Protein Synthesis by D-Peptidases

DOI: 10.1002/anie.200800340

D-Amino Acid Specific Proteases and Native All-L-Proteins: A Convenient Combination for Semisynthesis**

Nicole Wehofsky, Andreas Pech, Sandra Liebscher, Stephanie Schmidt, Hidenobu Komeda, Yasuhisa Asano, and Frank Bordusa*

As a consequence of successful enzyme, medium, and substrate engineering, proteases have gained in importance as regio- and stereospecific catalysts in organic synthesis.^[1] Especially for applications based on their native hydrolysis activity, proteases are now generally recognized as established reagents. In principle, the same optimization methods also allow for extensions of the proteases' substrate specificity as well as minimization of undesired proteolytic cleavages; these are the main drawbacks when proteases are used for peptide synthesis.^[2] The application of these approaches has made possible their use for the synthesis of shorter-[3] and longerchain peptides^[4] up to full-length proteins^[5] by means of enzymatic fragment ligation. However, despite these successful examples, it is still the enzymes themselves that hinder their broad use for peptide and especially protein synthesis. While substantial improvements to overcome the specificity problem have been already attained through the development of substrate mimetics, [1] the permanent risk of proteolytic cleavage during the ligation reaction remains an unsolved

We report here on the first application of a D-amino acid specific protease for the synthesis of peptides and proteins entirely composed of L-amino acid moieties. In contrast to conventional protease-based syntheses, which exclusively use L-amino acid specific proteases, the D-specificity of the novel catalyst naturally excludes any undesired proteolysis of the L-configured reactants or products. This strategy was inspired by our recent finding that L-proteases are ironically much better catalysts for the synthesis of all-D-peptides, although

[*] S. Liebscher, Prof. Dr. F. Bordusa
Institute of Biochemistry/Biotechnology
Martin-Luther-University Halle-Wittenberg
Kurt-Mothes-Strasse 3, 06120 Halle/Saale (Germany)
Fax: (+49) 345-551-1972

E-mail: bordusa@enzyme-halle.mpg.de
Dr. N. Wehofsky, A. Pech, S. Schmidt
Max-Planck Research Unit for Enzymology of Protein Folding
06120 Halle/Saale (Germany)

Dr. H. Komeda, Prof. Dr. Y. Asano Biotechnology Research Center Toyama Prefectural University 5180 Kurokawa, Kosugi, Toyama 939–0398 (Japan)

[**] We acknowledge support from the DFG (SFB 610, SPP1170). The authors also thank Prof. G. Fischer for wt-Par10 and S. Kaufmann for technical assistance, and Dr. A. Schierhorn, C. Gersching, Dr. M. Liebscher, and Dr. C. Mrestani-Claus for recording mass, CD, and NMR spectra.

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

nature designed these enzymes to be specific for L-amino acids. [6] Acceptance of the nonspecific all-L-peptide donors by the D-amino acid specific protease is likewise achieved by using the substrate-mimetic strategy. The combination of this approach with the use of a D-protease has finally enabled the synthesis of all-L-peptides and even a native enzyme (Figure 1), compounds for which the D-protease is completely

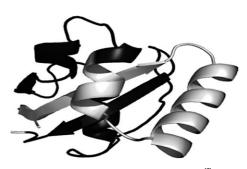


Figure 1. Three-dimensional structure of E. coli Par10.^[7] Black/white contrasts relate to the N- and C-terminal peptide fragments used for the ADP-catalyzed ligation reaction. white: Par10(1–35); black: Par10-(36–92).

nonspecific. Mild reaction conditions, high flexibility towards the sequence at the site of ligation, and the possibility of using side-chain-unprotected peptide reactants are further hall-marks of this approach which, in principle, does not require either organic solvents or artificial mutant enzymes.

First and of particular importance, a suitable D-protease had to be identified. The selection of appropriate candidates was hampered, however, hampered by the significantly small number described in the literature. Presently, only a very few microbial proteases acting on D-peptides are known, and most of them, for example, D-aminopeptidase and carboxypeptidase DD, are exopeptidases active exclusively towards single N- or C-terminal amino acid moieties.[8] Of the remaining enzymes with an endopeptidase activity, which is essentially required for the ligation of peptide fragments, we finally selected "alkaline D-peptidase" (ADP) from Bacillus cereus, which was originally discovered in previous screening studies as a D-Phe specific protease. [9] In addition to being an endopeptidase, ADP is also a serine protease. Thus is can be used for kinetically controlled peptide synthesis which can be combined with the substrate-mimetic approach.^[1] ADP itself can be easily obtained by recombinant expression, while its fusion with a C-terminal His6 tag simplifies purification. In this way a pure enzyme fraction was obtained with a specific activity of 5.3 U mg⁻¹ (see the Supporting Information).

The general acceptance of substrate mimetics by ADP was evaluated by using seven types of esters, which were empirically chosen by adapting their structures to the aromatic phenyl ring system of D-Phe, for which ADP is specific. A more rational selection was not possible since the enzyme's three-dimensional structure is not known. Some experimental hints of possible relevance were available from chymotrypsin, which can be considered to be the "mirror image" of ADP in terms of its specificity towards L-Phe. For chymotrypsin, a distinct substitution pattern of the phenyl ring was found to promote substrate binding mainly by the formation of additional hydrogen bonds.[10] Assuming that ADP behaves similarly, we selected in addition to the benzyl and phenyl esters also 4-carboxyphenyl, 4-carboxamidophenyl, (4-hydroxyphenyl)sulfanyl, 2-carboxy-1*H*-indol-5-yl, and in particular 4-guanidinophenyl derivatives, the latter of which were already found to be accepted by chymotrypsin^[10] (see the Supporting Information). As a nonspecific acyl residue, Boc-L-alanine, the simplest L-amino acid, was chosen.

Following the chemical synthesis of the Boc-L-Ala-O/SR esters, the function of each leaving group as a recognition site for ADP was evaluated by hydrolysis kinetic studies. To guarantee for ester-group-based recognition, Boc-L-Ala-OMe with its simple and nonaromatic methyl ester functionality was included as a control. The results document that all aromatic esters were specifically hydrolyzed by ADP (see the Supporting Information). Remarkably, despite the structural diversity of the ester moieties, the rates of hydrolysis are almost in the same range and differ by less than a factor of three; the (4-hydroxyphenyl)sulfanyl and phenyl ester have the slowest hydrolysis rates while the benzyl and 4-guanidinophenyl counterparts have the fastest rates. In contrast to this putative broad specificity of ADP, no hydrolysis was found for Boc-L-Ala-OMe proving the essential requirement for the aromatic leaving groups to mimic substrates. Further studies using the slightly preferred Boc-Xaa-OGp esters showed that the mimicking effect of the OGp group is not restricted to L-Ala. In fact, of the Boc-Xaa-OGp esters tested (Xaa = L-Glu, L-Lys, L-Gln, L-Leu, Gly, L-Phe), all were accepted by ADP, including the achiral Gly and even L-Phe as the enantiomer of the originally specific D-Phe moiety. Remarkably, for the most specific ester Boc-Gly-OGp, the reaction rate was only 12-fold lower than that for the standard substrate Ac-D-Phe-OMe ($v = 52 \text{ min}^{-1} \text{ vs. } 629 \text{ min}^{-1}$).

The utility of ADP to couple L-amino acid derived reactants was initially investigated by model peptide syntheses using the specific OGp esters as the acyl donors and a representative set of twelve L-amino acid amides as the acyl acceptors. To allow for sensitive HPLC analysis of the reactions, analogous N^{α} -Bz-protected acyl donors were used rather than the original N^{α} -Boc-substituted counterparts. Additionally, similar reactions with the standard substrate Ac-D-Phe-OMe were included as a point of reference. Finally, parallel reactions without ADP were analyzed for each reactant combination to control for spontaneous hydrolysis and aminolysis of the acyl donor esters. From the latter, nonenzymatic aminolysis could be ruled out, and the extent of spontaneous hydrolysis was found to be less than 5%. The results obtained for the ADP-catalyzed reactions are summarized in Table 1. Remarkably, besides a fast conversion of all acyl donors within a reaction time of less than 15 min, in all cases the formation of the desired dipeptide products was observed. Since this holds equally for the Bz-L-Xaa-OGp esters, the data clearly prove that in combination with suitable substrate mimetics D-proteases are generally capable of coupling entirely L-configured amino acid moieties. This novel finding is unprecedented to the best of our knowledge. On analysis of the reaction yields, significant effects of both the acyl donor and acyl acceptor became apparent. Generally, the influence of the latter on the reaction yield directly reflects the individual S'-subsite specificity of ADP. Accordingly, the enzyme prefers larger hydrophobic aliphatic and aromatic amino acid moieties at its S₁' binding site, while smaller residues, polar, and in particular negatively charged amino acids are less favored. Although this tendency also holds for all acyl donors, significantly higher product yields of up to 93% were consistently found for the OGp esters. This surprising behavior goes along with our previous finding that substrate mimetics affect the enzyme's specificity not only

Table 1: Yields [%] of the "alkaline D-peptidase" catalyzed synthesis of all-L dipeptides from the standard substrate Ac-D-Phe-OMe and 4-guanidinophenyl esters derived from nonspecific L-amino acids.^[a]

Acyl acceptor	Acyl donor						
	Ac-D-Phe-OMe ^[b]	Bz-L-Phe-OGp ^[a]	Bz-L-Ala-OGp ^[c]	Bz-L-Leu-OGp ^[c]	Bz-L-Lys-OGp ^[d]	Bz-Gly-OGp ^[e]	Bz-L-Glu-OGp ^[d]
H-L-Pro-NH ₂	0.46	4.51	3.75	5.26	6.03	9.73	10.8
$H-L-Glu-NH_2$	2.45	24.9	15.6	23.6	26.5	35.4	37.4
H-L-Ser-NH ₂	3.5	26.2	16.4	28.5	31.6	41.1	45.7
H-L-Ala-NH ₂	12.5	39.7	30.0	37.6	42.0	59.8	63.7
H-Gly-NH ₂	15.7	38.8	34.1	37.9	57.9	62.8	64.6
H-L-Lys-NH ₂	15.1	33.9	30.0	50.9	61.8	69.1	72.4
H-L-Val-NH ₂	28.7	35.0	49.7	51.0	65.9	82.2	83.7
H-L-Arg-NH ₂	29.2	43.0	60.3	65.6	73.1	84.3	87.2
H-L-Ile-NH ₂	35.4	45.7	69.0	71.4	77.1	88.1	89.8
H-L-Phe-NH ₂	43.0	55.1	78.7	78.2	80.3	89.6	90.9
H-L-Leu-NH ₂	47.0	57.0	75.1	83.2	80.9	92.3	91.6
H -L- M et- NH_2	51.4	65.4	82.6	83.3	84.7	91.7	93.1

[a] Conditions: 30 °C, 0.1 м HEPES buffer pH 8.0, 10 vol% DMF, [acyl donor] = 2 mм, [acyl acceptor] = 20 mм, [enzyme]: [b] 0.27 μм, [c] 43.6 μм, [d] 21.8 μм, [e] 2.22 μм. Reaction time: 15 min. Errors are less than 5% (\pm 2.5%).

Communications

towards the carboxy but also the amino component. [6] This general substrate-mimetic effect is, however, affected by the individual amino acid of the OGp ester. Interestingly, with the exception of the less bulky Gly and L-Ala moieties, this yield-improving effect appears to be inversely related to the enzyme's primary specificity. Thus, the L-Glu-derived OGp ester, the slowest substrate in the hydrolysis studies, gave the highest yields in the synthesis reactions. Vice versa, the readily hydrolyzed L-Phe ester shows the lowest synthesis efficiency of all OGp esters tested. But even here, the yields with specific acyl acceptors are in a synthetically useful range.

After proving the acceptance of short amino acid derived reactants, we evaluated the scope of ADP for coupling elongated peptides. For this purpose, reactions with selected OGp esters covering 1, 3, and 10 amino acids and acyl acceptors with 1, 10, and 16 amino acids were performed. To exclude an inference by the nature of the amino acids at the site of ligation, both the C-terminal amino acid of the acyl donor (Gly) and the acyl acceptors' N-terminal residue (Leu) were kept constant. The reactions, including all control reactions, were performed under conditions similar to those described for the initial dipeptide syntheses; however, for synthesis-economy reasons, the concentration of the acyl acceptors was reduced from 20 to 10 mм. While the control reactions lacking ADP did not indicate any spontaneous aminolysis or considerable hydrolysis of the acyl donor esters, the enzyme-catalyzed syntheses were very rapid, resulting in complete conversion of all esters to the desired peptide products within 15 min (Table 2). As for the efficiency of the reaction using Bz-Gly-OGp und H-L-Leu-NH2, the lower yield can be explained by the reduced acyl acceptor concentration. Considering this, the elongation of the acceptor's chain length leads only to a slight reduction in the product yield. The reactions with the extended substrate mimetics support this finding, indicating that ADP tolerates longer acyl acceptor peptides without a significant loss of its synthesis activity. In contrast, chain length of the acyl donor has a more apparent effect on the product yields. Elongation of Bz-Glv-OGp to Bz-Phe-Gly-Gly-OGp results in a decrease in the yields of about 30% nearly regardless of the bulkiness of the acceptor component. However, an additional elongation of the acyl donor to Bz-AYLDAYVKAG-OGp does not lead to further considerable reductions in the synthesis efficiencies;

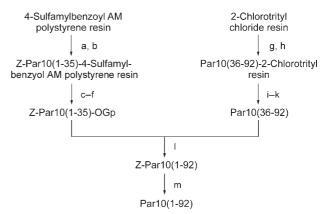
Table 2: Yields of the "alkaline D-peptidase" catalyzed ligation of elongated all-L-peptide fragments.^[a]

Acyl donor	Acyl acceptor	Yield [%]
Bz-Gly-OGp	H-L-Leu-NH ₂	86.1
	H-LGSVKASAYK-OH	77.5
	H-LIVDAVLEPVKAAGAY-OH	83.2
Bz-Phe-Gly-Gly-OGp	H-L-Leu-NH ₂	54.0
, , ,	H-LGSVKASAYK-OH	45.6
	H-LIVDAVLEPVKAAGAY-OH	49.0
Bz-AYLDAYVKAG-OGp	H-L-Leu-NH ₂	46.9
·	H-LGSVKASAYK-OH	40.1
	H-LIVDAVLEPVKAAGAY-OH	43.1

[a] Conditions: 30°C, 0.1 M HEPES buffer pH 8.0, 10 vol % DMF, [acyl donor] = 2 mM, [acyl acceptor] = 10 mM, [enzyme] = $(5.5-8.7) \times 10^{-6}$ M. Reaction time: 15 min. Errors are less than 5% (± 2.5 %).

this indicates that ADP tolerates even longer substrate mimetics. Importantly, besides hydrolyzed peptide esters, no further side products could be detected even after extended reaction times of 24 h. Clearly, competing peptide bond cleavages did not occur (see the Supporting Information).

Finally, we evaluated the function of ADP for the synthesis of native proteins. As the synthesis target we selected the peptidyl prolyl *cis-trans* isomerase (PPIase) parvulin 10 from *E. coli* (Par10). The primary catalytic function of PPIases is to facilitate *cis-trans* isomerizations of Xaa-Pro bonds; they are involved in the folding of newly synthesized proteins, in the control of the cell cycle control, and in the immune system in higher developed organisms.^[11] According to Scheme 1, the synthetic route to Par10 follows a



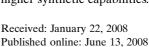
Scheme 1. General course of the synthetic route to full-length E. coli Par10. a, b) Resin loading and solid-phase peptide synthesis; c-f) linker activation, peptide release, side-chain deprotection, and purification; g, h) resin loading and solid-phase peptide synthesis; i-k) simultaneous peptide release/deprotection and purification; l) ADP-catalyzed ligation; m) N-terminal deprotection. Z-Par10(1–35)-OGp: Z-AKTAAALHILVKEEKLA LDLLEQIKNGADFGKLAK-OGp; Par10(36–92): KHSICPSGKRGGDLGEFRQGQMVPAFDKVVFSCPV LEPTGPLHTQFGY-HIIKVLYRN. Conditions: 30 °C, 0.1 m HEPES buffer pH 8.0, 20 vol % DMF, 1 mm TCEP, [acyl donor] = 0.2 mm, [acyl acceptor] = 0.1 mm, [enzyme] = 2.0 × 10⁻⁵ m. TCEP = tris(2-carboxyethyl) phosphine.

single-step enzymatic coupling of an OGp ester of N^{α} protected Par10(1-35) and the remaining 57-mer Par10(36-92) fragment. The ligation site at Lys35-Lys36 was chosen for purely chemical reasons (the efficiency of the solid-phase peptide synthesis of the two fragments). CD-spectroscopic studies reveal a that the structure of the longer Par10(36–92) fragment is similar to that of the full-length protein with corresponding α -helical and β -sheet regions (Figure 1). Thus, the site of ligation can be assumed to be an α -helical motif, which is known to be hardly recognized by proteases. Thus, in the present synthesis Par10 is a difficult sequence for protease-based ligations. The starting fragments were prepared by standard solid-phase Fmoc chemistry using either 2chlorotrityl chloride resin^[12] for the preparation of Par10(36– 92) or Kenner's 4-sulfamylbenzoyl aminomethyl (AM) safety-catch resin^[13] for synthesizing Z-Par10(1-35)-OGp (see the Supporting Information). Peptide release, deprotection, and purification led to the desired fragments which for solubility reasons were subsequently used in concentrations of 0.2 mm (Z-Par10(1-35)-OGp) and 0.1 mm (Par10(36-92)) for the ADP-catalyzed ligation. In spite of these relatively low reactant concentrations, the addition of 20 vol % DMF was required to ensure sufficient solubility of the two fragments. Additionally, 1 mm TCEP had to be added to avoid undesired oxidation reactions. After thermal equilibration of the assay mixture and subsequent enzyme addition, a reaction time of 30 min led to a complete ester consumption and formation of the ligation product. In contrast, no reaction was found in the control synthesis lacking ADP. Detailed quantitative analyses of the ADP-catalyzed syntheses by HPLC and SDS-PAGE (Figure 2d) reveal a product yield of 61%, which nearly matches the yield found for the synthesis of Bz-L-Lys-L-Lys-NH₂ (Table 1) and indicates the high flexibility of ADP towards the length of the acyl donor and acceptor moiety. Furthermore, MALDI-TOF MS measurements confirmed the molecular weight of the ligation product as that of native Par10 (Figure 2b). Importantly, the nonquantitative product formation results exclusively from a competing hydrolysis of Z-Par10(1-35)-OGp. Undesired cleavages, which inevitably occur when L-amino acid specific proteases are used as catalysts, could be entirely excluded even after reaction times ten times longer than usual (Figure 2d). From the lack of any protein products of a higher molecular weight than Par10, it can be further expected that no additional acylation at nucleophilic side chains take place.

Following purification and refolding, we finally analyzed the structure and function of the synthetic Par10 by CD and enzyme kinetic studies. As a point of reference, recombinant wt-Par10 was used. On analyzing the CD spectra (Figure 2c), we noted a striking similarity of the secondary structural elements of recombinant and synthetic Par10. The same appears to be true for the stability of both Par10 species in variable-temperature CD studies (see the Supporting Information). Finally, only minor differences were found in the catalytic function of the two Par10 preparations (Figure 2a).

In summary, our results clearly show that naturally occurring proteases, in contradiction to the general doctrine in the field, are capable of synthesizing native proteins in aqueous solution without any risk of proteolytic side reactions. Just by taking the distinct stereospecificities of the enzymes into account, irreversible peptide ligations can be achieved in an straightforward and highly flexible manner. Restrictions toward the nature of the amino acid moieties at the site of ligation, like that known for purely chemical approaches such as native chemical ligation, are of much less importance. The substrate-mimetic esters required for this novel ligation technique do not significantly differ from those needed for other approaches. Thus, protocols for their

chemical synthesis are already established. Moreover, owing to the broad tolerance of ADP towards the structure of the ester leaving group, a number of ester types more common than OGps can be used. This finding may further simplify their preparation. Finally, one has to keep in mind that with ADP only one single D-protease was evaluated for its synthesis behavior and it was more or less selected by chance. Thus, broader screenings with more enzymes of this type have a great potential for identifying further candidates with even higher synthetic capabilities.



Keywords: peptide ligation · proteases · protein semisynthesis · substrate mimetics

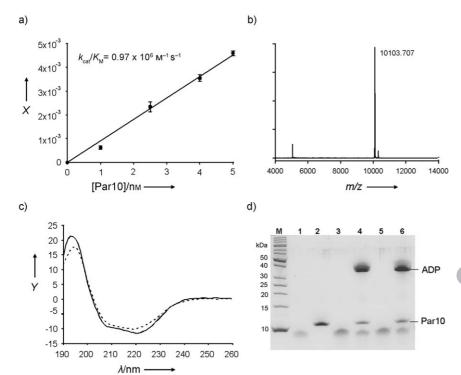


Figure 2. Characterization of synthetic *E. coli* Par10. a) Specificity constants k_{cat}/K_M of synthetic Par10 (Error: $\pm 0.05 \times 10^6 \, \text{m}^{-1} \, \text{s}^{-1}$) and wt-Par10 (2.0×10⁶ $\pm 0.1 \times 10^6 \, \text{m}^{-1} \, \text{s}^{-1}$), *X*: reaction rate in s⁻¹; b) MALDI-TOF mass spectrum of synthetic Par10 (mass calcd: 10101); c) CD spectra of synthetic (dashed line) and wt-Par10 (solid line), *Y*: ellipticity in mdeg; d) SDS-Page. *M*: molecular weight marker, 1: Par10(36–92), 2: recombinant wt-Par10, 3: ligation reaction without ADP (reaction time: 30 min), 4: ligation reaction with ADP (reaction time: 30 min), 5: ligation reaction without ADP (reaction time: 5 h), 6: ligation reaction with ADP (reaction time: 5 h); intensity of band staining is related to the differences in the size of ADP and Par10 and does not relate to their assumed concentrations.

a) F. Bordusa, Chem. Rev. 2002, 102, 4817-4867; b) D. Kumar, T. C. Bhalla, Appl. Microbiol. Biotechnol. 2005, 68, 726-736; c) D. Goradia, J. Cooney, B. K. Hodnett, E. Magner, Biotechnol. Prog. 2006, 22, 1125-1131.

^[2] M. Bergmann, H. Fraenkel-Conrat, J. Biol. Chem. 1937, 119,707-720; W. Kullmann, Enzymatic Peptide Synthesis, CRC, Boca Raton, 1987.

Communications

- [3] a) N. Wehofsky, N. Koglin, S. Thust, F. Bordusa, J. Am. Chem. Soc. 2003, 125, 6126–6133; b) K. J. Doores, B. G. Davis, Chem. Commun. 2005, 2, 168–170.
- [4] a) V. Cerovksy, J. Kockskämper, H. G. Glietsch, F. Bordusa, ChemBioChem 2000, 1, 126-129; b) Z. Machova, R. Eggelkraut-Gottanka, N. Wehofsky, F. Bordusa, A. G. Beck-Sickinger, Angew. Chem. 2003, 115, 5065-5067; Angew. Chem. Int. Ed. 2003, 42, 4916-4918; Z. Machova, R. Eggelkraut-Gottanka, N. Wehofsky, F. Bordusa, A. G. Beck-Sickinger, Angew. Chem. 2003, 115, 5065-4067.
- [5] D. Y. Jackson, J. Burnier, C. Quan, M. Stainley, J. Tom, J. A. Wells, *Science* 1994, 266, 243–247.
- [6] N. Wehofsky, S. Thust, J. Burmeister, S. Klussmann, F. Bordusa, Angew. Chem. 2003, 115, 701–704; Angew. Chem. Int. Ed. 2003, 42, 677–679.
- [7] A. Kuhlewein, G. Voll, B. Hernandez Alvarez, H. Kessler, G. Fischer, J. U. Rahfeld, G. Gemmecker, *Protein Sci.* 2004, 13, 2378–2387.

- [8] a) Y. Asano, A. Nakazawa, Y. Kato, K. Kondo, J. Biol. Chem.
 1989, 264, 14233-14239; b) Y. Asano, T. L. Lübbehüsen, J. Biosci. Bioeng. 2000, 89, 295-306; c) D. H. Baek, S.-J. Kwon, S.-P. H., M.-S. Kwak, M.-H. Lee, J. J. Song, S.-G. Lee, K.-H. Yoon, M.-H. Sung, Appl. Environ. Microbiol. 2003, 69, 980-986.
- [9] Y. Asano, H. Ito, T. Dairi, Y. Kato, J. Biol. Chem. 1996, 271, 30256-30262.
- [10] F. Bordusa, Curr. Protein Pept. Sci. 2002, 3, 159-180.
- [11] C. Scholz, J. Rahfeld, G. Fischer, F. X. Schmid, J. Mol. Biol. 1997, 273, 752-762; C. Schiene-Fischer, C. Yu, FEBS Lett. 2001, 495, 1-6; G. Fischer, T. Aumüller, Rev. Physiol. Biochem. Pharmacol. 2003, 148, 105-150.
- [12] K. Barlos, D. Gatos, J. Kallitsis, G. Papaphotiu, P. Sotiriu, Y. Enqing, W. Schäfer, *Tetrahedron Lett.* 1989, 30, 3943–3946.
- [13] a) G. W. Kenner, J. R. McDermott, R. C. Sheppard, J. Chem. Soc. Chem. Commun. 1971, 636-637; b) B. J. Backes, J. A. Ellman, J. Org. Chem. 1999, 64, 2322-2330.