Protein Mimics

DOI: 10.1002/anie.200705718

Synthesis of Protein Mimics with Nonlinear Backbone Topology by a Combined Recombinant, Enzymatic, and Chemical Synthesis Strategy**

*Stephan Pritz, Oliver Kraetke, Annerose Klose, Jana Klose, Sven Rothemund, Klaus Fechner, Michael Bienert, and Michael Beyermann**

Dedicated to Professor Louis A. Carpino on occasion of his 80th birthday

The development of powerful methods for protein synthesis provides new ways to understand protein function, either by regioselective incorporation of modified amino acids as physical probes into natural proteins or de novo protein design to mimic some of the structural and functional properties of native proteins. Expressed protein ligation and protein *trans* splicing have emerged as efficient semisynthetic approaches for the preparation of engineered proteins with normal linear backbone structure.[1] However, the access to engineered proteins with an unnatural, multiply branched, and multicyclic backbone topology requires combinations of recombinant, enzymatic, and chemical synthesis (CRECS). Such proteins with nonlinear backbone topology which mimic extracellular binding domains (ECDs) of G protein-coupled receptors (GPCRs) may provide valuable insight into binding mechanisms of biologically important peptide ligands.

GPCRs are targeted by most pharmaceuticals today,[2] but understanding how ligands bind to GPCRs at the molecular level is still hampered by a lack of high-resolution structural information. The extracellular domains of peptide receptors, namely the receptor N-terminus (ECD1) and the three loops (ECD2, 3, 4), often play a crucial role for ligand binding. Soluble nonglycosylated ECD1 of the corticotropin-releasing factor receptor type 1 (CRF_1), which belongs to the B1 subfamily of GPCRs, has been reported to be the major site of ligand interaction, as it shows high affinity to the natural peptide agonist urocortin 1 and the peptidic antagonist astressin.[3] Interestingly, we observed no binding of such a soluble ECD1 of CRF_1 to ^{125}I-labeled sauvagine,[4] another natural peptide agonist that binds like urocortin 1 with high affinity to wild-type CRF receptors.[5] This finding is indicative of a different contribution of the extracellular loops to high-affinity binding of the full-length receptor to the two ligands. To address this question, and to demonstrate the potential of a CRECS strategy with the first preparation of a protein mimic consisting of the four ectodomains of a GPCR, we synthesized such a mimic of CRF_1, and determined the binding behavior of this receptor mimic for urocortin 1 and sauvagine.

The template-assembled synthetic proteins (TASP) concept introduced by Mutter and Vuilleumier[6] proposes the use of topological templates as structure-supporting scaffolds, and was applied also to the assembly of a construct with three receptor loops.[7] As no experimental data are available on the topology of class B receptors, we initially placed no structural restrictions on the receptor model, but rather designed the template to allow a high degree of flexibility for the domains. Thus, it could be determined whether self-organization is possible in the construct by specific interactions between each other. Our primary efforts were aimed at developing a simple strategy for preparing single receptor domains (the cyclic forms of the loop sequences and the long N-terminus with its three disulfide bridges) in high purity, and their regioselective attachment to an appropriate peptide template. To obtain a uniform construct, it was necessary to purify and analyze each of the domains before the regioselective coupling to the template. The peptide template consists of glycine, glutamic acid, and lysine, whereby the glutamic acid side chains improve solubility in aqueous systems, and the lysine amino groups serve as anchoring points for domains. Additionally, the template contains biotin, which will allow the construct to be bound to avidin-coated beads for a scintillation proximity assay in ligand-binding studies (Figure 1).

The regioselective attachment of domains to the peptide template was accomplished using selectively removable amino protecting groups (Z = benzyloxycarbonyl, Dde = (4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl) for the step-

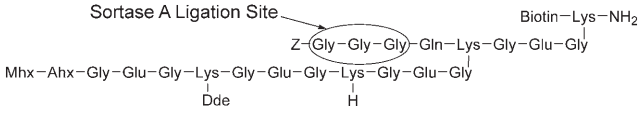

Figure 1. Structure of the template used for regioselective domain attachment. Mhx = maleimidohexanoic acid, Ahx = ε-Aminohexanoic acid.

[*] S. Pritz, Dr. O. Kraetke, A. Klose, Dr. J. Klose, Dr. S. Rothemund, Dr. K. Fechner, Prof. M. Bienert, Dr. M. Beyermann
Abteilung für Peptidchemie und -biochemie
Leibniz-Institut für Molekulare Pharmakologie
Robert-Rössle-Strasse 10, 13125 Berlin (Germany)
Fax: (+49) 30-9479-3159
E-mail: beyermann@fmp-berlin.de
Homepage: http://www.fmp-berlin.de

[**] We thank Gabriela Vogelreiter for technical assistance, and Dr. H. Berger for critical reading of the manuscript. This work was supported by the DFG (SFB 449, TP A6) and the European Commission (QLK3-CT-2002-01989).

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

wise incorporation of maleimidohexanoic acid residues, which were used for thiol ligation of the loops, and enzymatic ligation of the receptor N-terminus to the triglycine chain of the template (Figure 2).

The two naturally occurring cysteine residues found in ECD2 and ECD3 of CRF_1 were replaced by alanine or serine, because this replacement has been shown to have no effect on the high affinity of natural peptide ligand.[8] Thiol groups required for ligation to the template were introduced into linear loop sequences by an extra N-terminal cysteine, which served also for cyclization by the native chemical ligation (NCL)[9] of corresponding peptide thiol esters. To prevent epimerization, and for fast subsequent cyclization, an additional C-terminal glycine residue was introduced. The maleimido group for the first loop attachment (ECD4) was incorporated during the template synthesis. As ECD4 contains no lysine, after attachment of ECD4, the single amino group of the resulting construct was reacted with maleimidohexanoic acid-*N*-hydroxysuccinimide ester (Mhx-OSu). The resulting maleimido group was subsequently used for thiol ligation of ECD2, which like ECD4 has no lysine residue. Thus, after removal of the Dde group from the two-loop construct by hydrazine treatment, the resulting single amino group was reacted with Mhx-OSu. ECD3 was finally coupled by thiol ligation. The subsequent removal of the Z group with trifluoromethanesulfonic acid resulted in the three-loop construct (Figure 3B) ready for coupling to the receptor N-terminus (for LC-MS data of all intermediates, see the Supporting Information).

For the preparation of soluble CRF receptor N-terminus, we followed our recently described procedure[4] for recombinant synthesis in *E. coli*, in-vitro folding, and RP-HPLC purification, but the N-terminus was C-terminally elongated by the peptide sequence LPKTGGRR, a cleavage site for the transpeptidase sortase A. The ECD1-LPKTGGRR was obtained in high purity (Figure 3A) and its correct disulfide pattern was shown by LC-MS analysis of fragments after chymotryptic degradation (see the Supporting Information). Sortase A cleaves LPXTG motifs between T and G, and transfers the N-terminal protein fragment formed by a thiol ester onto N-terminal oligoglycine chains, just like the GGG-motif of the three-loop template construct. Our optimized conditions for sortase-A-mediated transpeptidation[10] were successfully applied herein (Figure 3C,D). The desired prod-

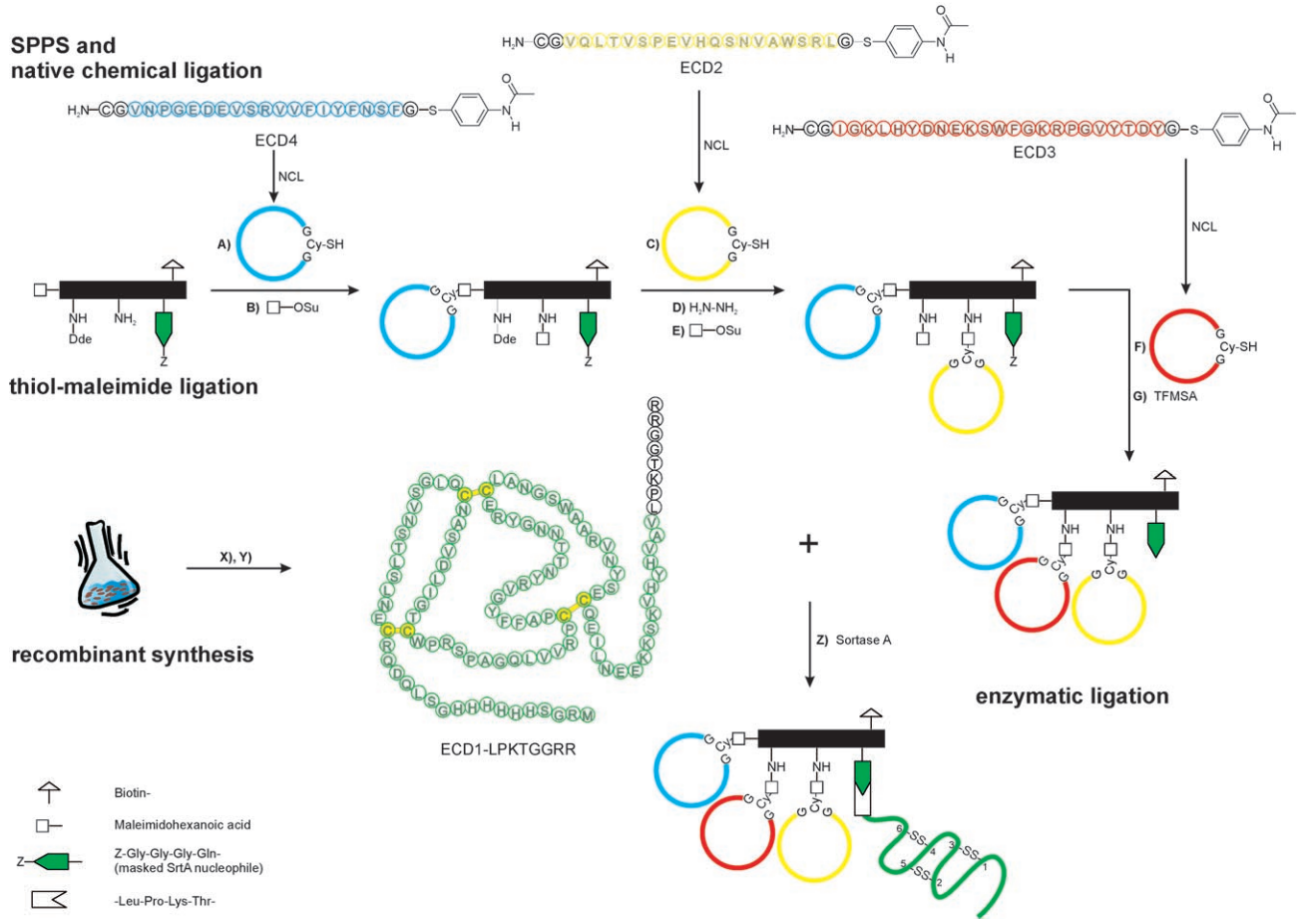

Figure 2. CRECS strategy for assembling a protein mimic consisting of ectodomains of the G protein-coupled receptor CRF_1: A, C, F) thiol-maleimide ligation, B, E) incorporation of maleimidohexanoic acid by OSu ester, D) removal of the Dde group with hydrazine, G) removal of the Z group with TFMSA/TFA, X) overexpression of ECD1-LPKTGGRR in *E. coli*, Y) oxidative folding/formation of the disulfide pattern of ECD1-LPKTGGRR, Z) sortase A-mediated ligation of the three-loop construct and ECD1-LPKTGGRR (see also the Supporting Information). Z = benzyloxycarbonyl, Dde = (4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl.

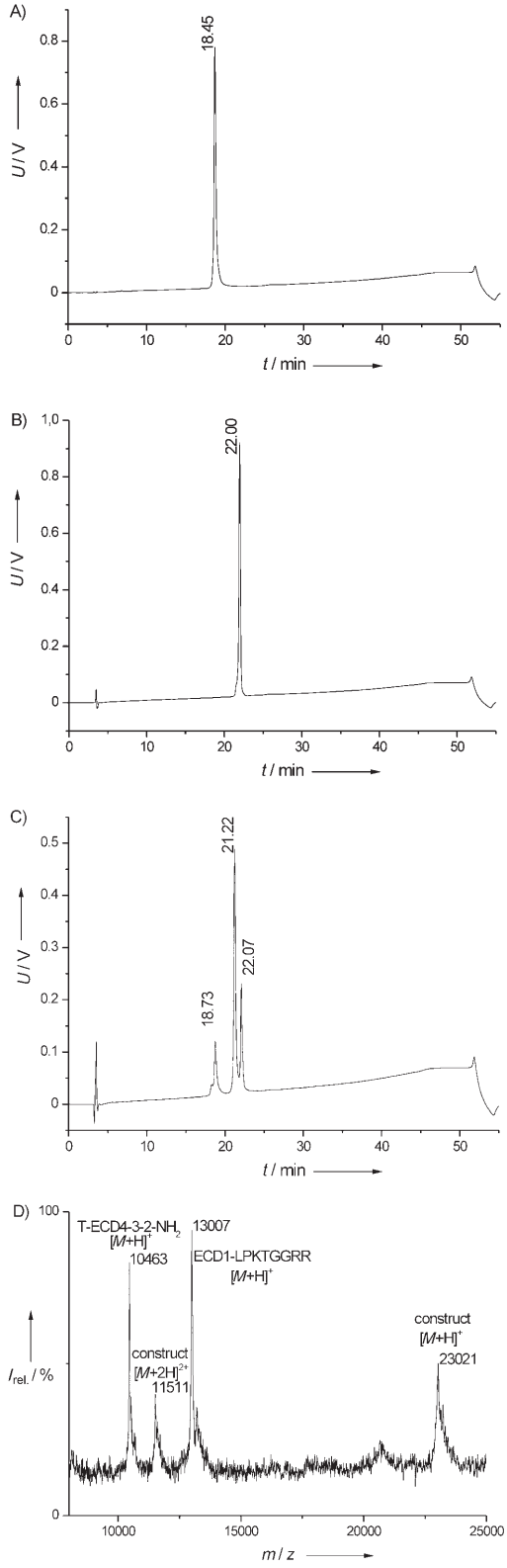

Figure 3. Sortase A-mediated ligation; HPLC profiles (220 nm) of ECD1-LPKTGGRR (A), the three-loop construct (B), ligation mixture after 2 days (C), and MALDI-MS spectrum of the ligation mixture (D).

uct was purified by RP-HPLC (Figure 4A). MALDI-MS analysis showed the correct mass of the protein mimic, which

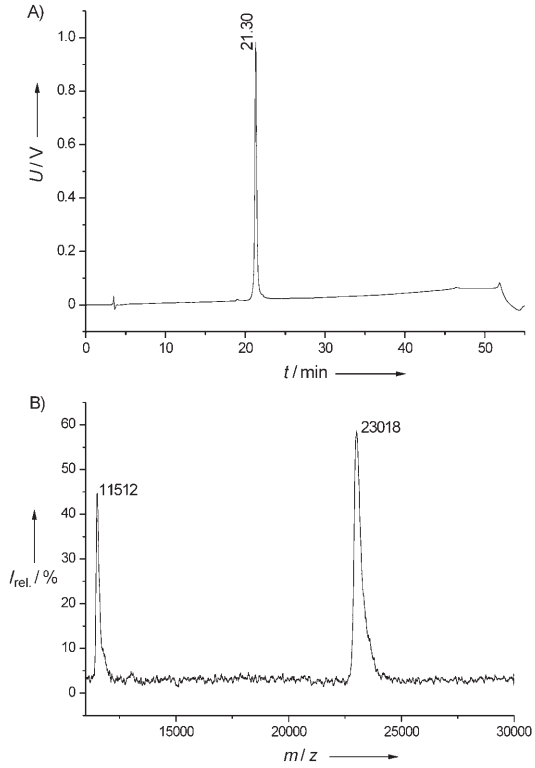

Figure 4. Characterization of the protein mimic consisting of the ectodomains of the corticotropin-releasing factor receptor type 1: A) HPLC profile (220 nm) and B) MALDI-MS results: $[M+H]^+$ calcd: 23024, found: 23018 ± 10; $[M+2H]^{2+}$ calcd: 11512, found: 11512 ± 10.

consists of an ensemble of the four ectodomains of CRF_1 (Figure 4B).

Unlike the receptor N-terminus,[4] for our receptor mimic we observed high-affinity binding to both sauvagine and urocortin 1. Thus ^{125}I-sauvagine could be used as tracer (0.5 nM), and urocortin 1 also showed a high-affinity binding (Figure 5). The mimic without the N-terminus, that is, the three-loop construct, exhibited no ligand binding at all (data not shown), but loops clearly contribute to the binding of sauvagine, as shown by the receptor mimic. On the contrary, the binding of urocortin 1, which binds with high affinity to the N-terminus ($K_d \approx 130$ nM[4]), is not increased much for the receptor mimic.

The affinity of urocortin 1 for CRF_1 stably expressed in HEK cells was in the nanomolar range (uncoupled state $K_d \approx 4$ nM[11]). As expected, this high affinity was not attained for our first receptor mimic (urocortin 1 $K_d \approx 70$ nM), but it provides an interesting starting point for investigation of the ligand binding by varying the topology of the mimic, for example, using cyclic peptide templates or incorporating the native disulfide bridge between ECD2 and ECD3. Moreover, such GPCR mimics may provide valuable tools for ligand binding studies because of their small size compared with that of the full-length receptors, and a water solubility that facilitates investigations without the presence of cells, membranes, or detergents.

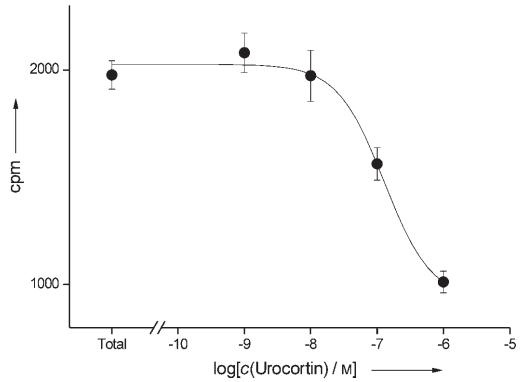

Figure 5. Binding of Urocortin 1 to the CRF_1 receptor mimic. Tracer: ^{125}I-Sauvagin (0.5 nM).

CRECS strategies, such as the presented example of a combined recombinant, enzymatic, and chemical synthesis strategy, provide an entry to a new generation of engineered proteins with nonlinear backbone architecture. Use of orthogonal protecting groups together with state of the art techniques in chemical ligation opens the way for assembling of very complex protein structures. The synthesis and stepwise attachment of purified and well characterized cyclic peptides to a regioselectively addressable template are easily performed in aqueous solution using NCL and thiol-maleimide ligation. In particular, by exploiting the specific transpeptidase activity of enzymes such as sortase A, the efficient peptide and protein transfer at highly specific sites, without significant competition by hydrolysis of the substrate or proteolysis of the product, can be achieved. The mild reaction conditions of enzymatic techniques facilitate the incorporation of chemically very labile peptide and protein units, such as in our case a disulfide-containing protein or, moreover, glycosylated or sulphated ones. This will allow design and synthesis of appropriate protein mimics to investigate the role of pivotal structural elements of proteins for molecular recognition processes and functional activity.

Experimental Section

ECD2 (cyclo(CGVQLTVSPEVHQSNVAWSRLG)), ECD3 (cyclo-(CGIGKLHYDNEKSWFGK RPGVYTDYG)), and ECD4 (cyclo-(CGVNPGEDEVSRVVFIYFNSFG)), and the template (Biotin-K[Z-GGGQK(Mhx-Ahx-GEGK(Dde)GEGKGEG)GEG]-amide) were synthesized using a standard solid-phase synthesis with the Fmoc strategy (Applied Biosystems, 433a). After cleavage from peptide resin, peptide thiol esters of protected peptides were prepared with carbodiimide in the presence of *p*-acetamidothiophenol.[12] After deprotection, peptide cyclization was carried out by NCL. ECD2, ECD3, and ECD4 were consecutively incorporated by thiol-maleimide ligation. The N-terminus was prepared by recombinant synthesis in *E. coli* and oxidative folding,[4] and then coupled to the three-loop construct by sortase A-catalysed transpeptidation. The purity and identity of the receptor mimic was confirmed by HPLC and mass spectrometry. Scintillation proximity assay to determine ligand binding properties was carried out as previously described.[4]

Received: December 13, 2007
Published online: March 28, 2008

Keywords: chemical ligation · enzymatic ligation · protein design · receptor mimics

[1] V. Muralidharan, T. W. Muir, *Nat. Methods* **2006**, *3*, 429–438.
[2] J. Drews, *Science* **2000**, *287*, 1960–1964.
[3] M. H. Perrin, W. H. Fischer, K. S. Kunitake, A. G. Craig, S. C. Koerber, L. A. Cervini, J. E. Rivier, J. C. Groppe, J. Greenwald, S. M. Nielsen, W. W. Vale, *J. Biol. Chem.* **2001**, *276*, 31528–31534.
[4] J. Klose, K. Fechner, M. Beyermann, E. Krause, N. Wendt, M. Bienert, R. Rudolph, S. Rothemund, *Biochemistry* **2005**, *44*, 1614–1623.
[5] J. Vaughan, C. Donaldson, J. Bittencourt, M. H. Perrin, K. Lewis, S. Sutton, R. Chan, A. V. Turnbull, D. Lovejoy, C. Rivier, J. Rivier, P. E. Sawchenko, W. Vale, *Nature* **1995**, *378*, 287–292.
[6] M. Mutter, S. Vuilleumier, *Angew. Chem.* **1989**, *101*, 551–571; *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 535–554.
[7] M. Mutter, P. Dumy, P. Garrouste, C. Lehmann, M. Mathieu, C. Peggion, S. Peluso, A. Razaname, G. Tuchscherer, *Angew. Chem.* **1996**, *108*, 1588–1591; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1482–1485.
[8] L. J. Qi, A. T. Leung, Y. Xiong, K. A. Marx, A.-B. Abou-Samra, *Biochemistry* **1997**, *36*, 12442–124448.
[9] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.
[10] S. Pritz, Y. Wolf, O. Kraetke, J. Klose, M. Bienert, M. Beyermann, *J. Org. Chem.* **2007**, *72*, 3909–3912.
[11] M. Beyermann, N. Heinrich, K. Fechner, J. Furkert, W. Zhang, O. Kraetke, M. Bienert, H. Berger, *Br. J. Pharmacol.* **2007**, *151*, 851–859.
[12] R. von Eggelkraut-Gottanka, A. Klose, A. G. Beck-Sickinger, M. Beyermann, *Tetrahedron Lett.* **2003**, *44*, 3551–3554.

