PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 9/12, 15/54, C12P 19/62, 29/00, 21/04, 37/00, 35/00, 19/50, C12Q 1/48, C12N 1/21

A2

(11) International Publication Number: **WO 97/13845**

(43) International Publication Date: 17 April 1997 (17.04.97)

(21) International Application Number: PCT/US96/16202

(22) International Filing Date: 11 October 1996 (11.10.96)

(30) Priority Data:

60/005,152	13 October 1995 (13.10.95)	US
60/021,650	12 July 1996 (12.07.96)	US

(71) Applicants: PRESIDENT AND FELLOWS OF HARVARD COLLEGE [US/US]; 17 Quincy Street, Cambridge, MA 02138 (US). UNIVERSITY OF CALIFORNIA, SAN FRANCISCO [US/US]; Surge 104-Box 0541, San Francisco, CA 94143 (US).

(72) Inventors: LAMBALOT, Ralph, H.; 5 Pond Street, Wrentham, MA 02093 (US). GEHRING, Amy, M.; Harvard Medical School, Dept. of Biological Chemistry and Molecular Pharmacology, 240 Longwood Avenue, Boston, MA 02115 (US). REID, Ralph; University of California, San Francisco, Biomolecular Resource Center, Surge 104-Box 0541, San Francisco, CA 94143 (US). WALSH, Christopher, T.; Wellesley College, President's House, 735 Washington Street, Wellesley, MA 02181 (US).

(74) Agents: MANDRAGOURAS, Amy, E. et al.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PHOSPHOPANTETHEINYL TRANSFERASES AND USES THEREOF

(57) Abstract

The invention pertains to isolated phosphopantetheinyl transferases, such as the E. coli acyl carrier protein synthase, which transfer a phosphopantetheinyl group onto a substrate. The enzyme can be purified from a natural source, produced recombinantly, or synthetically. Accordingly, the invention provides compositions and kits including phosphopantetheinyl transferases and host cells expressing phosphopantetheinyl transferases. The invention also provides nucleic acids encoding phosphopantetheinyl transferases and vectors comprising such nucleic acids. The invention further provides methods for phosphopantetheinylating a substrate *in vitro* or *in vivo* and methods for producing antibiotics *in vitro* or *in vivo*.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

PHOSPHOPANTETHEINYL TRANSFERASES AND USES THEREOF

Background of the Invention

Acyl carrier protein (ACP) is a small acidic protein (8,800 Da) responsible for acyl group activation in fatty acid biosynthesis. The gene encoding ACP (*acpP*) has been cloned and overexpressed (Rawlings, M. and Cronan, J. E., Jr. (1992) *J. Biol. Chem.*, 267, 5751-5754; Jones, A. L., et al. (1993) *Biochem. Soc. Trans.*, 21, 202S) and the solution structure of ACP has been solved by NMR spectroscopy (Holak, T. et al. (1988) *Eur. J. Biochem.* 175:9-15). Homologs of *E. coli* ACP exist throughout nature in two forms; either as an integral domain of a much larger multifunctional enzyme (type I) or as a discrete protein capable of associating with several other enzymes constituting a multienzyme synthase complex (type II). In these two forms, ACPs play central roles in a broad range of other biosynthetic pathways that depend on iterative acyl transfer steps, including polyketide (Shen, B., et al. (1992) *J. Bacteriol.* 174:3818-3821), non-ribosomal peptide (Baldwin, J. E., et al. (1991) *J. Antibiot.* 44:241-247), and depsipeptide biosynthesis (Rusnak, F., et al. (1991) *Biochemistry* 30:2916-2927) as well as in the transacylation of oligosaccharides (Geiger, O., et al. (1991) *J. Bacteriol.* 173:2872-2878) and proteins (Issartel, J. P., et al. (1991) *Nature* 351:759-761).

A definitive feature of ACP is the 4'-phosphopantetheine (4'-PP) prosthetic group (Majerus, P. W. et al. (1965) *Proc. Natl. Acad. Sci. USA* 53:410-417). 4'-PP is attached through a phosphodiester linkage to a conserved serine residue found in all ACPs. Acyl groups of the many substrates recognized by type I and type II ACPs are activated for acyl transfer through a thioester linkage to the terminal cysteamine thiol of the 4'-PP moiety. The ß-alanyl and pantothenate portions of the 4'-PP structure are believed to serve as a tether between the phosphodiester-ACP linkage and the terminal thioester, suggesting that 4'-PP may function as a swinging arm, shuttling growing acyl chains between various active sites, e.g. as in the sequential addition of 11 amino acids by the 800 kDa cyclosporin synthetase (Lawen, A. and Zocher, R. (1990) *J. Biol. Chem.* 265:11355-11360).

Holo-ACP synthase (holo-ACPS) transfers the 4'-PP moiety from Coenzyme A (CoA) to Ser-36 of apo-ACP to produce holo-ACP and 3',5'-ADP in a Mg^{2+} dependent reaction. The (acyl carrier synthase protein) ACPS from *E. coli* was partially purified 780-fold from crude extracts (Elovson, J. and Vagelos, P. R. (1968) *J. Biol. Chem.* 243:3603-3611), and the ACPS from spinach has been partially purified (Elhussein, S. A.,et al. (1988) *Biochem. J.* 252:39-45), but remarkably little has been shown about the mechanism or specificity of this post-translational phosphopantetheinylation process. A mutant of *E. coli* conditionally defective in the synthesis of holo-ACP has been identified and the mutant phenotype attributed to an altered holo-ACP synthase activity (Polacco, M. L. and Cronan, J. E., Jr. (1981) *J. Biol. Chem.* 256:5750-5754).

Summary of the Invention

This invention pertains to isolated and purified natural and recombinant phosphopantetheinyl transferases, e.g., acyl carrier protein synthases (ACPSs), from eucaryotes, procaryotes, or plants. Also within the scope of the invention are active fragments of phosphopantetheinyl transferases, modified phosphopantetheinyl transferases, and modified active fragments of phosphopantetheinyl transferases. These forms of phosphopantetheinyl transferase are preferably purified to at least about 60% purity, more preferably to at least about 70% purity, more preferably to at least about 80% purity, more preferably to at least about 90% purity and even more preferably to at least about 95% purity. The phosphopantetheinyl transferase of the invention can be used for *in vitro* phosphopantetheinylation of substrates, such as acyl carrier proteins (ACPs), which have, for example, been produced by overexpression in a host cell. Kits including the phosphopantetheinyl transferase described herein are also within the scope of the invention.

The invention also provides host cells modified to express at least one nucleic acid encoding at least one phosphopantetheinyl transferase or active fragment thereof. In one embodiment, the host cells of the invention are further modified to express at least one nucleic acid encoding at least one substrate of a phosphopantetheinyl transferase. Such host cells may further express nucleic acids encoding other components associated with the ACP. Modified host cells of the invention can be used for the production of antibiotics or other compounds whose synthesis requires an ACP.

Brief Description of the Drawings

Figure 1 is a schematic representation of the transfer of a 4'-PP moiety from CoA to Ser-36 of apo-ACP by ACPS to produce holo-ACP and 3',5'-ADP in a Mg^{2+} dependent reaction. Holo-ACP can then activate acyl groups for acyl transfer through a thioester linkage to the terminal cysteamine of the 4'-PP moiety.

Figure 2 Panel A shows the results of a native SDS-PAGE analysis showing *in vitro* formation of holo-ACP using recombinant ACPS (Lane 1, apo-ACP; lane 2, holo-ACP standard (Sigma), reduced with DTT; lane 3, ^3H-labeled holo-ACP formed *in vitro* using recombinant ACPS, reduced with DTT; lane 4, holo-ACP standard, not reduced with DTT; lane 5, ^3H-labeled holo-ACP formed *in vitro* using recombinant ACPS, not reduced with DTT).

Figure 2 Panel B represents an autoradiogram of the gel shown in Figure 2 Panel A which confirms the introduction of [^3H]phosphopantetheine into holo-ACP *(I)*, holo-ACP-CoA mixed disulfide *(II)*, and $(holo\text{-}ACP)_2$ disulfide *(III)*.

Figure 3 shows the results of Tris-Tricine SDS-PAGE analysis of fractions from purification of wild-type and recombinant ACPS (lane 1, crude lysate of *E. coli* K-12; lane 2, DE-52 supernatant; lane 3, SP-Sepharose pool; lane 4, apo-ACP affinity column pool (0.5

mL sample concentrated 20-fold by acetone precipitation); lane 5, SP-Sepharose purified recombinant ACPS).

Figure 4 is a graphic representation of the number of pmol of holo-Dcp formed per number of pmol of apo-Dcp incorporated in an *in vitro* phosphopantetheinylation reaction with recombinant *E. coli* ACPS.

Figure 5 is a schematic representation of the transfer of an acyl group onto the sulfhydryl group of an ACP by fatty acid synthases (FASs) and polyketide synthases (PKSs) and the transfer of an amino-acyl group onto the sulfhydryl group of an ACP by peptide synthesases.

Figure 6 represents an alignment of amino acid sequences of yeast fatty acid synthases Fas 2, Gsp, Lpa, Sfd, EntD, and *E. coli* ACPS. Conserved residues are boxed and indicated by stars.

Figure 7 is a schematic diagram of homologous regions (shaded) among fatty acid synthases (FAS) from S. cervisiae, S. pombe, C. albacans, A. nidulans, and P. patulum, the EntD homolgoy family including Sfp, Gsp, HetI, Lys5, and O195, and *E. coli* AcpS. Ketoreductase (KR), ketosynthase (KS), acyl transferase (AT), enoyl reductase (ER), dehydratase (DH), and malonyl/palmitoyl transferase (MT/PT).

Figure 8 represents an alignment of amino acid sequences of *E. coli* ACPS, entD, sfp, gsp, and the yeast fatty acid synthase fas2 showing the amino acid sequence homology between these proteins. The two areas showing the strongest homology are boxed and shaded.

Figure 9 Panel A is a diagram showing the location of the proposed P-pant transferase domains and the location of consensus sequences in the fungal fatty acid synthases (FAS), the Sfp/Gsp/EntD/o195 homology family, and *E. coli* ACPS. Component FAS activities are abbreviated as AT, acyl transferase; ER, enoyl reductase; DH, dehydratase; MT/PT malonyl/palmitoyl transferase; ACP, acyl carrier protein; KR, ketoreductase; KS, ketosynthase.

Figure 9 Panel B shows amino acid sequence alignments of the consensus sequences of the P-pant transferase enzyme superfamily. Highly conserved residues are boxed.

Figure 10 is a schematic diagram of the protein TycA and the region of 112 amino acids from TycA used for preparing the substrate His6-peptidyl carrier protein (His6-PCP).

Figure 11 represents a histogram showing Sfp, EntD, and o195 mediated incorporation of [3H]-4'-Ppant into ACP and His6-PCP.

Figure 12 represents a histogram showing EntD and Sfp mediated incorporation of [3H]-4'-Ppant into EntF and SrfB1, respectively.

Figure 13 is a graphical representation of the amount of [3H]-4'-Ppant incorporated into PCP-His6 during a 15 minute reaction with 13 nM Sfp and different concentrations of the substrate.

Figure 14 Panel A is a diagram representing the amount of Holo-EntF formed as a function of time upon incubation of Apo-EntF with EntD, ACPS, or o195.

Figure 14 Panel B is a diagram representing the amount of Holo-ACP formed as a function of time upon incubation of Apo-ACP with EntD, ACPS, or o195.

Figure 15 is a histogram representing the extent of [^{14}C]Valine activation by holo-SrfB1 preincubation of SrfB1 with CoA in the absence of Sfp (- Sfp + [14C]Val) or in the presence of Sfp (+ Sfp + [14C]Val) before subsequent incubation with [^{14}C]-L-Valine, or in the presence of Sfp before subsequent incubation with [^{14}C]-L-Aspartate and ATP (+ Sfp + [14C] Asp).

Detailed Description of the Invention

This invention provides isolated and purified phosphopantetheinlyating enzymes, also termed phosphopantetheinyl transferases, such as acyl carrier protein synthases (ACPS) or active fragments thereof. The term "phosphopantetheinyl transferase" and "phosphopantetheinyl transferase enzyme" and "phosphopantetheinylating enzyme" is intended to include a molecule, e.g. enzyme, which catalyzes the transfer of a 4'-phosphopantetheine group from a donor compound, such as CoA, to a substrate. Accordingly, phosphopantetheinyl transferases include natural enzymes, recombinant enzymes, or synthetic enzymes, or active fragments thereof, which are capable or phosphopantetheinylating a compound. Preferred phosphopantetheinyl transferases include enzymes, such as the *E. coli* ACPS, which modify acyl carrier proteins (ACPs) and preferably result in activation of an ACP. The term phosphopantetheinyl transferase also includes enzymes which modify non-acyl carrier proteins, which require a 4'-phosphopantetheine group, for example, for enzymatic activity. The terms "isolated" and "purified" are used interchangeably herein and refer to a phosphopantetheinyl transferase or active fragment thereof that is substantially free of other components of the host organism with which it is associated in its natural state. The terms "isolated" and "purified" are also used to refer to a phosphopantetheinyl transferase or active fragment that is substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized.

In one embodiment, the phosphopantetheinyl transferase or active fragment thereof is purified from a cell naturally expressing the synthase, such as a procaryotic or eucaryotic cell. Since ACPs are involved in fatty acid biosynthesis, it is likely that all species having fatty acids would require phosphopantetheinyl transferases for their synthesis. Procaryotic cells which produce phosphopantetheinyl transferase include bacteria, such as *E. coli*. Phosphopantetheinyl transferases can also be isolated from eucaryotic cells, for example mammalian cells, yeast cells, and insect cells. Other sources for phosphopantetheinylating enzymes include plant tissues, such as spinach and pea leaves (Elhussein, et al. *supra*).

Purification of a phosphopantetheinyl transferase from a cell naturally expressing the enzyme can be accomplished using standard techniques. However, it is preferred that the process of purification of a phosphopantetheinyl transferase from a native source include an affinity purification step over an apo-acyl carrier protein column, such as is further detailed in the Exemplification Section. Use of this affinity purification step allows purification of a phosphopantetheinyl transferase to at least about 60% purity, more preferably to at least about 70% purity, and most preferably to at least about 80% purity. Purification of a phosphopantetheinyl transferase to at least about 90% purity, 95% purity, 97% purity, 98% purity or 99% purity by this technique is preferred. In addition, purification of a phosphopantetheinyl transferase from a natural source by the techniques described herein will preferably result in an enrichment of the phosphopantetheinyl transferase by at least about 800 fold or 1,000 fold, more preferably at least about 10,000 fold, more preferably at least about 50,000 fold, and even more preferably at least about 70,000 fold. It is also preferred that a purified phosphopantetheinyl transferase have a specific activity of about 250 mU/mg, more preferably about 400 mU/mg, and even more preferably about 500 mU/mg of protein. The specific activity of a phosphopantetheinyl transferase is defined in mU/mg with one Unit being equal to 1 μmol of substrate, e.g., holo-ACP, formed per minute at 37 °C in an *in vitro* assay, such as described in the Exemplification section.

Another embodiment of the invention provides an isolated and purified phosphopantetheinyl transferase, such as an ACPS, or active fragment thereof produced by recombinant techniques. Recombinant phosphopantetheinyl transferase or active fragment thereof can be produced by expression of a nucleic acid encoding the synthase or fragment thereof in an appropriate host cell. Host cell, expression vectors and techniques for expression of heterologous nucleic acids are known to those skilled in the art. For example, using recombinant techniques, a phosphopantetheinyl transferase can be expressed as an intracellular protein, a membrane associated protein, or as a secreted protein. In one embodiment of the invention, a phosphopantetheinyl transferase or active fragment thereof is expressed in a procaryotic cell, such as a bacterial cell, preferably *E. coli*. Alternatively, the phosphopantetheinyl transferase or fragment thereof can be expressed in a eukaryotic cell, such as a yeast cell, a mammalian cell, such as the CHO or COS cells, or in an insect cell (baculovirus system). The phosphopantetheinyl transferase or fragment thereof can also be expressed in a plant cell. Nucleic acid regulatory elements required for the expression of nucleic acid encoding a phosphopantetheinyl transferase or fragment thereof in the various systems, and methods for introducing one or more nucleic acids in these host cells are well known in the art. Such techniques are routine and are described, for example in Sambrook et al., (*Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory textbooks.

If a phosphopantetheinyl transferase of the invention is produced recombinantly or synthetically, it may not be nessary to purify the enzyme significantly. A recombinantly produced phosphopantetheinyl transferase may have an activity identical to that of its natural counterpart, or it may have an activity which varies more or less from that of its natural counterpart. For example, a recombinantly produced phosphopantetheinyl transferase can vary in the efficiency of catalysis. It is possible to modify the amino acid sequence of the enzyme to, e.g., improve its efficiency, or to change its substrate specificity, or both.

Using the techniques described herein, a phosphopantetheinyl transferase or active fragment thereof can be "overexpressed" in an appropriate host cell. The term "overexpression" of a phosphopantetheinyl transferase is intended to include expression of the nucleic acid encoding a phosphopantetheinyl transferase to levels which are about 100 fold higher, more preferably about 1,000 fold higher, more preferably about 10,000 fold higher, and even more preferably about 100.000 fold higher than the level of the expression of the endogenous phosphopantetheinyl transferase.

Various methods can be used for isolating and purifying recombinant phosphopantetheinyl transferase or fragment thereof from a host cell, which are preferably adapted to the specific host system used for expression. A preferred method of purification of a phosphopantetheinyl transferase or fragment thereof comprises an apo-acyl carrier protein affinity purification step. A method of purification of phosphopantetheinyl transferase including such an affinity purification step is described in the Exemplification section. Purification of a recombinant phosphopantetheinyl transferase or fragment thereof by this technique results in a preparation of phosphopantetheinyl transferase or active fragment thereof having a purity of at least about 90% or greater, preferably at least about 95% purity, more preferably at least about 97% purity, more preferably at least about 98% purity, and even more preferably at least about 99% purity. The purified recombinant phosphopantetheinyl transferase or fragment thereof preferably has a specific activity of about 200 mU/mg of protein, more preferably about 250mU/mg of protein, and even more preferably about 300 mU/mg of protein.

This invention further pertains to isolated or purified "active fragments" of a phosphopantetheinyl transferase which are capable of phosphopantetheinylating a substrate. An "active fragment" of a phosphopantetheinyl transferase is intended to include a fragment of a phosphopantetheinyl transferase that is capable of phosphopantetheinylating a substrate. Active fragments of phosphopantetheinyl transferases are included in the term phophopantetheinyl transferases as used herein, since active fragmenst of phosphopantetheinyl transferase are capable of transfering a phosphopantetheine to a substrate. Active fragments can be produced using the techniques previously described herein for the preparation of purified phosphopantetheinyl transferase, recombinant phosphopantetheinyl transferase, or chemically synthesized using known techniques. For example, a peptide fragment of a phosphopantetheinyl transferase can be obtained by

expressing nucleic acid fragments of a nucleotide sequence encoding a phosphopantetheinyl transferase in an appropriate host cell and screening the resulting peptide fragments for activity using known techniques. Various methods are known in the art to prepare libraries of clones expressing various fragments of a phosphopantetheinyl transferase. These clones can then be screened to identify those which express an active fragment capable of phosphopantetheinlyating a substrate. In one method, a fragment of a phosphopantetheinyl transferase is tested in an *in vitro* assay in which apo-acyl carrier protein and CoA having a radiolabelled 4'-phosphopantetheine group is incubated with the test phosphopantetheinyl transferase fragment in an appropriate buffer. The amount of radiolabel incorporated in the holo-acyl carrier protein is then measured and is representative of the activity of the fragment of the phosphopantetheinyl transferase. Further details regarding this *in vitro* test are provided in the Exemplification section.

Other compounds within the scope of the invention include modified phosphopantetheinyl transferase or modified active fragments thereof. The term "modification" is intended to include addition, deletion, or replacement of one or more amino acid residues in the phosphopantetheinyl transferase or active fragment thereof, such that the phosphopantetheinylating activity of the enzyme is either maintained, increased, or decreased. "Modified phosphopantetheinyl transferase" or "modified active fragment" also includes a phosphopantetheinyl transferase or fragment in which the substrate specificity has been altered. For example, a modified phosphopantetheinyl transferase or modified active fragment thereof may be capable of phosphopantetheinylating a different acyl carrier protein than the non-modified phosphopantetheinyl transferase. Alternatively, the modified phosphopantetheinyl transferase may have a different range of specificity. In particular, modification of the enzyme can result in a phosphopantetheinyl transferase capable of phosphopantetheinylating additional substrates. On the other hand, modification of a phosphopantetheinyl transferase can result in a phosphopantetheinyl transferase with a more restricted range of substrates. It is possible to determine the target(s) of a phosphopantetheinyl transferase by, for example, performing *in vitro* phosphopantetheinylation tests, such as described above or in the examples and replacing the *E. coli* apo-acyl carrier protein with the substrates that are of interest.

Modified phosphopantetheinyl transferases or modified active fragments thereof further include those in which the stability or solubility, or both has been altered. Other modified phosphopantetheinyl transferases include those wherein a "tag" is attached to the phosphopantetheinyl transferase or active fragment thereof, such as for facilitating purification of the protein from the host cell producing it. Such tags are well known in the art and include polypeptide recognized by specific antibodies.

This invention further provides homologs of phosphopantetheinyl transferases, e.g., those identified herein including the *E. coli* ACPS described in the Exemplification section. The term "homolog" of a phosphopantetheinyl transferase is intended to include

phosphopantetheinyl transferases from species other than the *E. coli* ACPS described herein which are capable of phospopantetheinylating the same or different substrates from those of *E. coli* ACPS. Accordingly, homologs of *E. coli* ACPS are intended to include any enzyme capable of phosphopantetheinylating a substrate. Phosphopantetheinylation is intended to include the transfer of a phosphopantetheine group from one substrate to another substrate, e.g., the transfer of phosphopantetheine from CoA to an ACP, such as is depicted in Figure 1. It has been shown, for example, that plant tissues have a phosphopantetheinyl transferase, and it is highly likely that most cells, whether eucaryote or procaryote also contain at least one enzyme which is capable of phosphopantetheninylating substrates. Homologs of the *E. coli* *ACPS* transferase can be isolated using the nucleic acid encoding the *E. coli* ACPS identified herein. Thus, a nucleic acid encoding a homolog of the *E. coli* ACPS can be obtained by hybridization of the nucleic acid encoding the *E. coli* ACPS, or fragment thereof, e.g., a portion encoding motif 1 and motif 2 (as described herein), to libraries of clones containing nucleic acids from specific sources under low or high stringency conditions. Homologs can also be cloned by other methods known in the art, such as by PCR methods using degenerate oligonucleotides having a nucleic acid sequence derived from that of *E. coli* ACPS. Yet other methods for isolating homologs of the *E. coli* ACPS include those employing an antibody specific for the *E. coli* ACPS. Antibodies for use in these methods can be prepared according to methods known in the art.

Accordingly, nucleic acids encoding a phosphopantetheinyl transferase, or fragment thereof, or homolog thereof are also within the scope of the invention.

Homologs of *E. coli* ACPS, within the scope of the invention can vary in the degree of amino acid sequence homology or identity with the amino acid sequence of *E. coli* ACPS (Lam et al. (1992) *J. Bacteriol.* 174:1554; Takiff et al. (1992) *J. Bacteriol.* 174:1544: SEQ ID NO: 2), so long as the homolog enzyme is capable of phosphopantetheinylating a substrate. Homologs of *E. coli* ACPS within the scope of the invention can be encoded by a nucleic acid having any degree of nucleic acid sequence homology with a nucleic acid encoding *E. coli* ACPS (Lam et al. *supra;* Takiff et al. *supra*). Homology, also termed herein "identity" refers to sequence similarity between two proteins (peptides) or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequences is occupied by the same nucleotide base or amino acid, then the molecules are homologous, or identical, at that position. A degree (or percentage) of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. A degree or percentage of "identity" between amino acid sequences refers to amino acid sequence similarity wherein conserved amino acids are considered to be identical for the purpose of determining the degree or percentage of similarity. A conserved amino acid substitution is, e.g., substitution of one amino acid having a negative side chain for another amino acid having a negative side chain.

In a specific embodiment, preferred *E. coli* ACPS homologs or members of the phosphopantetheine transferase superfamily, have an overall amino acid sequence identity or similarity of at least about 50%, more preferably at least about 60%, more preferably at least about 70 %, more preferably at least about 80%, and most preferably at least about 90% with an amino acid sequence shown in SEQ ID NO: 2 or of the phosphopantetheinylating enzymes and/or enzymes having the conserved amino acid motifs described herein. Peptides having an overall amino acid sequence identity or similarity of at least about 93%, more preferably at least about 95%, and most preferably at least about 98-99% with a sequence set forth in SEQ ID NO: 2 or of the phosphopantetheinylating enzymes and/or enzymes having the conserved amino acid motifs described hereinare also within the scope of the invention.

In yet other embodiments of the invention, *E. coli* ACPS homologs have less than about 50% overall amino acid sequence identity or similarity with an amino acid sequence shown in SEQ ID NO: 2. It has in fact been shown, as described herein, that specific proteins or peptides having limited overall amino acid sequence homology with *E. coli* ACPS are capable of phosphopantetheinylating substrates. In fact, it has been shown herein, that the *B. subtilis* protein Sfp, involved in the biosynthesis of surfactin, a cyclic lipopeptide antibiotic, is a phosphopantetheinyl transferase. Another protein, EntD from *E. coli*, which is involved in the biosynthesis of enterobactin, a secreted iron-scavenging dihydroxylbenzoyl-serine trilactone, is also capable of phosphopantetheinylating substrates. Further, the *E. coli* protein o195 of unknow function was also shown herein to phosphopantetheinylate substrates. Even though EntD, Sfp, and o195 have only limited overall amino acid sequence homology with *E. coli* ACPS, the C-terminus of EntD and Sfp have 2 regions of conserved amino acid residues with *E. coli* ACPS. The regions of homology between the amino acid sequences are represented in Figures 6 and 8. Further sequence comparison revealed that these regions of conserved amino acid residues are also present in the following proteins: the Fas2 protein from several organisms, including *S. cerevisiae*, which is involved in fatty acid synthesis, the Gsp protein from *B. brevis* which is involved in the synthesis of the peptide antibiotic gramicidin, and the Lpa protein from *B. subtilis* (Figure 9). Additional phosphopantetheinyl transferase homologs, listed in Figure 9, were found to be involved, e.g., in cyanobacterial heterocyst formation and lysine biosynthesis in yeast. Based at least in part on these amino acid sequence homologies, Fas2, Gsp, and Lpa are also likely to be phosphopantetheinyl transferases. Mutagenesis studies have further provided evidence that these conserved regions are required for phosphopantetheinylation of substrates.

Accordingly, preferred phosphopantetheinyl transferases of the invention have at least one or more conserved amino acids, such as the asterisked amino acids shown in Figure 6 or the regions shown in Figure 9, or regions of conserved amino acids, such as the boxed amino acid regions in Figures 6, 8, and 9. Even more preferred phosphopantetheinyl transferases of the invention have one or more of the following amino acid sequence motifs, wherein N

represents any amino acid residue and two amino acids separated by "/" respresents a residue which can be either amino acid:

G-N-D-N-N-E (motif 1a)	(SEQ ID NO: 3)
G-N-D (motif 1b)	(SEQ ID NO: 4)
W-S-A-K-E-N-N-N-K-N-N-G (motif 2a)	(SEQ ID NO: 5)
F/W-N-N-K/R-E-S/A-N-N-K (motif 2b)	(SEQ ID NO: 6)

In an even more preferred embodiment, a phosphopantetheinyl transferase of the invention comprises at least one motif 1 (a or b) and at least one motif 2 (a or b). Yet further preferred phosphopantetheinyl transferases of the invention comprise a motif 1 (a or b) and motif 2 (a or b) separated by at least about 5, preferably at least about 10, more preferably at least about 15, 20, 25, 30, 35, 40, 45, 50, and 55 amino acid residues. In a preferred embodiment, the two motifs are separated by at least about 30 to 45 amino acid residues. Other prefered phosphopantetheinyl transferases of the invention include those comprising amino acid sequences which are significantly homologous or similar to the amino acid sequence of motifs 1a, 1b, 2a, or 2b. Also within the scope of the invention are phosphopantetheinyl transferases which contain a motif 1 and a motif 2 which contain amino acid substitutions, deletions, or additions. Preferred amino acid substitutions are conserved amino acid substitutions, e.g, a substitution of one amino acid for another having a similar characteristic.

Other phosphopantetheinyl transferases within the scope of the invention comprise one or more motifs 1 and 2 having the following amino acid sequences or having one or more conserved amino acid substitutions in these sequences, i.e, substitution of one amino acid with a similar amino acid, or having a deletion of one or more amino acids:

<u>motif 1 sequences:</u>

GTDIVEIARI	(SEQ ID NO: 5)
GIDIEEIFSV	(SEQ ID NO: 6)
GIDIEKTKPI	(SEQ ID NO: 7)
GIDIERISEI	(SEQ ID NO: 8)
GVDVELITSI	(SEQ ID NO: 9)

Motif 2 sequences:

FAVKEAAAKAFG (SEQ ID NO: 10)
FSAKESAFKASE (SEQ ID NO: 11)
WSMKESFIKQEG (SEQ ID NO: 12)
WTIKESYIKAIG (SEQ ID NO: 13)
WSAKEAVFKSLG (SEQ ID NO: 14)

Also within the scope of the invention are phosphopantetheinyl transferases which contain a motif 1 having an amino acid sequence listed above or indicated in the left column of Figure 9 and a motif 2 having an amino acid sequence listed above or indicated in the right column of Figure 9. phosphopantetheinyl transferases having amino acid substitutions, e.g., conservative substitutions, deletions or amino acids are also within the scope of the invention. Several assays can be used to confirm that a phosphopantetheinyl transferase having modified or non-modified motif 1 and 2 sequences has a phosphopantetheinylating activity and are described, e.g., in the Examples. Such assays are preferably peformed on various substrates since certain phophopantetheine transferases show substrate specificity.

Preferred phophopantetheinylating enzymes have an amino acid sequence identity or similarity of at least about 50%, more preferably at least about 60%, more preferably at least about 70 %, more preferably at least about 80%, and most preferably at least about 90% with an amino acid sequence of a motif 1 and a motif 2 shown above or in any Figure. Peptides having an amino acid sequence identity or similarity of at least about 93%, more preferably at least about 95%, and most preferably at least about 98-99% with an amino acid sequence of a motif 1 and a motif 2 shown above or in any Figure are also within the scope of the invention.

Based on the existence of amino acid sequence homologies between phosphopantetheinylating enzymes, it is possible to identify additional phosphopantetheinylating enzymes. Such phosphopantetheinylating enzymes are also within the scope of the invention. Several methods can be used to identify additional phosphopantetheinylating enzymes: amino acid or nucleic acid sequence comparisons and *in vitro* assays, as described herein.

Enzymes phosphopantetheinlyated according to the methods of the invention may be involved in transfer of various groups onto the newly introduced sulfhydryl (SH) group of the phosphopantetheine prosthetic group which acts as a nucleophile. Acyl-CoAs can be used for fatty acid synthesis (FAS) and all the polyketide synthases (PKS), whereas aminoacyl-AMPs can be used for the peptide synthetases (Figure 5). In the PKS complexes the acyl-ACPs undergo capture by carbanion nucleophiles for carbon skeleton assembly in polyketide construction while in peptide and depsipeptide synthetases, the aminoacyl-S-ACPs are attacked by nitrogen and oxygen nucleophiles in amide and ester bond forming steps.

A phosphopantetheinyl transferase protein, active fragment thereof, modified phosphopantetheinyl transferase or modified active fragment, and phosphopantetheinyl transferase homologs purified from their natural source or produced as recombinant protein, e.g., according to the techniques described herein can then be used for *in vitro* phosphopantetheinylation of a substrate, such as an acyl carrier protein. In a specific embodiment, the substrate is a "heterologous" substrate, i.e., a substrate from a species different from the species from which the phosphopantetheinyl transferase originates. It has in fact been shown that a phosphopantetheinyl transferase from one species is capable of phosphopantetheinylating a substrate from another species. For example, several plant enzymes phosphopantetheinylated *E. coli* ACP as effectively as plant ACP (Elhussein, et al. *supra*). Furthermore, it has also been shown that *E. coli ACPS* is capable of phosphopantetheinylating acyl carrier proteins other than *E. coli* ACP: (1) *Lactobacillus* D-Alanyl Carrier Protein (DCP) which is active in the biosynthesis of the bacterial cell wall component lipoteichoic acid, (2) *Rhizobial* NodF which is active in the nodulation process of leguminous vegetables, and (3) *Streptomycete* ACPs involved in actinorhodin, granaticin, tetracenomycin, frenolicin, oxytetracycline, and tetracenomycin polyketide antibiotic biosynthesis.

In one embodiment, a phospopantetheinylating enzyme within the scope of the invention is capable of phosphopantetheinylating predominantly a single substrate. In another embodiment of the invention, the phophopantetheinlyting enzyme, e.g. *E. coli ACPS*, is capable of phosphopantetheinylating a limited number of substrates. In yet another embodiment, the phosphopantetheinylating enzyme has a wide range of substrates.

This invention further provides kits for *in vitro* phosphopantetheinylation of substrate proteins. Such kits include an isolated or purified phosphopantetheinyl transferase or active fragment thereof as described herein and instructions for use. The phosphopantetheinyl transferase or active fragment thereof can be produced by recombinant technique, chemically synthesized, or purified from a native source. Such phosphopantetheinyl transferase or fragment thereof preferably has a specific activity of at least 250 mU/mg. The phosphopantetheinyl transferase or fragment thereof can be packaged, such that it retains substantially all of the phosphopantetheinylation activity for at least one month, preferably for at least 2 months, even more preferably for at least 3 months. Even more preferably, the phosphopantetheinyl transferase is stable for at least 6 months and most preferably for at least 1 year. For example, the phosphopantetheinyl transferase or fragment thereof can be maintained at -20 °C in a Tris-based buffer containing magnesium and about 20% glycerol. Alternatively, the phosphopantetheinyl transferase or fragment thereof can be provided as a lyophilized powder. The kits can further include a 4'-phosphopantetheinylate group providing reagent, such as Coenzyme A (CoA). In addition, an appropriate reaction buffer for the phosphopantetheinylation reaction, preferably in a concentrated form can be included in the kit. The reaction buffer preferably comprises magnesium ions and Tris buffer.

Compositions comprising a phosphopantetheinyl transferase and an appropriate buffer are also within the scope of the invention. The buffer can be a buffer which maintains or increases the stability of the transferase or which permits the transferase to be maintained in a form under which it maintains its activity or in which its activity can be recovered once the transferase is incubated in appropriate conditions. For example, the buffer can contain an appropriate amount of glycerol, or any other solution allowing the enzyme to be frozen and to be active upon thawing of the transferase. The phosphopantetheinyl transferase can also be in a lyophilized form.

Another embodiment of the invention provides host cells modified to express at least one nucleic acid encoding a phosphopantetheinyl transferase or fragment thereof. Host cells expressing or overexpressing a phosphopantetheinyl transferase have been described herein. In addition, this invention provides host cells transformed with a nucleic acid encoding a non-secreted form of a phosphopantetheinyl transferase or fragment thereof and at least one additional nucleic acid encoding at least one polypeptide. In a preferred embodiment, the at least one polypeptide is a component of the substrate of the phosphopantetheinyl transferase. It is known that expression of heterologous 4'-phosphopantetheine synthases in *E. coli* is limited by the ability of *E. coli* to phosphopantetheinylate large amounts of overproduced protein *in vivo*. Thus, in one embodiment, the invention provides a method for expressing an active, recombinant form of an enzyme requiring a 4'-phosphopantetheine group by expressing or overexpressing both the enzyme and a phosphopantetheinyl transferase in the same host cell. The enzyme coexpressed in the host cell can be any enzyme which requires phosphopantetheinylation to be acitve in a reaction. For example, a host cell can be modified to express both the *E. coli ACPS* and the *E. coli* acyl carrier protein or a heterologous substrate, such as NodF or the D-Alanyl Carrier Protein, such that it produces active *E. coli* acyl carrier protein, NodF, or D-Alanyl Carrier Protein, respectively.

In a preferred embodiment of the invention, host cells expressing or overexpressing a phosphopantetheinylating enzyme, e.g., phosphopantetheinyl transferase, or active fragment thereof are used for the production of compounds whose synthesis have at least one step requiring a phosphopantetheinylating enzyme. Such compounds include polyketides (both aromatic and macrolide) and non-ribosomally produced peptides, including depsipeptides, lipopeptides, and glcyopeptides. Biosynthetic pathways that can be modulated with the phophopantetheinylate transferases of the invention include fatty acids, lipopeptide sufactants, and antibiotics. Preferred antibiotics are those non-ribosomally produced peptides, which can be synthesized by a thiotemplate mechanism. Much preferred non-ribosomally produced peptide antibiotics include (a) linear peptides, e.g, Edeine, ACV, Gramacidin, and Alamethicinecyclic peptides, (b) cyclic peptides, e.g., Cyclopeptin, Enterochelin, Ferrichrome, Gramacidin S, Tyrocidine, and Mycobacillin, (c) lactones, e.g., Destruxin, Actinomycin, Etamycin, and Echinomycin, (d) branched cyclopeptides, e.g, Polymyxin and Bacitracin, and (e) Depsipeptides, e.g., Enniatin and Beauvericin (Kleinlauf

and von Doehren (1990) *Eur. J. Biochem.* 192:1). Other antibiotics whithin the scope of the invention include erythromycin, clarythromycin, oxytetracycline, bacitracin, cyclosporin, penicillins, cephalosporins, vancomycin. Additional antibiotics within the scope of the invention include those well known in the art, which can be found, e.g., in pharmacology catalogs.

The synthesis of coumpounds of the invention, e.g, antibiotics can be catalyzed by enzymes including polyketide synthases (which are involved for example in the synthesis of erythromycin and tetracyclin), non-ribosomal peptide synthetases, and depsipeptide synthetases. These enzymes belong to the family of acyl carrier proteins, which are homologs of *E. coli* ACP, and require a 4'-phosphopantetheine group for their activity. The ACPs are either type I or type II acyl carrier proteins. The invention thus provides methods for producing compounds, e.g., antibiotics, whose synthesis is catalyzed by either a type I or the type II ACP.

Type I ACPs (also termed type I synthase) are multifunctional enzymes, also termed "multienzyme polypeptides" containing in addition to a domain which is capable of phosphopantetheinylating a substrate, at least some, or all, catalytic activities necessary for peptide formation and activation. Type I ACPs include erythromycin, rapamycin, cyclosporin, avernectin, and tetracyclin synthases. Thus, this invention provides a method of producing a compound, e.g., an antibiotic, whose synthesis is catalyzed by a type I ACP, by use of a host cell that has been modified to express a type I ACP and a phosphopantetheinyl transferase to activate the type I ACP. The modified host cell is incubated in a medium containing the necessary substrates for the multienzyme polypeptide. Alternatively, compounds requiring a type I ACP for their synthesis can be produced *in vitro* with purified enzymes or with recombinantly produced enzymes.

Type II ACPs (also termed type II synthase) are discrete proteins which are capable of associating with several other enzymes to form a multienzyme synthase complex. Type II ACPs are components of tetracyclin, gramicidin, tyrocidin, anthrocyclin, and bacitracin synthetases. Thus, this invention provides a method for producing an antibiotic whose synthesis is catalyzed by a type II ACP by introducing into, and expressing in, a host cell the nucleic acids encoding at least some or all the subunits of the multienzyme complex and a nucleic acid encoding a phosphopantetheinyl transferase and incubating such a modified host cell in a medium containing the appropriate substrates for the enzymes. Alternatively, compounds requiring a type II ACP in their synthesis can be produced *in vitro*, by contacting purified phosphopantetheinyl transferase(s) and the appropriate enzymes necessary for synthesis of the compound, in an appropriate buffer.

Preferred host cells for producing antibiotics include cells for which the antibiotic is non toxic. Alternatively, a modified form of the antibiotic that is non toxic to the cell can be expressed and the modified form altered *in vitro* or *in vivo* to obtain the active form of the antibiotic.

Also within the scope of the invention are organisms which naturally produce antibiotics and which have been modified to express or overexpress a phosphopantetheinyl transferase. Such modified organisms may produce additional quantities of an antibiotic by assuring that the type I or type II ACPs required for their synthesis are in an activated form.

In type I and type II ACPs, the multicomponent system contains a small (e.g., about 80 to 100 amino acids) protein subunit or domain that functions as a carrier protein for the growing acyl chain. These acyl carrier proteins, recognizable by the conserved sequence signature motif D(N,D)FFX(L,I)GG(H,D)S(L,I)X(A,G,C)XX(L,V,M) (SEQ ID NO: 15) (Stein et al. (1995) *Biochemistry* 34:4633) or (L,V)(G,L)(G,A,F,Y)(D,H,K,E)S(L,Q)(D,A,G) (Schlumbohm, W. et al.(1991) *J. Biol. Chem.* 266:23135) (SEQ ID NO:) are converted from inactive apo forms to functional holo forms by postranslational modification involving attack of the conserved serine ß-CH_2OH side chain on the pyrophosphate linkage of CoA, resulting in transfer of the 4-phosphopantetheine moiety of CoA onto the attacking serine. The newly introduced -SH of the phosphopantetheine (P-pant) prosthetic group now acts as a nucleophile for acylation by a specific substrate, i.e. acyl-CoA and malonyl-CoA derivatives for the fatty acid and polyketide synthases (PKS) and aminoacyl-AMPs for the peptide and depsipeptide synthetases. In the PKS complexes the carboxy-activated malonyl-ACP derivative then undergoes decarboxylation, forming a nucleophilic carbanion species which attacks a second acyl thiolester to yield a new carbon-carbon bond polyketide biosynthesis. In peptide and depsipeptide synthetases, the aminoacyl-ACPs or hydroxyacyl-ACPs serve as nucleophiles in amide and ester bond-forming steps respectively. Thus, phosphopantetheinylation of apo-ACP domains plays a central role in activation of multienzyme synthases responsible for the biogenesis of a vast array of natural products.

Nucleic acids encoding type I and type II ACPs and components associated therewith have been isolated and are described, for example, in the following references: Donadio, S. et al. (1991) *Science* 252:675; Gocht, M. et al. (1994) *J. Bacteriol.* 176:2654; MacCabe, A. et al. (1991) *J. Biol. Chem.* 266:12646; Schweke, T. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:7839; Ye, J. et al. (1994) *J. Bacteriol.* 176:6270; and Zhang et al. (1995) *J. Bacteriol.* 177: 4009.

In a specific embodiment, the invention provides a phosphopantetheinyl transferase, e.g., an enzyme in which this enzymatic activity is the main function of the enzyme. *E. coli ACPS* is likely to be such an enzyme. In another embodiment, the invention provides phosphopantetheinyl transferase molecules which have additional enzymatic activities, such as a molecule which has both a phosphopantetheinyl transferase and an ACP activity. For example, the yeast fatty acid synthase subunit II (FAS II) has been shown herein to have a domain homologous to *E. coli ACPS* and Sfp, indicating that this enzyme could act intramolecularly to add a phosphopantetheinyl unit to Ser-180, the putative ACP domain of this polyprotein. Accordingly, the invention provides polyfunctional or polyenzymatic

proteins comprising as one of the various enzymatic functions, the ability to transfer phosphopantetheinyl groups. Such transfer can be onto the same or another molecule.

The invention further provides proteins which do not normally possess the ability to phosphopantetheinylate, but which may be modified to obtain this enzymatic activity. For example, a peptide capable of transfering a phosphopantetheinyl group can be grafted onto, or linked to, another protein by, e.g., chemical cross-linking or recombinant methods. In a preferred embodiment, at least one phosphopantetheinyl transferase or active fragment thereof is linked to a protein which is an ACP. The transferase activity that is grafted can be from *E. coli ACPS*, from Sfp, or any other protein having such catalytic activity.

Other compounds whose synthesis require a type I and type II ACP (also termed type I and type II synthases, respectively) and which thus require phosphopantetheinylation by a phosphopantetheinyl transferase include immunosuppressant agents (e.g., FK506), antifungal agents (e.g., amphotericin), antiparasitic agents, cardiovascular agents (e.g., lovostatin) and antitumor agents (e.g., anthracyline) agents among others. Thus, also within the scope of the invention are methods for producing such compounds using the phosphopantetheinyl transferases, active fragments thereof and expression systems described herein.

Methods for producing purified phosphopantetheinyl transferase are also within the scope of the invention. In one embodiment, a phosphopantetheinyl transferase or active fragment thereof is isolated from a cell naturally producing the enzyme, such as *E. coli* and purified by a factor of at least about 1,000 fold, more preferably at least about 10,000 fold and even more preferably at least about 50,000 fold. Methods for isolating the enzyme preferably include an affinity purification step over a column containing apo-acyl carrier protein. Such methods are further described in detail herein. The purified phosphopantetheinyl transferase preferably has a specific activity of 250 mU/mg of protein, more preferably 400 mU/mg, and even more preferably 500 mU/mg of protein. The invention also provides methods for producing recombinant forms of phosphopantetheinyl transferases. Such methods comprise transforming a host cell with a nucleic acid encoding a phosphopantetheinyl transferase, under conditions appropriate for expression of the protein and isolating the synthase from the host cells and culture media. A purified recombinant phosphopantetheinyl transferase preferably has a specific activity of 200 mU/mg, and more preferably 250 mU/mg.

Also within the scope of the invention are methods for identifying inhibitors of a phosphopantetheinyl transferase. An inhibitory compound is defined as a compound that reduces or inhibits phosphopantetheinylation of a substrate by a phosphopantetheinyl transferase. Inhibitors of phosphopantetheinyl transferase can be identified by using an *in vitro* or an *in vivo* test. In one embodiment, potential inhibitory compounds are screened in an *in vitro* phosphopantetheinylation assay, such as an assay described in the Exemplification section. In another embodiment, potential inhibitory compounds are contacted with a microorganism having a phosphopantetheinyl transferase and the phosphopantetheinylation

of substrate is monitored. Inhibitory compounds can be used, for example, for blocking specific pathways requiring phosphopantetheinylation in microorganisms, such as ACP dependent pathways. ACPs are involved, for example, in the transacylation of proteins, such as haemolysin, a pathogenic factor of *E. coli*. The identification of the yeast protein Lys 2, involved in a primary lysine biosynthetic pathway, as a phosphopantetheinylating enzyme, provides a method for killing yeast. Accordingly, inhibitors of phosphopantetheinylating enzymes can be used as anti-fungal agents. Moreover, many membrane-associated proteins are acylated and blocking or reducing acylation of these proteins in a microorganism can potentially affect the viability of the microorganism. ACPs have also been involved in transacylation of cell wall components, which are popular therapeutic targets. Other reactions involving ACPs include transacylation of oligosaccharides, involved for example in the nodulation factors of Rhizobia genus. Transacylation of oligosaccharides are an essential step in nodule formation in nitrogen fixing plants and has agricultural applications.

Yet other embodiments within the scope of the invention include screening methods to uncover novel antibiotics. In one embodiment of the invention, novel antibiotics are identified by coexpressing in a host cell a library of various forms of type I or type II ACPs, such as modified ACPs obtained by random mutagenesis and/or a phosphopantetheinyl transferase, or mutated form thereof, which is required for phosphopantetheinylation and thus activation of the ACPs. Similarly, the invention provides methods for isolating novel immunosuppressant, antifungal, antihelminthic, antiparasitic, and antitumor agents.

It has been shown herein that the *E. coli* EntD gene which encodes a product necessary for the synthesis of the iron-chelating and transport molecule enterobactin (Ent) shows significant amino acid sequence homology with the phosphopantetheinyl transferases *E. coli ACPS* and Sfp, and that it phosphopantetheinylates at least two substrates, ACP and PCP. Accordingly, the invention provides methods for modulating iron uptake by bacteria. In one embodiment, the invention provides a method for improving iron uptake of bacteria, such as by expressing in bacteria the gene EntD. Improved iron uptake can result in more efficient bacterial growth, which will be useful in cases where bacteria are grown for the production of a beneficial compound. The invention also provides methods for reducing iron uptake by bacteria, such as by growing bacteria in the presence of an inhibitor of phosphopantetheinylation. Such an inhibitor can be isolated as described above and in the exemplification section. Accordingly, the invention provides a method for inhibiting bacterial growth or killing of bacteria by iron deprivation.

The invention will also be useful for the design of strategies for heterologous production of functional polyketide and polypeptide synthetases, e.g., in combinatorial biosythesis of "unnatural" natural products. For example, new drug molecules can be made by genetically altering gene clusters in an organism, such as Streptomyces, that synthesize polyketides. Then the modified genes, e.g., produced in E. coli can be reinserted in the Streptomyces where new polyketides are expressed. Such a process can be applied to the

synthesis of a variety of compounds in addition to polyketide antibiotics. Thus, processes for producing new compounds in a host or *in vitro*, by using modified, e.g., mutated phosphopantetheinylating enzymes, are within the scope of the invention. Preferred biosynthetic pathways that can be modified according to this method include those in involved in the biosynthesis of erythromycin, the anticancer drug daunorubicin, and the immunosuppressant, rapamycin.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents, and published patent applications cited throughout this application are hereby incorporated by reference.

Example 1: Cloning and overproduction of the *Escherichia coli* Holo-Acyl Carrier Protein Synthase

Materials and Methods:

Preparation of [^{3}H]Coenzyme A. CoA (220 mg) was labeled by tritium gas exposure (New England Nuclear) to yield 600 mCi of crude material. This material was added to unlabeled CoA and the mixture was acylated and purified as described by Elovson and Vagelos (*supra*). In this manner, [^{3}H]CoA with specific activities as high as 7×10^{14} dpm/mol and having 70% of the ^{3}H-label in the phosphopantetheine portion was prepared.

Assay of phosphopantetheinyl transferase Activity. In a typical assay, 100 µM [^{3}H]CoA, 50 µM apo-ACP, 10 mM $MgCl_2$, 50 mM Tris·HCl, pH 8.8, and ACPS in a final volume of 100 µL were incubated at 37 °C for 30 min in a 1.5 mL microcentrifuge tube. Reactions were quenched with 800 µL 10% TCA. BSA (20 µL of a 25 mg/mL solution) was then added to facilitate precipitation of radiolabeled protein. The 1.5 mL tubes were centrifuged at 12,000 x g for 5 min. Supernatants were removed and the pellets were rinsed with 3 x 900 µL of 10% TCA. Residual TCA was collected by centrifugation, and the pellets were resuspended in 150 µL of 1 M Tris base. The resuspended pellets were transferred to scintillation vials, 2.5 mL of scintillation cocktail (Packard) was added and the amount of ^{3}H-labeled holo-ACP formed was quantified by liquid scintillation counting.

Confirmation of in vitro Holo-ACP Formation by native-PAGE, Autoradiography and Mass Spectrometry. ACPS assays were incubated for 12 h at 37 °C. Control assays were worked up in the usual manner, and holo-ACP formation was confirmed by liquid scintillation counting. Assay mixtures for native-PAGE were not quenched with 10% TCA. The ACPS assay mixture was divided into two equal portions. To one 50 µL portion was added 20 µL of 5 x native-PAGE sample buffer (Ausubel, F. M., et al. (1992) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York) which contained DTT whereas the other sample was not reduced with DTT. These samples were analyzed by 20% native gel electrophoresis followed by Coomassie staining. The stained gels were soaked in Amplify

(Amersham) for 15 min before drying under vacuum. The dried gels were autoradiographed at -80 °C followed by photographic development (Figure 2). Holo-ACP-SH migrates slightly faster than apo-ACP on 20% native gels whereas holo-ACP dimer migrates considerably slower (Rock, C. O. and Cronan, J. E., Jr. (1981) *Methods Enzymol.* 71:341-351).

Overproduction and Purification of Apo-ACP. E. coli DK554, an apo-ACP overproducer strain, was provided by Prof. John E. Cronan, Jr. (Department of Microbiology and Biochemistry, University of Illinois at Urbana-Champaign). Cultures grown in Terrific Broth supplemented with 50 mM glucose, 25 μM pantothenate and 50 μg/mL kanamycin were induced with 1 mM IPTG at an O.D. of 0.8. Cells were lysed by two passages through a French pressure cell at 10,000 psi. The majority of overproduced ACP was present in the apo-form. Minor amounts of holo-ACP were converted to apo-ACP using endogenous holo-ACP hydrolase by incubating the lysate with 10 mM $MgCl_2$ and 2 mM $MnCl_2$ for 60 min at 25 °C with stirring (Fischl, A. S. and Kennedy, E. P. (1990) *J. Bacteriol.* 172:5445-5449). Apo-ACP was then purified following the procedure of Rock and Cronan (*supra*) to yield 60 mg per L culture.

Purification of ACPS from E. coli K-12. A 500 g frozen block of *E. coli* K-12 cells (ATCC 14948) grown to 3/4 log phase (University of Alabama Fermentation Facility) was broken into smaller pieces with a mallet and added to 1 L of 50 mM Tris, 10 mM $MgCl_2$, 1 mM DTT, 1 mM PMSF, 1 mM benzamidine, 50 μM CoA, and 5% (w/v) glycerol, titrated to pH 8.1 with 1 M MES. The cells were lysed by a single passage through an Amicon French pressure cell at 8,000 - 16,000 psi. Cellular debris was removed by centrifugation at 8,000 x g for 30 min to yield 1.5 L of crude extract. The supernatant was added to 150 g of DE-52 slurry (Whatman) in 50 mM Tris·HCl, pH 8.0 and mixed gently for 15 min at 4 °C. DE-52 was removed by centrifugation and the supernatant was treated once again with 150 g of DE-52, pH 8.0. After the removal of the DE-52 resin the supernatant was clarified further by centrifugation at 16,000 x g for 30 min.

The clarified extract (1.3 L) was titrated to pH 6.5 with a saturated MES solution and was then loaded at a flow rate of 10 mL/min onto a 3 x 30 cm SP-Sepharose column (Pharmacia) which had been pre-equilibrated with 50 mM MES, 10 mM $MgCl_2$, 5% (w/v) glycerol, pH 6.1 (Buffer A). After the extract was loaded, the column was washed with 750 mL of Buffer A while collecting 25 mL fractions. The column was then eluted with a linear 0 to 1 M NaCl gradient (1 L) in Buffer A. Active fractions were pooled to yield 190 mL. This 190 mL SP-Sepharose purified material was next loaded at a flow rate of 2 mL/min onto a 2.5 x 4.0 cm Affi-Gel 15 apo-ACP affinity column (BioRad. Hercules, CA, prepared following the manufacturer's instructions) while collecting 25 mL fractions. The column was washed with 100 mL of Buffer A, and ACPS was then eluted with 50 mL of 6 M guanidinium·HCl in 50 mM MES, pH 6.1 while collecting 8 mL fractions. phosphopantetheinyl transferase activity was reconstituted by diluting the guanidinium·HCl to a final concentration < 2M in the assay mixture. Active fractions were pooled to yield 16

mL which was then dialyzed against 2 x 1L of Buffer A to yield about 0.2 mg protein of apparent 70,000-fold purity (Table 1). Tris-Tricine SDS-PAGE analysis (Schagger, H. and von Jagow, G. (1987) *Annal. Biochem.* 166:368-379) revealed the presence of only a few major bands (Figure 3). Previous purifications had demonstrated that the 14 kDa band copurified with the phosphopantetheinyl transferase activity (data not shown). Aliquots (500 µL) of the affinity purified protein were concentrated by acetone precipitation. The precipitated protein was resolved by 16% T, 6% C Tris·Tricine SDS-PAGE and then electroblotted to a Pro-Blot membrane (Applied Biosystems Inc., Foster City, CA) following the manufacturer's instructions. Proteins were visualized by staining briefly with 0.1% amido black in 1% acetic acid. The 14 kDa protein was excised and submitted for N-terminal sequencing.

Table 1. Purification of wild-type ACPS from *E. coli* K-12

fraction	protein mg	activity mU[a]	specific activity mU·mg^{-1}	-fold	% Overall Yld.
Crude extract	50,000	420	8.4 10^{-3}	-	-
- DE-52	13,000	340	2.6 x 10^{-2}	3	80
SP-Sepharose	260	320	1.2	140	75
Apo-ACP affinity	~0.2	120	~600	~70,000	30

[a] One unit of activity produces 1 µmol holo-ACP per minute.

Cloning and Overexpression of the dpj gene. The *dpj* gene was amplified using a freshly-grown single colony of *E. coli* strain BW13711 as template in the polymerase chain reaction (PCR). *E. coli* strain BW13711, a gift from Professor Barry Wanner of Purdue University, has a lacX74 deletion of the entire lac operon but is otherwise wild-type *E. coli* K-12 that has been cured of lambda and the F factor. The forward primer incorporated an *NdeI* restriction site at the start codon: 5'-TGTACCTCAGACCATATGGCAATATTAGGTTTAGGCACGG-3' (SEQ ID NO: 16). The reverse primer incorporated a *HindIII* restriction site after the stop codon: 5'-TGATGTCAGTCAAGCTTAACTTTCAATAATTACCGTGGCA-3' (SEQ ID NO: 17). The resulting PCR product was subcloned into the *NdeI/HindIII* site of the pET22b expression plasmid (Novagen) using standard molecular biology procedures and designated pDPJ (Asubel, F.M., et al. *supra*). *E. coli* BL21(DE3) was transformed with supercoiled pDPJ.

Three 1 L cultures of *E. coli* BL21(DE3)pDPJ in 2xYT media supplemented with 50 μg/mL ampicillin were grown at 37 °C, 250 rpm to an O.D. of 0.8 - 1.0 before transferring the cultures to 30 °C, 250 rpm and inducing with 100 μM IPTG. Cultures were grown at 30 °C for an additional 3 h and were then harvested by centrifugation. Cells were resuspended (5 mL/ g wet cell mass) in 50 mM Tris·HCl, 10 mM $MgCl_2$, 5% glycerol, pH 8.0 (Buffer B) and lysed by two passages through a French pressure cell at 10,000 - 15,000 psi. Cellular debris was removed by centrifugation at 16,000 x g for 30 min. The cell free extract was then treated twice with an equal volume of DE-52 slurry (pH 8.0). The DE-52 supernatant was adjusted to pH 6.5 with a saturated MES solution and loaded onto a 3 x 30 cm SP-Sepharose column which had been pre-equilibrated with Buffer A. The column was washed with 250 mL Buffer A. ACPS was then eluted with a linear 500 mL, 0 to 1 M NaCl gradient.

Results:

In order to initiate purification of ACPS, a rapid and reliable assay for monitoring phosphopantetheinyl transferase activity through the purification process was sought. Of the several methods described for the *in vitro* determination of phosphopantetheinyl transferase activity (Elovson, J. and Vagelos, P.R. *supra*; Elhussein et al. *supra*; Polacco, M.L. and Cronan, J. E., Jr. *supra*) the direct-discontinuous radioassay employing [^{3}H]-(*pantetheinyl*)-CoA and apo-ACP was chosen. The amount of phosphopantetheinyl transferase activity is measured by monitoring the rate at which radiolabeled pantetheine gets incorporated into holo-ACP. Radiolabeled holo-ACP is quantified by co-precipitation with BSA using 10% TCA followed by liquid scintillation counting of the protein pellet. The formation of ^{3}H-labeled holo-ACP was confirmed by autoradiography of 20% native polyacrylamide gels.

Earlier reports had indicated that ACPS is a basic protein which would not bind to anion exchange resins (Elovson, J. and Veglas, P.R. *supra*). A 3-fold purification was thereby quickly achieved by batchwise DE-52 treatment. The DE-52 supernatant was then adsorbed onto the cation exchanger resin SP-Sepharose. Following an extensive wash, the column was eluted using a linear NaCl gradient (0 - 1 M). ACPS activity eluted at approximately 0.35 M NaCl. The submicromolar K_m of ACPS for apo-ACP previously measured with 780-fold purified enzyme (Elovson, J. and Veglas, P.R. *supra*) suggested a very tight binding interaction suitable for affinity chromatography. Apo-ACP was linked to an Affi-Gel 15 matrix and it was found that phosphopantetheinyl transferase activity was indeed tightly retained by the apo-ACP affinity column. phosphopantetheinyl transferase activity did not elute with either high salt or low pH, although it did elute with apo-ACP. Unfortunately, apo-ACP elution was not suitable for subsequent purification steps since separation of the apo-ACP from ACPS was difficult and trace contaminants in the apo-ACP preparation prevented identification of the low abundance ACPS protein. The elution requirement was satisfied when it was shown that the ACPS could be refolded and its activity reconstituted following elution with chaotropes under denaturing conditions by subsequent

dilution of the denaturant. While both urea and guanidinium·HCl proved suitable for this purpose, guanidinium·HCl (6M) was chosen as the preferred eluant to minimize the risk of N-terminal carbamylation of the target protein. Guanidinium·HCl elution of ACPS activity from the apo-ACP affinity column yielded an apparent 70,000-fold purified preparation. Tris-Tricine SDS-PAGE analysis (Schagger, H. and von Jagow, G. *supra*) revealed the presence of only a few major bands (Figure 3). Previous purifications had demonstrated that the 14 kDa band copurified with the ACPS activity (data not shown). Using the standard Tris·glycine SDS-PAGE analysis the 14 kDa protein migrated with the buffer front, greatly hampering initial detection of a candidate band for N-terminal sequence analysis. Tris·Tricine SDS-PAGE analysis offers greater resolution of low molecular weight proteins and was therefore used for all subsequent analyses of ACPS containing fractions. Superdex-75 gel filtration chromatography of the ACPS preparation indicated a native molecular weight of approximately 30,000, suggesting that the native enzyme is a homodimer (data not shown). The 14 kDa protein was electroblotted and submitted for N-terminal sequencing. Twenty-five cycles of N-terminal sequencing yielded a primary sequence of AILGLGTDIVEIARIEAVIARSGDR (SEQ ID NO: 18). A BLAST search (Altschul, S. F., et al. (1990) *J. Mol. Biol.* 215:403-410) of the non-redundant protein database revealed that the 14 kDa protein is encoded by *dpj* (downstream of Pyridoxal J), the second gene in the *pdxJ* operon (Takiff, H. E., et al. (1992) *J. Bacteriol.* 174:1544-1553).

The *dpj* gene was amplified from the *E. coli* genome by PCR and subcloned into the *NdeI/HindIII* restriction site of the pET22b vector (Novagen). Induction of *E. coli* BL21(DE3)pDPJ and purification of the phosphopantetheinyl transferase activity yielded 50 mg of protein with > 95 % purity and 320 mU/mg specific activity (Table 2). This corresponds to at least one-half the specific activity of the partially pure wild-type preparation. This difference is most probably due to errors associated with quantification of the dilute wild-type protein preparation using the Bradford protein assay. DNA sequencing of the pDPJ construct confirmed the recombinant sequence was correct (Dana-Farber Molecular Biology Core Facility, Boston, MA). The 14 kDa overproduced recombinant protein was blotted and submitted for N-terminal sequencing which confirmed the first ten residues as the *dpj* gene product. Mass spectrometric analysis indicated a molar mass of 13,950 within 0.2% of the calculated mass of 13,922. Incorporation of the [^{3}H]phosphopantetheine moiety into apo-ACP by recombinant enzyme was again confirmed by 20% native gel electrophoresis followed by autoradiography (Figure 2). Holo-ACP migrates slightly faster than apo-ACP on 20% native-PAGE (14). Furthermore, mass spectral analysis of unlabeled enzymatic holo-ACP product indicated a MW of 8841 (calculated 8847) in contrast to an observed MW of 8518 (calculated 8508) for the apo-ACP substrate. Steady-state kinetics on recombinant ACPS using the [^{3}H]CoA radioassay yielded a K_m value of 50 μM for CoA. As previously reported with partially purified ACPS (11), we observed substrate inhibition at apo-ACP concentrations greater than 2 μM. However, we

were able to assign an upper limit of ~1 μM to the K_m value for apo-ACP. An apparent k_{cat} value of about 10 min^{-1} was measured at saturating CoA and 50 μM apo-ACP. Identical K_m values were obtained for apo-ACP and CoA with wild-type ACPS under the same assay conditions. Differences between our kinetic constants and those reported previously, 0.4 μM and 150 μM for K_m(apo-ACP) and K_m(CoA) respectively (Elovson, J. and Vagelos, P. R. *supra*), are most likely attributable to variations in the apo-ACP and CoA substrate preparations and the assay conditions employed.

Table 2. Purification of recombinant ACPS from *E. coli* BL21(DE3)pDPJ

fraction	protein mg	activity U[a]	specific activity mU·mg^{-1}	-fold	% Overall Yld.
Crude extract	600	120	200	-	-
DE-52	160	40	250	1.3	27
SP-Sepharose	50	16	320	1.6	13

[a] One unit of activity produces 1 μmol holo-ACP per minute.

In vitro phosphopantetheinylation of NodF with recombinant E. coli ACPS.

This example shows that *E. coli* ACPS phosphopantetheinylates the heterologous substsrate *Rhizobial* NodF.

The conditions of the *in vitro* phosphopantetheinylation assay used in this example are the same as those described above using the *E. coli* ACP as a substrate. The efficiency of phosphopantetheinylation of wild-type NodF, an ACP-NodF chimera, and wild-type ACP by *E. coli* ACPS are presented below:

NodF	5 μL of 1.9 mg/mL stock per assay	
+ ACPS	236 000 dpm	47% conversion
- ACPS	15 500 dpm	

ACP::NodF	5 μL of 1.2 mg/mL stock per assay	
+ ACPS	122 000 dpm	% conversion
- ACPS	15 900	

ACP	5 μL of 8.5 mg/mL stock per assay	
+ ACPS	1 320 000 dpm	46% conversion
- ACPS	22 900	

The values are averages of 3 replicate assays. The associated errors were 5%, 20%, and 10% for the ACP, ACP::NodF, and NodF assays, respectively. The % conversion was calculated using the specific activity for the pantetheine moiety of the radiolabeled CoA, which is 4.8 x 10^8 dpm/umol.

In vitro phosphopantetheinylation of D-Alanyl carrier protein with recombinant E. coli ACPS

This example shows that *E. coli* ACPS is capable to phosphopantetheinylate *in vitro* D-Alanyl carrier protein (Dcp), such as *Lactobacillus casei* Dcp.

The conditions of the *in vitro* phosphopantetheinylation reaction were essentially the same as those described in the previous examples except for the 36 hour incubation time and an increase in the initial CoA concentration. Again, all assays were run in triplicate at 37°C. The reactions were quenched with 800µL 10% TCA and 500µg BSA was added. The 1.5 mL eppendorf tubes were mixed throroughly by inversion and the protein was pelleted by centrifugation at 12,000 x g. The pellets were washed 3x with 900µL 10% TCA and then resolubized in 150µL 1 M Tris base. The resolubized protein was then added to 3 mL Packard liquid scintillation cocktail and the amount of radioactivity incorporated into the protein fraction was quantified.

Dcp:ACPS assays (3x with ACPS, 3x without ACPS)

per 100µL assay	final concentration
5 µL apo-Dcp, 0.8 nmol/µL	40 µM
5 µL 0.2 M $MgCl_2$	10 mM
5 µL 50 mM DTT	2.5 mM
5 µL 1.5 M Tris-HCl, pH 8.5	75 mM
30 µL [^{3}H]CoA, 0.7 mM	210 µM
~7.0 x 10^{14} dpm/mole overall	
~4.9 x 10^{14} dpm/mole w.r.t. pantetheine	
40 µL H_2O	
10 µL 18 µM ACPS, 1.5 mUnits	1.8 µM

ACP:ACPS assays (3x with ACPS, 3x without ACPS)

per 100μL assay	final concentration
5 μL apo-ACP, 1 mM	50 μM
5 μL 0.2 M $MgCl_2$	10 mM
5 μL 50 mM DTT	2.5 mM
5 μL 1.5 M Tris-HCl, pH 8.5	75 mM
30 μL [^{3}H]CoA, 0.7 mM	210 μM

-7.0 x 10^{14} dpm/mole overall
-4.9 x 10^{14} dpm/mole w.r.t. pantetheine

50 μL H_2O	
10 μL 18 μM ACPS, 1.5 mUnits	1.8 μM

The results are presented in Table III:

Table III:

Dcp + ACPS	1.04 x 10^6 (± 0.93%)
"	1.00 x 10^6 (± 0.77%)
"	0.98 x 10^6 (± 0.95%)
Dcp - ACPS	6.91 x 10^4 (± 1.92%)
"	6.75 x 10^4 (± 1.95%)
"	6.69 x 10^4 (± 1.95%)
ACP + ACPS	2.13 x 10^6 (± .65%)
"	2.20 x 10^6 (± .64%)
"	2.21 x 10^6 (± .63%)
ACP - ACPS	3.40 x 10^4 (± 1.94%)
"	3.75 x 10^4 (± 1.99%)
"	3.60 x 10^4 (± 1.96%)

Thus, these results indicate that *E. coli* ACPS is capable of phosphopantetheinylating Dcp.

In another example, varying concentrations of apo-Dcp were incubated with ACPS at 37°C to determine extent of phosphopantetheinylation.

Final concentrations. stock solids

1.8 μM	Recombinant ACPS	10 μL	18 μM
vary	apo-Dcp	5 μL	10, 20, 50, 100, 200, 300, 400, 500μM
210 μM	[^{3}H]CoA	30	0.7mM
10 mM	$MgCl_2$	5	0.2mM
23 mM	DTT	5	50mM
75 mM	Tris HCl pH8.0	5	1.5 M
	H_2O	40	

The assays were incubated overnight at 37°C.

The [^{3}H] holo-Dcp formed was quantitated by dividing dpm by specific activity of pantetheinyl moiety of CoA which is equal to 4.83 x 10^{14} dmp/mol. The results are presented in Figure 4. These results indicate that the amount of holo-Dcp formed increases linearily with the amount of apo-Dcp in the reaction.

It has also been shown that *E. coli* ACPS modifies the apo-form of *Streptomycete* ACPs involved in frenolicin, granaticin, oxytetracycline, and tetracenomycin polyketide antibiotic biosynthesis.

Discussion:

Wild-type holo-ACP synthase (ACPS) was purified to homogeneity from *E. coli* and the N-terminal peptide sequence was used to identify *dpj* as the gene which encodes ACPS. ACPS appears to be a homodimer with a native molecular weight of 28,000. Overexpression of *dpj* has allowed the isolation of > 10 mg of active recombinant ACPS. Surprisingly, a first search of the Genbank databases including the recently reported *Haemophilus influenzae* genome (Fleischmann, R.D., et.al. (1995) *Science* 269:496-512) revealed no known genes which share significant homology with *dpj*. It is likely that *dpj* will be useful for the cloning of other phosphopantetheinyl transferases and will assist in the heterologous overproduction of appropriately modified 4'-PP requiring enzymes, such as PhbC (Gerngross, T. U., et al. (1994) *Biochemistry* 33:9311-9320), Dcp (Heaton, M. P. and Neuhaus, F. C. (1994) *J. Bacteriol.* 176:681-690; Perego, M., et al. (1995) *J. Biol. Chem.* 270:15598-15606), TcmM (Shen, B. et al. *supra*), and NodF (Geiger, O. et al. *supra*) thereby greatly facilitating the

production of acyl activating enzymes involved in macrolide, polyketide, depsipeptide, and non-ribosomal peptide biosynthesis as well as ACP-dependent transacylase activities.

Example 2: Identification of additional phosphopantetheinylating enzymes

BLAST searches (Basic Local Alignment Search Tool) (Altschul, S.F. et al. (1990) *J. Mol. Biol.* 215:403-410) with the 125 aa *E. coli* ACPS protein sequence revealed marginal similarities to the C-terminal region of five fungal fatty acid synthases, suggesting that phosphopantetheinylation activity may have been subsumed as a domain in these polyenzymes (Figures 6 and 7). In particular, similarities of 15-22% over 120 residues to the C-terminal region of three fungal fatty acid synthases were found. Genetic evidence supports a scheme in which the C-terminus of the FAS2-subunit could be responsible for the auto-phosphopantetheinylation of the FAS2 N-terminal ACP domain (Kuhn, L. et al. (1972) *Eur. J. Biochem.* 24:492-497; Schweizer, E. et al. (1973) *Eur. J. Biochem.* 39:353-362; Schweizer, E. (1977) *Naturwissenschaften* 64:366-370; Schweizer, E. et al. (1987) *Fat Sci. Technol.* 89:570-577; Werkmeister, K. et al. (1980) *Biochem. Biophys. Res. Comm.* 96:483-490). For example the C-terminal 121 amino acids of the 1894 amino acid (aa) yeast fatty acid synthase subunit II (yFASII) could act intramolecularly to add a P-pant unit to Ser-180, the putative ACP domain of this polyprotein. Schweizer's group has previously reported intrallelic complementation and *in vitro* reactivation of mutant FAS, one at S180 and one at G1777, which are by themselves inactive, consistent with this proposal (Kuhn, L. et al. (1972) *Eur. J. Biochem.* 24:492; Schweizer, E. et al. (1973) Eur. J. Biochem. 39:353; Schweizer, E. (1977) *Naturwissenschaften* 64:366; Schweizer, E.. et al.. (1987) *Fat Sciences Technology* 89:570; Werkmeister, K. et al. (1980) (1980) *Biochem. Biophys. Res. Commun.* 96:483; Schorr, R. et al. (1994) *J. Plant Physiol.* 143:407).

From the homology between fungal FAS2 C-termini and ACPS, further sequence comparisons revealed homology of *E. coli* ACPS to three bacterial proteins, EntD (*E. coli*), Sfp (*B. subtilis*), and Gsp (*B. brevis*) (Figures 6 and 8). The specific biochemical functions of entD, sfp and gsp were unknown. Ent D had been shown to be required for production of the Fe^{III}-chelating siderophore enterobactin in *E. coli* (Coderre, P.E. and Earhart, C.F. (1989) *J. Gen. Micorbiol.* 135:3043-3055). Sfp was isolated as a locus required for production of the lipopeptide antibiotic surfactin in *B. subtilis* (Nakano, M.M.. et al. (1992) *Mol. Gen. Genet.* 232:313-321). Gsp is a protein required for the synthesis of gramicidin by the multidomainal gramicidin synthase complex GrsAB (Borchert, S. et al. (1994) *J. Bacteriol.* 176:2458). In addition to EntD, Sfp and Gsp further BLAST searches revealed that a third *E. coli* ORF (in addition to ACPS and EntD) of unknown function designated o195 is also homologous to ACPS.

EntD, Sfp, and Gsp have previously been identified as orthologs of high three way homology (Grossman, T.H. et al. (1993) *J. Bacteriol.* 175:6203). In fact, *E. coli* EntD can complement sfp mutants consistent with equivalent functionality (Grossman, T. H. et al. *supra*; Borchert, S. et al. (1994) *J. Bacteriol.* 176:2458).

Figures 6, and 8 show alignments of amino acid sequences of the yeast fatty acid synthases, *E. coli* ACPS, EntD, Sfp, Gsp, and Lpa. These alignments show 2 regions of conserved amino acids in these proteins. In addition to EntD, Sfp, Gsp, and Lpa, further BLAST searches revealed several other proteins which share potential homology with ACPS (Figure 9), among them a third *E. coli* ORF (in addition to ACPS and EntD) of unknown function designated o195 as well as proteins involved in cyanobacterial heterocyst differentiation and fungal lysine biosynthesis. Local sequence alignments of the putative P-pant transferase domains reveal two sequence motifs containing several highly conserved residues. The highly conserved residues in these regions of homology are boxed in Figure 9. It is highly likely that these regions are involved in transfer of the phopsphopantetheine group. Mutagenesis studies of this domain confirm in fact that the conserved regions are involved in phosphopantetheinyl transfer.

Example 3: Sfp, EntD, and o195 phosphopantetheinylate substrates *in vitro*

This example demonstrates that the proteins Sfp, EntD, and o195 shown to have sequence homology to the *E. coli* ACPS are capable of phosphopantetheinylating one or more substrates *in vitro*. Each of these proteins were overproduced and purified, and were shown to be able to transfer tritium-labeled 4'-Ppant from CoA to the *E. coli* FAS apo-ACP, the apo-ACP domain from the multifunctional peptide synthetase, TycA, EntF, and/or SrfB1, depending on the transferase.

Overproduction and purification of Sfp, EntD, and o195 was performed as follows. Sfp (26.1 kD) was overproduced and purified following previously published procedures (Nakano, M.M. et al. (1992) *Mol. Gen. Genet.* 232).

EntD (23.6 kD) had previously been cloned, but its overproduction had proven difficult, presumably due to the frequency of rare codons and an unusual UUG start codon (Coderre, P.E. & Earhart, C.F. (1989) *J. Gen. Microbiol.* 135:3043). Therefore the UUG start codon was changed to AUG and the codon usage for the first six residues was optimized. *EntD* was PCR-amplified from wild-type *E. coli* K-12 by colony PCR using the forward primer 5'-ATTATATCCATGGgtTCcTCcGTtTCcAAcATGGTCGATATGAAAACT ACGCA-3' (SEQ ID NO: 19) and the reverse primer 5'-GATGTCAAGCTTATTAATCGTG TTGGCACAGCGTTAT-3' (SEQ ID NO: 20) (IDT). The forward primer introduced an NcoI restriction site (underlined) which allowed mutation of the TTG start to an ATG start and inserted a Gly codon (GGT) after the Met initiator. In addition the forward primer optimized codon usage for the first six codons of the *entD* gene (modified bases shown in lower case). The reverse primer incorporated a HindIII restriction site (underlined). The NcoI/HindIII

digested PCR product was cloned into pET28b (Novagen) and transformed into competent *E. coli* DH5α. Competent cells of the overproducer strain *E. coli* BL21(DE3) were then transformed with the supercoiled pET28b-entD plasmid. Induction of a 2 L culture BL21(DE3)pET28b-entD with 1 mM IPTG followed by growth at 25 °C for 5 hours yielded predominantly inclusion bound EntD, however a modest amount of the overproduced protein was soluble. The induced cell paste was resuspended in 50 mM Tris, 1 mM EDTA, 5% glycerol, pH 8.0 (40 mL) and lysed by two passages through the French press at 15,000 psi. Cellular debris and inclusion bound protein was removed by centrifugation at 8,000 x g for 30 minutes. Pulverized ammonium sulfate was added to 35%, 65% and 80% saturation. The 35% fraction containing the largest fraction of EntD was applied to a 2.5 x 115 cm Sephacryl S-100 column. The column was eluted at a flow rate of 1 mL/min using the same buffer as above collecting 8 mL fractions to obtain homogeneously pure EntD.

Similarly, o195 (Sofia et al. (1994) *Nucleic Acids Research* 22:2576-2586) was PCR amplified from wild-type *E. coli* K-12 by colony PCR using the forward primer 5'-ATTATATCCATGGgtTAcCGGATAGTTCTGGGGAAAGTT-3' (SEQ ID NO: 21) and the reverse primer 5'-TGATGTCAAGCTTATCAGTTAACTGAATCGATCCATTG-3' (SEQ ID NO: 22) (IDT). The forward primer with its NcoI restriction site (underlined) gave insertion of a Gly codon (GGT) after the Met initiator codon of the o195 sequence. Codon usage for the succeeding codon was also optimized (base change shown in lower case). The reverse primer incorporated a HindIII restriction site (underlined). The NcoI/HindIII-digested PCR product was cloned into pET28b (Novagen) and transformed into competent *E. coli* DH5α. Competent cells of the overproducer strain *E. coli* BL21(DE3) were then transformed with the supercoiled pET28b-o195 plasmid. Induction of a 2 L culture (2 x YT media) of BL21(DE3)pET28-o195 with 1 mM IPTG followed by growth at 37 °C for 3.5 h yielded predominantly inclusion-bound o195 protein. The cell paste was resuspended in 50 mM Tris·HCl, 1 mM EDTA, 5% glycerol, pH 8.0 (40 mL) and lysed by two passages through a French pressure cell at 15,000 psi. Cellular debris and inclusion-bound protein was pelleted by centrifugation at 27,000 x g for 30 minutes (min.). The inclusion-bound protein pellet was resuspended in 30 mL of 50 mM Tris·HCl, pH 8.0, 1 mM EDTA, and 5% glycerol and incubated for 30 minutes at room temperature with 10 mg lysozyme and 30 mg deoxycholate. The pellet was reobtained by centrifugation for 15 min at 27,000 x g and solubilized in 30 mL of 8 M urea, 50 mM Tris·HCl, pH 8.0, 10 mM DTT. Residual solid material was removed by centrifugation for 15 min at 27,000 x g. The urea-solubilized solution (30 mL) was then applied to a 2.5 x 10 cm Q-Sepharose column equilibrated with 8 M urea, 50 mM Tris·HCl, pH 8.0. The column was washed with 50 mL of the equilibration buffer and then a gradient of 250 mL 0-0.25 M NaCl in 8 M urea, 50 mM Tris·HCl pH 8.0 followed by 200 mL of 0.25 - 1 M NaCl in the same buffer was applied. The o195 protein eluted at approximately 200 mM NaCl as determined by 15% SDS-PAGE. The purified o195 was renatured by diluting a portion of it 10-fold in 8 M urea, 50 mM Tris·HCl, pH 8.0, 10 mM

DTT and dialyzing overnight at 4 °C against 10 mM Tris·HCl, pH 8.0, 1 mM DTT. Two liters of culture grown in 2 x YT media yielded 3.1 g of cells from which approximately 80 mg of o195 protein was obtained.

The substrates were overexpressed and purified as follows. The *E. coli* fatty acid synthase ACP was overproduced and purified in its apo-form from *E.* coli strain DK554 (Crosby, J. et al. (1995) *Biochemica et Biophysica Acta* 1251:32) following the procedure of Rock and Cronan (Rock, C.O. & Cronan, J.E., Jr. (1981) *Methods Enzymol.* 71:341-351) with the exception that following cell disruption and centrifugation (30 min at 28,000 x g), the crude extract containing 10 mM $MgCl_2$ and 10 mM $MnCl_2$ was incubated for 60 min at room temperature. In this manner, minor amounts of holo-ACP were hydrolyzed to the apo-form using the endogenous *E. coli* ACP phosphodiesterase (Fischl, A.S. & Kennedy, E.P. (1990) *J. Bacteriol.* 172, 5445-5449).

The peptidyl carrier protein (PCP) domain of TycA (Figure 10) was overproduced with a hexa-histidine tag using *E. coli* strain SG13009(pREP4)/pQE60-PCP (Stachelhaus et al. (1996) *Chemistry & Biology*, in press). Following lysis of the induced culture the His_6-tagged protein was purified by nickel-chelate chromatography.

Apo-SrfB1 was cloned from plasmid p120-21E (Nakano, M.M. et al. (1991) *J. Bacteriol.* 173:1770-1778). Briefly, p120-21E was digested with EcoRV to release a 3,648 base pair fragment encoding the SrfB1, valine-activating domain of surfactin synthetase. This fragment was inserted into StuI-cleaved pPROEX-1 (Gibco/BRL Life Sciences Technologies) to give plasmid pML118 which codes for a N-terminal hexa-hisitidine-tagged SrfB1 domain (142.7 kD). His6-SrfB1 was overproduced using E. coli strain AG1574 (Frisby, D. & Zuber, P. (1991) *J. Bacteriol.* 173:7557-7564). Cells were grown at 25 °C in 2 x YT media (2L) to an O.D. of 0.4 at which point they were induced with 1 mM IPTG and allowed to grow for an additional 4 hours. Cells were harvested by centrifugation (3 g), resuspended in 35 mL of 5 mM imidazole, 500 mM NaCl, 20 mM Tris·HCl, pH 7.9 and lysed by two passages through a French pressure cell. This crude extract was clarified by centrifugation for 30 min at 27,000 x g. More than 50% of the overexpressed SrfB1 was obtained in the soluble fraction as determined by 6% SDS-PAGE. Hexa-histidine tagged SrfB1 was purified on His-Bind resin (Novagen) following the manufacturer's recommendations.

EntF was overproduced and purified as described in Keating, D.H. et al. (1995) *J. Biol. Chem.* 270:22229 and Reichert, J. et al. (1992) *Prot. Sci.* 1:549.

Phosphopantetheinyl transferase activity toward a substrate was assayed by monitoring the transfer of [^{3}H]-4'-phosphopantetheine from [^{3}H]-(*pantetheinyl*)-CoASH in the presence of the putative P-pant transferase enzyme by radioassay. Enzyme preparations were incubated with 75 mM Tris•HCl, pH 8.5, 10 mM MgCl2, 5 mM DTT, 160 μM [3H]-(pantetheinyl)-CoA, 25 μM ACP or His6-PCP for 60 min at 37° C in a final volume of 100 μL. The incubations were quenched with 10% TCA and 500 μg BSA was added as a carrier.

The protein was precipitated by centrifugation, washed 3X with 10% TCA, and the protein pellet was solublized with 150 µL 1M Tris base. The resuspended protein was added to 3 mL liquid scintillation cocktail and the amount of [3H]-phosphopantetheine incorporated into the substrate protein was quantified by liquid scintillation counting.

Specific covalent attachment of 4'-Ppant to the ACP and His6-PCP substrates was confimed by autoradiography. These assays were performed as described above except 20 µM [^{3}H]-(*pantetheinyl*)-CoASH (2.6 x 10^6 dpm total activity) was used in the assay, no BSA was added to the TCA precipitate, and pellets were solubilized in SDS or native PAGE sample buffer titrated with 1 M Tris base. Assays using apo-PCP as substrate were resolved by 15% SDS-PAGE, assays using *E. coli* ACP were resolved by 20% native PAGE, and assays using SrfB1 or EntF were resolved on 8% SDS-PAGE. Gels were Coomassie-stained, soaked 30 minutes in Amplify (Amersham), dried at 80 °C under vacuum and exposed to X-ray film for 24 to 150 hours at -70 °C. The autoradiograms were scanned using a digital scanner and relative intensities of the radiolabeled bands were quantified using NIH Image 1.59 software (National Institutes of Health, USA).

The results of the phosphopantetheinylation assays performed in the presence of Sfp, EntD, or o195 are presented in Figure 11. These results show that Sfp, EntD and o195 are capable of transferring phosphopantetheine groups to a substrate such as ACP and His6-PCP. It was further shown that [3H]-phosphopantetheine is in fact specifically and covalently attached to ACP and His6-PCP following the phosphopantetheinylation reaction mediated by Sfp, EntD, or o195.

Figure 12 shows that EntD is capable of phosphopantetheinylating an EntF fragment and Sfp is capable of phosphopantetheinylating an SfrB1 fragment.

Further confirmation that the tritium radioactivity incorporated into the apo-proteins represented transfer of the intact phosphopantetheinyl group was assessed by mass spectrometry (Lambalot, R.H. and Walsh, C. T. (1995) *J. Biol. Chem.* 270:24658). Mass spectrometric analysis (MALDI-TOF) of unlabeled enzymatic holo-PCP indicated a molecular weight of 13,431 (calculated 13,459) in contrast to an observed molecular weight of 13,130 (calculated 13,120) for the apo-PCP substrate. Thus, these data establish that EntD, Sfp, and o195 are enzymes that catalyze the transfer of P-pant to the serine side chain of an acyl carrier protein.

For the determination of Km for apo-ACP and His6-PCP using Sfp, the linear range of the assay was determined by monitoring the percentage conversion of apo-substrate to holo-substrate versus time at various enzyme concentrations using the radioassay described above. Having determined the linear range of the assay, the initial velocity of phosphopantetheinyl transfer was measured using a fixed concentration of Sfp and varying the concentration of the substrate (ACP or His6-PCP).

The results are presented in Figure 13. These show that Sfp is capable of phosphopantetheinylating a substrate, such as PCP, in a wide range of concentrations.

A time course of EntD-catalyzed incorporation of radiolabel into EntF was performed. As shown in Fig. 14A, EntF is modified effectively by EntD (100 nM), whereas EntF undergoes almost no modification in the presence of 15-fold higher concentrations of ACPS and o195, clearly demonstrating the specificity of EntD for EntF. In contrast, Fig. 14B shows almost exclusive modification of apo-ACP by ACPS, confirming that ACPS is the P-pant transferase which activates the type II fatty acid synthase and EntD is the P-pant transferase which activates the type I enterobactin synthetase in *E. coli*.

Thus, this analysis provided *in vitro* evidence of at least two partner-specific P-pant transfer reactions occurring within *E. coli*: ACPS specifically catalyzes transfer of P-pant to apo-ACP while EntD is the partner transferase for EntF. o195 is likely to has a unique specificity for a third unknown substrate in *E. coli*. It is likely that P-pant transfer to this unknown protein would be catalyzed efficiently only by o195 and not by ACPS or EntD. One possible partner for o195 is the unknown 35 kDa protein in *E. coli* which has been observed to incorporate [^{3}H]ß-alanine *in vivo* (Gerngross, T. U. et al. (1994) *Biochemistry* 33:9311).

Sfp, on the other hand, would appear to be non-specific with respect to the two *Bacillus* derived type I peptide synthetase domains, PCP and SrfB1, and the *E. coli* type II fatty acid synthase ACP subunit. As shown above, Sfp efficiently catalyzes modification of all three substrates and in addition can catalyze modification of EntF. Based on this evidence, Sfp would appear not to discriminate between type I peptide synthetase domains and type II fatty acid subunits suggesting that there may be cross-talk between Sfp and fatty acid synthase at least when expressed in *E. coli*.

The establishment of phosophopantetheinyl transfer activity for Sfp in the studies described here clearly assigns a catalytic loading function to Sfp for posttranslational modification of the conserved serine in the first subsite of SrfB responsible for valine activation. The srf operon consists of four open reading frames in which srfA, srfB, and srfC encode for the activites that activate and assemble the seven component amino acids and branced chain ß-hydroxy fatty acid of surfactin. It is likely that Sfp is able to modify the consensus serine residue in all seven amino acid activating sites in SrfABC.

By extension, Gsp is responsible for the posttranslational modification of the five amino acid activating sites in GrsA and GrsB, allowing for the sequential activation and polymerization of amino acids in the thiotemplate mechanism for non-ribosomal peptide bond assembly.

Likewise, demonstration of 4'-phosphopantetheinyl (4'-Ppant) transferase activity in EntD assigns a biochemical function to EntD. EntF, a 140 kDa component in the pathway had previously been cloned, sequenced, and purified and shown to activate L-serine and to contain phosphopantetheine (Rusnak, F. et al. (1991) *Biochemistry* 30:7740). Given that

EntD is required for enterobactin biosynthesis *in vivo* and, as set forth above, shows high activity for *in vitro* P-pantetheinylation of pure apo-EntF, EntD can be defined as the specific P-pant transferase that makes active holo-EntF from apo-EntF *in vivo*. Pure *E. coli* ACPS will not significantly posttranslationally modify EntF, consistent with the hypothesis that protein-protein recognition controls the specificity of phosphopantetheinylation *in vivo*. It is likely, that incubations of EntD and the enterobactin synthetase components with CoASH, L-serine and dihidroxybenzoate reconstitute *in vitro* enterobactin production. At 140 kDa, EntF is the typical size of an aa-activating module in multidomain polypeptide synthetases (Stachelhaus, T. & Marahiel, M.A. (1995) *FEMS Microbiol. Lett.* 125:3). Its efficient modification *in vitro* by EntD shows that P-pant addition can occur after translation of the apo-protein rather than exclusively co-translationally prior to folding of the apo-protein into its native structure. The NMR structure of *E. coli* apo-ACP shows the nucleophilic Ser-36 is in an accessible β-turn (Holak, T.A. et al. (1988) *Eur. J. Biochem.* 175:9), possibly a common architectural scaffolding for ACP domains in polyketide and polypeptide synthases which may play a role in recognition by P-pant transferases.

The genetics argue strongly for specific partner peptide synthetase recognition by a given posttranslational modification enzyme and this may well be a general theme in non-ribosomal peptide antibiotic biosynthesis. It has been previously observed that pure *E. coli* ACP synthase will not posttranslationally modify EntF, consistent with protein-protein recognition to control specificity.

Example 4: Holo-SrfB1 is competent for activation of its cognate amino acid, L-valine.

This Example shows that holo-SrfB1 produced *in vitro* from modification of apo-SrfB1 with Sfp is competent for activation of its cognate amino acid, L-valine.

The reaction was performed as follows. Apo-SrfB1 (2 μM) was incubated with 200 μM CoASH, 75 mM Tris·HCl pH 8.0, 10 mM $MgCl_2$, 25 mM DTT and 1.3 μM Sfp for 15 min at 37 °C to generate holo-SrfB1. To the SrfB1-Sfp reaction mixture, unlabeled amino acid (valine or aspartic acid) was added to 90 μM final concentration. ATP was added to a final concentration of 2 mM, followed by 0.5 μCi [^{14}C]Val (281 Ci/mol) or [^{14}C]Asp (207 Ci/mol). The reaction (115 μL) was incubated for 15 min at 37 °C, then stopped by the addition of 800 μL 10% TCA with 15 μL of a 25 mg/mL BSA solution as carrier. The precipitate was collected by centrifugation, washed with 10% TCA, dissolved in 150 μL Tris base, and then counted by liquid scintillation.

The results are presented in Figure 15. The histogram presented in this figure shows that apo-SrfB1 undergoes very little acylation when incubated with [^{14}C]-L-valine indicating a small amount of contamination by holo-SrfB1. However, following incubation with Sfp the level of [^{14}C]-L-valine-holo-SrfB1 covalent complex formed in the complete incubation mixture increases about 14-fold, consistent with an increase in the amount of holo-SrfB1

present. This arises from valyl-AMP formation by the amino acid-activating domain of holo-SrfB1 followed by intramolecular acyltransfer to the SH group of the P-pant moiety in the adjacent PCP domain. Finally, holo-SrfB1 cannot be covalently acylated by the non-cognate L-aspartate residue, the fifth amino acid to be activated by SrfABC, consistent with the absence of an aspartate specific adenylation domain on SrfB1.

Thus, this Example shows that the holo-SrfB1 formed following incubation with Sfp and CoASH has both an active adenylation domain and a functional holo-peptidyl carrier protein domain. Furthermore, it shows that the action of Sfp on the 143 kD SrfB1 fragment in conversion of the apo to holo-form generates a phosphopantetheinylated SrfB1 competent to undergo specific recognition and acylation by the amino acid L-valine, residue 4 in surfactin. Sfp and other phosphopantetheinyl transferases should therefore be useful reagents, e.g., to probe peptide bond-forming steps between adjacent sites of multienzyme, multiple thiotemplate synthases.

Example 5: Identification of additional phosphopantetheinylating enzymes

Using the EntD/Sfp/Gsp family as a base for further database searches has led to the identification of several additional candidates that are likely 4'-P-pant transferase family members (Table IV and Figure 9). A hypothetical protein, HIO152, in *H. influenzae* has been identified as a putative P-pant transferase thereby satisfying the apparent lack (using ACPS-based searches) of a P-pant transferase in the *Haemophilis* genome. HIO152 is positioned directly upstream of the *H. influenzae* fatty acid synthase gene cluster, consistent with a function for its protein product in fatty acid biogenesis. Also identified in Table IV and Figure 9 are two additional proteins in cyanobacteria and in yeast for which prior genetic evidence accords with the putative functions proposed herein. In *Anabaena*, the genes *HetI*, *HetM*, and *HetN* have been implicated in the production of a proposed secondary metabolite which inhibits heterocyst differentiation, a process occurring under low fixed nitrogen conditions in which a subset of cyanobacterial cells differentiate into the specialized heterocysts which have the ability to fix nitrogen (Black, T.A. & Wolk, C.P. (1994) *J. Bacteriol.* 176:2282). Sequence analysis suggests HetN is a NAD(P)H-dependent oxidoreductase such as those involved in the biosynthesis of polyketides and fatty acids while HetM has an ACP domain. HetI, with its similarity to Sfp/Gsp/EntD, is thus likely to be the HetM-specific phosphopantetheinyl tranferase in the synthesis of the hypothesized secondary metabolite.

Another candidate of the phosphopantetheinylating family of enzymes is the 272 amino acid Lys5 protein involved in the yeast lysine biosynthetic pathway. Yeast and other fungi synthesize lysine via the unique α-aminoadipate pathway, an eight-step pathway beginning with homocitrate and proceeding via α-aminoadipate to saccharopine to lysine (Bhattacharjee, J.K. (1985) *CRC Critical Reviews in Microbiology* 12:131). Genetic analysis has implicated Lys2 and Lys5 by complementation to be involved in the same step in this

pathway, the reduction of α-aminoadipate to aminoadipate semialdehyde (Storts, D.R. & Bhattacharjee, J.K. (1989) *Biochem. Biophys. Res. Commun.* 161:182). This reaction appears to proceed through an α-aminoadipoyl-AMP intermediate as indicated by labeled pyrophosphate exchange experiments (Sagisaka, S. & Shimura, K. (1960) *Nature* 188:1191; Sinha, A.K. & Bhattacharjee, J.K. (1971) *Biochem. J.* 125:743). Recent sequence analysis (Morris, M.E. & Jinks-Robertson, S. (1991) *Gene* 98:141) shows Lys2 to be a 155 kDa protein with homology to amino acid-activating peptide synthetases including TycA, GrsAB, and SrfA. Analogous to these peptide synthetases, Lys2 is believed to cleave ATP to AMP and PPi thereby activating α-aminoadipate as the α-aminoacyl-AMP which is then reduced by NADPH to the aldehyde. A search for a consensus P-pant attachment site in Lys2 reveals the signature motif LGGHS around Ser-880. Therefore, it is likely that Lys2 and Lys5 form a two subunit enzyme (Storts, D.R. and Bhattacharjee, J.K. (1989) *Biochem. Biophys. Res. Commun.* 161:182), that the 272 aa Lys5 is a specific phosphopantetheinyl transferase for Ser-880 in Lys2. The thiol of the newly-introduced P-pant prosthetic group on Lys2 would attack the aminoadipoyl-AMP to give aminoadipoyl-S-pant-Lys2, in close analogy to sequential formation of aminoacyl-AMP to aminoacyl-S-pant-TycA in the homologous tyrocydine synthetase A subunit. At this point, hydride addition to the acyl-S-pant-Lys2 would yield a thiohemiacetal which would readily decompose to aldehyde product and HS-pant-Lys2. This sequence has precedent in the reverse direction in the oxidation of glyceraldehyde-3-P to the acyl-S-enzyme in GAP dehydrogenase catalysis via a cysteinyl-S-enzyme hemithioacetal (Walsh, C.T. (1979) *Enzymatic Reaction Mechanisms*. W.H. Freeman and Company, New York).

The *bli* and *lpa-14* gene products most likely play an equivalent role, iterative P-pantetheinylation of each aa-activating domain in *B. licheniformis* bacitracin synthetase (Gaidenko, T.A., et al. (1992) *Biotechnologia* 13) and *B. subtilis* iturin A synthetase respectively (Huang, C.-C. et al. (1993) *J. Ferment. Bioeng.* 76:445).

Table IV. ACP Synthase Homologs

pathway	gene	organism	protein size
Surfactin Biosynthesis	Sfp	*B. subtilis*	224 aa
	Psf-1	*B. pumilus*	233 aa
Iturin A Biosynthesis	Lpa-14	*B. subtilis*	224 aa
Gramicidin S Biosynthesis	Gsp	*B. brevis*	237 aa
Enterobactin Biosynthesis	EntD	*E. coli*	209 aa
		Salmonella typhimurium	232 aa
		Salmonella austin	232 aa
		Shigella flexneri	209 aa
Heterocycst Differentiation	HetI	*Anabaena sp.*	237 aa
	SYCCPNC	*Synechocystis sp.*	246 aa
Lysine Biosynthesis	LYS5	*S. cerevisiae*	272 aa
Fatty Acid Biosynthesis	AcpS (Dpj)	*E. coli*	126 aa
	HI0152	*H. influenzae*	235 aa
	FAS2	*S. cerevisiae*	1894 aa
		Candida albicans	1885 aa
		Penicillium patulum	1857 aa
		S. pombe	1842 aa
		Emericella nidulans	1559 aa
Bacitracin	Bli	*B. licheniformis*	225 aa
Nosiheptide	NshC	*S. actuosis*	253 aa
Proteins of Unknown Function	o195	*E. coli*	195 aa
	1314154	*S. pombe*	258 aa
	CELTO4G9	*C. elegans*	297 aa

Thus, these results provide evidence for the existence of a family of more than a dozen proteins with catalytic posttranslational modification activity. It is likely, that there are P-pantetheinyl transferases having CoASH as a common substrate, but that they show specificity, directed by protein-protein interactions, for the conserved serine motif in specific partner proteins. It is further likely that most if not all of the multienzyme peptide synthetases that use the multiple thiotemplate scaffolding to make peptide antibiotics nonribosomally will follow this paradigm of a specific posttranslational modifying enzyme to covalently arm the swinging arm thiol group required to enable acyl transfers. Compared to

the 126 amino acid *E. coli* subunit of ACPS, the phosphopantetheinyl transferase homologs shown in Figures 6, 8, and 9 have an extra 50 -150 amino acid residues which may be specificity-conferring regions for partner proteins.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

CLAIMS

1. A composition comprising an appropriate buffer and an isolated phosphopantetheinyl transferase in an amount sufficient to phosphopantetheinylate a substrate.

2. The composition of claim 1, wherein the phosphopantetheinyl transferase is a procaryotic phosphopantetheinyl transferase.

3. The composition of claim 2, wherein the phosphopantetheinyl transferase is an E. coli phosphopantetheinyl transferase.

4. The composition of claim 1, wherein the phosphopantetheinyl transferase is a eucaryotic phosphopantetheinyl transferase.

5. The composition of claim 1, wherein the phosphopantetheinyl transferase has a specific activity of at least 250 mU/mg.

6. The composition of claim 1, wherein the buffer contains glycerol in an amount sufficient for freezing the composition without significantly affecting the activity of the phosphopantetheinyl transferase.

7. An isolated phosphopantetheinyl transferase in a lyophilized form.

8. A kit for *in vitro* phosphopantetheinylation of a substrate, comprising an isolated phosphopantetheinyl transferase and instructions for use.

9. The kit of claim 8, further comprising at least one reagent necessary for *in vitro* phosphopantetheinylation of the substrate.

10. The kit of claim 9, wherein the phosphopantetheinyl transferase is lyophilized.

11. A recombinant expression vector comprising a gene encoding a phosphopantetheinyl transferase having at least about 50% amino acid sequence identity with an amino acid sequence of an E. coli acyl carrier protein synthase having an amino acid sequence of SEQ ID NO: 2.

12. The recombinant expression vector of claim 11, wherein the phosphopantetheinyl transferase has an amino acid sequence of SEQ ID NO: 2.

13. A host cell transformed with a recombinant expression vector of claim 11.

14. A host cell transformed with a recombinant expression vector of claim 12.

15. The host cell of claim 13, further transformed with an expressible form of at least one gene encoding a substrate of the phosphopantetheinyl transferase.

16. The host cell of claim 15, further transformed with an expressible form of at least one gene encoding a component of a multienzyme complex involved in the synthesis of an antibiotic.

17. A eukaryotic host cell transformed with an expressible form of a gene encoding a phosphopantetheinyl transferase.

18. A method for *in vitro* phosphopantetheinylating a substrate, comprising contacting an isolated phosphopantetheinyl transferase with the substrate, to thereby phosphopantetheinylate the substrate.

19. The method of claim 18, wherein the phosphopantetheinyl transferase is purified at least about 800 fold.

20. The method of claim 19, wherein the phosphopantetheinyl transferase is a procaryotic phosphopantetheinyl transferase.

21. The method of claim 20, wherein the phosphopantetheinyl transferase is from E.coli.

22. The method of claim 19, wherein the phosphopantetheinyl transferase has an amino acid sequence comprising at least the amino acid sequences GND and (W/F)NNKE(A/S)NNK, wherein N is any amino acid.

23. The method of claim 22, wherein the phosphopantetheinyl transferase has at least about 50% amino acid sequence identity with an amino acid sequence of an E. coli acyl carrier protein synthase having an amino acid sequence of SEQ ID NO: 2.

24. The method of claim 23, wherein the phosphopantetheinyl transferase has an amino acid sequence of SEQ ID NO: 2.

25 The method of claim 18, wherein the phosphopantetheinyl transferase is selected from the group consisting of Sfp, EntD, o195, and active fragments thereof.

26. The method of claim 18, wherein the substrate is a type I acyl carrier protein.

27. The method of claim 19, wherein the substrate is a type II acyl carrier protein.

28. The method of claim 18, wherein the substrate is involved in the biosynthesis of fatty acids, polyketides, and non-ribosomally produced peptides.

29. The method of claim 28, wherein the substrate is involved in the biosynthesis of an antibiotic.

30. The method of claim 28, further comprising contacting the phosphopantetheinyl transferase and the substrate with at least one additional component involved in the biosynthesis of fatty acids, polyketides, and non-ribosomally produced peptides.

31. A method for phosphopantetheinylating a substrate in a cell or increasing phosphopantetheinylation of a substrate in a cell, comprising transforming the cell with an expressible form of a gene encoding a phosphopantetheinyl transferase, such that the substrate is phosphopantetheinylated or that phosphopantetheinylation of the substrate in the cell has increased compared to the non-transformed cell.

32. The method of claim 31, wherein the phosphopantetheinyl transferase is selected from the group consisting of Sfp, EntD, o195, E. coli acyl carrier protein synthase, and active fragments thereof.

33. The method of claim 31, further comprising transforming the cell with an expressible form of a gene encoding a substrate of the phosphopantetheinyl transferase.

34. The method of claim 33, wherein the substrate is a type I acyl carrier protein.

35. The method of claim 33, wherein the substrate is a type II acyl carrier protein.

36. A method for producing an antibiotic *in vitro*, comprising contacting an isolated phosphopantetheinyl transferase with a substrate involved in the synthesis of the antibiotic in the presence of an appropriate buffer and appropriate reagents, such that the antibiotic is produced *in vitro*.

37. A method for producing an antibiotic in a cell, comprising transforming a cell with an expressible form of a first gene encoding a phosphopantetheinyl transferase and an expressible form of a second gene encoding a substrate of the phosphopantetheinyl transferase, such that the antibiotic is produced in the cell.

38. The method of claim 37, further comprising transforming the cell with an expressible form of at least a third gene encoding a protein necessary for the synthesis of the antibiotic in the cell.

39. The method of claim 36, wherein the antibiotic is selected from the group consisting of erythromycin, charythromycin, oxytetracycline, bacitracin, cyclosporin, penicillins, cephalosporins and vancomycin.

FIG.1

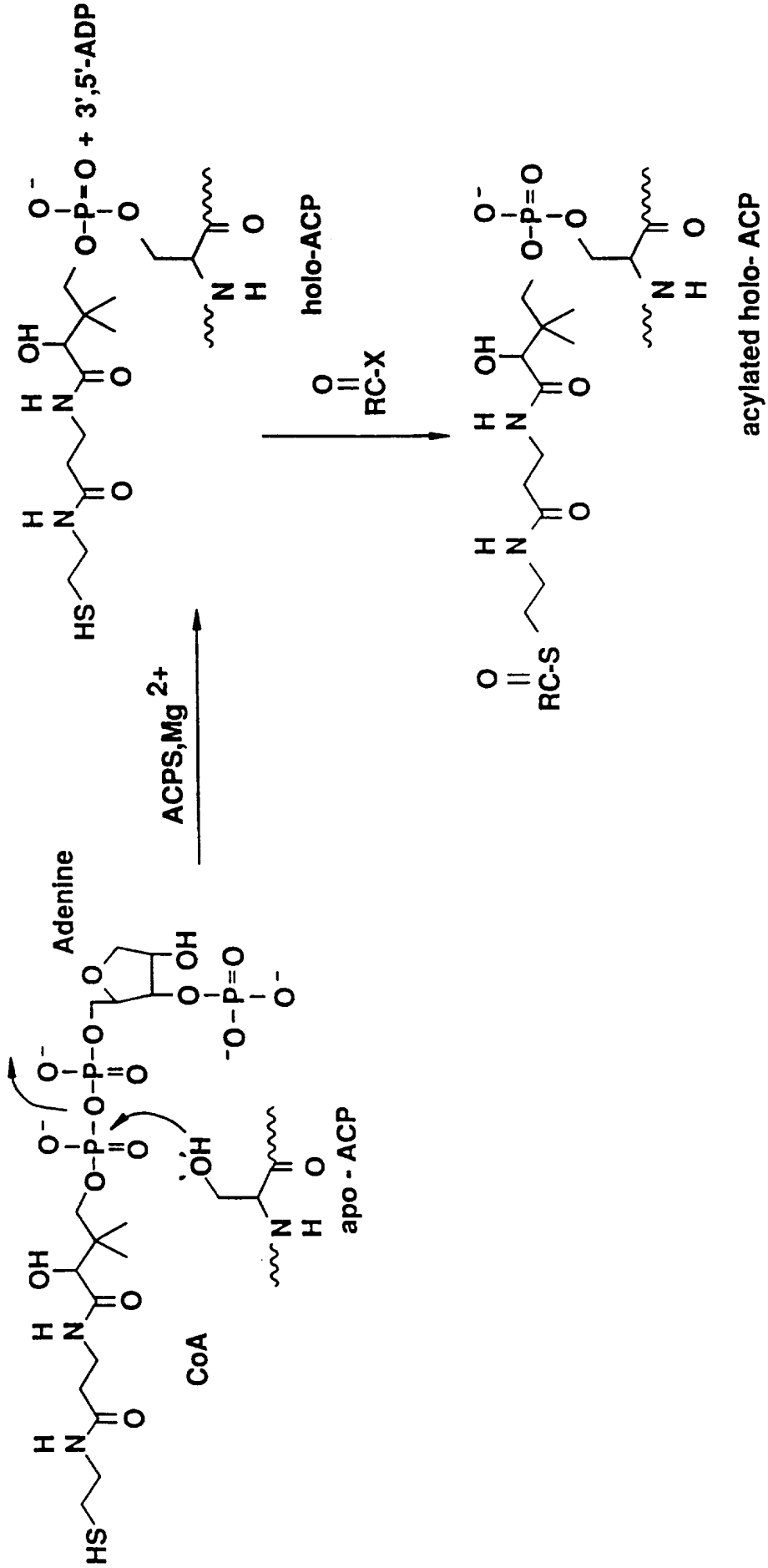

Adenine
ACPS,Mg 2+
CoA
apo - ACP
holo-ACP
RC-X
RC-S
acylated holo- ACP
+ 3',5'-ADP

FIG.2A

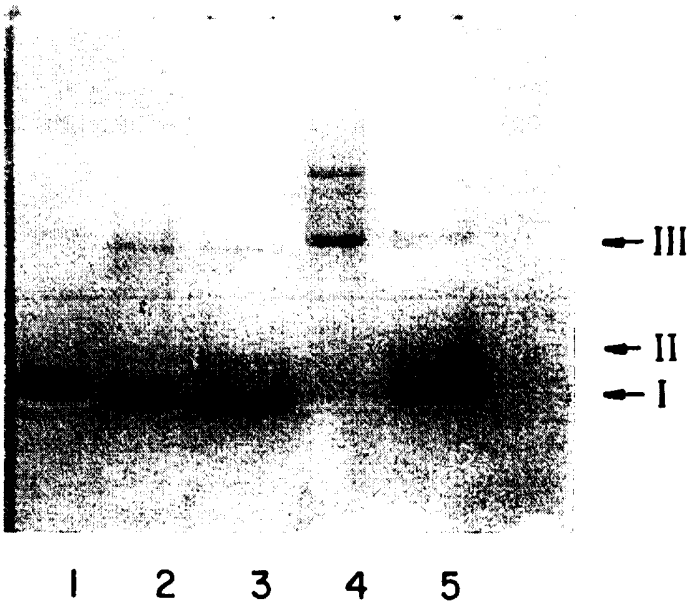


FIG.2B

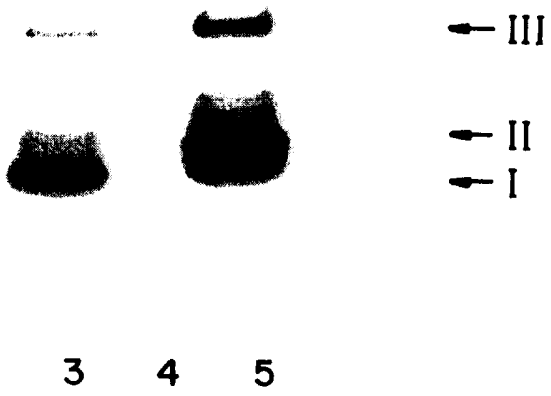


FIG.3

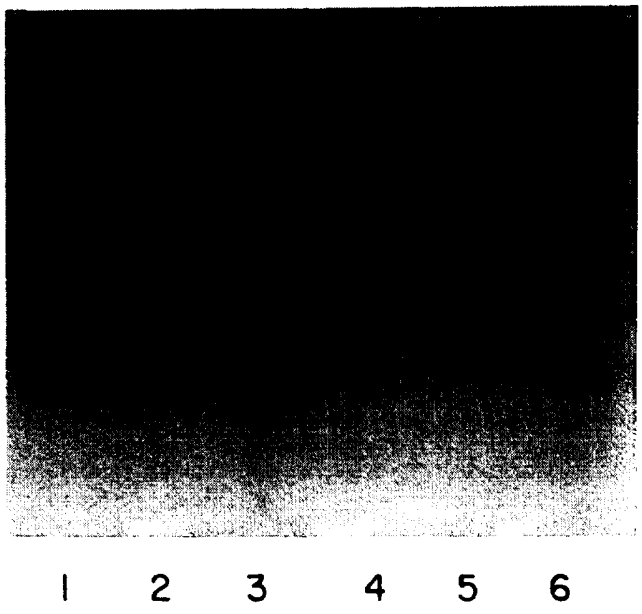
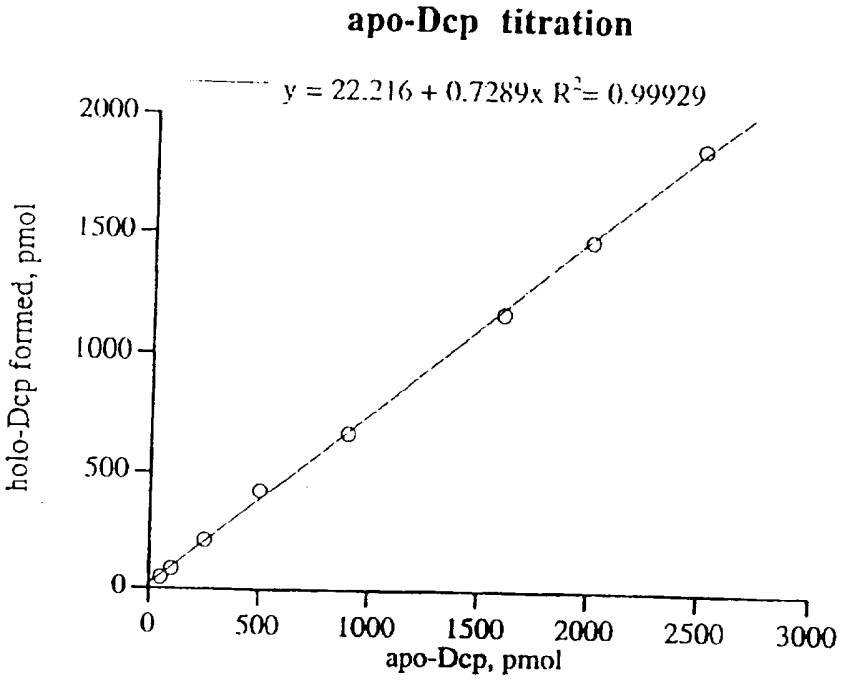


Figure 4


apo-Dcp titration
y = 22.216 + 0.7289x R²= 0.99929
holo-Dcp formed, pmol
2000
1500
1000
500
0
0
500
1000
1500
2000
2500
3000
apo-Dcp, pmol

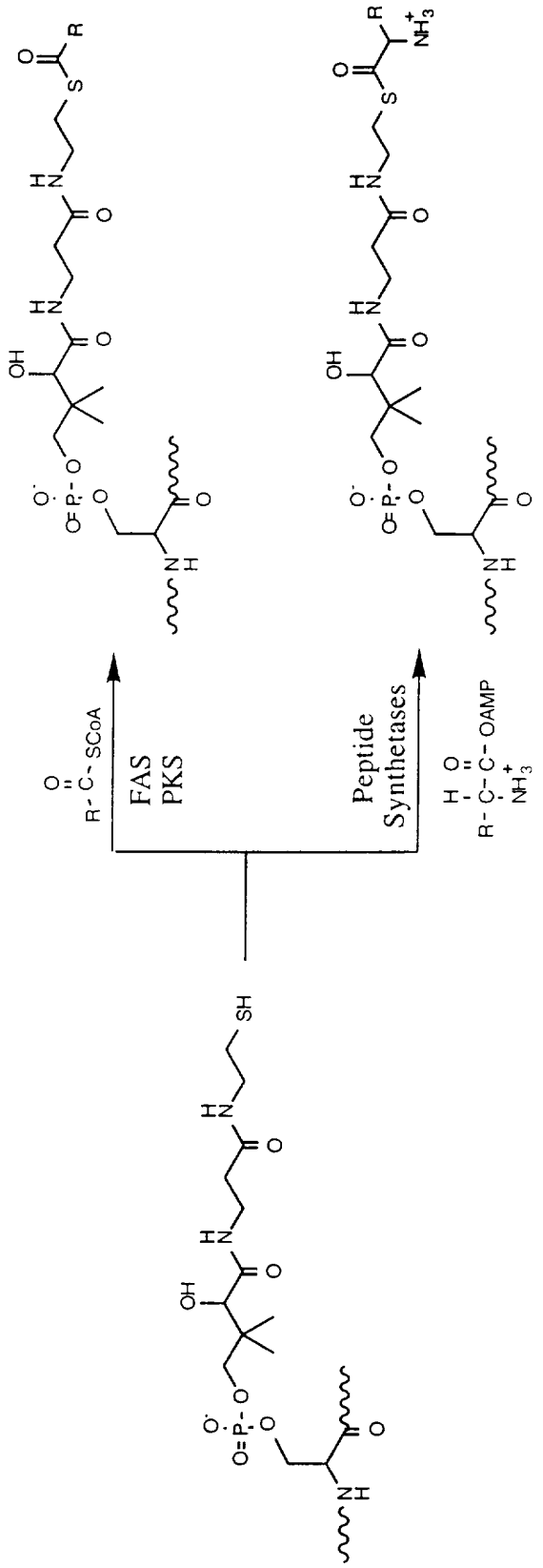
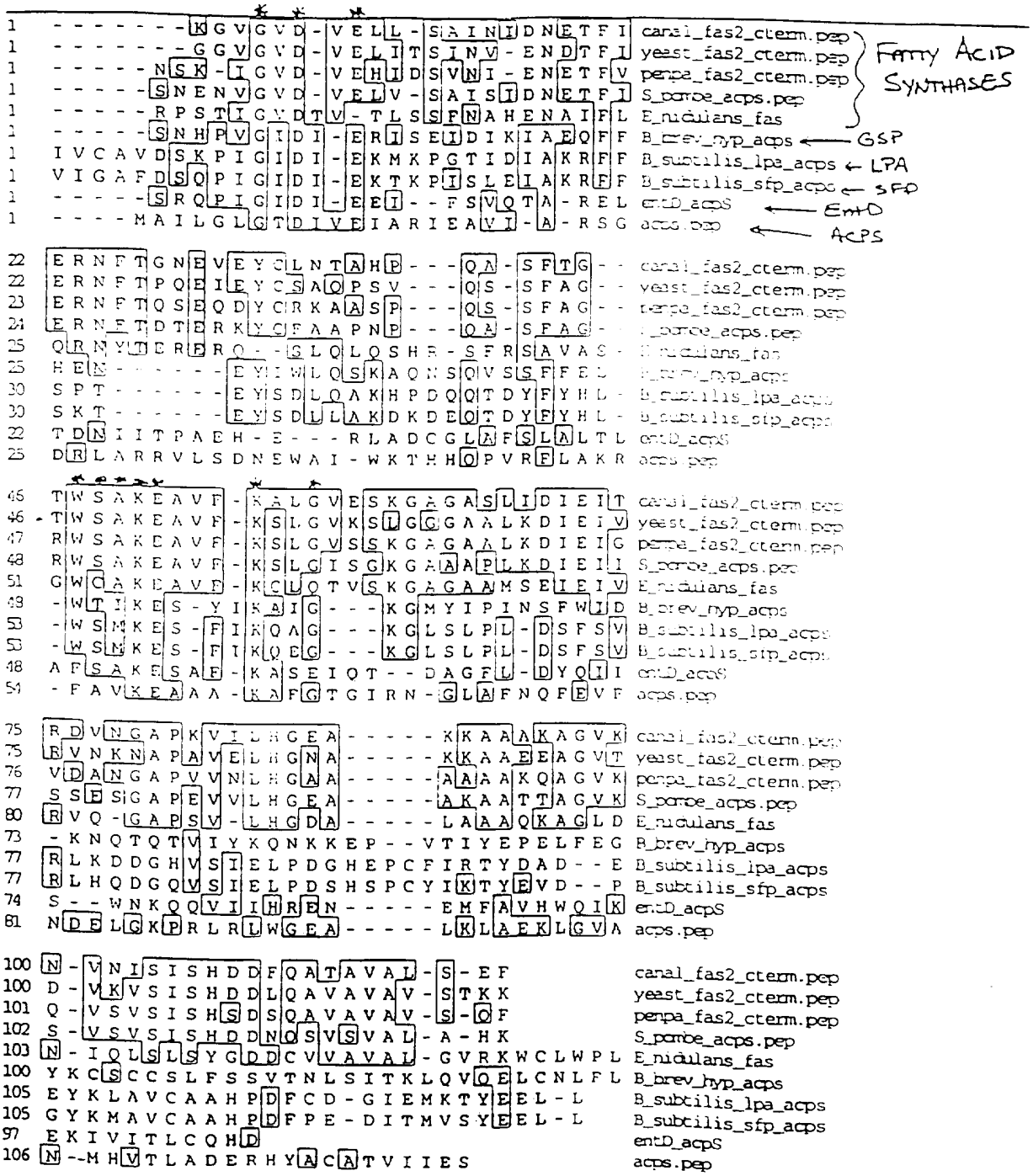


Figure 5

Figure 6


canal_fas2_cterm.pep
yeast_fas2_cterm.pep
penpa_fas2_cterm.pep
S_pombe_acps.pep
E_nidulans_fas
B_brev_hyp_acps
B_subtilis_lpa_acps
B_subtilis_sfp_acps
entD_acpS
acps.pep
Fatty Acid Synthases
GSP
LPA
SFP
EntD
ACPS

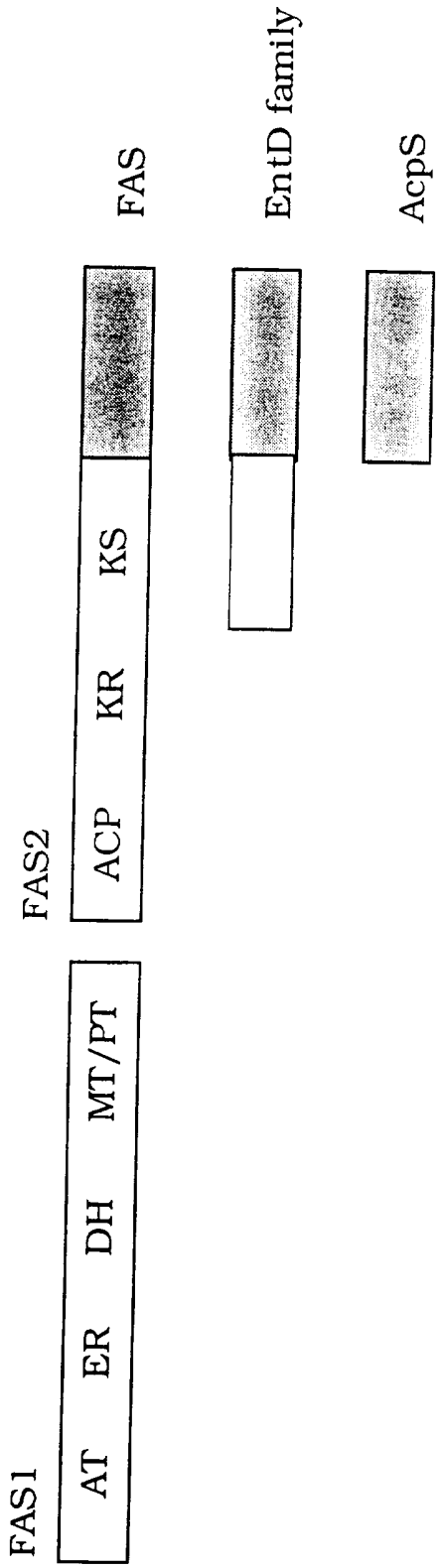


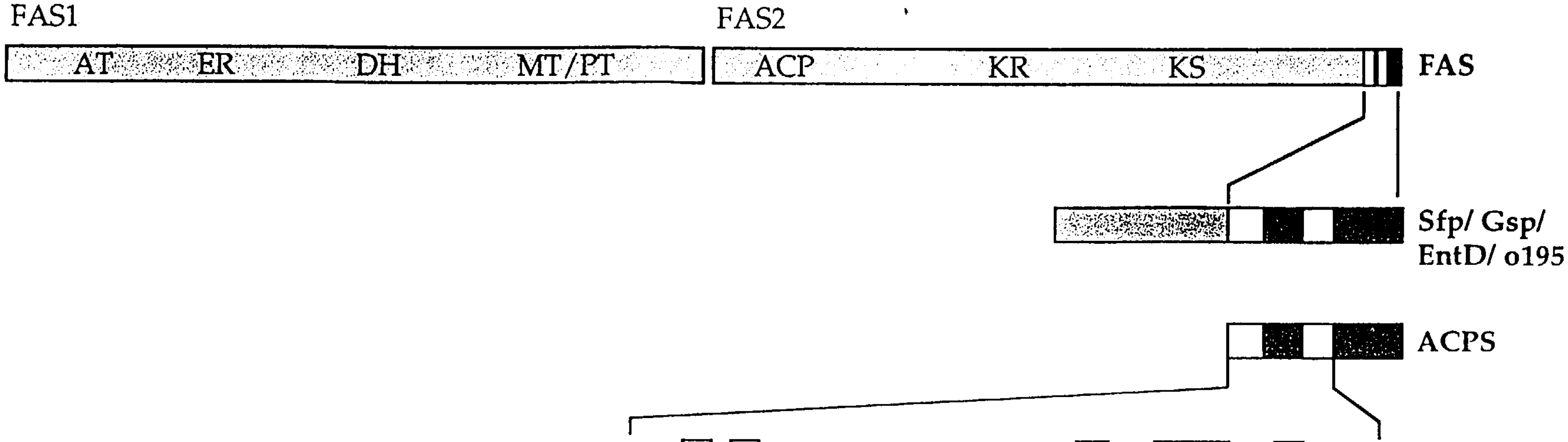
Figure 7

Figure 8

```
2   - - - - - A I L G L G T D I V E I A R I E A V I A R S G D R L A R R V L S D N E W A I W K T H H Q P  acpS E. coli
1   - - - - - S R Q P I G I D I E E I F S V Q T A R E L T D N I I T P A E H E R L - - - - A D C G L A F  entD E. coli
1   V I G A F D S Q P I G I D I E K T K P I S L - - E I A K R F F S K T E Y S D L L A K D K D E Q T D Y  sfp B. subtilis
1   - - - - - S N H P V G I D I E R I S E I D I - - K I A E Q F F H E N E Y I W L Q S K A Q N S Q V S S  gsp B. brevis
2   - - - - - - - G G V G V D V E L I T S I N V E - - - - N D T F I E R N F F T P Q E I E Y C S A Q P S V  fas2 S. cerevisiae

47  V R F L A K R F A V K E A A A K A F G T - G I R N G L A F - - - - - - N Q F E V F N D E L G K P R  acpS E. coli
42  S L A L T L A F S A K E S A F K A S E I Q T D A G F L D Y Q I I S W N K Q Q V I I H R E N - - - - -  entD E. coli
49  F Y H L - - - - W S M K E S F I K Q E G K G L S L P L D S F S V R L H Q D G Q V S I E L P D S H S P C  sfp B. subtilis
44  F F E L - - - - W T I K E S Y I K A I G K G M Y I P I N S F W I D K N Q T - Q T V I Y K Q N K K E P -  gsp B. brevis
41  Q S S F A G T W S A K E A V F K S L G V K S L G G G A A L - - - - - - K D I E I V R V N K N A P A  fas2 S. cerevisiae

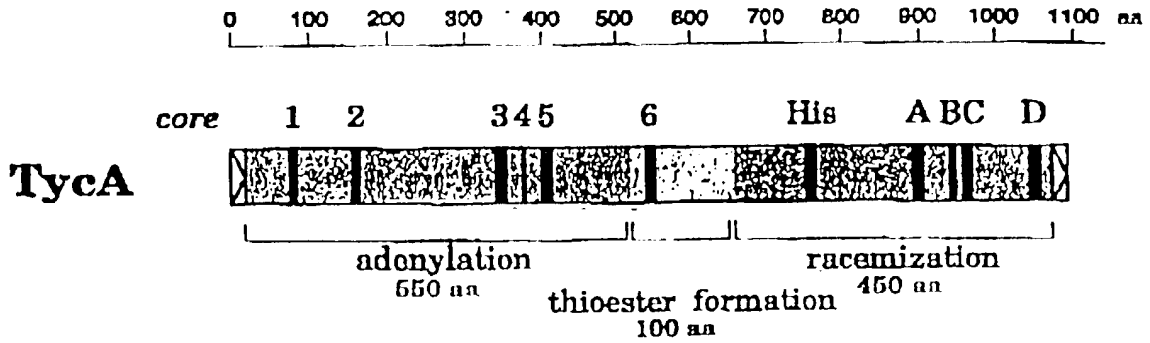
89  - L R L W G E A L K L A E K L G V A N M H V T L A D - - E R H Y A C A T V I I E S  acpS E. coli
87  - - - - - - E M F A - - - - - - - - - - V H W Q I K E K I V I T L C Q H D  entD E. coli
96  Y I K T Y E V D - - P G Y K M A V C A A H P D F P E - D I T M V S Y E E  sfp B. subtilis
89  - V T I Y E P E L F E G Y K C S C C S L F S S V T N L S I T K L Q V Q E L C N L F  gsp B. brevis
84  - V F L H G N A K K A A E E A G V T D V K V S I S H - - D - D L Q A V A V A V S T  fas2 S. cerevisiae
```



ACPS	*E. coli*	5	G L G T D I V E I	54	F A V K E A A A K A F G
EntD	*E. coli*	106	P I G I D I E E I	151	F S A K E S A F K A S E
EntD	*S. flexneri*	106	P I G V D I E E I	151	F S A K E S A F K A S E
EntD	*S. typhimurium*	102	R I G I D I E K I	146	F S A K E S V Y K A F Q
EntD	*S. austin*	102	R V G V D I E K I	146	F S A K E S V Y K A F Q
Sfp	*B. subtilis*	103	P I G I D I E K T	147	W S M K E S F I K Q E G
Psf-1	*B. pumilus*	106	P V G I D I E E I	150	W S M K E A F I K L T G
Gsp	*B. brevis*	102	P V G I D I E R I	146	W T I K E S Y I K A I G
Lpa-14	*B. subtilis*	103	P I G I D I E K M	147	W S M K E S F I K Q A G
NshC	*S. actuosus*	124	G V G I D A E P H	169	F C V K E A V F K A W Y
o195	*E. coli*	87	E V G C D I E V I	133	W T R K E A I V K Q R G
HetI	*Anabaena* sp.	124	Q I G I D L E Y L	170	W T C K E A Y L K A T G
SYCCPNC	*Synechocystis* sp.	106	E I G V D L Q I M	152	W T A K E A F L K A T G
Lys5	*S. cerevisiae*	132	D V G I D I A S P	178	W S L K E S Y T K F T G
CELT04G9	*C. elegans*	118	K V G V D V M R L	171	W C L K E A I L K A T G
HI0152	*H. influenzae*	108	A V G I D I E F P	151	W C L R E A V L K S Q G
unknown	*S. pombe*	122	N I G V D I V E C	169	W T C K E A I L K A L G
FAS2	*E. nidulans*	1428	T I G V D T V T L	1476	W C A K E A V F K C L Q
FAS2	*S. pombe*	1724	N V G V D V E L V	1769	W S A K E A V F K S L G
FAS2	*C. albicans*	1767	G V G V D V E L L	1812	W S A K E A V F K A L G
FAS2	*P. patulum*	1739	K I G V D V E H I	1784	W S A K E A V F K S L G
FAS2	*S. cerevisiae*	1775	G V G V D V E L I	1820	W S A K E A V F K S L G

Figure 9

Thioester formation module (Peptidyl-Carrier-Protein, PCP)




```
  1  MPLTPNDKID RKALPEPDLT ANQSQAAYHP PRTETESILV SIWQNVLGIE   50
 51  KIGIRDNFYS LGGDSIQAIQ VVARLHSYQL KLETKDLLNY PTIEQVALFV  100
101  KSTTRSHHHH HH   112
```

(SEQ ID NO. _____)

FIGURE 10

Figure 11

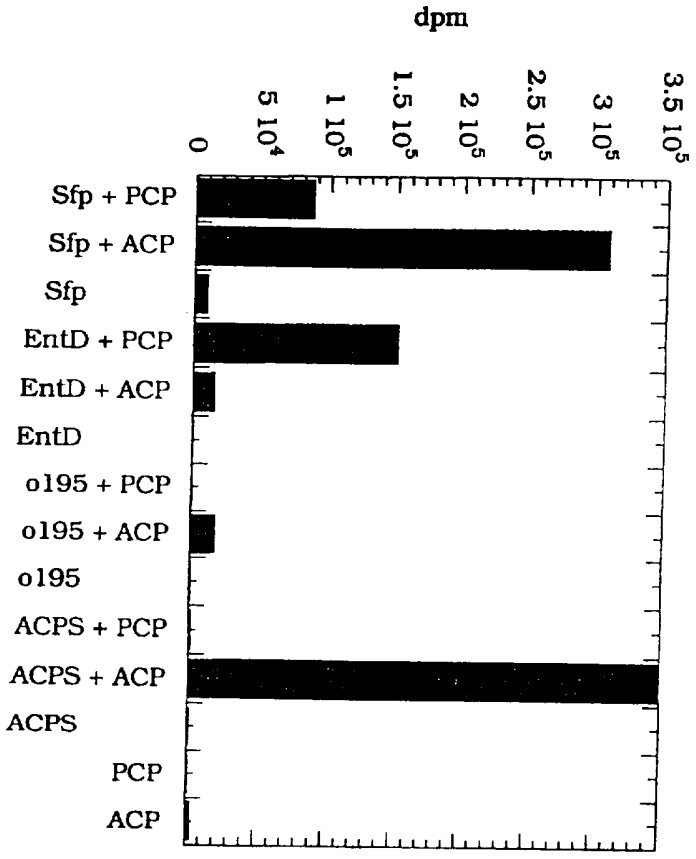

dpm
0
5 10^4
1 10^5
1.5 10^5
2 10^5
2.5 10^5
3 10^5
3.5 10^5
Sfp + PCP
Sfp + ACP
Sfp
EntD + PCP
EntD + ACP
EntD
o195 + PCP
o195 + ACP
o195
ACPS + PCP
ACPS + ACP
ACPS
PCP
ACP

Figure 12

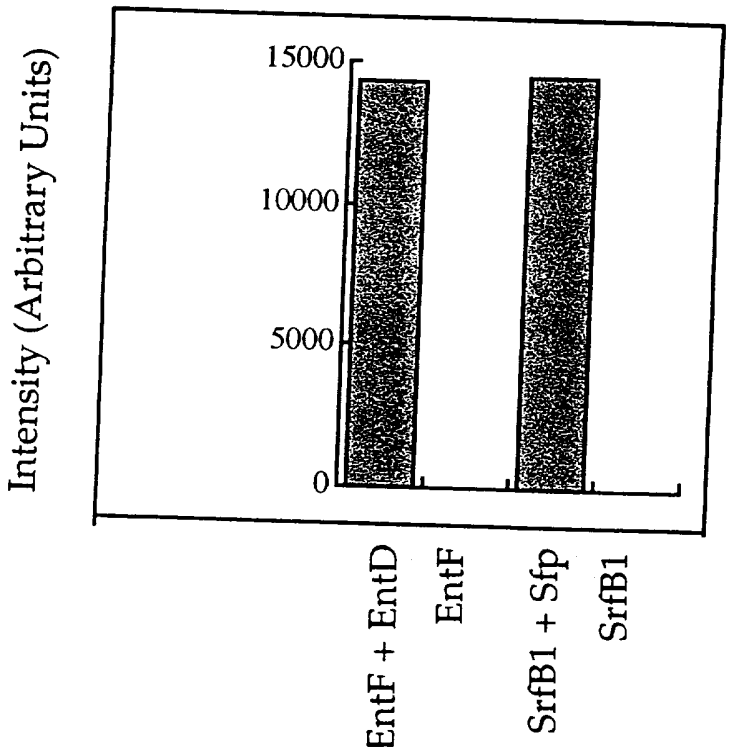

Intensity (Arbitrary Units)
15000
10000
5000
0
EntF + EntD
EntF
SrfB1 + Sfp
SrfB1

Figure 13

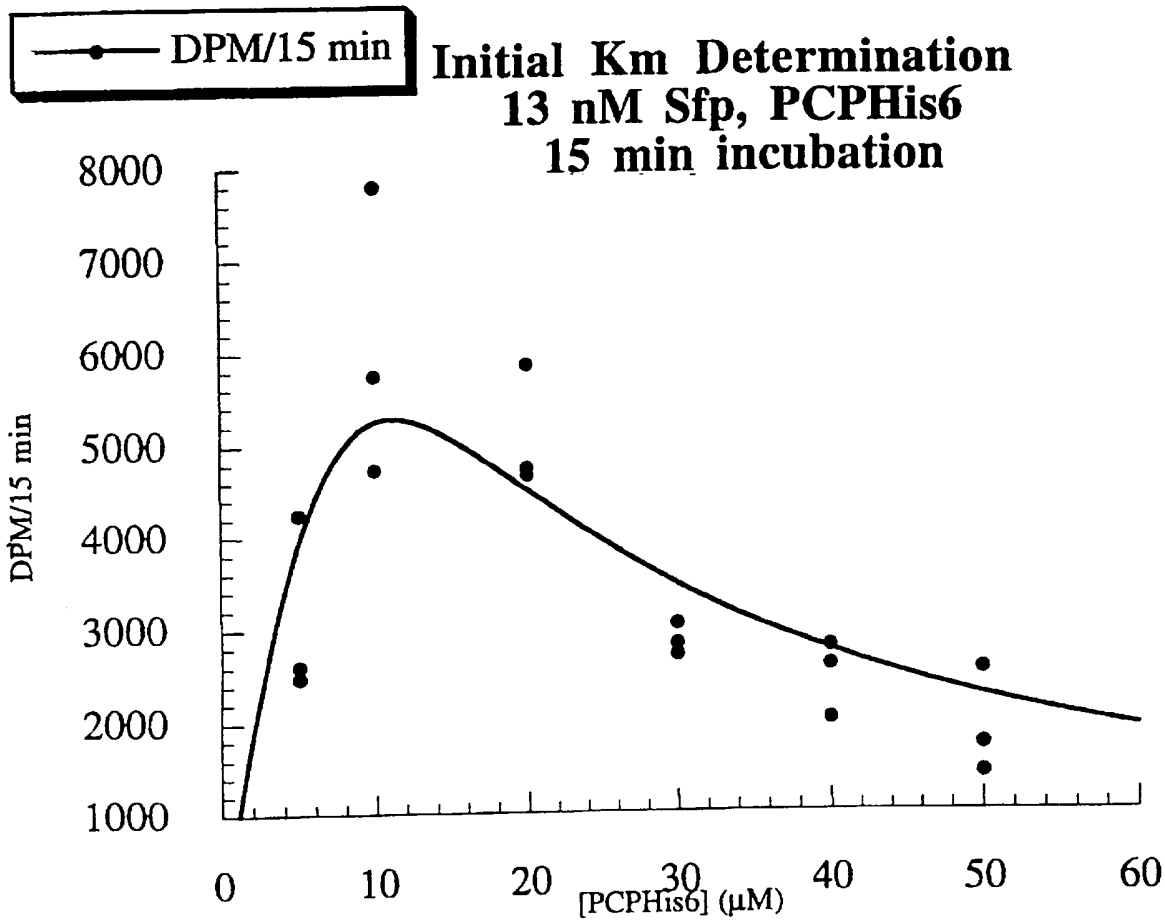

DPM/15 min
Initial Km Determination
13 nM Sfp, PCPHis6
15 min incubation
DPM/15 min
8000
7000
6000
5000
4000
3000
2000
1000
0
10
20
30
40
50
60
[PCPHis6] (μM)

FIG.14A

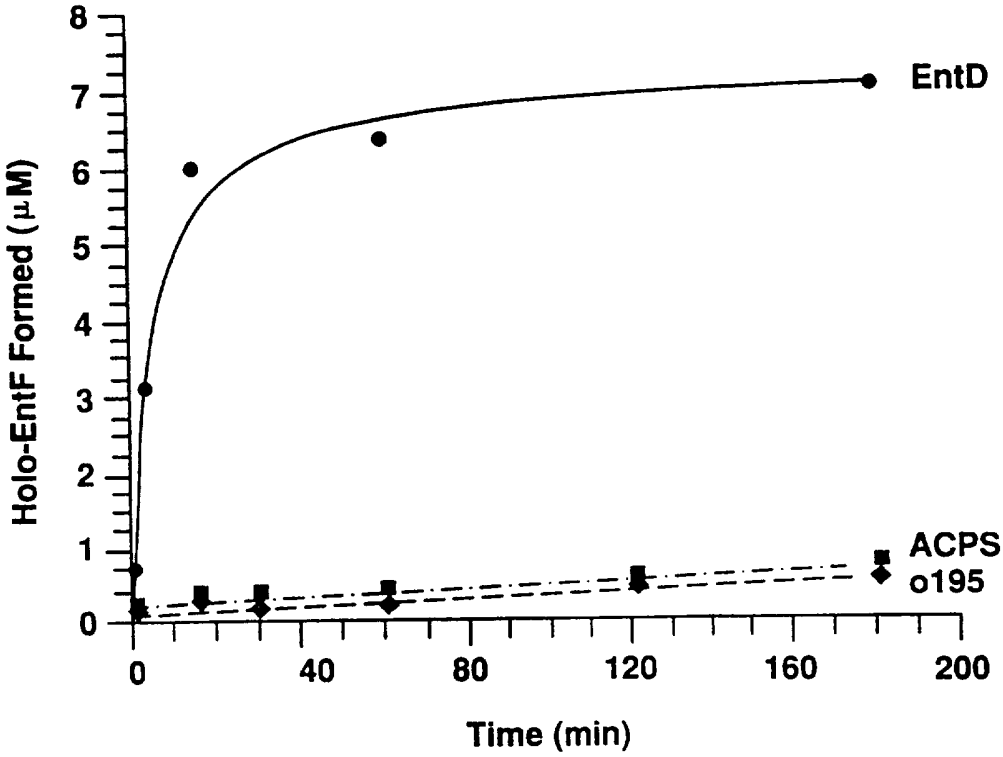


FIG.14B

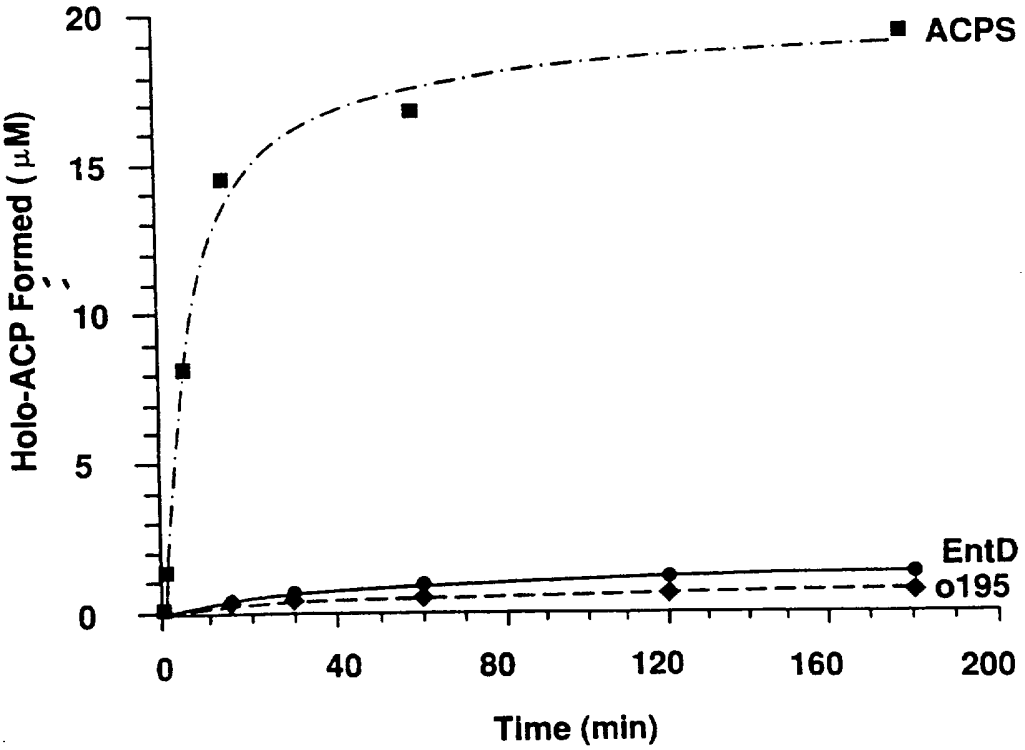
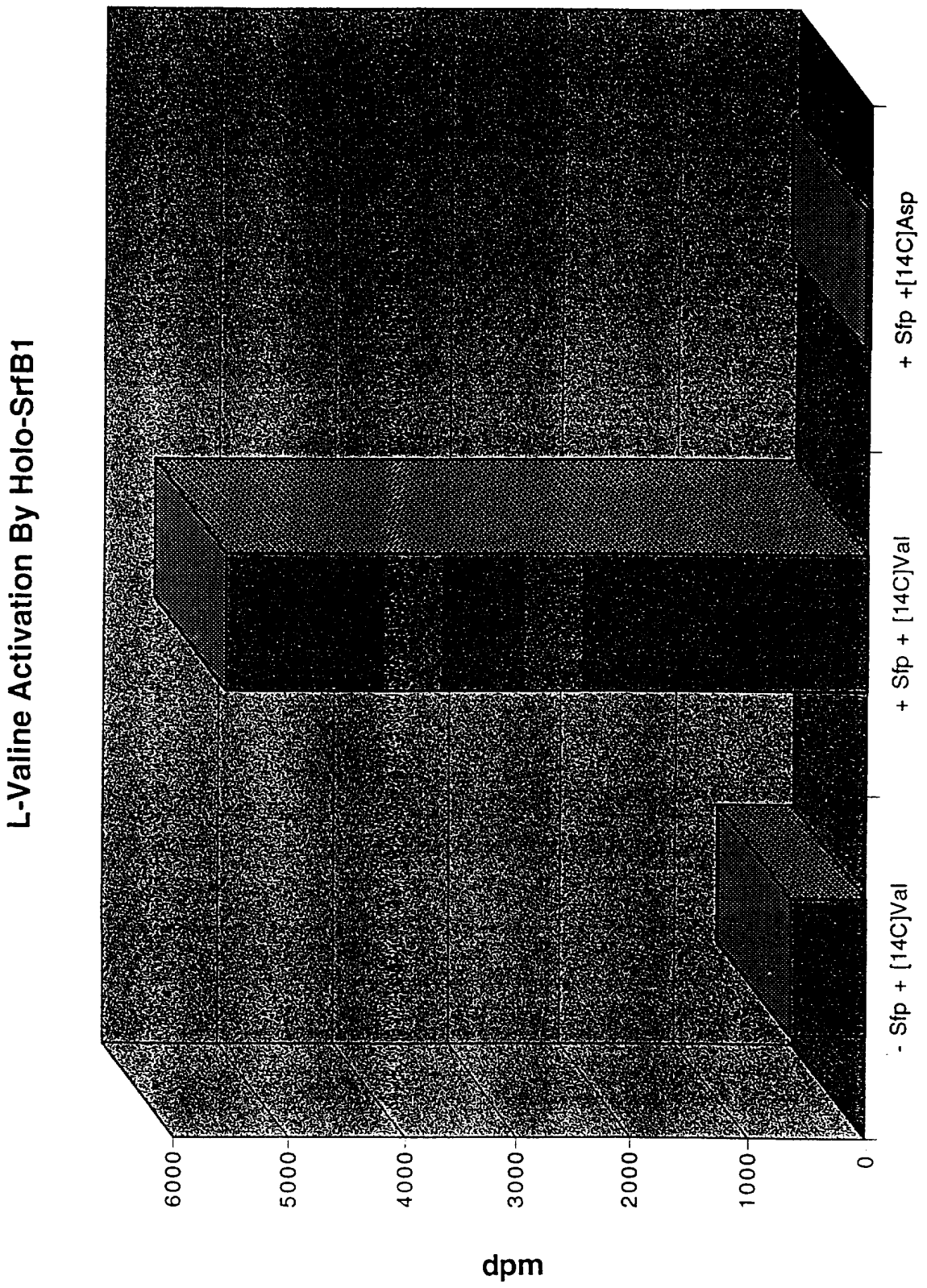


Figure 15


L-Valine Activation By Holo-SrfB1
- Sfp + [14C]Val
+ Sfp + [14C]Val
+ Sfp +[14C]Asp
0
1000
2000
3000
4000
5000
6000
dpm