Chemistry & Biology 13, 869–879, August 2006 DOI 10.1016/j.chembiol.2006.06.011

Interdomain Communication between the Thiolation and Thioesterase Domains of EntF Explored by Combinatorial Mutagenesis and Selection

Zhe Zhou,[1] Jonathan R. Lai,[1] and Christopher T. Walsh[1,*]
[1] Department of Biological Chemistry and Molecular Pharmacology
Harvard Medical School
240 Longwood Avenue
Boston, Massachusetts 02115

Summary

Thiolation (T) domains are protein way stations in natural product assembly lines. In the enterobactin synthetase, the T domain on EntF is recognized in cis by its catalytic partners: the EntF condensation (C), adenylation (A), and thioesterase (TE) domains. To assess surface features of the EntF T domain recognized by C, A, and TE, regions of the EntF T domain were submitted to shotgun alanine scanning and Ent production selection, which revealed residues that could not be substituted by Ala. EntF mutants bearing Ala in such positions were assayed in vitro for Ent production with EntEB, and for A-T, C-T, and T-TE communications. We concluded that G1027A and M1030A are specifically defective in acyl transfer from T to TE. These residues define an interaction surface between these two in cis domains in an NRPS module.

Introduction

Polyketides (PKs) and nonribosomal peptides (NRPs) are two classes of pharmaceutically interesting natural products that are biosynthesized by large, multimodular enzymes known as polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) [1]. Examples of important PK and NRP compounds include the antibiotics erythromycin (PK) [2], rifamycin (PK) [3], and vancomycin (NRP) [4], the immnosuppresants FK506 (PK) [5] and rapamycin (hybrid PK/NRP) [6], and the antitumor agents epothilone [7] and bleomycin (both are hybrid PK/NRP) [8].

Central to the multidomain architecture of PKSs and NRPSs are carrier proteins, autonomously folding domains of 80–100 amino acid residues (8–10 kDa) that carry the growing acyl or peptidyl chains during substrate elongation cycles [9]. Carrier proteins from PKS assembly lines have been termed acyl carrier proteins (ACPs), and those from NRPSs have been termed peptidyl carrier proteins (PCPs); both are homologous to the ACPs from primary metabolism (fatty acid synthases, FASs). Because ACP and PCP domains require posttranslational priming to become active, by installation of a phosphopantetheinyl prosthetic group whose terminal thiol becomes the site of covalent tethering of growing acyl chains, the ACP and PCP domains are often referred to as thiolation (T) domains to emphasize the pantetheinyl thiol functionality [9]. This convention is used here.

In type I synthases/synthetases, with multiple domains organized into monomer-processing modules and up to eight modules per protein subunit, a T domain is embedded in each module and is serviced in cis by the adjacent catalytic domains [10–12]. In type II FAS, PKS, and NRPS multienzyme systems, the T domains are most often free-standing subunits, interacting in trans with catalytic partners [11, 13, 14]. Whether in cis or in trans, the multiple cycles of loading and elongation that result in production of the growing acyl or peptidyl chain require that T domains participate in well-timed communication events with the domains that activate the monomer to be incorporated (adenylation domains, A, in NRPSs and acyl transferase domains, AT, in PKSs) and the domains that catalyze chain elongation (condensation domains, C, and ketosynthase domains, KS) [15, 16]. In addition, the T-domain-tethered substrates can undergo various modifications, such as oxidation [17], methylation [18], epimerization [19], and halogenation [20]. Therefore, a great amount of information is encoded in the 8–10 kDa T domains for directing enzymatic reactions in a spatially and temporally controlled fashion.

The siderophore enterobactin, produced by gram-negative enteric bacteria in times of iron starvation, is a cyclic trilactone comprised of three *N*-(2,3-dihydroxybenzoyl)-serine (DHB-Ser) residues [21]. The DHB moieties in enterobactin provide three catecholic groups to form a high affinity hexadentate complex with ferric iron. Enterobactin is assembled by the two module NRPS enterobactin synthetase [22], spreading over three proteins EntEBF. EntE is a DHB-selective adenylation domain, EntB is a two-domain protein, one domain of which is a free-standing T domain, and EntF is a four-domain NRPS elongation/termination module (C-A-T-TE) [23–27]. The EntEB pair functions together as a type II NRPS, while the EntF module has type I organization. The A domain of EntF activates L-Ser and installs it on the phosphopantheine free thiol of the EntF holo-T domain [27]. The C domain condenses the DHB acyl group (presented as the thioester tethered to EntB via its phosphopantetheine group, DHB-*S*-EntB) onto the amino group of Ser-*S*-EntF [28]. Finally the TE domain of EntF acts as an elongation and cyclotrimerization catalyst, releasing the trilactone siderophore [29]. A scheme for the catalytic steps for enterobactin production is shown in Figure 1A.

The two T domains in the Ent assembly line thus reflect the two distinct contexts, in cis versus in trans, for recognition by partner domains. Ample evidence exists from natural product assembly lines with in trans components that many T domains are differentiated by potential partner catalytic domains. An example of this, T domain selectivity is demonstrated by the PKS systems that produce frenolicin (FrnN/J) and R1128 (ZhuG/N) [30]. Only cognate pairwise interactions of T domains and KSs occur in both initiation and elongation cycles of these type II synthases. In the NRPS system for coronatine production, the nonheme iron halogenase

*Correspondence: christopher_walsh@hms.harvard.edu

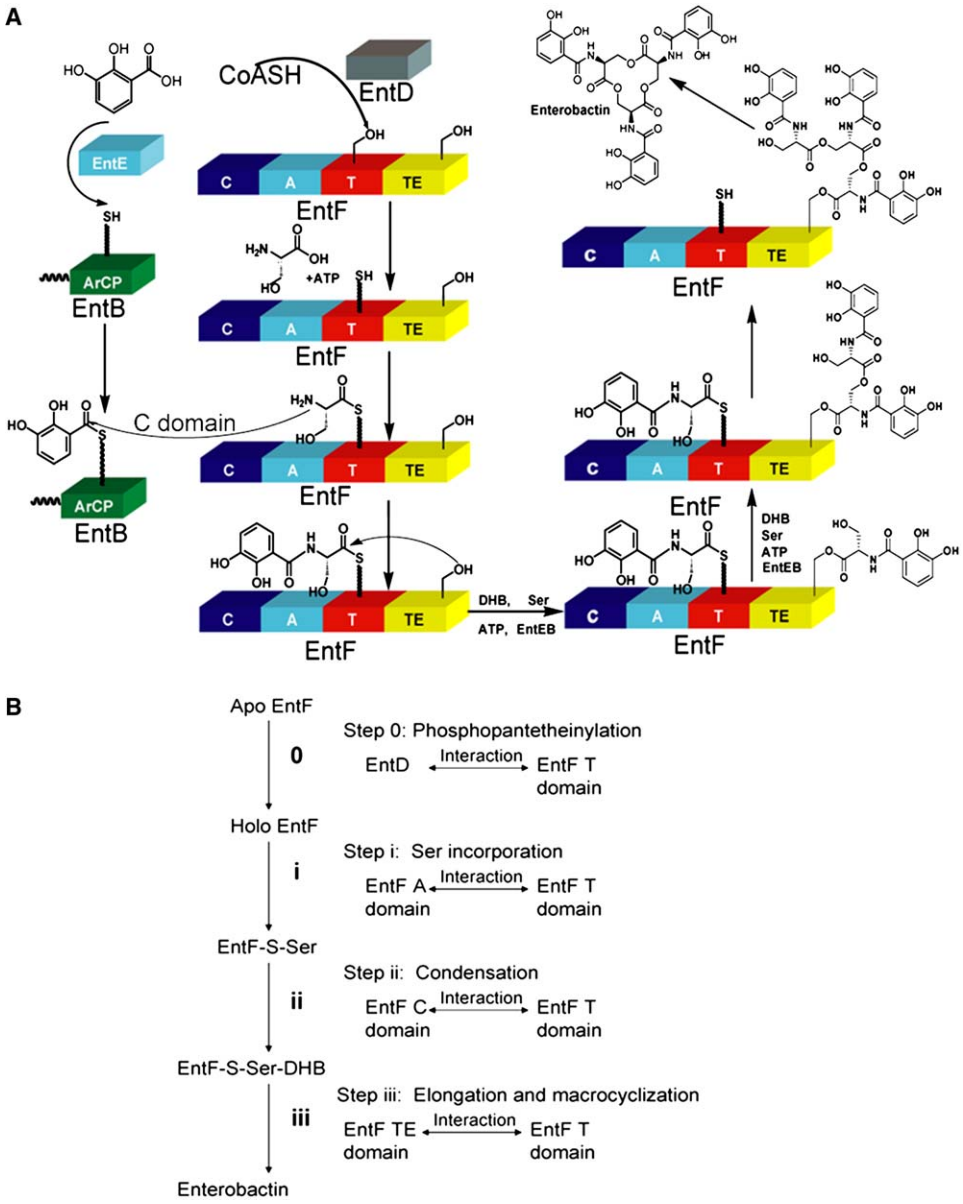

Figure 1. Enterobactin Biosynthesis Scheme

(A) EntD is a PPTase that primes EntB and EntF. EntE loads DHB onto holo-EntB-ArCP. EntF is a four-domain NRPS elongation/termination module (C-A-T-TE). The A domain of EntF catalyzes loading of the T domain with Serine. The C domain then mediates condensation of Serine with DHB loaded on EntB ArCP. The TE domain elongates DHB-Ser and macrocyclizes the DHB-Ser trimer to form enterobactin.

(B) The cascade of enterobactin biosynthesis reactions that involve the EntF T domain. The interdomain interactions that must occur for each step are shown.

that installs a chlorine group onto alloisoleucine (CmaB) recognizes the alloisoleucine substrate only when presented on the T domain of CmaD but not CmaA [20]. It has been much more difficult to evaluate whether in cis T domains of NRPS modules have such specific recognition by their juxtaposed neighboring catalytic domains.

In any prototypic enzymatic assembly line, there are chain initiation, elongation, and termination cycles [1]; in type I NRPS systems, these would comprise an A-T module, a C-A-T module, and a C-A-T-TE module, respectively. In an initiation module the T domain must be recognized by: (1) the dedicated phosphopantetheinyl transferase (PPTase) that primes it from inactive apo to the active holo form; (2) the A domain that selects, activates, and loads a specific acyl/aminoacyl group on the holo-T domain; (3a) the downstream C domain in the first elongation module that catalyzes bond formation between the activated initiation acyl/aminoacyl group and the downstream amino acid. In elongation modules there is: (4) recognition of an acyl/aminoacyl-*S*-T domain by both the upstream as well as the downstream C domains during a condensation/chain transfer step. Finally, when the full-length acyl chain has reached the termination module, the downstream C domain is typically replaced by the TE domain; this scenario would