

Inteins

Prepared by Juliet Williams

Basics

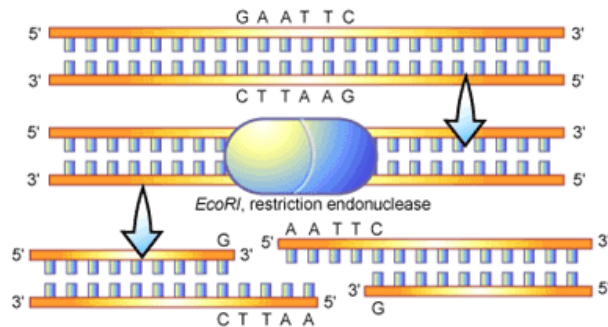
- Definition: a segment of a protein that is able to excise itself and rejoin the remaining portions with a peptide bond (Anaraku).
 - AKA Protein introns.
 - Unlike introns, inteins are transcribed and translated with their host proteins (Khan).
- Found in all living organisms, but mostly archea.
- Found mainly in enzymes involved in DNA replication and repair (Liu 2000).
 - Generally in highly conserved regions.
- Ratio of intein's size and host protein varies greatly.

Basics

- Genes that code for them are considered “selfish” or “parasitic” genes because they contain an endo-nuclease domain that makes them able to insert themselves in the genome

Parasitic because it doesn't contribute to the fitness of the organism

- Inteins are preserved in organisms because:
 - They do not interrupt the function of the expressed protein.
 - They are not excised because they are located in vital proteins.



<http://www.scq.ubc.ca/restriction-endonucleases-molecular-scissors-for-specifically-cutting-dna/>

First Intein

- First intein discovered 1987 by Anaruku while studying carrot and *Neurospora crassa* vacuolar ATPases and a putative Ca²⁺ pumping ATPase.
 - The N- and C-terminal segments of the Ca²⁺ pumping ATPase gene were very similar to corresponding regions in the vacuolar ATPases.
 - The central region of the Ca²⁺ pumping ATPase gene had no similarity to any known ATPase, but was instead similar to that associated with an endonuclease.

First Intein

- The Ca²⁺ pumping ATPase had also been shown to make yeast resistant to calmodulin antagonist trifluoperazine.
- The gene isolated for the vacuolar ATPase was shown to have the same sequence, including the central sequence, however, no trifluoperazine resistance was present
- Both proteins were run through a gel and it was shown the vacuolar protein had a molecular weight of the protein without the central region expressed
 - This region must have been spliced out since it was present in the mRNA but not the final product

Split Intein

- Intein of the precursor protein may come from two different genes
 - When the two protein come together the two intein splice out can the two exteins remain and are now attached

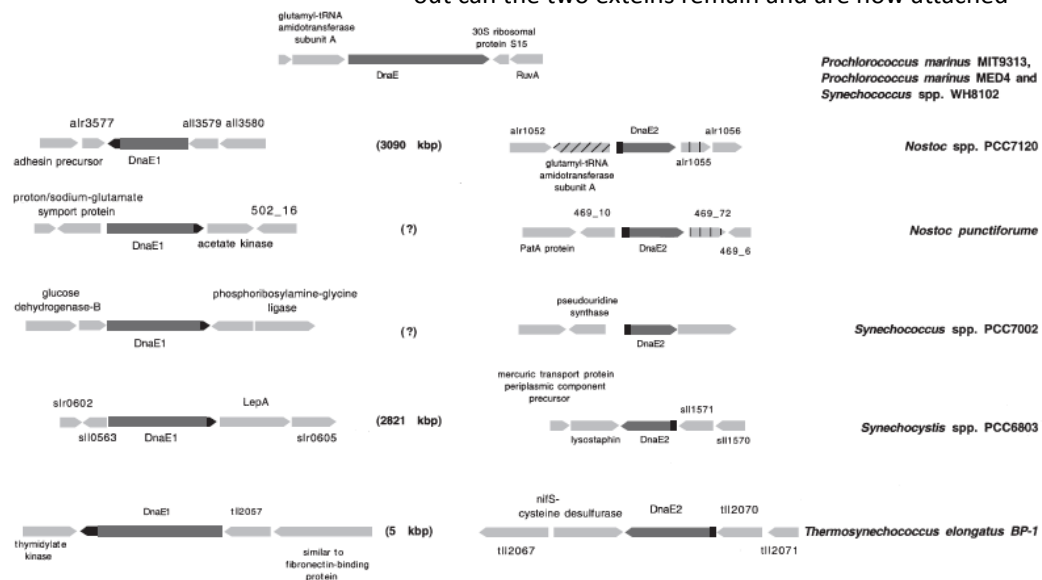
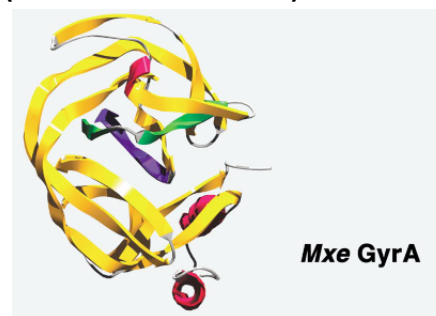
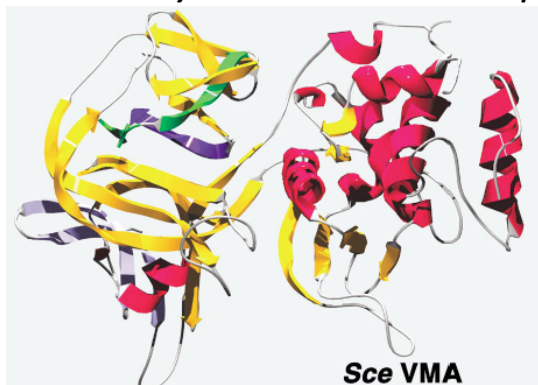


Fig. 1. Cyanobacterial *dnaE* gene loci. Each line shows the *dnaE* gene locus or the loci of *dnaE1* and *dnaE2* split genes from one species, except for the top locus that is identical in all three listed species. Gene protein-coding regions are shown as rectangles with an arrowhead at their 3' ends. The *dnaE* genes are shown in dark grey with the split intein parts in black. Other homologous genes are indicated by similar patterns – there are only two such pairs, between *Nostoc* species PCC7120 and *Prochlorococcus marinus* MED4, and *Nostoc* species PCC7120 and *Nostoc punctiforme*. Gene names or functions are indicated where known. The distance between the split *dnaE* genes in each species is indicated where known.

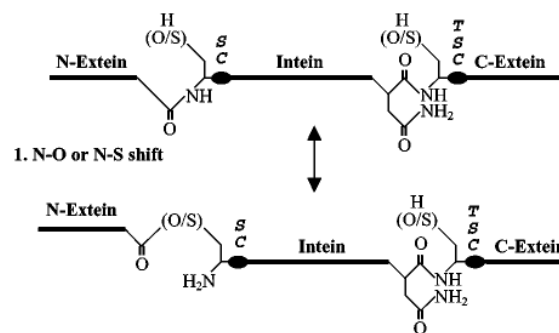
3D Structure Comparison

- Endonuclease domain on the right, self splicing domain on the left (Duan 1997)
 - They are clearly distinct
 - Splicing domain similar in structure to mini-intein in *Mycobacterium xenopi* (Klabunde 1998)



Mechanism of Protein Splicing

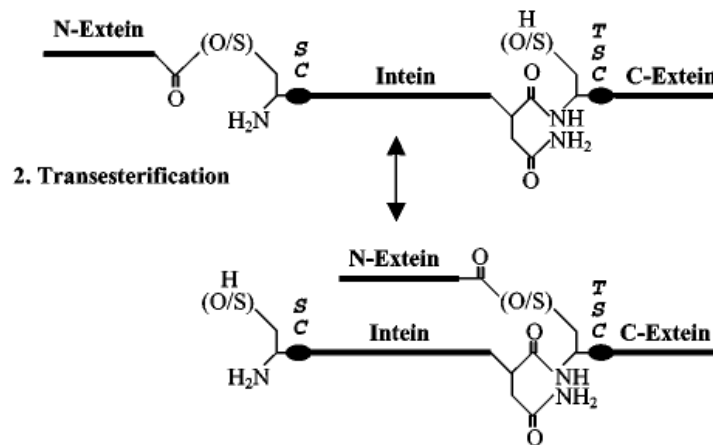
- 4 steps (Perler 1997)
- 1. The amino-terminal splice junction of the intein is ac



By and NO that leads to an ester intermediate	By and NS that leads to a thioester intermediate
Rearrangement occurs	
N-extein binds to the oxygen of a serine at the amino terminal splice junction	N-extein binds to sulfur of a cysteine at the amino terminal splice junction

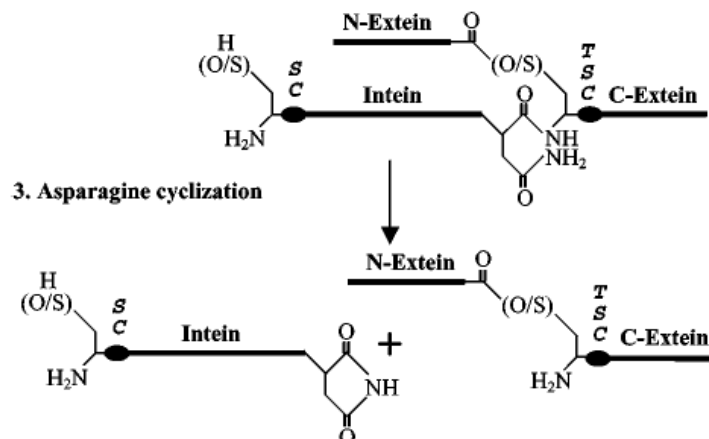
Mechanism of Protein Splicing

2. Cleavage of the ester or thioester at the amino-terminal splice junction occurs via nucleophilic attack from a residue in the C-terminal splice junction.
 - this transesterification results in a branched protein intermediate.



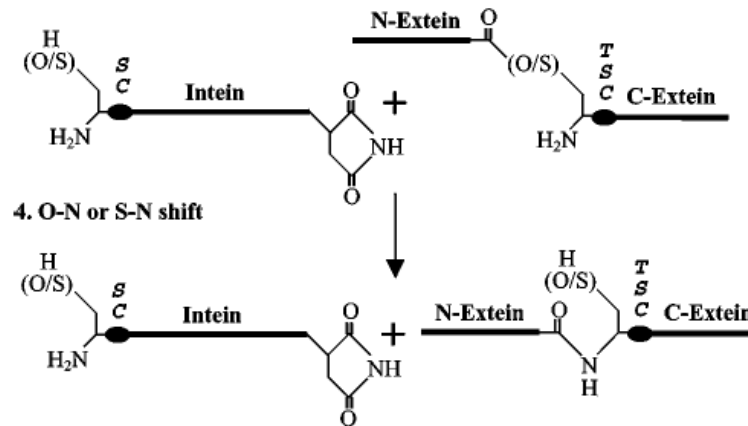
Mechanism of Protein Splicing

3. Cleavage occurs via Asparagine or Glutamine cyclization.
 - This causes excision of the intein and splice of the two exteins by an ester/thioester bond.



Mechanism of Protein Splicing

- Rearrangement results in the formation of a peptide bond between the two exteins.



Life Cycle

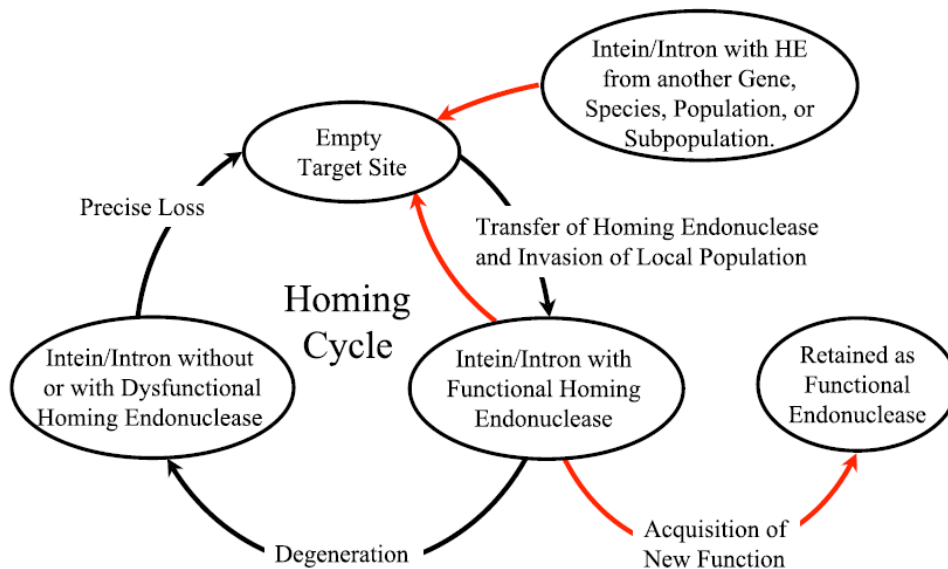


Figure 1
Homing cycle of a parasitic genetic element (modified from [3, 13]). Recent findings suggest that due to complex population structure the cycle might not operate in synchrony in different subpopulations. The red arrows indicate the trajectory of the functioning HE and the black arrows the fate of the host gene. The precise loss can occur through recombination with an intein or intron free allele, or, in case of introns, through recombination with a reverse transcript of the spliced mRNA [39, 40].

Uses for Inteins

- Inteins are very efficient proteins splicers
 - This makes them very useful for biotechnological applications
- Hopes for gene therapy
 - Inserted in to genes to prevent function
 - Insertion of a hydrophilic intein into a hydrophobic protein could aid in uptake into organelles such as the mitochondria (Grey 2000)

Inteins in Biotechnology

- Protein purification
 - Traditional techniques involve the incorporation of a tag into the gene
 - The tag causes the protein to bind in an affinity chromatography column
 - The protein is released when the column is washed with an external agent that breaks the protein from the tag

Inteins in Biotechnology

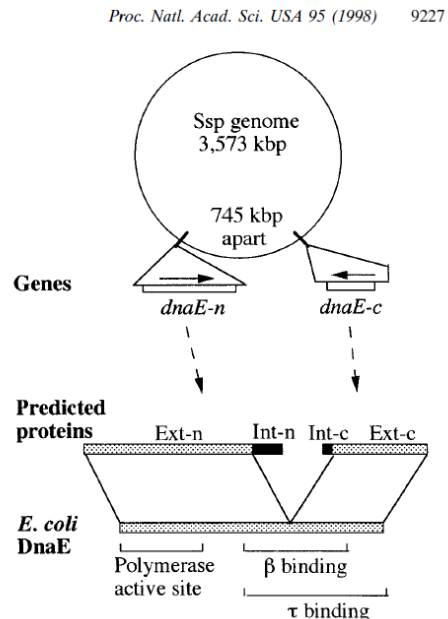
- Protein purification
 - The *N*- or *C*-terminus of an intein gene can be mutated and the desired gene can have the intein incorporated into it
 - After expression of the gene the fusion protein is purified by utilizing properties of the intein
 - The chemical environment is changed in a unique way that triggers excision of the intein
 - New England Biolabs

Inteins in Biotechnology

- In vitro trans-splicing was used to introduce NMR labels into a large protein
 - This allows structural analysis of large proteins (over 50kD) by NMR (Otomo 1999)

Example

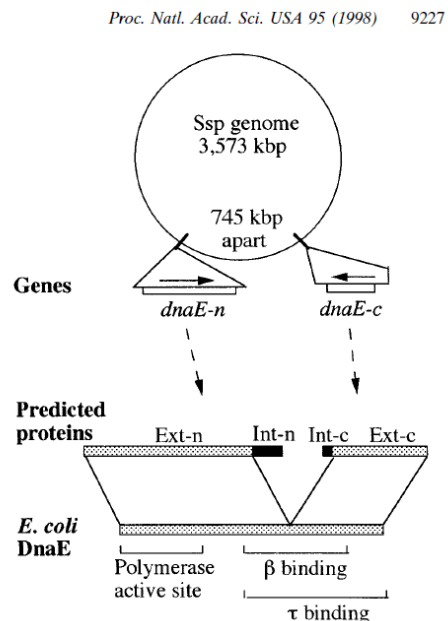
- Classic Example
- The most commonly used naturally occurring split intein is the *Ssp* DnaE intein of *Synechocystis* sp. PCC6803
 - The N-terminal of the intein has 134 amino acids and the C-terminal has 36
 - *Synechocystis* has its catalytic α subunit of the DNA polymerase III (DnaE) split
 - Gene *dnaE-n* encodes for the first 774 amino acids
 - Gene *dnaE-c* encodes for the remaining 423 amino acids



Wu, H *et al.* (1998) "Protein trans-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803" *Proc. Natl. Acad. Sci. U.S.A.* 95, 9226-9231
 Reviewed by Khan 2005

Example

1. The two portions of DnaE are transcribed from opposite strands of the genome.
 2. Each gene is transcribed and translated with part of the DnaE and part of the split intein.
 3. The intein undergoes self excision and the exteins are ligated using normal DNA replication/repair machinery.
- This is called a trans-splicing reaction.
 - There is also a *Ssp* DnaB mini intein capable of trans-splicing.



A Biomedical Use for Inteins

- De Grey, A. (2000) Mitochondrial gene therapy: an area for the biomedical use of inteins. *Trends Biotech.* 18, 394-399
- Mitochondrial DNA mutations may play a role in some rare diseases as well as the normal aging process.
- Genetic engineering seems to most plausible fix for mtDNA mutations