Regulation of protein activity with small-molecule-controlled inteins

GEORGIOS SKRETAS AND DAVID W. WOOD
Department of Chemical Engineering, Princeton University, Princeton, New Jersey 08544, USA
(RECEIVED July 18, 2004; FINAL REVISION October 22, 2004; ACCEPTED October 27, 2004)

Abstract
Inteins are the protein analogs of self-splicing RNA introns, as they post-translationally excise themselves from a variety of protein hosts. Intein insertion abolishes, in general, the activity of its host protein, which is subsequently restored upon intein excision. These protein elements therefore have the potential to be used as general molecular “switches” for the control of arbitrary target proteins. Based on rational design, an intein-based protein switch has been constructed whose splicing activity is conditionally triggered in vivo by the presence of thyroid hormone or synthetic analogs. This modified intein was used in Escherichia coli to demonstrate that a number of different proteins can be inactivated by intein insertion and then reactivated by the addition of thyroid hormone via ligand-induced splicing. This conditional activation was also found to occur in a dose-dependent manner. Rational protein engineering was then combined with genetic selection to evolve an additional intein whose activity is controlled by the presence of synthetic estrogen ligands. The ability to regulate protein function post-translationally through the use of ligand-controlled intein splicing will most likely find applications in metabolic engineering, drug discovery and delivery, biosensing, molecular computation, as well as many additional areas of biotechnology.

Keywords: protein function regulation; intein; protein splicing; allosteric enzyme; domain insertion; small molecule

The development of efficient tools for controlling protein function in living cells is one of the greatest challenges of chemical biology. Traditionally, genetic approaches such as regulation of transcription and gene knockouts have been used extensively as they enable control of gene function in a highly specific and generalizable manner (Shogren-Knaak et al. 2001). Genetics, though, are hindered by transcriptional and translational delays before a functional product is formed as well as by compensation phenomena which can occur when the genetic information of a cell is perturbed (Shogren-Knaak et al. 2001). An alternate strategy is chemical genetics, where a small molecule is used to directly activate or inactivate a particular protein target (Schreiber 2003). Although this method can provide dose-dependent control over gene products of interest at the post-translational level, a different small-molecule regulator must be discovered for each individual protein target (Shogren-Knaak et al. 2001).

Inteins are protein elements that are found as in-frame insertions within the sequences of particular host genes (Noren et al. 2000; Paulus 2000). They possess the ability to excise themselves post-translationally from their protein hosts and ligate the flanking peptide sequences (exteins) via an efficient self-catalyzed reaction called protein splicing (Noren et al. 2000; Paulus 2000). In general, inteins activate their host proteins but restore their activity upon splicing (Daugelat and Jacobs 1999; Wood et al. 1999; Lew and Paulus 2002; Zeidler et al. 2004). Furthermore, several inteins have been shown to retain their activity when transferred to nonnative genes or cell hosts (Derbyshire et al. 1997b; Wu et al. 2002; Gangopadhyay et al. 2003; Zeidler et al. 2004). The only host protein requirement for efficient splicing is a Cys, Ser, or Thr residue as the first amino acid of the C-terminal extein (Noren et al. 2000). This flexibility...
suggests that inteins could be used as molecular “switches” for the control of arbitrary target proteins. However, construction of a practical in vivo protein switch requires that the intein splicing reaction be readily controlled by an easily manipulated signal that does not affect the general cellular metabolism.

Previous attempts to generate inteins with tunable behavior include the evolution of temperature-sensitive intein alleles (Adam and Perler 2002; Cann et al. 2004; Zeidler et al. 2004), inteins incorporating nonnative amino acids whose activity is regulated by photolysis (Cook et al. 1995) as well as ligand-controlled reconstitution and splicing of split inteins (Mills and Paulus 2001; Mootz and Muir 2002; Mootz et al. 2003). Although all of these approaches provide valuable insights on how to engineer controllable protein splicing they have not yet demonstrated a general ability to activate different proteins in vivo through controllable intein splicing. Very recently, two studies have shown that controllable intein splicing in cis can be used to achieve this goal. In the first case, regulation of protein activity was accomplished with the isolation and use of temperature-sensitive variants of the *Saccharomyces cerevisiae* vacuolar ATPase subunit intein (Zeidler et al. 2004) and, in the second case, with the construction of a novel chimeric intein whose splicing activity is triggered by the addition of 4-hydroxytamoxifen (Buskirk et al. 2004).

This work describes a rational protein engineering approach for generating inteins whose splicing activity is regulated in vivo by the presence of human thyroid hormone. Designed according to other engineered allosteric enzyme prototypes (Baird et al. 1999; Doi and Yanagawa 1999; Tucker and Fields 2001; Guntas and Ostermeier 2004), an artificial intein chimera was created by fusing a thyroid–hormone-binding domain within a previously engineered mini-intein. The insertion of the binding domain abolished the splicing activity of the intein, but allowed it to be later restored by addition of thyroid hormone or synthetic analogs. The resulting allosteric intein was then used to conditionally activate a variety of different proteins in *Escherichia coli* in a dose-dependent manner. Finally, a combination of directed evolution and genetic selection was employed to engineer an additional controllable intein whose splicing activity is inhibited by the presence of an orthogonal set of synthetic estrogen ligands.

**Results**

*Design and construction of a ligand-activated intein*

The previously engineered ΔI-SM mini-intein (Wood et al. 1999) was chosen to be the protein splicing element of the chimeric fusion. This artificial mini-intein was originally derived from the *Mycobacterium tuberculosis* RecA wildtype intein (Mtu RecA) through the deletion of its entire central endonuclease domain (ΔI mini-intein) (Derbyshire et al. 1997b). A single amino acid substitution was then used to enhance its stability and restore splicing efficiency back to the level of the wild-type protein (ΔI-Splicing Mutant) (Wood et al. 1999). The Mtu RecA intein and the ΔI-SM mutant have been shown to retain high levels of splicing efficiency when inserted within a variety of nonnative gene hosts (Derbyshire et al. 1997b; Daugelat and Jacobs 1999; Wood et al. 1999; Lew and Paulus 2002; Wu et al. 2002; Gangopadhyay et al. 2003). The ligand-binding domain of the human thyroid hormone receptor β (TR) was selected as the small-molecule-binding protein. This binding protein, a member of the nuclear receptor superfamily, has been shown to undergo a conformational change as well as a stability enhancement upon thyroid hormone binding (Apriletti et al. 1995; Wagner et al. 1995; Ribeiro et al. 1998). The solved crystal structure of TR reveals that its N- and C-termini are in relative proximity (Wagner et al. 1995), a factor that may play an important role in avoiding large steric constraints in the constructed chimera. TR was inserted into the ΔI-SM intein in place of its deleted endonuclease domain. This decision was based on a priori knowledge that this is a permissive site for insertions of short polypeptides (Gangopadhyay et al. 2003) as well as entire folded protein domains (Fitzsimons Hall et al. 2002; Wu et al. 2002). Furthermore, it has been shown that mutations in the endonuclease domain of the full-length Mtu RecA intein can affect splicing efficiency (unpubl. results). This implies that this structure allows for significant functional cross-talk between the two comprising domains. The resulting chimeric intein is referred to as ΔI-SM<sup>TR</sup> (Fig. 1A).

In order to evaluate the splicing performance as well as the controllability of the ΔI-SM<sup>TR</sup> intein, the chimeric construct was inserted into the sequence of the bacteriophage T4 thymidylate synthase gene (TS) of the plasmid pKT to yield pKT:ΔI-SM<sup>TR</sup>. The pKT vector was developed as part of a previous genetic selection system where intein insertion was shown to abolish TS activity and restore it upon splicing (Derbyshire et al. 1997b; Wood et al. 1999; Fig. 1B). Intein activity can therefore be monitored through the growth phenotype of TS-knockout cells expressing these fusions in thymineless growth media. Because the requirement for active TS increases rapidly with increasing temperature, observation of TS phenotypes at different temperatures provides semi-quantitative information about intein splicing efficiency (Derbyshire et al. 1997b). This genetic selection system has been calibrated in such a way that TS<sup>−</sup> phenotypes at low temperatures imply very low splicing activity, whereas TS<sup>+</sup> phenotypes at high temperatures correspond to a highly active intein (Derbyshire et al. 1997b; Wood et al. 1999).

*Escherichia coli* D1210ΔthyA cells transformed with pKT:ΔI-SM<sup>TR</sup> and grown on thymineless agar plates exhib-
with pKT:

hormone (T₃) or synthetic analogs (Triac). A thymine-rich medium containing (TTM) agar plates for 24 h at 37°C yielded strong TS⁺ phenotypes (reporter system for evaluation of its splicing efficiency and controllability. /H9004 row) and incubated on thymineless (+THY) was used as a positive control. Remarkable, the addition of the natural thyroid hormone 3,3',5'-triiodo-L-thyronine (T₃) or synthetic analog 3,3',5'-triiodothyroacetic acid (Triac) had an inhibitory effect on cell growth (Fig. 1C). Furthermore, in a thymine-rich medium supplemented with the dihydrofolate reductase inhibitor trimethoprim, where selection against cells that carry active TS takes place (Belfort and Pedersen-Lane 1984), the addition of T₃ or Triac had an inhibitory effect on cell growth (Fig. 1C). This provides strong evidence that the observed thyroid–hormone-dependent phenotypes arise from ligand-triggered splicing activity and not from a general splicing-activating or growth-enhancing property of thyroid hormone.

Figure 1. Construction of a thyroid–hormone-controlled intein. (A) Design of the ligand-binding domain/intein chimera. ΔI: Mycobacterium tuberculosis RecA artificial mini-intein; ΔI-SM: fast-splicing mutant of the ΔI intein; TR: ligand-binding domain of the human thyroid hormone β; ΔI-SMTR: chimeric intein. (B) Insertion of the ΔI-SMTR intein into the TS reporter system for evaluation of its splicing efficiency and controllability. (C) Growth phenotypes of E. coli D1210ΔthyA cells transformed either with pKT:ΔI (top row), pKT:ΔI-SM (middle row), or pKT:ΔI-SMTR (bottom row) and incubated on thymineless (+THY) and trimethoprim-containing (TTM) agar plates for 24 h at 37°C with and without 50 μM thyroid hormone (T₃) or synthetic analogs (Triac). A thymine-rich medium (+THY) was used as a positive control.

To directly visualize the hormone-regulated splicing activity suggested by the TS growth phenotypes, the chimeric intein was inserted within a tripartite fusion composed of a maltose-binding protein (M) as the N-terminal extein, fused to the ΔI-SMTR intein, fused to the C-terminal DNA-binding domain of the I-TevI homing endonuclease (C) (Derbyshire et al. 1997a) as the C-terminal extein (Fig. 2A). The high solubility of the M and C domains and the fact that the intein is inserted as an end-to-end fusion allows for the overexpression of a soluble precursor protein and, thus, for the direct analysis of the splicing reaction products via SDS-PAGE. Furthermore, the identities of the products resulting from intein splicing in this context have been verified by Western blot analysis and mass spectrometry in previous work (Derbyshire et al. 1997b). Stable and soluble MΔI-SMTRC precursor protein was initially overexpressed in the absence of hormone for 1 h, at which point chloramphenicol was added to arrest expression of additional precursor. Cells containing the overexpressed MΔI-SMTRC precursor were then incubated at 37°C in the presence and absence of 40 μM T₃, and samples were analyzed for appearance of the spliced ΔI-SMTR intein and the ligated MC extein product via Coomassie-stained SDS-PAGE (Fig. 2B). Although some premature splicing was observed during overexpression, no additional formation of ligated MC or spliced ΔI-SMTR was observed in the absence of thyroid hormone. However, the addition of T₃ resulted in increased formation of the spliced intein as well as of the ligated extein product. In the presence of hormone, product accumulation could clearly be detected 1 h after the addition of T₃ and the reaction had gone effectively to completion after ~4 h as judged by loss of precursor (data not shown). It was noted that the precursor protein was also consumed in the absence of hormone, presumably by isolated N- and C-terminal cleavage reactions or by simple degradation, although no additional spliced products following overexpression were formed in this case.
Thyroid hormone-induced activation of other test proteins

The general ability of the ΔI-SMTR intein to controllably activate additional host proteins was investigated by creating insertional fusions of this intein with β-lactamase and β-galactosidase (β-gal). The chimeric intein was initially inserted into the sequence of TEM-1 β-lactamase immediately upstream of Cys121. The initial 23-residue periplasmic signal peptide was removed to ensure cytoplasmic accumulation of the precursor protein (Zlokarnik et al. 1998). The constructed intein-β-lactamase fusion was expressed in the E. coli strain BL21(DE3) in the presence and absence of 40 μM T3 and cell lysates were evaluated for β-lactamase activity using the chromogenic substrate nitrocefin (Galarneau et al. 2002). In cells grown in the presence of thyroid hormone, lysate addition resulted in a rapid color change from yellow to red, thus revealing the presence of active β-lactamase (Fig. 3A). In contrast, no significant nitrocefin color change could be observed in the absence of hormone over the same time period.

The ΔI-SMTR intein was additionally inserted upstream of Cys11 of the lacZα gene in the vector pUC18. The resulting intein fusion was expressed in the E. coli strain XL1-Blue with and without 40 μM T3. This strain constitutively expresses the ω-fragment of β-gal and successful α-complementation can be detected with a number of colorimetric assays (Shuman and Silhavy 2003). Yellow color formation resulting from the hydrolysis of the chromogenic compound o-nitrophenyl-β-D-galactopyranoside (ONPG) was observed only when hormone had been added (Fig. 3B). Thyroid–hormone-triggered β-gal activity was also observed when a white-blue assay was performed with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (data not shown). This intein fusion, as well as the β-lactamase fusion, could similarly be activated by addition of Triac (data not shown).

Protein activation is dosable

The ability to regulate the function of a target protein with a single small-molecule activator in a dosable manner would offer remarkable flexibility. The previously observed hormone-dependent TS phenotypes were used to evaluate the potential of using the ΔI-SMTR intein to achieve controlled levels of protein activation. E. coli D1210ΔthyA cells were transformed with pKT:ΔI-SMTR and grown in liquid...
thymineless medium at 37°C in the presence of increasing concentrations of T₃. Once again, cell survival relied on the presence of hormone while levels of cell growth were graded with respect to the amount of added T₃ (Fig. 4A). The generated dose–response curve revealed an apparent thyroid hormone detection limit of ~2.5 μM and a half-maximal effective concentration (EC₅₀) of ~12 μM. Cell growth appeared to saturate at ~80 μM T₃ but concentrations >100 μM were not tested as the solubility of the hormone in the growth medium became limiting. Thyroid hormone concentrations between 5 and 40 μM appear to provide a good linear range of protein activation. As expected, control experiments with unmodified inteins in cells harboring either the original pKT:ΔI or pKT:ΔI-SM plasmids indicated that these inteins are insensitive to thyroid hormone in this concentration range (Fig. 4A; data not shown). Furthermore, the ΔI-SM₄₃⁻²⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-~-
plates in the presence and absence of a 50-µM concentration of the high-affinity synthetic estrogen analog diethylstilbestrol. After 48 h of incubation, survivors were identified on all plates, thus implying Δl-SM<sup>ER</sup> variants with splicing activity. A number of mutants from both plates were isolated and evaluated for estrogen-regulated splicing activity by TS phenotype in the presence and absence of estrogen and synthetic analogs. Three of the mutants that exhibited estrogen sensitivity were selected for further characterization. When expressed in E. coli D1210ΔthyA cells and incubated in liquid thymineless medium at 34°C, cell growth was inhibited in the presence of synthetic high-affinity ER binders (Fig. 5). A control TS phenotype experiment with cells harboring either the pKT:Δl or pKT:Δl-SM vectors showed that the addition of the estrogenic compounds tested does not have a significant general impact on cell growth (Fig. 5; data not shown). The TS phenotypes exhibited by the isolated Δl-SM<sup>ER</sup> mutants reveal that, in contrast to the Δl-SM<sup>IR</sup> intein, the evolved variants are functional in the absence of synthetic estrogen but can be inactivated upon tight ligand binding. All of the mutants were found to include multiple mutations in the estrogen-binding as well as in the splicing domain of the fusion (Table 1). The specific role of these mutations is currently under investigation but the ones found in the ligand-binding domain appear to be clustered in areas either important for hormone binding (Wrenn and Katzenellenbogen 1993; Ekena et al. 1996; Brzozowski et al. 1997) or in the vicinity of the contact points between the binding and the splicing domains.

![Figure 5](image-url) TS phenotypes of an estrogen-regulated intein. Growth of E. coli D1210ΔthyA cells transformed either with pKT:Δl-SM<sup>IR</sup> mutant 2 (gray) or pKT:Δl (white) and grown in 5-mL cultures of liquid thymineless medium at 34°C for 14 h. Cells transfected with pKT:Δl-SM yielded a saturated culture after ~6 h of incubation time (data not shown). 17β-Estradiol, diethylstilbestrol, hexestrol, and dienestrol were added at a concentration of 40 µM and 4-hydroxytamoxifen at 10 µM. All cultures contained an equal amount of ethanol that was kept below 1%. The numbers above the growth bars represent the relative binding affinity of each compound for ER (17β-Estradiol is arbitrarily set to 100) as reported by Kuiper et al. (1997). Experiments were carried out in duplicate. Error bars represent one standard deviation from the mean value.

**Discussion**

Based on simple rules of protein design, an engineered self-splicing protein element (intein) has been constructed that is conditionally triggered in vivo by the presence of natural or synthetic thyroid hormone. The potential of this controllable intein to be used as a conditional activator for arbitrary target proteins was investigated by inserting the engineered intein into the sequences of a number of different proteins. It was shown in all cases that these proteins are functional only in the presence of hormone after ligand-induced splicing takes place. Protein splicing is triggered rapidly, allowing detection of the ligated extein product in <60 min by SDS-PAGE. It is also demonstrated that intein insertion can disrupt the assembly and activation of homo- and heteromultimeric protein complexes (TS and α-complementation of β-gal, respectively), suggesting a possible use in modulating protein–protein interactions in more complex systems. The activation of intein splicing is conveniently dosable, and thus the amount of functional target protein can be easily specified in vivo by adding the required amount of the activator ligand. This offers an advantage over systems under transcriptional control where the concentration of inducer can at best regulate transcription levels and not necessarily the specific amount of functional protein. The described approach can also be combined with random mutagenesis and selection to evolve inteins which are regulated by other small molecules, thus enabling the simultaneous control of different protein targets in the same cell.

The engineered intein was constructed by fusing the ligand-binding domain of the human thyroid hormone receptor within a previously developed minimal splicing domain. This hormone-binding domain is known to undergo a conformational change and a stability enhancement upon ligand binding (Apriletti et al. 1995; Wagner et al. 1995; Ribeiro et al. 1998). The insertion of the foreign domain into the splicing domain is designed to cause a structural perturbation, and most likely affects its stability as well. These effects abolish the splicing activity of the intein, but allow it to be restored upon hormone binding. A number of similarly engineered allosteric enzymes have been previously reported, including a dihydrofolate reductase enzyme regulated by estrogen or FK506 in yeast (Tucker and Fields 2001), and the recently reported insertional fusion of β-lactamase with a maltose-binding domain (MBD) where β-lactamase activity can be activated by a variety of carbohydrates (Guntas and Ostermeier 2004). Interestingly, from a number of tested MBD binders in the latter example, only those that induce conformational changes upon binding were found to have an impact on the enzymatic activity of the chimeric fusion (Guntas and Ostermeier 2004). In other reported artificial chimeric fusions, the enzymatic activity of one domain is allosterically controlled by the binding of small ions (Baird et al. 1999), specific antibodies (Brennan et al. 1995;
Benito et al. 1996), or entire folded protein domains (Doi and Yanagawa 1999). Similar mechanisms of allosteric modulation may be responsible for the ligand sensitivity of the constructed chimeric intein.

Another possible explanation for the thyroid–hormone-triggered protein activation mediated by the ΔI-SM<sup>ER</sup> intein is that the presence of hormone stabilizes the intein–extein fusion precursor. According to this mechanism, a precursor protein is formed with the ability to splice by itself in the absence of ligand, but hormone addition stabilizes this fusion, thus resulting in an increased amount of released extein product. Although this simple mechanism would still enable our desired regulation of protein activity in vivo, this mode of control is not consistent with some of the results presented in this study. First, a highly soluble MΔI-SM<sup>TR</sup>C precursor protein could be overexpressed in the absence of hormone (Fig. 2), while the addition of T<sub>3</sub> or Triac was found to have an insignificant impact on the expression levels of the precursor protein (data not shown). These data suggest that a moderately stable and well-folded precursor protein can be expressed with this intein regardless of the presence of hormone. Furthermore, the post-translational splicing activation shown in Figure 2 indicates the presence of "splicing competent" precursor protein at least 2 h after expression. Although some background splicing was observed during overexpression, the precursor was unable to continue splicing without thyroid hormone. This leads us to hypothesize that some splicing can be attributed to a possible folding intermediate that may be able to splice at low frequency. In addition, the highly optimized MΔI-SM<sup>TR</sup>C splicing context is likely to exacerbate any splicing leakage. In contrast, there is no evidence of background splicing in the absence of thyroid hormone with all the insertional reporter fusions tested (TS, β-gal and β-lactamase), as revealed by phenotypic and enzymatic assays. Finally, the identification of variants of the ΔI-SM<sup>ER</sup> intein, whose splicing activity is inhibited by tight binding of estrogen analogs, indicates that ligand-induced stabilization is not sufficient to explain the behavior of the constructed inteins.

A similar chimeric intein that can be triggered by the presence of 4-hydroxytamoxifen in yeast has been very recently reported (Buskirk et al. 2004). In this case, a mutant estrogen-binding domain was inserted into a minimal Mtu RecA intein and the resulting chimera was evolved for allosteric activity using a combination of positive and negative selections. Similar to the ΔI-SM<sup>ER</sup> chimera, the evolved ER-intein was then used to post-translationally activate a number of unrelated proteins in the yeast <i>S. cerevisiae</i> in a dose-dependent way. These results, taken together with our findings, demonstrate that the described protein engineering approaches can be used to evolve families of small-molecule-controlled inteins that respond to different com-

<table>
<thead>
<tr>
<th>Mutant</th>
<th>N-terminal</th>
<th>C-terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I29M</td>
<td>L119P; E131G</td>
</tr>
<tr>
<td>2</td>
<td>Y79F</td>
<td>L119Q; M127V</td>
</tr>
<tr>
<td>3</td>
<td>A42V</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Summary of the mutations found in the evolved ΔI-SM<sup>ER</sup> inteins

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Splicing domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I29M, Y79F, A42V</td>
</tr>
<tr>
<td>2</td>
<td>L119P, E131G, L119Q, M127V</td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Estrogen-binding domain

<table>
<thead>
<tr>
<th>N-term</th>
<th>Helix 3</th>
<th>Helix 4/5</th>
<th>Helix 6/7</th>
<th>Helix 8/9</th>
<th>Helix 10/11</th>
<th>Helix 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L310P</td>
<td>N359Y</td>
<td>W383R</td>
<td>L448P</td>
<td>H501R</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y331C</td>
<td></td>
<td></td>
<td>L479Q</td>
<td>Q502R</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>T311S</td>
<td>M343K</td>
<td>V392D</td>
<td>H398Y</td>
<td>G415E</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D332E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>S305N</td>
<td>V368A</td>
<td>I475N</td>
<td>K520E</td>
<td>L541Q</td>
<td>L544M</td>
</tr>
</tbody>
</table>

Estrogen Contact Points

<table>
<thead>
<tr>
<th></th>
<th>M343</th>
<th>W383</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L346</td>
<td>M384</td>
</tr>
<tr>
<td></td>
<td>M421</td>
<td>G521</td>
</tr>
<tr>
<td>2</td>
<td>T347</td>
<td>L384</td>
</tr>
<tr>
<td></td>
<td>I424</td>
<td>M522</td>
</tr>
<tr>
<td>3</td>
<td>L349</td>
<td>M388</td>
</tr>
<tr>
<td></td>
<td>F425</td>
<td>H525</td>
</tr>
<tr>
<td></td>
<td>L540</td>
<td></td>
</tr>
</tbody>
</table>

* Taken from Brozozowski et al. 1997.
pounds, and that these chimeric inteins can be used to regulate protein activity in different host organisms. Ultimately such intein-based molecular protein switches could be applied to the post-translational control of protein activity in higher eukaryotes.

An interesting contrast between the evolved ER-intein variants described by Buskirk et al. (2004) and the ΔI-SME\textsuperscript{ER} inteins reported here is that the former are activated by 4-hydroxytamoxifen, whereas the latter are inhibited upon ligand binding. Although the design of the ER-splicing domain fusions were very similar in the two studies, the specifics of how they were constructed and of how the genetic selection was carried out might be responsible for the very different outcomes. First, in this work, residues 301–553 of ER were used to replace residues 110–383 of the Mtu RecA intein, whereas Buskirk et al. (2004) used residues 304–551 of the receptor sequence along with six-aa-long linkers flanking the receptor’s termini to replace residues 95–382 of the intein. It is very likely that the length of the linkers that join the two domains together is important for the structural and functional communication of the two domains in the resulting chimera. Second, our ΔI-SME\textsuperscript{ER} fusion was created by inserting the wild-type ER sequence into a splicing mutant of a minimal Mtu RecA intein, whereas Buskirk et al. (2004) inserted a mutant ER, that binds 4-hydroxytamoxifen but not natural estrogen, into a wild-type minimal Mtu RecA intein. Starting from different points in sequence space may have guided the evolution experiments in different directions with different results.

The way the genetic selection was carried out might also explain why no estrogen-inhibited ER-intein variants were identified in the Buskirk et al. (2004) study. In this case, round one variants were selected for active splicing only in the presence of 4-hydroxytamoxifen, so mutants with the ability to splice only in the absence of ligand would not have survived. Finally, our work took place entirely in E. coli, while the evolved ER-intein reported by Buskirk et al. (2004) was used only in yeast. An important difference between these two organisms is the molecular chaperone Hsp90, which has been shown to form complexes with the ligand-binding domains of steroid receptors (Pratt and Toft 1997). As verified by Buskirk et al. (2004), Hsp90 increases expression and apparent stability of their ER-intein precursors. Although a bacterial homolog of Hsp90 exists in E. coli, it has many functional differences (Buchner 1999), and in contrast to steroid receptors expressed in mammalian cell lines (Pratt and Toft 1997), it does not copurify with ΔI-SME\textsuperscript{ER} in the absence of estrogen (unpubl. results). The Hsp90 chaperone is only one example of the many differences between yeast and E. coli that might explain differences between the behaviors of these two inteins. An important point to note is simply that controllable inteins could be generated in both E. coli and yeast, suggesting widespread future applications.

Small-molecule-controlled inteins are likely to find utility as conditional activators for gene products of unknown function in proteomics studies as well as in numerous other applications in biotechnology. For example, ligand-induced intein splicing out of a strong reporter protein could be used as a highly specific biosensor. In principle, since intein splicing can be used to activate any host protein, the reporter protein can be chosen according to convenience. Because the activation is irreversible, the reporter would provide high sensitivity to the presence of the ligand. In drug discovery, allosteric activation of a reporter based on ligand binding to a drug target could become a rapid method for chemical library screening. This method could be combined with in vivo combinatorial biosynthesis and a genetic selection to directly identify cells that synthesize potential drug compounds against the target. Finally, these inteins might find applications in cellular and molecular computation. Groups of allosteric inteins that react to various secondary metabolites might be organized in complex networks to regulate gene expression, enzyme activity, or the production of various reporter proteins. In this case, the inteins could provide a new type of computational tool, or an advanced control strategy for cellular metabolism.

Materials and methods

Reagents and strains

All experiments were carried out with E. coli strains. XL1-Blue cells (Stratagene) were used in all cases for plasmid constructions as well as for the expression of the intein-LacZα fusion. The strain D1210ΔthyA::KanR [F’Δ(gpt-proA)62 leuB6 supE44 ara-14 galK2 lacY1 Δ(mcrC-mrr) rpsL2 (Str') yfl-5 mit-1 recA13 lacP'] was used for the determination of TS growth phenotypes (Derbyshire et al. 1997b). BL21(DE3) (Novagen) was used for the expression of the intein-β-lactamase fusion, while ER2566 (New England Biolabs) was used for the overexpression of the tripartite ΔI-SME\textsuperscript{ER}C fusion. Estrogen, thyroid hormone, and their synthetic analogs (17-β-estradiol, diethylstilbestrol, hexestrol, dienestrol, 4-hydroxytamoxifen, 3,3’5’-triiodo-L-thyronine and 3,3’5’-triiodothyroacetic acid) were purchased from Sigma. All compounds were prepared as 10 mM stock solutions in EtOH. Nitrocefin was purchased from Becton-Dickinson and X-gal from Fisher Scientific. Both compounds were dissolved in dimethylsulfoxide to form 2 mM and 1% solutions, respectively.

Plasmid construction in the TS context

The coding sequence corresponding to residues Glu203–Asp461 of the human thyroid hormone receptor β were amplified from the vector pCMV-hTRβ (provided by Paul Webb, Metabolic Research Unit, University of California, San Francisco), while the one coding for residues Ser301–Thr553 of the human estrogen receptor α were derived from Gal4-ER-VP16 (a gift from Randall Morse, Wadsworth Center). In each case, PCR primers were used to add a BssHII restriction site at each of the N- and C-termini of the hormone-binding domains. The ligand-binding domains of the thyroid hormone and estrogen receptors were then inserted within the
BssHII site of the mini-intein of the vector pKT:ΔI-SM to replace residues 110–383 of the Mtu RecA intein (Wood et al. 1999). The sequences of the constructed mini-intein/binding domain chimeras were verified with nucleotide sequencing.

**Determination of TS phenotypes**

Cells were grown overnight in LB medium containing 200 μg/mL ampicillin and supplemented with 50 μg/mL thymine. For experiments in liquid media these saturated cultures were used with a 1:200 dilution to inoculate 5 mL of defined thymineless medium (Belfort and Pedersen-Lane 1984) with 200 μg/mL ampicillin and the specified concentrations of each of the hormone ligands. Stock solutions of various concentrations of hormone analogs in ethanol were prepared to ensure a constant final ethanol concentration in phenotype tests. In all growth experiments ethanol concentrations were kept below 1%. Levels of cell growth were measured as OD_{600} on a GENESYS 2 spectrophotometer. Thymine-rich media contained at least 50 μg/mL thymine, and trimethoprim was added to a 10-μg/mL concentration. All experiments were carried out at least in duplicate.

**Splicing kinetics with SDS-PAGE**

The MΔI-SM^{TR}C fusion was constructed by replacing the NTT383 mini-intein of the previously described pMIC vector (Derbyshire et al. 1997b) with the ΔI-SM^{TR} intein. Cells carrying the resulting plasmid were grown in LB medium with 200 μg/mL ampicillin to mid-log phase (OD 600 ∼ 1) and overexpression of the mini-intein was induced by the addition of 1 mM isopropyl-β-D-thiogalactoside (IPTG) for 1 h at 37°C. Chloramphenicol was then added to a concentration of 70 μg/mL to terminate protein production, and the specified concentrations of thyroid hormone were supplied. The progress of the splicing reaction was monitored by analyzing clarified cell lysates via SDS polyacrylamide gel electrophoresis.

**β-Lactamase and β-galactosidase colorimetric assays**

Silent SacI and PstI restriction sites flanking Cys121 of the TEM-1 β-lactamase gene of the vector pGEM-11Zf(+) (Promega) were introduced with the GeneEditor in vitro site-directed mutagenesis system (Promega). These sites were used to insert the ΔI-SM^{TR} intein immediately upstream of the Cys121 residue of this gene. The constructed intein-β-lactamase fusion was then transferred to the polylinker of pET-24a(+) (Novagen) using the available EcoRI and HindIII restriction sites, along with a N-terminal maltose-binding protein affinity tag. Cells carrying the resulting plasmid were grown overnight in LB medium with 20 μg/mL kanamycin, and this saturated culture was used with a 1:100 dilution to inoculate fresh 5-mL cultures of LB media with 20 μg/mL kanamycin and the specified concentrations of thyroid hormone. Cell were harvested after ~6 h of incubation, washed with PBS, and resuspended in 1 mL of 100 mM phosphate buffer at pH 7.0. The cells were then lysed by sonication and cell debris was removed by centrifugation at 4°C for 5 min. Nitrocefin was then added to a final concentration of 100 μM and its color change was monitored as the change in OD_{492} on a GENESYS 2 spectrophotometer.

Cells transfected with the vector carrying the intein-LacZα fusion were grown overnight in LB medium with 200 μg/mL ampicillin and diluted 1:100 into fresh 5-mL cultures of LB media with 200 μg/mL ampicillin. When the cell density reached an OD_{600} of 0.2–0.3, 1 mM IPTG was added to the cultures along with the specified concentrations of thyroid hormone. Cells were harvested after ~2 h of incubation and ONPG hydrolysis was monitored according to a β-galactosidase assay kit (Stratagene).

**ΔI-SM^{ER} intein mutagenesis and selection**

Mutants of the ΔI-SM^{ER} intein were generated with error-prone PCR as previously described (Wood et al. 1999) except that 0.1 mM MnCl2 was added in four reactions each one containing a reduced concentration of one of the four nucleotides from 200 to 80 μM. A library of ~10,000 variants was constructed and selected for functional mutants.

**Acknowledgments**

We thank Paul Webb and Randall Morse for providing plasmids with the TR and ER coding sequences, respectively. This work was supported by startup funds provided to D.W.W. by the Department of Chemical Engineering at Princeton University.

**References**


Doi, N. and Yanagawa, H. 1999. Design of generic biosensors based on green