ATP Binding in Peptide Synthetases: Determination of Contact Sites of the Adenine Moiety by Photoaffinity Labeling of Tyrocidine Synthetase 1 with 2-Azidoadenosine Triphosphate†

Maja Pavela-Vrancic,‡ Eva Pfeifer,‡ Henk van Liempt,‡ Hans-Jochen Schäfer,§ Hans von Döhren,*,‡ and Horst Kleinkauf‡

Institut für Biochemie und Molekulare Biologie, Technische Universität Berlin, 10587 Berlin, FRG, and Institut für Biochemie, Johannes-Gutenberg-Universität, 55128 Mainz, FRG

Received November 17, 1993; Revised Manuscript Received February 24, 1994®

ABSTRACT: Characterization of the nucleotide binding domain in peptide synthetases was approached by photoaffinity labeling of tyrocidine synthetase 1 (TY1) with 2-azidoadenosine triphosphate (2-azido-ATP). Exposure of TY1 in the presence of photolabel to irradiation with ultraviolet light resulted in a time-dependent covalent modification of the enzyme with a concomitant loss of catalytic activity. Inactivation was not observed if incubation was performed in the absence of either light or the nucleotide analogue. Specificity of labeling was indicated by the ability of 2-azido-ATP to serve as a substrate in the amino acid activation reaction. The modified protein was subjected to tryptic digestion, and the fragments labeled by the nucleotide analogue were purified by reverse-phase high-performance liquid chromatography. Sequence analysis identified three tryptic peptides corresponding to residues G373–K384, W405–R416, and L483–K494, derived from the N-terminal half of the TY1 sequence. As this region shows similarity to strongly conserved regions in other peptide synthetases and acyl-CoA synthetases, it is considered to be the region catalyzing aminoacyl adenylate formation. The identified sequences appear to define components of the nucleotide binding domain found in close proximity to the adenine ring in ATP. Conservation of primary structure and homology to other carboxyl-activating enzymes of this superfamily, including peptide synthetases, insect luciferases, and acyl-CoA synthetases, is discussed.

Small peptides, peptidolactones, depsipeptides, and peptide moieties of structurally diverse secondary metabolites, produced by bacteria and fungi, are generally synthesized nonribosomally with the aid of multienzyme complexes (Kleinkauf & von Döhren, 1990). In this multienzymatic system, only the protein structure determines substrate selection and product sequence. In the ribosomal system, amino acid selection and peptide sequence are determined by protein–nucleic acid interactions in tRNA aminoacylation and codon–anticodon interactions within the ribosomal framework. Studies on the enzymology of the cyclic peptides tyrocidine and gramicidin S produced by *Bacillus brevis* and the β-lactam antibiotic precursor peptide δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV)[1] from *Aspergillus nidulans*, *Cephalosporium acremonium*, and *Streptomyces clavuligerus* revealed the general principles of the reaction mechanism (Kleinkauf & von Döhren, 1990; MacCabe et al., 1991; van Liempt et al., 1991; Schwecke et al., 1992; Turgay et al., 1992; Marahiel, 1992; Aharonowitz et al., 1993). The constituent carboxyl compounds (amino, imino, and hydroxy acids) are activated by adenylation and successively transferred to specific thiols, yielding covalently bound thioester intermediates. The elongation reaction is thought to be mediated by intrinsic transport systems involving an enzyme-bound cofactor, 4′-phosphopantetheine (Kleinkauf et al., 1971; Schlumbohm et al., 1991), with subsequent release of the peptide product by cyclization, hydrolysis, or terminal modification.

Peptide synthetases are isofunctional enzymes with high diversity in terms of substrate specificity and molecular weight. The molecular masses range from 123 kDa for tyrocidine synthetase 1 to 1400 kDa for cyclosporin synthetase (Schmidt et al., 1992). The large size and complexity of these proteins have made structural studies difficult. Sequence alignment has shown that peptide synthetases are composed of highly conserved functional units or modules, whose occurrence is associated with the number of amino acids incorporated into the peptide product (Turgay et al., 1992). A sequence of approximately 1000 amino acid residues composes the activation unit for one amino acid, including such partial reactions as covalent binding, epimerization, and possible further modifications. The implication from this type of building block arrangement would be that any peptide could be synthesized by providing a gene containing the appropriate number of activating units with the correct specificity. Prerequisites for the successful design of new, more efficient, and less toxic peptide products are a better understanding of the structure–function relationship within the polypeptide chain and the identification of specific residues involved in substrate activation as possible targets for modification and alteration of specificity. Inspection of the primary structure

† This work was supported by a postdoctoral fellowship from the Federation of European Biochemical Societies (FEBS), by a Wood/Whelan Research Fellowship from the International Union of Biochemistry and Molecular Biology (IUBMB) to M.P.-V., and by grants from the Deutsche Forschungsgemeinschaft (Scha 344/1-3 and Do 270/5-2).

* Author to whom the correspondence should be addressed. Telephone: (030) 314 22697. Fax: (030) 314 24783. e-mail: hans@cicsg.tk.tu-berlin.de.

‡ Technische Universität Berlin.

§ Johannes-Gutenberg-Universität.

® Abstract published in *Advance ACS Abstracts,* April 1, 1994.

[1] Abbreviations: ACV, δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine; ATP, adenosine triphosphate; 2-azido-ATP, 2-azidoadenosine triphosphate; DTE, dithioerythritol; TPCK, *N*-tosyl-L-phenylalanyl chloromethyl ketone; TY1, tyrocidine synthetase 1; EDTA, ethylenediaminetetraacetic acid; D-Phe, D-phenylalanine, SDS, sodium lauryl sulfate; HPLC, high-pressure liquid chromatography.

0006-2960/94/0433-6276$04.50/0 © 1994 American Chemical Society

of various modules revealed the presence of several motifs carrying highly conserved core sequences, some of which have been proposed to be involved in adenylate formation and aminoacylation (MacCabe et al., 1991; van Liempt et al., 1991; Marahiel, 1992). Peptide synthetases are structurally related to several carboxyl-activating enzymes utilizing ATP for substrate activation (Babbitt et al., 1992), but are surprisingly dissimilar from aminoacyl-tRNA synthetases (Cramer & Freist, 1992; Pavela-Vrancic et al., 1994a). In this superfamily of adenylate-forming enzymes, sequence information has been obtained by now for more than 40 amino acid or hydroxy acid activating domains and 20 acyl-CoA synthetases (von Döhren, 1993), but information on catalytic and substrate binding sites is still lacking.

Tyrocidine synthetase 1 (TY1) was the first peptide synthetase to be sequenced (Weckermann et al., 1988) and the first to be cloned and expressed in large quantities in *Escherichia coli* (Marahiel et al., 1985). As a component of a multienzyme complex composed of three complementary enzymes (tyrocidine synthetases 1, 2, and 3), it catalyzes the first step in the biosynthesis of the decapeptide antibiotic tyrocidine in *B. brevis* (Lee et al., 1973). It activates a single amino acid (L-Phe) as an aminoacyl adenylate, followed by aminoacylation and epimerization into D-Phe. It is a representative model of a peptide synthetase module containing all required functions integrated in one polypeptide chain of 123 kDa. Because of its simplicity and ease of purification, it was used as a model system for this entire family of enzymes.

In order to locate the ATP binding site, we have chosen an affinity labeling approach, since photosensitive substrate analogues have been used successfully in labeling the active sites of several nucleotide binding proteins. The most commonly used probes have been 2- and 8-azidopurine derivatives (Scheurich et al., 1978; Garin et al., 1986; Cross et al., 1987; Lunardi et al., 1987; Davis et al., 1990; Rush & Konigsberg, 1990; Salvucci et al., 1992; Grammer et al., 1993) and 3′-ribose-coupled aryl azides (Schäfer & Dose, 1984; Schäfer et al., 1989; Maruta et al., 1990). These compounds are stable in the dark, but irradiation with ultraviolet light converts them into highly reactive nitrenes that can covalently insert into the nearby amino acid side chain or the polypeptide backbone. Photoaffinity labeling with azidopurine derivatives enables the identification of peptides in close proximity to the adenine ring of ATP. A requirement for successful labeling is that the analogue is a substrate in the dark or at least a competitive inhibitor. Data obtained from analysis of the nucleotide binding sites of known structure (Fry et al., 1986; Serrano, 1988; Pai et al., 1989; Zheng et al., 1993) and from substrate specificity studies on peptide synthetases using nucleotide analogues (Pavela-Vrancic et al., 1994a) are consistent with a hydrophobic pocket for the adenine ring with strict steric constraints and the requirement of a hydrogen-bond donor at the N-6 position. Peptide synthetases have been shown to accept only 2-substituted ATP derivatives as substrates, contrary to the 8-substituted analogues that are not recognized (Pavela-Vrancic et al., unpublished results). The failure of 8-azido-ATP to displace the natural substrate is attributed to the *syn* conformation about the N-glycosidic bond induced by substitution at the C-8 position (Sarma et al., 1974). However, nucleotides substituted at the C-2 position tend to adopt the *anti* conformation typical of the natural substrate (Czarnecki, 1984). Substitution at the ribose ring causes complete loss of activity, rendering impossible photoaffinity studies with nucleotide analogues modified at the sugar moiety.

Thus, 2-azidoadenosine derivatives may be used to study nucleotide binding sites that are not accessible to other photoaffinity probes. A major constraint is the lability of the labeled photoprobe to experimental workup due to the instability of the N-glycosidic bond and certain covalent bonds formed upon photoinsertion (Lewis et al., 1989; Haley, 1991). 2-Azido-ATP was prepared and used to photolabel the nucleotide binding site of TY1. Fragments of the modified protein were isolated and identified, providing new insights into the adenine binding domain of peptide synthetases.

MATERIALS AND METHODS

Chemicals. ATP, dithioerythritol (DTE), iodoacetamide, urea, and trypsin (TPCK-treated) were purchased from Sigma; mercaptoethanol and HPLC-grade acetonitrile were from Merck. [^{32}P]PP$_i$ was a product of DuPont.

Enzyme Preparation. The *E. coli* strain XL1-blue (Bullock et al., 1987) was used as the host for plasmid pGC12. The plasmid pGC12 contains a 4.6-kb *Hin*cII fragment (containing the *tycA* region) from pBT2 under the control of the lac promoter from pUC18 (Marahiel et al., 1985). The cells were grown in Luria–Bertani medium containing 50 μg/mL ampicillin. The protein purification procedure will be reported elsewhere (E. Pfeifer, unpublished data).

Synthesis of 2-Azido-ATP. 2-Azidoadenosine 5-triphosphate was synthesized as described by van Dongen et al. (1986), according to Schaeffer and Thomas (1958), Sowa and Ouchi (1975), and Hoard and Ott (1975).

Substrate Specificity. Activity measurements were performed in an assay mixture containing, in a final volume of 70 μL, 50 mM Tris-HCl (pH 7.4), 2.5 mM Mg^{2+}, 1 mM DTE, 0.1 mM EDTA, 0.07 mM PP$_i$, 0.1 μCi of [^{32}P]PP$_i$, and 0.7 mM D-Phe, at varying concentrations of the nucleotide substrate (0.02–0.5 mM). The reaction was started by the addition of 20 μL of the enzyme solution (final concentration 22 μg/mL) and allowed to proceed for 15 min at 37 °C.

Enzyme Inactivation. TY1 (200 μg/mL) was suspended in 50 mM Tris-HCl (pH 7.2) containing 2.5 mM Mg^{2+} and varying concentrations of the photoprobe. The concentration of the 2-azido-ATP stock solution was determined spectrophotometrically using an extinction coefficient of 10 300 M^{-1} cm^{-1} at 270 nm. After a 20-min incubation in the dark at room temperature, the reaction mixture was irradiated for 45 min with long-wavelength ultraviolet light from a Mineralight UV lamp positioned 3 cm from the solution surface. The progress of the reaction was monitored by measuring the isotope ATP–[^{32}P]PP$_i$ exchange reaction with 2-azido-ATP as the nucleotide substrate. Aliquots were withdrawn at intervals and assayed for remaining activity. Activity measurements were performed in an assay mixture containing, in a final volume of 70 μL, 50 mM Tris-HCl (pH 7.4), 0.15 mM ATP, 0.75 mM Mg^{2+}, 1 mM dithioerythritol, 0.1 mM EDTA, 0.1 mM PP$_i$, 0.1 μC$_i$ of [^{32}P]PP$_i$, and 1 mM D-Phe. The sample was diluted 20-fold in 50 mM Tris-HCl (pH 7.4). The reaction was started by the addition of 20 μL of the enzyme solution and allowed to proceed for 15 min at 37 °C. In order to assess the effect of UV light on enzyme activity and substrate specificity, either enzyme or label was exposed to irradiation and subsequently included into the assay mixture at concentrations equal to those used in the previous experiment.

Photolabeling of TY1 with 2-Azido-ATP. Large-scale photolabeling was performed in a plastic Petri dish by irradiation with ultraviolet light in the presence of 0.25 mM 2-azido-ATP and 2.5 mM Mg^{2+} for 20 min. A maximum degree of labeling was achieved by increasing the concentration

of the photoprobe to 0.5 mM, followed by an additional illumination period of 40 min. Labeled protein was separated from most of the unincorporated probe by gel-filtration chromatography on a Sephadex G-50 column previously equilibrated with 10 mM ammonium bicarbonate (pH 8.3). The eluant was evaporated in a SpeedVac concentrator and stored at −20 °C until digestion.

Tryptic Digestion. Labeled TY1 was resuspended in 8 M urea in 0.4 M NH_4HCO_3 (pH 8.3). Reduction was performed by the addition of 4 mM DTE followed by incubation at 37 °C for 1 h. Iodoacetamide was added to yield a concentration of 8 mM, and the solution was incubated for 1 h in the dark at room temperature. Enough water was added to bring the urea concentration to 2 M. The sample was subjected to tryptic digestion by incubation with trypsin for 2 h at 37 °C at a trypsin to protein weight ratio of 1:17. The course of the digestion was monitored by SDS–polyacrylamide gel electrophoresis (Laemmli, 1970).

Purification of Labeled Peptides. Tryptic peptides were purified by reverse-phase high-pressure liquid chromatography. The tryptic digest was applied onto a Waters Deltapack C_{18} analytical column developed with a gradient of solvent B (0.05% trifluoroacetic acid in 80% acetonitrile) in solvent A (0.05% trifluoroacetic acid) as follows: 2% B for 10 min, followed by a linear increase to 10% B over 20 min, isocratically at 10% B for 30 min, followed by a linear increase to 60% B over 155 min at a flow rate of 0.8 mL/min. The absorbance of the eluate was monitored at 214 and 254 nm to detect the ultraviolet absorbance of adenosine. Peak fractions were partially dried by a speed vacuum and reinjected onto the C_{18} column developed in a different gradient of solvent B: 2% B for 5 min, followed by linear increases to 20% B over 15 min and 60% B over 165 min. The isolated peptides were subjected to sequencing on an Applied Biosystems protein sequencer (Model 473A).

RESULTS

Photoaffinity Labeling of Tyrocidine Synthetase 1. Photoaffinity labeling with azidoadenine nucleotide analogues provides a direct approach to identifying nucleotide binding domains in proteins. 2-Azido-ATP was synthesized and used to label the nucleotide binding site of tyrocidine synthetase 1. Since the nitrene group formed upon irradiation is directly located at the purine ring, it may be expected that the major part of the nucleotide will react immediately at the adenine binding site. The suitability of 2-azido-ATP as an affinity probe was demonstrated by its ability to act as a substrate in the absence of light, as evidenced by the ATP–[^{32}P]PP_i exchange reaction. It was accepted with almost the same catalytic efficiency (K_m = 0.11 mM and k_{cat} = 0.16 s^{-1} for 2-azido-ATP) as the natural substrate (K_m = 0.08 mM and k_{cat} = 0.33 s^{-1} for ATP) (Figure 1). The preference of the *anti* position of 2-azido-ATP is not surprising, because a large number of purine nucleotides bind to enzymes of known crystal structure in this conformation (Garin et al., 1986; Kabsch et al., 1990; Salvucci et al., 1992; Grammer et al., 1993). However, the potential of 2-azidoadenosine as an effective photoaffinity probe appears to be limited by its property of rearranging into non-photoreactive tetrazoles at neutral and basic pH (Czarnecki, 1984).

Irradiation of TY1 in the presence of 2-azido-ATP at pH 7.2 resulted in the time-dependent inactivation of enzymatic activity. The time course of inactivation is shown in Figure 2. The effect of UV light on the stability and nucleotide binding properties of the enzyme was determined. Prolonged

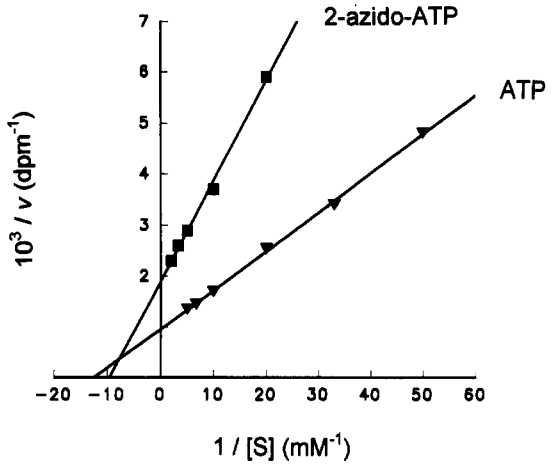


FIGURE 1: Determination of the kinetic constants (k_m and k_{cat}) for 2-azido-ATP and ATP. Activity was measured by the D-Phe-dependent 2-azido-ATP and ATP–[^{32}P]PP_i exchange reaction at a saturating concentration of D-Phe (0.7 mM), as described in Materials and Methods.

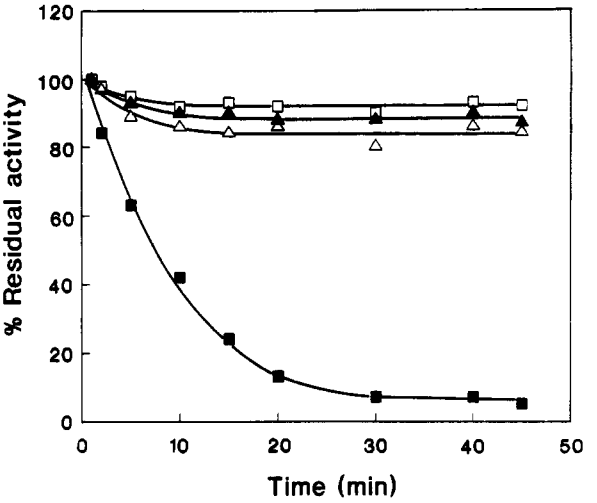


FIGURE 2: Time course of inactivation: (□) TY1 incubated with 0.5 mM 2-azido-ATP and 2.5 mM Mg^{2+} in the dark; (△) nucleotide substrate irradiated with UV light and included into the assay at concentrations equal to the concentrations used in the previous experiment; (▲) enzyme solution containing 2.5 mM Mg^{2+} and (■) enzyme plus 0.5 mM 2-azido-ATP and 2.5 mM Mg^{2+} photolyzed and the residual activity determined as described under Materials and Methods. Samples were irradiated for the times indicated and assayed for catalytic activity by measuring the ATP–[^{32}P]PP_i exchange reaction.

exposure of TY1 to short-wave UV light was deleterious to the enzyme; however, longer wavelengths (>300 nm) were not harmful to the protein material. In control experiments performed in the absence of light or by irradiation of the enzyme in the absence of the nucleotide analogue, enzymatic activity remained unchanged over the course of incubation. After 45 min of irradiation, less than 15% of the activity was lost compared to the nonirradiated control. A tight, noncovalent interaction with photolytic breakdown products of the probe might be the cause of inactivation during photolabeling. However, this was not the case since prephotolyzed probe showed constant activity when included in the assay at concentrations equal to those used in the photoaffinity labeling experiments. Therefore, any observed decrease in enzymatic activity must result from a selective interaction of the label with the substrate binding site. Maximal inhibition was reached upon a second round of photolabeling by increasing the concentration of label to 0.5 mM at a total incubation

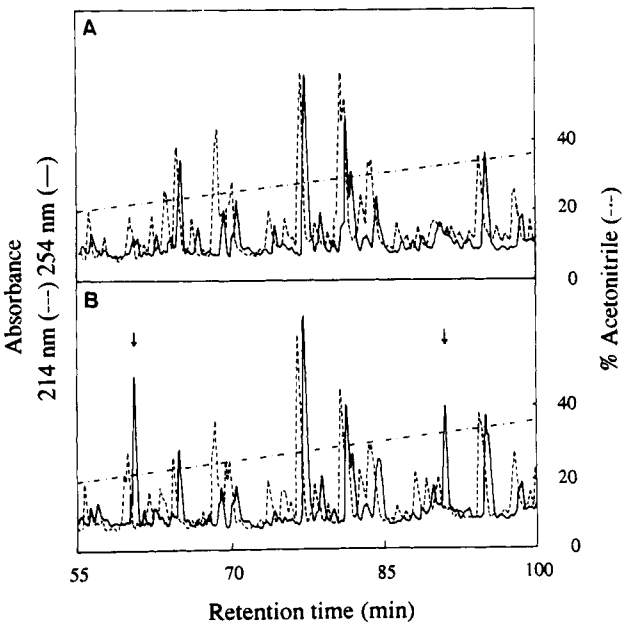


FIGURE 3: HPLC profiles of the tryptic digest of unmodified TY1 and the photolabeled enzyme preparation monitored at 214 and 254 nm. TY1 was irradiated in the absence (A) and presence (B) of 0.5 mM 2-azido-ATP and 2.5 mM Mg^{2+} at pH 7.2 and 37 °C for 60 min. The tryptic digest was fractionated on a C_{18} column using a gradient of acetonitrile in 0.05% trifluoroacetic acid, as described under Materials and Methods. Absorbance was monitored at 214 (full scale corresponding to 0.32 absorption units) and 254 nm (0.08 absorption units full scale). The two graphs have been shifted for better recognition by the equivalent of 1 min, with the absorbance at 214 nm appearing to the left of the absorbance at 254 nm.

time of 60 min. The unbound label was removed from the reaction mixture by gel-filtration chromatography over a Sephadex G-50 column.

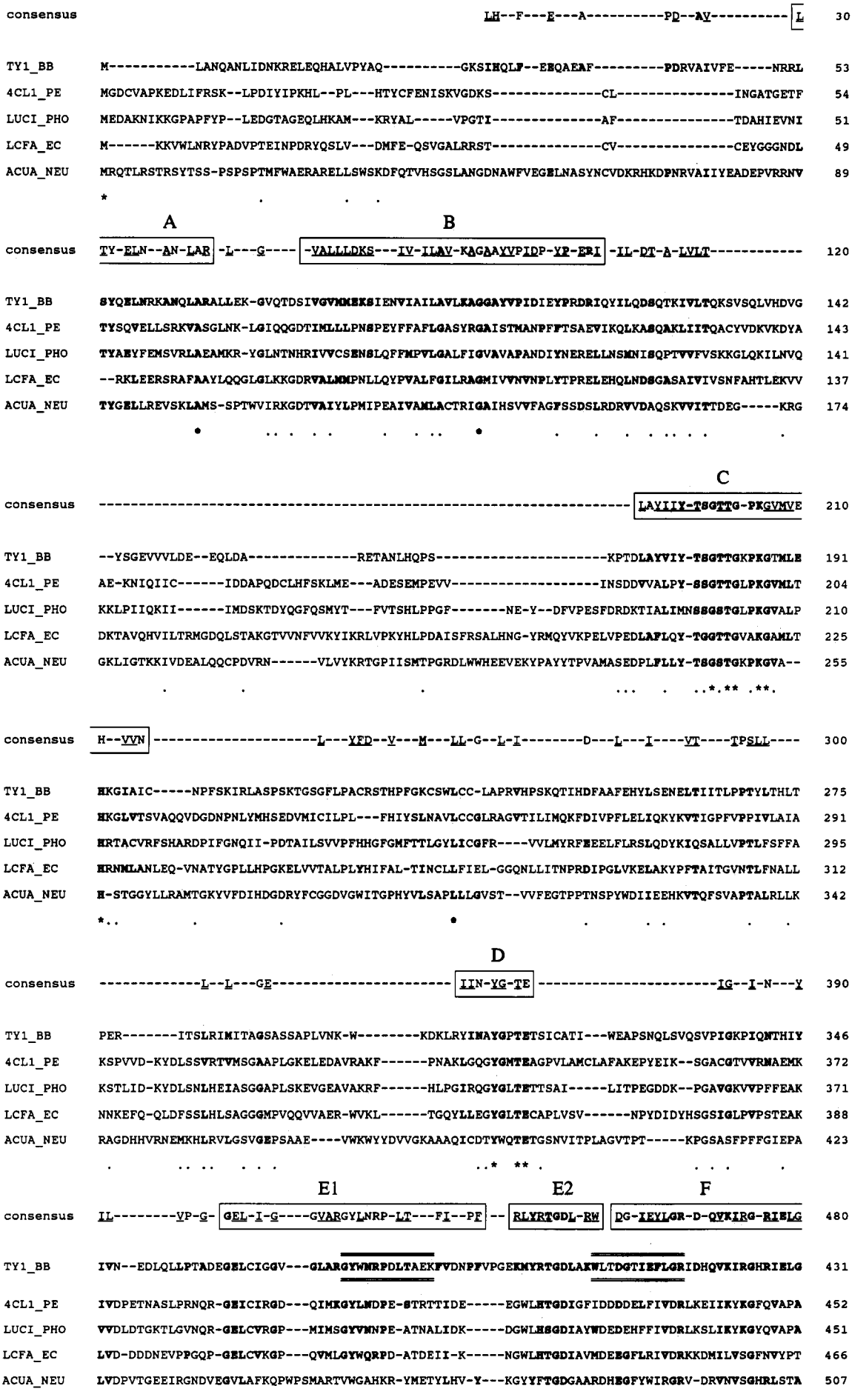
Isolation of Photolabeled Peptides. After the removal of unbound ligand followed by total proteolysis of the enzyme with trypsin, peptides were separated by reverse-phase HPLC using a C_{18} column eluted with a gradient of increasing acetonitrile concentration in 0.05% trifluoroacetic acid. Modified peptides could be readily detected by absorbance measurements at 254 nm (Davis et al., 1990), indicative of the label, relative to a reference sample containing the digest of the unmodified irradiated enzyme preparation. The absorbance of the eluent was monitored continuously at 214 and 254 nm. The elution profiles for the unmodified and labeled samples are presented in Figure 3A,B, respectively. Most of the label was recovered in the void volume, indicating significant instability of the photoinserted bond to subsequent working conditions. Two fractions showing increased absorbance at 254 nm eluted at 22 and 34% acetonitrile (Figure 3B). In the elution profile for the digest of the irradiated, unmodified enzyme sample (Figure 3A), the corresponding absorption signals were not present. In a control experiment performed with prephotolyzed label, the nucleotide analogue eluted with the void volume, substantiating the assumption that the increased absorbance in Figure 3B was due to specific modification of the tryptic fragments. The labeled samples were subjected to further purification by HPLC. The sample eluting at 22% acetonitrile gave one peptide, whereas the sample eluting at 34% acetonitrile generated two labeled peptides upon rechromatography. Sequence analysis of the purified fragments identified tryptic peptides corresponding to residues G373–K384, W405–R416, and L483–K494. They are distributed throughout the central region and span a total of 121 amino acid residues. No modified amino acids, indicative of the site of labeling, could be detected. A similar observation was made by other authors reporting on photoaffinity labeling with 2-azido-ATP (Davis et al., 1990; Haley, 1991; Salvucci et al., 1992).

DISCUSSION

Binding of 2-Azido-ATP. In a recent analysis of peptide synthetase binding sites using structural analogues of ATP, we have shown that 2-chloro-ATP is an effective substrate for various amino acid activation sites in peptide synthetases. On the other hand, 8-substituted analogues were not recognized as either substrate or inhibitor (Pavela-Vrancic et al., 1994). From these data, we concluded that 2-azido-ATP should be a suitable substrate analogue for photoaffinity labeling work, and it indeed proved to be an effective substrate (Figure 1). Apparent K_m values for 2-azido-ATP and ATP were comparable, with 0.11 and 0.08 mM, respectively, while the k_{cat} of the azido compound was about one-half that of ATP. From these results, we assumed that the attachment of the photolabel would correspond to the ATP binding site.

Sequence Comparison of Different Acyl Adenylate Forming Enzymes. The purpose of this study was to characterize the adenine binding domain of MgATP in peptide synthetases. Tyrocidine synthetase 1, a representative model enzyme for this class of proteins, was subjected to labeling with the photoreactive nucleotide analogue, 2-azido-ATP. The modified tryptic peptides correspond to a region of the TY1 primary structure previously recognized as highly conserved among peptide synthetases (Marahiel, 1992). Two of the peptides are located on both sides of the highly conserved TGD motif, which is detected not only in peptide synthetases but also in acyl-CoA synthetases (Figure 4). This motif is found within a stretch of highly conserved sequences, termed E1, E2, F, G, and H, which have been defined by the alignment of peptide synthetase data. Our third labeled peptide is located within the boxed sequence H (Figure 4). This likewise has a corresponding conserved region in acyl-CoA synthetases, as can be seen in an alignment of the amino acid sequence of tyrocidine synthetase 1 (Weckermann et al., 1988) with the consensus sequence derived from 38 peptide synthetase modules (Pavela-Vrancic et al., 1994a) and with the members of four classes of acyl-CoA ligases: 4-coumaryl-CoA ligase from parsley (Loyoza et al., 1988), firefly luciferase (De Wet et al., 1987), rat long-chain fatty acid CoA ligase (Suzuki et al., 1991), and *Neurospora crassa* acetyl-CoA ligase (Connerton et al., 1990).

Common to all enzymes is the activation of the carboxyl group at the expense of ATP, and any motif occurring in each group of enzymes is therefore likely to be involved in aminoacyl adenylate formation. Most of the conserved motifs in the consensus sequence correspond to the boxes described earlier (van Liempt et al., 1991), which were based on an analysis of only seven modules. These sequences have been indicated by the boxes A–I. Even though this analysis shows that the homologies sometimes extend farther than the boxes, we use this nomenclature here to simplify the discussion. It appears that the A and B motifs may be involved in amino acid recognition and binding, since they are absent from enzymes that do not have amino acid substrates. The signature sequence, (S,T)SG(T,S)TGxPKG (located within the C box), occurs in all peptide synthetases and related carboxyl-activating enzymes. The highly conserved lysine residue within this sequence may be involved in the binding of phosphate in ATP in analogy to the lysine residue found within the glycine-rich P-loop sequence of several nucleotide binding proteins of


consensus
TY1_BB
4CL1_PE
LUCI_PHO
LCFA_EC
ACUA_NEU
A
B
C
D
E1
E2
F

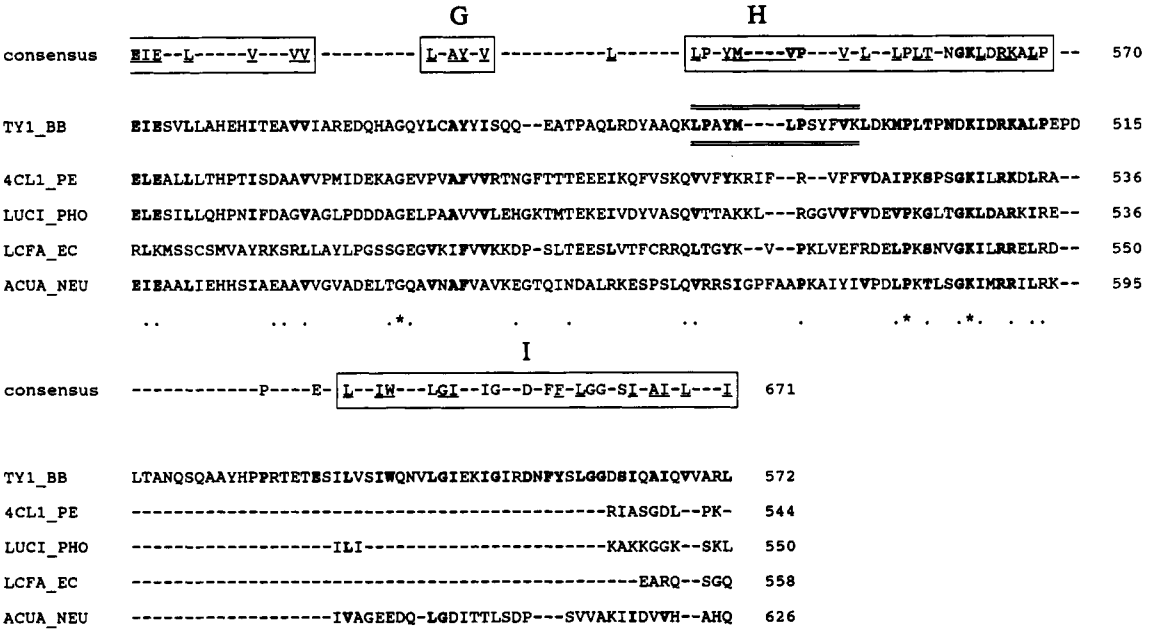


FIGURE 4: Sequence comparison between peptide synthetases and acyl-CoA ligases. A consensus sequence compiled from the 38 peptide synthetase modules published so far (Pavela-Vrancic et al., 1994a) is compared to the sequences of TY1 and the members of four classes of acyl-CoA ligases (coumaryl-CoA ligases, luciferases, long-chain fatty acyl-CoA ligases, and acetyl-CoA ligases). In the consensus sequence, underlined symbols are understood to be any of the group of similar residues, while boldface symbols represent positions at which all sequences (or all minus one) contain the same residue (or a member of a group of similar amino acids). A dash represents positions with less than 50% conservation. The five amino acid sequences [TY1 (Weckermann et al., 1988), 4-coumaryl-CoA ligase from parsley (Loyoza et al., 1988), firefly luciferase (De Wet et al., 1987), rat long-chain fatty acid CoA ligase (Suzuki et al., 1991), and *Neurospora crassa* acetyl-CoA ligase (Connerton et al., 1990)] were aligned with the CLUSTAL program. An asterisk below a position points indicates a fully conserved residue, while a dot signifies a strongly conserved position. Boldface symbols were chosen for residues that correspond to the consensus sequence for the peptide synthetase modules.

known structure (Walker et al., 1982; Fry et al., 1986; Pai et al., 1989). We propose that this residue may be essential for catalysis rather than binding (Pavela-Vrancic et al., 1993).

The conserved sequence within the E2 motif shows homology to a consensus sequence for the putative ATP binding region of a series of ATPases, TGD (Taylor & Green, 1989; Serrano, 1988), indicating possible involvement of the aspartic acid side chain in the coordination of magnesium in MgATP. The TGD motif has been missed in the alignments of peptide synthetases and several carboxyl-activating enzymes utilizing ATP as an energy source (Babbitt et al., 1992), thus erroneously leading to a subclass definition. The functional role of the TGD motif in peptide synthetases has not yet been established. Adenylate formation, as well as the inverse reaction, is known to depend on the presence of magnesium. Magnesium ions in nucleotide binding proteins generally have six ligands coordinated in an octahedral arrangement. A region comprising a hydrophobic stretch terminated by an aspartic acid ($GX_{4-8}D$) was found upon homology searches of adenylate kinase, leading Fry et al. (1986) to suggest the involvement of Asp-119 in Mg^{2+} binding. In the crystal structure of cAMP-dependent protein kinase, Asp-184 chelates the primary magnesium ion that bridges the β- and γ-phosphates in ATP (Zheng et al., 1993). In analogy to other enzyme classes, a similar structure may be proposed for the nucleotide binding site of peptide synthetases, with Lys and Asp involved in phosphate orientation and metal coordination.

Within the F motif, the conserved glycine corresponding to Gly-425 has recently been identified by mutant enzyme analysis and site-specific mutagenesis as a requirement for adenylate formation within the proline-activating module of gramicidin S synthetase 2 (Tokita et al., 1993). The extended I motif is only found in those enzymes catalyzing the formation of covalent amino acid–enzyme intermediates. It has been established that the covalent complex involves an enzyme-bound cofactor attached to the serine residue within the LGG-(H/D)S sequence (Schlumbohm et al., 1991).

The major part of the reports on nucleotide binding proteins deals with the phosphate binding sequence. Although it is very clear that other regions must be involved in the binding of the AMP moiety of ATP, very little information is available concerning the possible contact sites for the purine base and the ribose. For the D, E1, F, G, and H sequence motifs, no plausible function has been offered even though they are highly conserved among peptide synthetases but only partially conserved among other homologous carboxyl-activating enzymes. It is well-known that enzymes catalyzing the same reaction may display different structural frameworks. It seems reasonable to assume that the aspartate and glutamate residues in the acidic region of the E2/F motif may be involved in binding of the metal ion.

The identified fragments correspond to three separate regions of the TY1 primary structure. Investigations with other proteins have shown that more than one peptide from the binding domain may be modified in a single photoaffinity labeling experiment (Salvucci et al., 1992). Modification of chicken muscle adenylate kinase with 2- or 8-azido-ATP resulted in the labeling of several distinct sites spanning a region of 150 amino acid residues. Thus, it was proposed that the adenine binding domain for the MgATP site in adenylate kinase is flanked by residues near positions V30–Y35 and A157–Y165. The fact that all three peptides identified by photoaffinity labeling of TY1 are invariant throughout the peptide synthetase family verifies a potential role in binding or catalysis. Fragment G373–K384 composes a segment of the E1 consensus sequence. The GYxNxP sequence motif is highly conserved among peptide synthetases and related carboxyl-activating enzymes. Peptide W405–R416 is flanked

by the TGD and RIELGEIE core sequences (motif E2/F), whose structural features imply a possible role in magnesium coordination. Fragment L483–K494 is highly conserved among peptide synthetases, showing little similarity with the acyl-CoA synthetases. However, it is directly adjacent to a highly conserved region corresponding to the H motif of peptide synthetases. It is thus likely that the contact sites identified here are significant in both types of enzymes. Although the labeled peptides are far apart in the primary sequence, folding of the polypeptide chain may enable them to compose the active site in the tertiary structure and contribute to binding of the adenine and ribose moieties of the nucleotide.

TY1 Nucleotide Binding Site Compared to ATP Binding Sites of Known Structure. So far little is known about the structure of the active sites in peptide synthetases or about the chemical basis for their nucleotide specificity. In general, the binding of MgATP is associated with several attachment sites, including specific recognition of the adenine ring, selectivity for the ribose moiety, and coordination of Mg^{2+} interacting with the β,γ-phosphates of ATP. Biosynthesis studies with the tyrocidine system have shown that the incorporation of L-[^{14}C]leucine into tyrocidine is ineffective if ATP is replaced by other nucleoside triphosphates, indicating high selectivity in the purine binding region (Fujikawa et al., 1968). Some characteristic features of the binding site emerged from analogue studies. AMP is a competitive inhibitor of the amino acid activation reaction, suggesting that its binding is taking place at the nucleotide binding site. On the basis of the high binding affinity for AMP (K_d = 0.12 mM), relative to the affinity for ATP (K_d = 1.3 mM), it appears that the recognition of ATP is not mediated by the triphosphate group, but rather by the adenine ring and the ribose moiety. The structural properties of the substrate molecule argue for consistent and limited interaction sites and the involvement of similar amino acid residues in recognition and binding. In cAMP-dependent protein kinase, the adenine ring is enclosed in a hydrophobic pocket consisting of residues Leu-49, Val-57, Ala-70, Met-120, Tyr-122, Val-123, and Leu-173, the hydrophobic character of which is conserved within the kinase family (Zheng et al., 1993). The ATP binding site of rabbit muscle adenylate kinase places the adenine ribose moiety within a hydrophobic pocket near the side chains of the Ile-28, Val-29, His-36, Leu-37, and Leu-91 residues (Fry et al., 1986). In phosphoglycerate kinase, only the adenine is in a hydrophobic pocket with the ribose in a more exposed position, forming hydrogen bonds to Glu-341 and Asp-372 (Banks et al., 1979). dATP binds to the Klenow fragment of DNA polymerase I from *E. coli* near at least two hydrophobic amino acid residues thought to be some combination of Ile, Leu, and Val and one aromatic residue, probably Tyr (Rush & Konigsberg, 1990).

Although there is variation in the details of adenosine binding in known structures, it is expected that recognition of the adenine ring occurs in peptide synthetases, both by hydrophobic interactions and by hydrogen bonding. Prior substrate affinity studies have shown that two hydrogen bonds may serve to anchor the adenine ring in the active site. The primary hydrogen-bond contact with the N-6 amino group is quite evident from the remarkable reduction of activity upon substitution. Removal of the hydrogen-bond donor at this position is very deleterious. The second hydrogen bond is presumably formed with the N-1 nitrogen, as evidenced by complete loss of activity when N^1-methyl-ATP is used as the nucleotide substrate. Modification at position N-7 of the adenine ring causes only marginal reduction in affinity, suggesting that this nitrogen atom is not involved in binding, similar to the results obtained for aminoacyl-tRNA synthetases of class I (Cramer & Freist, 1993).

At present, we cannot draw any conclusions concerning the precise locations of the labeled residues. Due to the lability of the photoinserted bond toward subsequent working conditions, we were not able to identify the specific residues modified by photoaffinity labeling. The few cases where the site of labeling could be identified have usually involved photoinsertion into a Tyr residue (Garin et al., 1986; Cross et al., 1987; Rush & Konigsberg, 1990). Tyrosine may interact by hydrogen bonding with an N-atom of the adenine ring. Affinity labeling of beef heart F1-ATPase with 2-azidonucleotides led to derivatization of the catalytic site, Tyr-345, positioned near the adenine ring of ATP and of Tyr-368 from the noncatalytic site shown to bind ADP (Cross et al., 1987). It has been suggested that the adenine moiety of the nucleotide bound in the catalytic site interacts with the aromatic ring of the tyrosine residue (Zhuo et al., 1992). A high degree of sequence conservation is generally associated with a strategic function in binding or catalysis. Tyr-374 from the G373–K384 fragment, a highly conserved residue in both peptide synthetases and several related carboxyl-activating enzymes, may be a potential candidate for interaction with the adenine base in ATP. In fragment W405–R416 there are two highly conserved residues, Asp-408 and Arg-416. Aspartic acid residues have been found to be involved in the coordination of magnesium (Fry et al., 1986; Pai et al., 1989; Zheng et al., 1993) and in hydrogen bonding to the N-atoms of the purine base and to the ribose moiety of the nucleotide substrate (Banks et al., 1979; Pai et al., 1989). Several arginine residues have been shown to participate in nucleotide binding by adenylate kinase (Kim et al., 1990). Segment L483–K494 is conserved only among peptide synthetases, raising doubts as to the significance of this site in ATP binding.

The MgATP binding site in peptide synthetases has been assigned to the central region by sequence comparison. On the basis of the experimental data obtained by affinity labeling of TY1, we believe that the region flanked by residues near positions G373–K384, W405–R416, and L483–K494 may be involved in adenine and ribose binding. Further refinement of the adenine binding site beyond the present level requires the synthesis of a reactive nucleotide analogue capable of forming a stable covalent linkage to the amino acid residues within the binding site. Despite its limitations, affinity labeling usefully identifies candidate residues whose roles in substrate binding and catalysis can be further assessed by site-directed mutagenesis. However, the exact nature of the interaction awaits X-ray crystallographic analysis of the TY1–MgATP complex.

ADDED IN PROOF

This work has been extended using fluorescein 5′-isothiocyanate as an affinity label (Pavela-Vrancic et al., 1994b). Two tryptic peptides have been identified, I417–R424 and L495–R508, the positions of which are adjacent to those detected here.

ACKNOWLEDGMENT

We thank Dr. Werner Schröder (Freie Universität Berlin) for peptide sequencing. We appreciate the expert technical assistance of Gabriele Rathgeber.

REFERENCES

Aharonowitz, Y., Bergmeyer, J., Cantoral, J. M., Cohen, G., Demain, A. L., Fink, U., Kinghorn, J., Kleinkauf, H., MacCabe, A., Palissa, H., Pfeifer, E., Schwecke, T., van Liempt, H., von

Döhren, H., Wolfe, S., & Zhang, J. (1993) *Bio/Technology 11*, 807–810.
Babbitt, P., Kenyon, G. L., Martin, B. M., Charest, H., Slyvestre, M., Scholten, J. D., Chang, K.-H., Liang, P.-H., & Dunaway-Mariano, D. (1992) *Biochemistry 31*, 5594–5604.
Banks, R. D., Blake, C. C. F., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. W., Merrett, M., & Phillips, A. W. (1979) *Nature 279*, 773–777.
Bullock, W. O., Fernandez, J. M., & Short, J. M. (1987) *BioTechniques 5*, 376–381.
Connerton, I. F., Fincham, J. R. S., Sandeman, R. A., & Hynes, M. J. (1990) *Mol. Microbiol. 4*, 451–460.
Cramer, F., & Freist, W. (1993) *Angew. Chem. 105*, 198–209.
Cross, R. L., Cunningham, D., Miller, C. G., Xue, Z., Zhuo, J.-M., & Boyer, P. D. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5715–5719.
Czarnecki, J. J. (1984) *Biochim. Biophys. Acta 300*, 41–51.
Davis, C. B., Smith, K. E., Campbell, B. N., & Hammes, G. G. (1990) *J. Biol. Chem. 265*, 1300–1305.
De Wet, J. R., Wood, K. V., De Luca, M., Helsinki, D. R., & Subrami, S. (1987) *Mol. Cell. Biol. 7*, 725–737.
Eggen, R. I. L., Geerling, A. C. M., Boshoven, A. B. P., & de Vos, W. M. (1991) *J. Bacteriol. 173*, 6383–6389.
Fry, D. C., Kuby, S. A., & Mildvan, A. S. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 907–911.
Fujikawa, K., Suzuki, T., & Kurahashi, K. (1968) *Biochim. Biophys. Acta 161*, 232–246.
Garin, J., Boulay, F., Issartel, J. P., Lunardi, J., & Vignais, P. V. (1986) *Biochemistry 25*, 4431–4437.
Grammer, J. C., Kuwayama, H., & Yount, R. G. (1993) *Biochemistry 32*, 5725–5732.
Haley, B. E. (1991) *Methods Enzymol. 200*, 477–487.
Hoard, D. E., & Ott, D. G. (1965) *J. Am. Chem. Soc. 87*, 1785–1788.
Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F., & Holmes, K. C. (1990) *Nature 347*, 34–44.
Kim, H. J., Nishikawa, S., Tokutomi, Y., Takenaka, H., Hamada, M., Kuby, S. A., & Uesugi, S. (1990) *Biochemistry 29*, 1107–1111.
Kleinkauf, H., & von Döhren, H. (1990) *Eur. J. Biochem. 192*, 1–15.
Kleinkauf, H., Roskoski, R., Jr., & Lipmann, F. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 2069–2072.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Lee, S. G., Roskoski, R., & Lipmann, F. (1973) *Biochemistry 12*, 398–405.
Lewis, C. T., Haley, B. E., & Carlson, G. M. (1989) *Biochemistry 28*, 9248–9255.
Loyoza, E., Hoffmann, H., Douglas, M., Schulz, W., Scheel, D., & Hahlbrock, K. (1988) *Eur. J. Biochem. 176*, 661–667.
Lunardi, J., Garin, J., Issartel, J.-P., & Vignais, P. V. (1987) *J. Biol. Chem. 262*, 15172–15181.
MacCabe, A. P., van Liempt, H., Palissa, H., Unkles, S. E., Riach, M. B., Pfeifer, E., von Döhren, H., & Kinghorn, J. R. (1991) *J. Biol. Chem. 266*, 12646–12654.
Marahiel, M. A. (1992) *FEBS Lett. 307*, 40–43.
Marahiel, M. A., Krause, M., & Skarpeid H.-J. (1985) *Mol. Gen. Genet. 201*, 231–236.
Maruta, S., Burke, M., & Ikebe, M. (1990) *Biochemistry 29*, 9910–9915.
Pai, E. F., Kabsch, W., Krengek, U., Holmes, K. C., John, J., & Wittinghofer, A. (1989) *Nature 341*, 209–214.
Pavela-Vrancic, M., Pfeifer, E., Gocht, M., Marahiel, M. A., van Liempt, H., & von Döhren, H. (1993) in *Industrial Microorganisms: Basic and Applied Molecular Genetics* (Baltz, R. H., Hegeman, G. D., & Skatrud, P. L., Eds.) Abstract, p 292, American Society for Microbiology, Washington, D.C.
Pavela-Vrancic, M., van Liempt, H., Pfeifer, E., Freist, W., & von Döhren, H. (1994a) *Eur. J. Biochem. 220*, 535–542.
Pavela-Vrancic, M., Pfeifer, E., Schröder, W., von Döhren, H., & Kleinkauf, H. (1994b) Identification of the ATP Binding Site in Tyrocidine Synthetase 1 by Selective Modification with Fluorescein 5′-Isothiocyanate, *J. Biol. Chem.* (in press).
Rush, J., & Konigsberg, W. H. (1990) *J. Biol. Chem. 265*, 482–4827.
Salvucci, M. E., Chavan, A. J., & Haley, B. E. (1992) *Biochemistry 31*, 4479–4487.
Sarma, R. H., Lee, C. H., Evans, F. E., Yathindra, N., & Sundaralingam, M. (1974) *J. Am. Chem. Soc. 96*, 7337–7348.
Serrano, R. (1988) *Biochim. Biophys. Acta*, 1–28.
Schaeffer, H. J., & Thomas, H. J. (1958) *J. Am. Chem. Soc. 80*, 3738–3742.
Schäfer, H.-J., & Dose, K. (1984) *J. Biol. Chem. 259*, 15301–15306.
Schäfer, H.-J., Rathgeber, G., Dose, K., & Kagawa, Y. (1989) *FEBS Lett. 253*, 264–268.
Scheurich, P., Schäfer, H.-J., & Dose, K. (1978) *Eur. J. Biochem. 88*, 253–257.
Schlumbohm, W., Stein, T., Ullrich, C., Vater, J., Krause, M., Marahiel, M. A., Kruft, V., & Wittmann-Liebold, B. (1991) *J. Biol. Chem. 266*, 23135–23141.
Schmidt, B., Riesner, D., Lawen, A., & Kleinkauf, H. (1992) *FEBS Lett. 307*, 355–360.
Schwecke, T., Aharonowitz, Y., Palissa, H., von Döhren, H., & Kleinkauf, H. (1992) *Eur. J. Biochem. 205*, 687–694.
Sowa, T., & Ouchi, S. (1975) *Bull. Chem. Soc. Jpn. 48*, 2084–2090.
Suzuki, H., Kawarabayasi, Y., Kondo, J., Abe, T., Nishikawa, K., Kimura, S., Hashimoto, T., & Yamamoto, T. (1991) *J. Biol. Chem. 265*, 8681–8685.
Taylor, W. R., & Green, N. M. (1989) *Eur. J. Biochem. 179*, 241–248.
Tokita, K., Hori, K., Kurotsu, T., Kanda, M., & Saito, Y. (1993) *J. Biochem. 114*, 522–527.
Turgay, K., Krause, M., & Marahiel, M. A. (1992) *Mol. Microbiol. 6*, 529–546.
van Dongen, M. B. M., de Geus, J. P., Korver, T., Hartog, A. F., & Berden, J. A. (1986) *Biochim. Biophys. Acta 850*, 359–368.
van Liempt, H., Pfeifer, E., Schwecke, T., Palissa, H., & von Döhren, H. (1991) *Biomed. Biochim. Acta 50*, 256–259.
von Döhren, H. (1993) *Biochem. Soc. Trans. 21*, 214–217.
Walker, J. E., Seraste, M., Runswick, M. J., & Gay, N. J. (1982) *EMBO J. 1*, 945–951.
Weckermann, R., Fürbass, R., & Marahiel, M. A. (1988) *Nucleic Acids Res. 16*, 11841.
Zheng, J., Knighton, D. R., Ten Eyck, L. F., Karlsson, R., Xuong, N., Taylor, S. S., & Sowadski, J. M. (1993) *Biochemistry 32*, 2154–2161.
Zhuo, S., Garrod, S., Miller, P., & Allison, W. S. (1992) *J. Biol. Chem. 267*, 12916–12927.