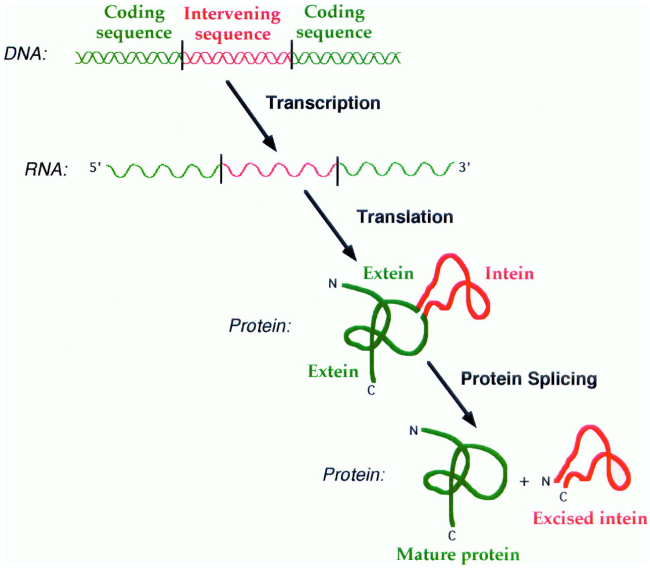
Two Proteins from One Polypeptide
Protein Splicing -- Exteins and Inteins

Andrew Pierce
Microbiology, Immunology and Molecular Genetics
University of Kentucky

MI/BCH/BIO 615

Protein Splicing



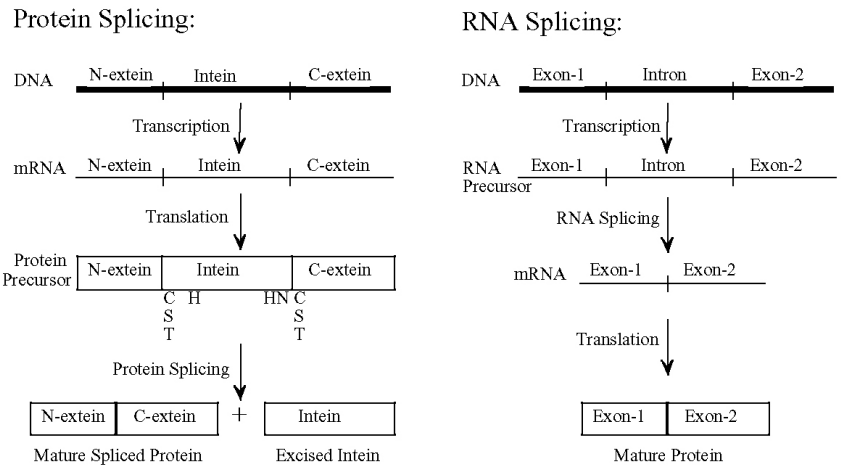
Annu Rev Biochem. 2000;69:447-96. Protein splicing and related forms of protein autoprocessing. Paulus H.

First Discovered

Science. 1990 Nov 2;250(4981):651-7.
Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase.
Kane PM, Yamashiro CT, Wolczyk DF, Neff N, Goebl M, Stevens TH.

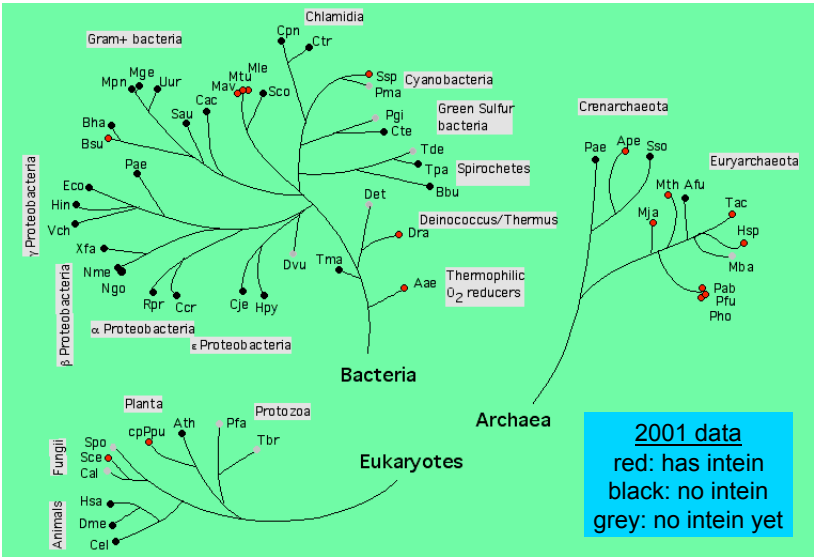
J Biol Chem. 1990 Apr 25;265(12):6726-33.
Molecular structure of a gene, VMA1, encoding the catalytic subunit of H (+)-translocating adenosine triphosphatase from vacuolar membranes of Saccharomyces cerevisiae.
Hirata R, Ohsumk Y, Nakano A, Kawasaki H, Suzuki K, Anraku Y.

Analogy to RNA Splicing



InBase at www.neb.com

Where are they found?



http://bioinformatics.weizmann.ac.il/~pietro/inteins/

Properties

1. Protein splicing is catalyzed entirely by amino acid residues contained in the intein.

3. Protein splicing is an intramolecular process (usually).

3. Protein splicing requires no coenzymes or sources of metabolic energy and therefore involves bond rearrangements rather than bond cleavage followed by resynthesis.

Annu Rev Biochem. 2000;69:447-96. Protein splicing and related forms of protein autoprocessing. Paulus H.

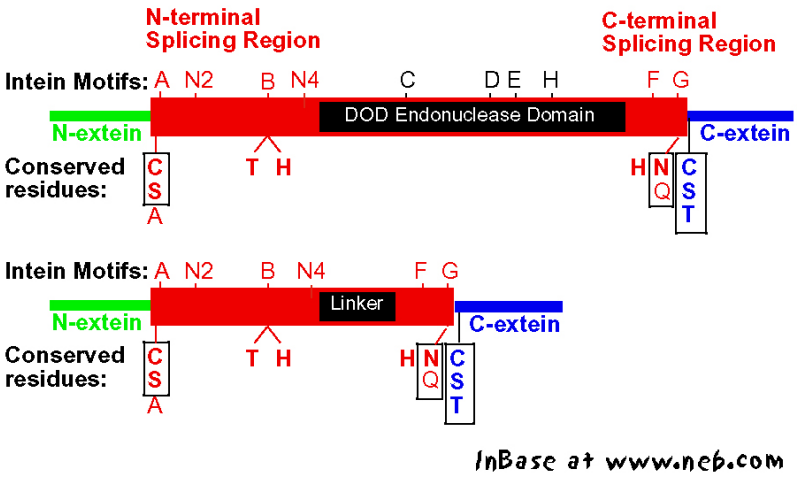
Totals: 216 inteins
46 types of protein family hosts

99 different species and strains.
24 eukaryotes (20 in unicellular organisms, 3 in plastids and 1 in viruses),
50 bacteria (4 in bacterio- and pro- phages),
25 archaea.

187 separate proteins with inteins:
150 with a single intein,
14 pairs each with one split intein,
18 with 2 inteins,
3 with 3 inteins,
1 with 4 inteins, and
1 pair: one with two and half inteins and the other with half of an intein
(the two halves being parts of a split intein)

http://bioinformatics.weizmann.ac.il/~pietro/inteins/ (2004 data)

What do they look like?

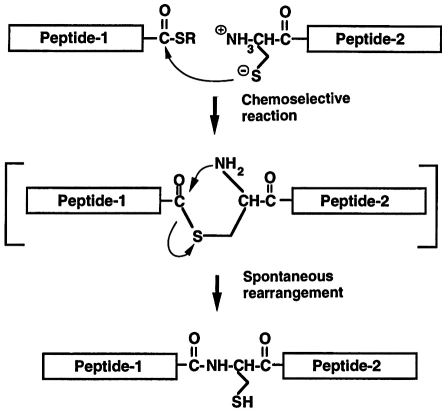


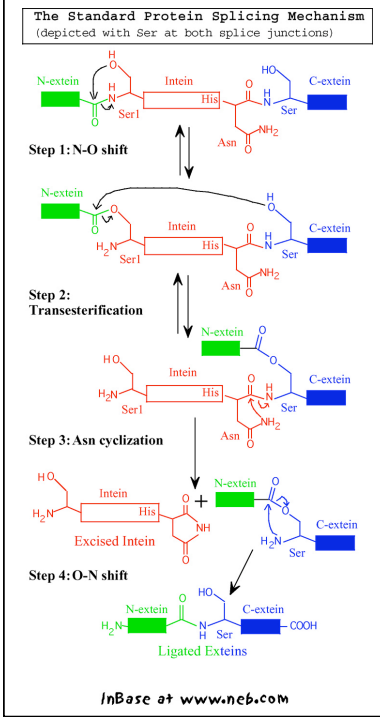
Small inteins are about 150 amino acids.
(the smallest is 134 amino acids, largest is 1650)

What kind of chemistry is involved?

Dawson PE, Muir TW, Clark-Lewis I, Kent SB.
Synthesis of proteins by native chemical ligation.
Science. 1994 Nov 4;266(5186):776-9.

Fig. 1. The principle of native chemical ligation. The synthetic segment, peptide 1, which contains a thioester at the α-carboxyl group, undergoes nucleophilic attack by the side chain of the Cys residue at the amino terminal of peptide 2 (R is an alkyl group). The initial thioester ligation product undergoes rapid intramolecular reaction because of the favorable geometric arrangement [involving a five-membered ring] of the α-amino group of peptide 2, to yield a product with a native peptide bond at the ligation site. Both reacting peptide segments are in completely unprotected form, and the target peptide is obtained in final form without further manipulation.





Step 1: formation of a linear ester intermediate by NO or NS acyl rearrangement involving the nucleophilic amino acid residue at the N-terminal splice junction;

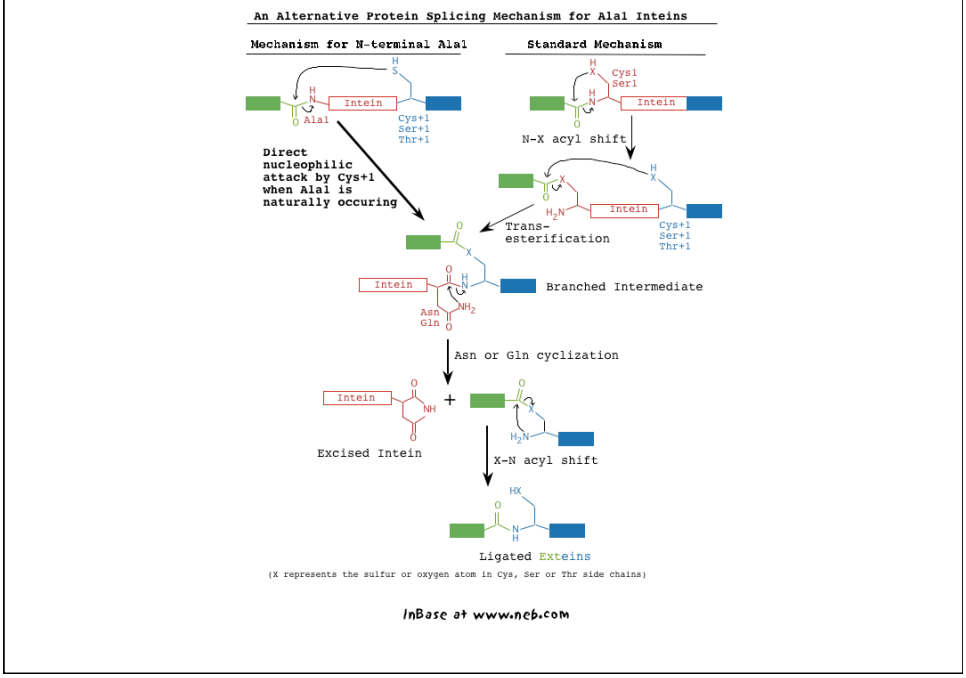
Step 2: formation of a branched ester intermediate by the attack of the nucleophilic residue at the C-terminal splice junction on the linear ester intermediate;

Step 3: cyclization of the asparagine residue adjacent to the C-terminal splice junction, coupled to cleavage of the branched ester intermediate to yield an excised intein with a C-terminal aminosuccinimide residue and the two exteins joined by an ester bond;

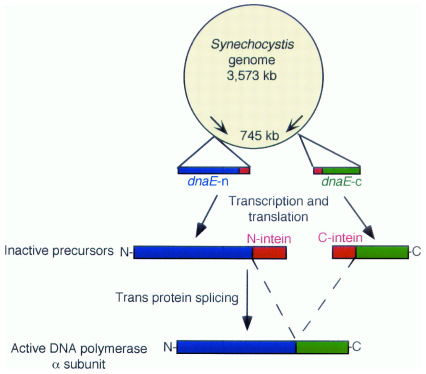
Step 4: spontaneous hydrolysis of the aminosuccinimide residue and rearrangement of the ester linking the exteins to the more stable amide bond.

The Asn cyclization is spontaneous and irreversible. The first two steps are catalyzed by the intein

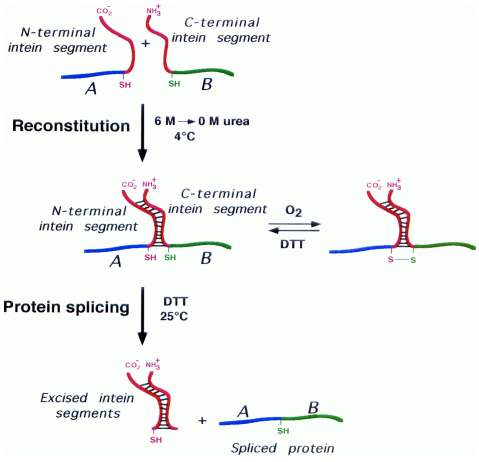
Annu Rev Biochem. 2000;69:447-96.
Protein splicing and related forms of protein autoprocessing.
Paulus H.



Intein-mediated Protein Splicing in *trans*



naturally occurring

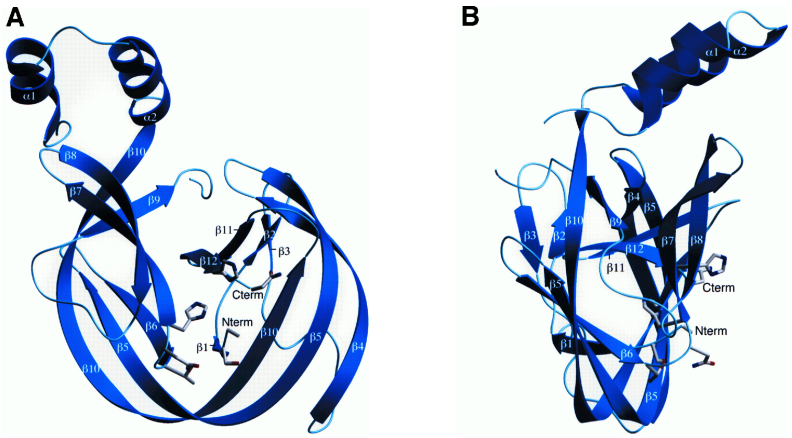


artificially created

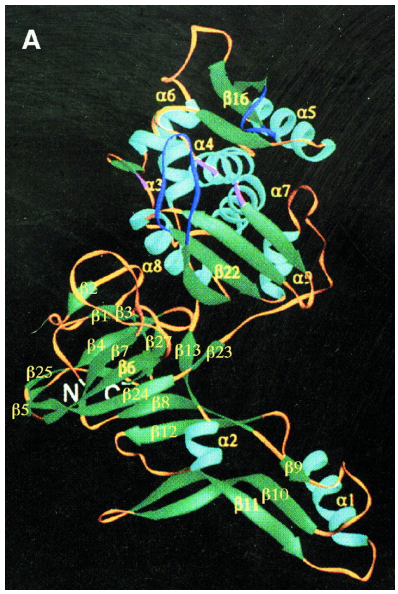
Proc Natl Acad Sci U S A. 1998 Aug 4;95(16):9226-31.
Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803.
Wu H, Hu Z, Liu XQ.

J Biol Chem. 1998 Jun 26;273(26):15887-90.
Protein splicing in vitro with a semisynthetic two-component minimal intein.
Lew BM, Mills KV, Paulus H.

Intein 3D Structure



Nat Struct Biol. 1998 Jan;5(1):31-6.
Crystal structure of GyrA intein from Mycobacterium xenopi reveals structural basis of protein splicing.
Klabunde T, Sharma S, Telenti A, Jacobs WR Jr, Sacchettini JC.

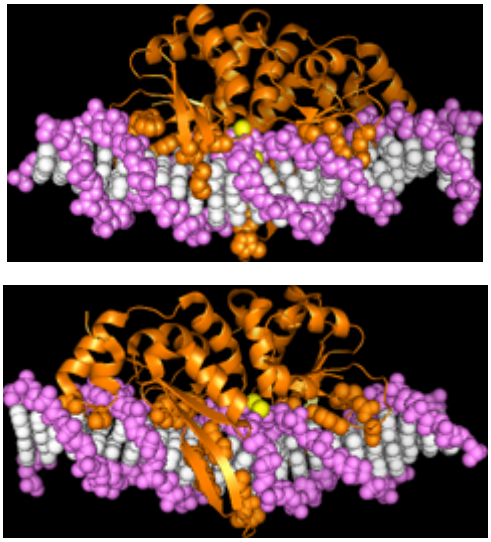


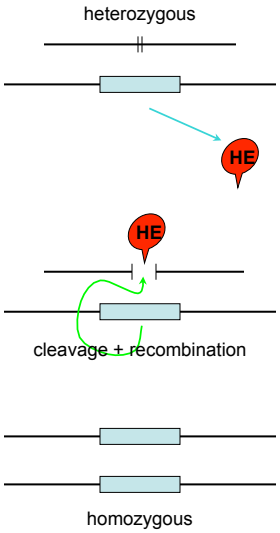
Intein

Homing Endonuclease

Cell. 1997 May 16;89(4):555-64.
Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity.
Duan X, Gimble FS, Quiocho FA.

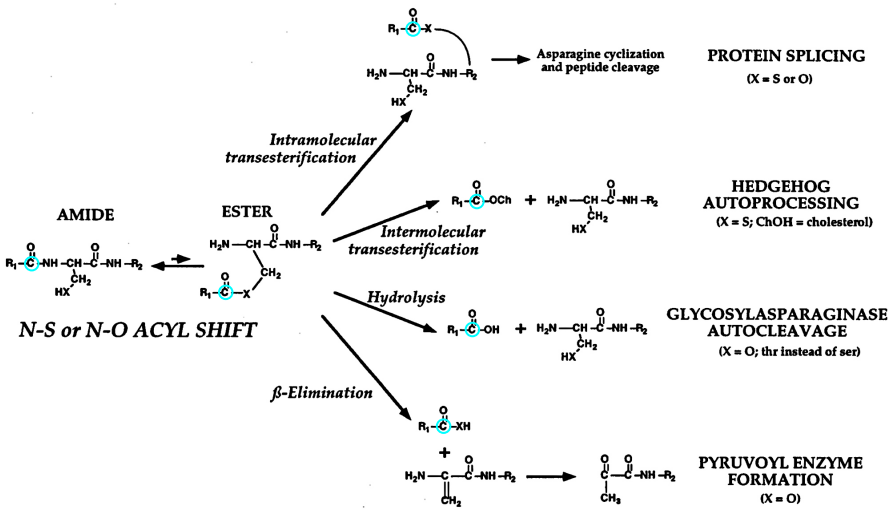
Intein-encoded Homing Endonucleases: Non-Mendelian Inheritance





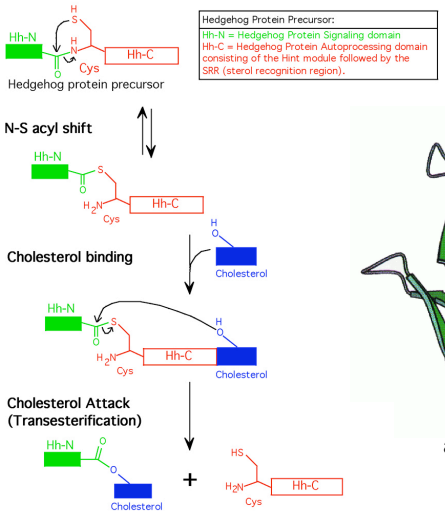
Moure CM, Gimble FS, Quiocho FA.
The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity.
J Mol Biol. 2003 Dec 5;334(4):685-95.

Related Biological Reactions



Annu Rev Biochem. 2000;69:447-96. Protein splicing and related forms of protein autoprocessing. Paulus H.

Hedgehog Protein Autoprocessing Mechanism

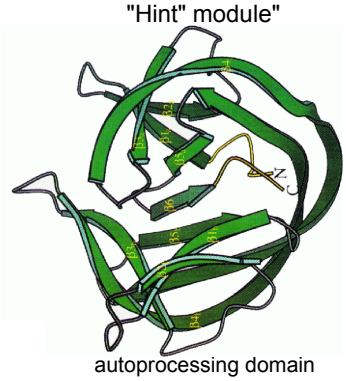


Cholesterol is covalently bound to the C-terminus of the Hedgehog protein signaling domain, resulting in attachment of the signaling domain to the cell surface.

Acids and bases assisting in catalysis are not shown.
See references by Porter, Beachy, and Leahy (Lee 1994, Porter 1995, Porter 1996, Hall 1997 & Beachy 1997).

InBase at www.neb.com

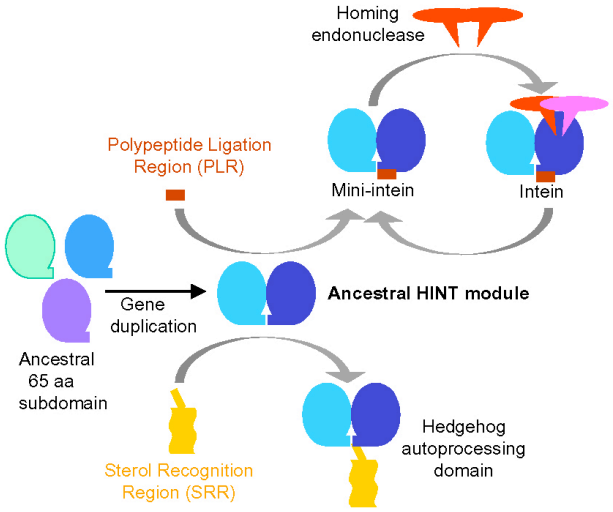
"Hint" module"



autoprocessing domain

Cell. 1997 Oct 3;91(1):85-97.
Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins.
Hall TM, Porter JA, Young KE, Koonin EV, Beachy PA, Leahy DJ.

Evolution of Inteins and Hedgehog Autoprocessing Domains



InBase at www.neb.com