into increased ligation yields. To this end, we prepared DAGK46–83 segments with two different C-terminal solubilization tags: a poly(ethyleneglycol)–polyamide (PPO) tag and a hexaarginine tag (see Figures S5 and S6 in the Supporting Information). Both of these tags have been used successfully to solubilize a variety of membrane peptides and also for DAGK peptide segments.[23–25] These tags lead to an increase in solubility in less complex solvent mixtures, as shown by shorter retention times on C4 reversed-phase HPLC columns (see Figure S5 and S6). However, the overall yields of DAGK46–83-PPO_2 and DAGK46–83-Arg_6 were lower (6 and 8 %, respectively) than those of the unmodified segments, which offset any advantages due to increased solubility.

The synthesis of the N-terminal segment DAGK1–45 was more challenging, initially as a result of the low quality of the crude peptide, which contained many deletion products. However, handling of the resulting peptide in any solvent system suitable for purification turned out to be even more problematic. We tried to overcome these problems by modifying our synthetic strategy. The DAGK1–45 segment was split into two parts: DAGK1–30 and DAGK31–35. This new synthetic approach involved four peptide segments and three NCL reactions. It was mainly based on the assumption that the amphiphatic peptide DAGK1–30 would be easier to synthesize than DAGK1–45. Furthermore, the membrane-spanning segment DAGK31–45 would be solubilized by the attachment of a hexaarginine or PPO tag, as described for DAGK46–83. The shorter segments were both obtained by Boc-based solid-phase synthesis. DAGK1–30 was successfully purified in sufficient yields (21 %) for the subsequent assembly of full-length DAGK (see Figure S4 in the Supporting Information). DAGK31–45 was also obtained as crude material; however, sufficient purification of this segment was not possible in our hands even after the attachment of a C-terminal polymer-based solubilizing tag or a polyarginine tag (see Figures S7 and S8 in the Supporting Information). Therefore, we further improved the automated synthesis and purification of DAGK1–45 and finally obtained about 42 mg of purified DAGK1–45 from a synthesis carried out on a 0.2 mmol scale (5 % yield; see Figure S3 in the Supporting Information).

The final native chemical ligation in 8 M urea buffered at pH 7.8 with 300 mM sodium phosphate in the presence of 20 mM octylglycoside gave full-length DAGK1–121 in 58 % yield after 48 h at room temperature. Final purification by RP HPLC on a C4 column at 60 °C with a water–isopropanol gradient (see the Supporting Information) provided DAGK1–121, for which a single peak and a well-defined charge series were observed in the RP HPLC chromatogram and ESI mass spectrum, respectively (Figure 1 B). Furthermore, SDS-PAGE analysis of synthetic DAGK revealed only a single band, which was in good agreement with the expected molecular weight of approximately 13 kDa (see Figure S9 in the Supporting Information).

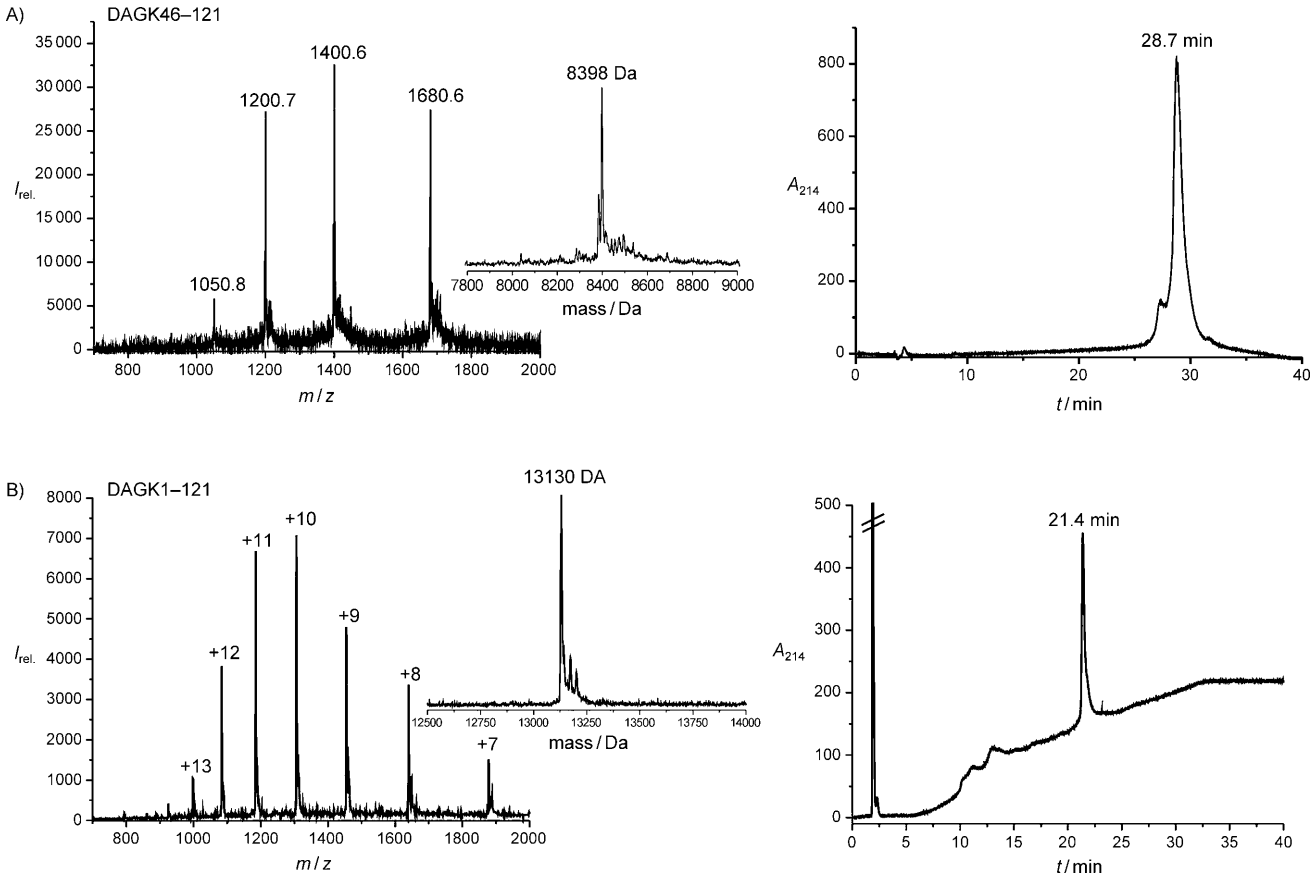


Figure 1. Analytical data for ligation products DAGK46–121 and full-length DAGK1–121. A) Left: Electrospray ionization (ESI) mass spectrum of DAGK46–121 after the first NCL reaction. The calculated molar mass of 8398.0 Da is in very good agreement with the observed mass for Acm-protected DAGK46–121. Right: RP HPLC of DAGK46–121 with a C4 column at 60 °C with a gradient of 5–80% buffer B in buffer A over 30 min. The HPLC trace indicates a purity of about 90%. B) Left: ESIMS of full-length DAGK1–121. The observed mass of 13130 Da is in very good agreement with the calculated mass of 13129.5 Da. Right: RP HPLC of DAGK1–121 with a C4 column at 60 °C with a gradient of 15% buffer C in buffer A to 100% C over 30 min. The HPLC trace indicates a purity of about 94%. The observed baseline shift always occurs with isopropanol-containing buffers (also with blank injections).

The purification of membrane peptides and proteins to high homogeneity remains challenging and needs to be optimized for individual peptides. The solid-phase peptide synthesis of DAGK segments 1–45, 46–83, and 84–121 in yields of more than 20 mg each provided the basis for this synthesis. Two subsequent NCL reactions with an overall yield of 29% (based on the C-terminal peptide segment DAGK84–121) gave a total of 4 mg of full-length DAGK. These results demonstrate the capabilities of Boc-based SPPS for the synthesis of amphiphatic and hydrophobic peptides, such as the three transmembrane domains and one amphiphilic cytoplasmic domain that constitute DAGK. In combination with RP HPCL and the judicious choice of solvents, peptides and the protein were purified without the need for solubilization tags, which often considerably improve peptide handling but can also impose new synthesis problems. Furthermore, such tags need to be removed after assembly of the full-length protein to exclude any influence on its function.[26]

The remaining challenge was now to fold the synthetic membrane protein into its functional quaternary structure. The unfolding and renaturation of *E. coli* wild-type (wt) DAGK has been studied previously, and it was demonstrated that folding is possible in detergent micelles without the need for lipid bilayers.[11] Synthetic DAGK was transferred directly from HPLC elution fractions containing the pure protein into trifluoroethanol (TFE) by adding 5 volume equivalents of TFE to DAGK-containing HPLC buffer and subsequent careful evaporation of the solvents. A similar procedure was described previously for the synthetic ion-channel protein MscL. It induced the formation of α helices in the previously unstructured protein.[4] Final folding was subsequently promoted by diluting DAGK concentrated in TFE into an aqueous buffer (150 mM NaCl, 25 mM NaP_i, pH 7.4) containing docecylmaltoside (DDM, 0.5%; see the Supporting Information). Folded DAGK was concentrated by ultrafiltration (molecular-weight cutoff: 5 kDa). These conditions for the folding of synthetic DAGK in detergent micelles consisting of DDM gave folded DAGK in about 65% yield.

The analysis of folded synthetic DAGK by CD spectroscopy revealed a mainly α-helical secondary structure, which is in very good agreement with previously reported data for

wt DAGK isolated from *E. coli* (Figure 2A). However, a predominantly α-helical secondary structure of monomeric DAGK can also be found for enzymatically inactive DAGK.[11] Therefore, the enzymatic activity of synthetic DAGK was tested in a coupled-enzyme assay.[11] We measured a k_{cat} value for the phosphoryl-transfer reaction of 8 s^{-1} (Figure 2B), which is in very good agreement with activities (ranging from 1 to 12 s^{-1}) reported for wt DAGK purified from *E. coli*.[6,27–29] The similarities in secondary structure and catalytic activity between synthetic DAGK and wt DAGK demonstrate that the chemically synthesized protein has adopted its native oligomeric structure.

The ability to synthesize, fold, and analyze the kinetics of integral membrane enzymes provides new possibilities for the detailed investigation of this class of proteins. Precise control over many different chemical modifications will allow direct insight into substrate binding, the catalytic mechanism of DAGK, its oligomeric assembly, and the role of specific amino acid side chains.[30–33] To this end, we prepared a first DAGK analogue containing a fluorescent probe (7-nitrobenz-2-oxa-1,3-diazole, NBD) linked to the side chain of L-diaminopropionic acid that replaces a native tryptophan residue in position 25 of DAGK and an N-terminal biotin moiety. This variant exhibits a similar k_{cat} value (7 ± 2 s^{-1}) to that measured for synthetic DAGK1–121 (see Figure S11 in the Supporting Information). The analogue synthesis is based on the robust procedure described herein and will be exploited to gain detailed mechanistic insight into DAGK function. This insight might become important not only for *E. coli* DAGK, but also for homologues from other bacteria, such as DAGK from *Streptococcus mutans*, the major cause of human dental caries. In *S. mutans*, the DAGK homologue and its kinase activity seem to be involved in virulence.[34] The correlation of functional groups in DAGK with their effect on folding and activity in this rather crude enzyme might also lead to general ideas about phosphotransferase reactions in more complex kinases, especially mammalian DAGKs.

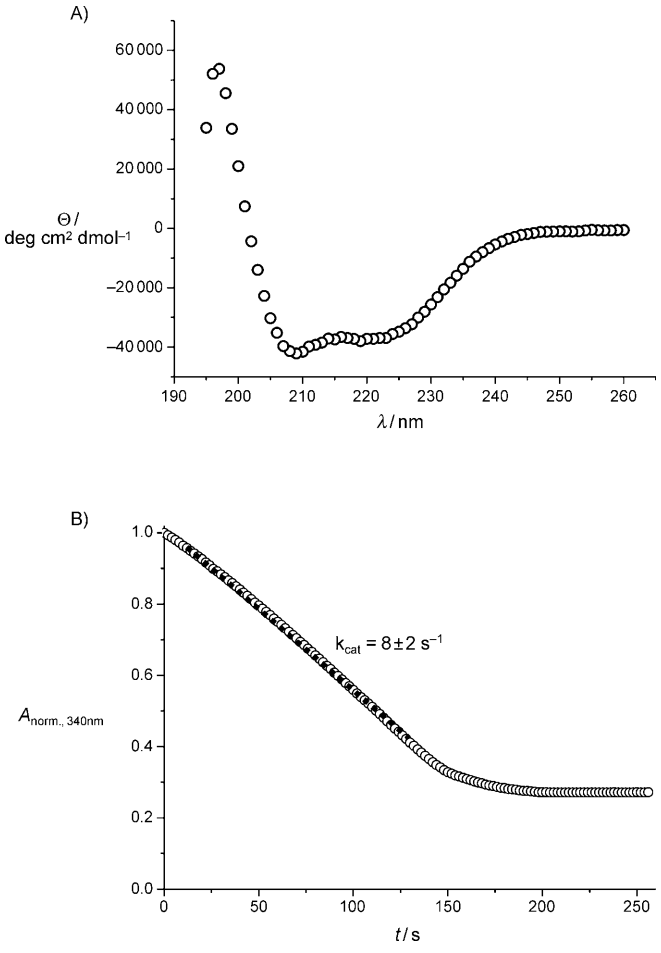


Figure 2. Analysis of folded, synthetic DAGK1–121 in a detergent-containing buffer (0.5% DDM). A) CD spectrum of DAGK1–121 at 25 °C and a concentration of 0.85 mg mL^{-1} in phosphate-buffered saline containing 0.5% DDM. B) Activity assay for DAGK1–121 with 1,2-dioleoyl-*sn*-glycerol as the substrate. Upon the addition of DAGK1–121 ($c = 0.09$ μM), the decrease in NADH absorption was measured and used to calculate the k_{cat} value of 8 ± 2 s^{-1} (NADH is the reduced form of nicotinamide adenine dinucleotide). Measurements with synthetic DAGK1–121 at concentrations between 0.015 and 0.15 μM provided similar k_{cat} values.

Received: October 25, 2010
Revised: February 14, 2011
Published online: March 23, 2011

Keywords: chemical protein synthesis · kinases · membrane proteins · native chemical ligation · protein folding

[1] G. G. Kochendoerfer, D. Salom, J. D. Lear, R. Wilk-Orescan, S. B. H. Kent, W. F. DeGrado, *Biochemistry* **1999**, *38*, 11905–11913.
[2] F. I. Valiyaveetil, R. MacKinnon, T. W. Muir, *J. Am. Chem. Soc.* **2002**, *124*, 9113–9120.
[3] D. Olschewski, C. F. Becker, *Mol. Biosyst.* **2008**, *4*, 733–740.
[4] D. Clayton, G. Shapovalov, J. A. Maurer, D. A. Dougherty, H. A. Lester, G. G. Kochendoerfer, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4764–4769.
[5] C. R. Raetz, K. F. Newman, *J. Biol. Chem.* **1978**, *253*, 3882–3887.
[6] P. Badola, C. R. Sanders, *J. Biol. Chem.* **1997**, *272*, 24176–24182.
[7] E. P. Kennedy, *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 1092–1095.
[8] C. R. Raetz, K. F. Newman, *J. Bacteriol.* **1979**, *137*, 860–868.
[9] J. P. Walsh, C. R. Loomis, R. M. Bell, *J. Biol. Chem.* **1986**, *261*, 11021–11027.
[10] J. P. Walsh, R. M. Bell, *Methods Enzymol.* **1992**, *209*, 153–162.
[11] F. W. Lau, J. U. Bowie, *Biochemistry* **1997**, *36*, 5884–5892.
[12] B. M. Gorzelle, J. K. Nagy, K. Oxenoid, W. L. Lonzer, D. S. Cafiso, C. R. Sanders, *Biochemistry* **1999**, *38*, 16373–16382.
[13] J. K. Nagy, C. R. Sanders, *Biochemistry* **2004**, *43*, 19–25.
[14] W. D. Van Horn, H. J. Kim, C. D. Ellis, A. Hadziselimovic, E. S. Sulistijo, M. D. Karra, C. Tian, F. D. Sonnichsen, C. R. Sanders, *Science* **2009**, *324*, 1726–1729.
[15] K. Oxenoid, F. D. Sonnichsen, C. R. Sanders, *Biochemistry* **2002**, *41*, 12876–12882.
[16] S. Yamaguchi, S. Tuzi, J. U. Bowie, H. Saito, *Biochim. Biophys. Acta Proteins Proteomics* **2004**, *1698*, 97–105.
[17] C. R. Sanders, J. K. Myers, *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 25–51.
[18] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.
[19] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193.
[20] C. F. W. Becker, D. Clayton, G. Shapovalov, H. A. Lester, G. G. Kochendoerfer, *Bioconjugate Chem.* **2004**, *15*, 1118–1124.

[21] D. Bang, S. B. Kent, *Angew. Chem.* **2004**, *116*, 2588 – 2592; *Angew. Chem. Int. Ed.* **2004**, *43*, 2534 – 2538.
[22] L. E. Canne, P. Botti, R. J. Simon, Y. J. Chen, E. A. Dennis, S. B. H. Kent, *J. Am. Chem. Soc.* **1999**, *121*, 8720 – 8727.
[23] E. C. B. Johnson, S. B. H. Kent, *Tetrahedron Lett.* **2007**, *48*, 1795 – 1799.
[24] C. F. Becker, M. Oblatt-Montal, G. G. Kochendoerfer, M. Montal, *J. Biol. Chem.* **2004**, *279*, 17483 – 17489.
[25] R. A. Melnyk, A. W. Partridge, J. Yip, Y. Wu, N. K. Goto, C. M. Deber, *Biopolymers* **2003**, *71*, 675 – 685.
[26] M. A. Hossain, A. Belgi, F. Lin, S. D. Zhang, F. Shabanpoor, L. Chan, C. Belyea, H. T. Truong, A. R. Blair, S. Andrikopoulos, G. W. Tregear, J. D. Wade, *Bioconjugate Chem.* **2009**, *20*, 1390 – 1396.
[27] F. W. Lau, X. Chen, J. U. Bowie, *Biochemistry* **1999**, *38*, 5521 – 5527.
[28] J. P. Walsh, L. Fahrner, R. M. Bell, *J. Biol. Chem.* **1990**, *265*, 4374 – 4381.
[29] J. Jittikoon, J. M. East, A. G. Lee, *Biochemistry* **2007**, *46*, 10950 – 10959.
[30] C. F. W. Becker, C. L. Hunter, R. Seidel, S. B. H. Kent, R. S. Goody, M. Engelhard, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5075 – 5080.
[31] J. Milić, R. Seidel, C. F. W. Becker, R. S. Goody, M. Engelhard, *Biopolymers* **2008**, *90*, 399 – 405.
[32] F. I. Valiyaveetil, M. Sekedat, R. MacKinnon, T. W. Muir, *J. Am. Chem. Soc.* **2006**, *128*, 11591 – 11599.
[33] S. Virdee, D. Macmillan, G. Waksman, *Chem. Biol.* **2010**, *17*, 274 – 284.
[34] Y. Shibata, J. R. van der Ploeg, T. Kozuki, Y. Shirai, N. Saito, M. Kawada-Matsuo, T. Takeshita, Y. Yamashita, *Microbiology* **2009**, *155*, 557 – 565.

