

Peptid-"tagging" bei der Proteinsynthese fehlerhafter mRNA

(Jentsch, 1996; Keiler *et al.*, 1996)

In der Mayakultur herrscht die Vorstellung, daß die ersten Menschen nicht wohlgeraten waren. Daher haben die Götter Tzakól (der Schöpfer) und Bitól (der Former) die ersten Geschöpfe schnell zerstört.

Dies kann man als eine erste Qualitätskontrolle ansehen, die nicht nur auf der makroskopischen Ebene, sondern auch im molekularen Bereich der Stoffwechselfvorgänge innerhalb der Zelle eine wichtige Rolle spielen.

Bei der **DNA Replikation** gibt es eine Korrekturkontrolle, bei der fehlerhaft eingebaute Nukleotide in den Tochterstrang durch ein sogenanntes '**prove reading**' und Reperatursystem korrigiert werden.

Seit langem diskutiert man die Frage, ob nicht auch die, während einer Proteinsynthese auftreten Fehler, erkannt und korrigiert werden können.

Bis vor etwa 2 Jahren blieb diese Frage unbeantwortet. 1994/95 tauchten erste Berichte mit konkreten Anhaltspunkten für eine Qualitätskontrolle bei der Translation. Fehlerhafte Proteine können entweder

- a) zerstört oder
- b) der Fehler geheilt werden.

Sie werden entweder durch **proteolytische Enzyme elimiert** oder durch **molekulare Chaperone** in einem faltungsaktiven Zustand gehalten.

Fehlerhafte Polypeptide, die jedoch von einer **defekten mRNA** synthetisiert werden, sind **irreparabel**; die Zerstörung ist die einzige Möglichkeit. Keiler und Mitarbeiter beschrieben 1996 einen neuen **Qualitätskontrollmechanismus** für bakterielle Proteine.

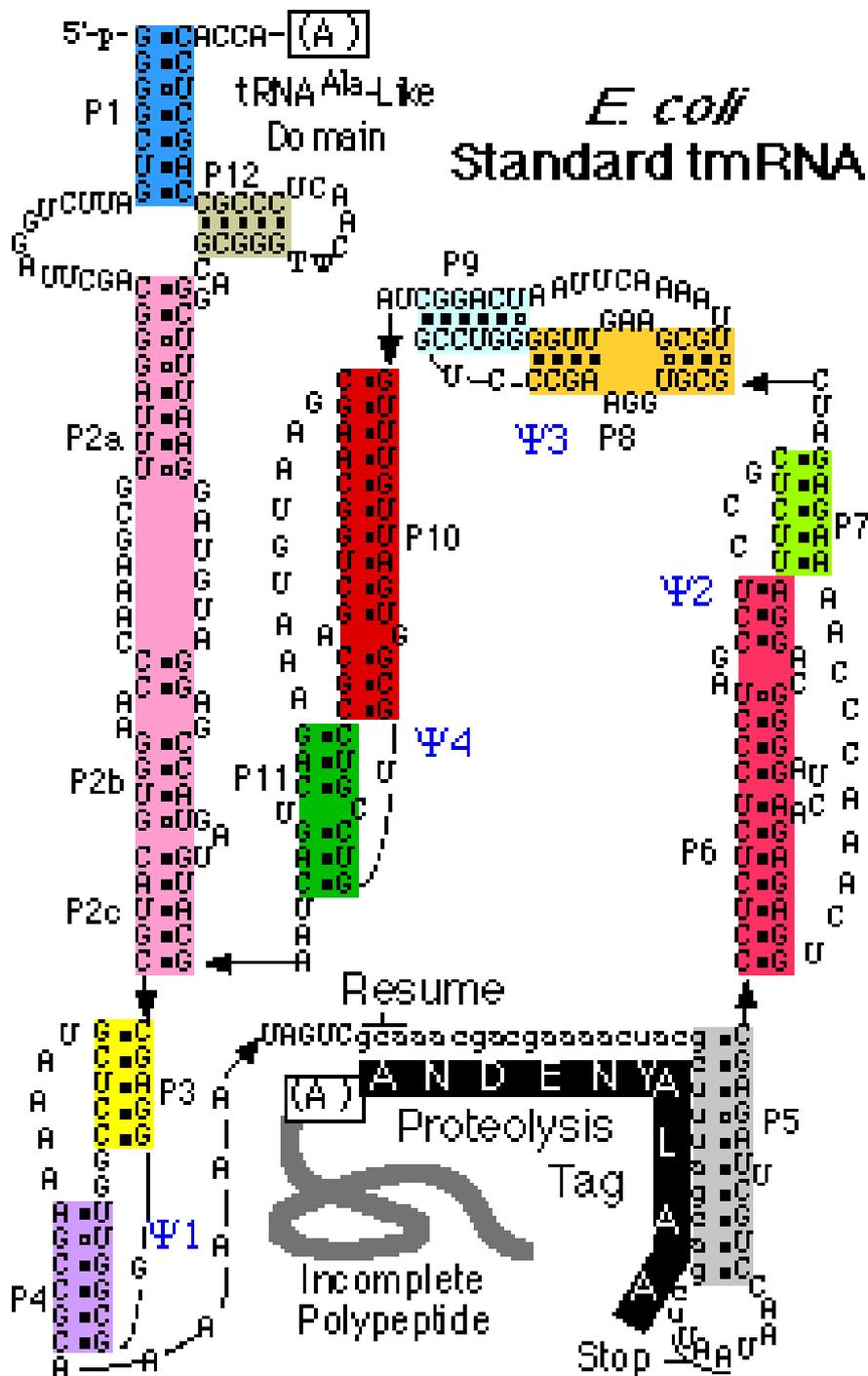
Sie entdeckten einen bislang völlig unbekanntem **ribosomalen Translationsmechanismus** bei dem naszierende Polypeptide, die von einer verkürzten fehlerhaften mRNA translatiert werden, durch eine **Peptidsequenz gekennzeichnet** werden, die wiederum **Erkennungssequenz für eine Protease** ist. Damit werden diese fehlerhaften Peptide dem **gezielten proteolytischen Abbau** zugeführt.

Diese Geschichte kam 1995 durch Tu und Mitarbeiter (Tu *et al.*, 1995) ins Rollen. Sie berichteten, daß ein Interleukin-6 Protein, das in *E. coli* exprimiert wurde, eine neue C-terminale Modifikation erhält. Die Proteinvarianten waren in unterschiedlicher Länge am C-terminalen Ende verkürzt, sie erhielten jedoch alle eine identische C-terminale Peptidsequenz ("tag"), deren Sequenz

AANDENYALAA

ist. Diese Sequenz ist **nicht in der IL-6 mRNA kodiert**. Diese neue Sequenz, mit Ausnahme des ersten Alaninrestes, ist auf einer kleinen (**362 nt**) **stabilen RNA** kodiert, die als **10Sa RNA** bekannt und von dem **ssrA Gen** kodiert wird.

Ein Teil dieser 10Sa RNA kann in eine tRNA ähnliche Struktur falten und mit Alanin beladen werden (Komine *et al.*, 1994).



The bacterial tmRNA is so-named for its dual tRNA-like and mRNA-like nature. (It's also known as 10Sa RNA or SsrA). tmRNA engages in a remarkable trans-translation process, adding a C-terminal peptide tag to the unfinished protein a stalled ribosome. The tmRNA-directed tag targets the unfinished protein for proteolysis.

Tu und Mitarbeiter haben daher vermutet, daß dieser Peptidschwanz von dem **Alanin** dieser 10Sa RNA und einem **Translationsprodukt von 10 Kodons**, ebenfalls von dieser 10Sa RNA herrührt. Keiler et al. (Keiler *et al.*, 1996) erkannten, daß dieser angefügte C-terminale Peptidschwanz ("tag") das Abbausignal für eine Schwanz-spezifische Protease, die als Tsp ("tail specific protease" oder Prc) bekannt ist.

Tsp wurde von Hara et al. (Hara *et al.*, 1991) beschrieben. Es ist im Periplasma von *E. coli* -Zellen lokalisiert und bindet gewisse apolare C-terminale Peptidschwänze von Substraten und spaltet die Proteine endeoproteolytisch (Parsell *et al.*, 1990). Im Cytoplasma gibt es ein Pendant zu dieser periplasmatischen Tsp Protease, die ebenfalls das Peptidende erkennt und die Proteine spaltet; dieses Enzym ist jedoch noch nicht charakterisiert.

Um die Hypothese des Abbausignals für eine Schwanz-spezifische Protease zu untermauern, wurden Transkriptionsterminatoren in ORFs eines periplasmatischen und eines cytosolischen Proteins inseriert (Keiler *et al.*, 1996). Dadurch kommt es zu verkürzten RNA Transkripten, die kein Stopkodon mehr aufweisen. Es stellte sich in der Tat heraus, daß die davon abgeleiteten Peptide am C-terminalen Ende, die oben erwähnte Schwanzsequenz, die durch die 10Sa RNA kodiert wird, besaßen. Das periplasmatische Fusionsprotein wurde durch Tsp abgebaut, und das cytosolische Protein durch den cytosolischen Gegenspieler.

Man muß sich fragen, was die **biologische Bedeutung** dieses ungewöhnlichen 'tagging'- Mechanismus ist. Durch die Aktivität von Nukleasen können mRNAs an ihrem 3' Ende abgebaut werden und damit das kodierte Stopkodon verlieren. Die Initiation der Translation dieser mRNA wird dadurch nicht beeinträchtigt, sie kommt jedoch zum Stillstand, wenn das Ribosom das 3' Ende der mRNA erhält.

In **Abwesenheit eines Stopkodons** kann der 'release'-Faktor (RF) die Dissoziation des naszierenden Polypeptids vom Ribosom nicht katalysieren. Alle Ribosomen, die in der Translation dieses Leserahmens involviert sind, bleiben an der mRNA arretiert. Eine Alanin geladene 10Sa RNA eilt zur Rettung herbei (**s. Abb.**). Die 10Sa RNA vereinigt sowohl tRNA-, als auch mRNA-Eigenschaften, so daß man diese RNA auch als **tmRNA** bezeichnen sollte. Keiler und Mitarbeiter argumentieren, daß die tRNA ähnliche Domäne es dem Ribosom ermöglicht, die nächste **Peptidyl-Transferreaktion** der naszierenden Kette auf die 10Sa RNA gebundenen Alaninrest zu katalysieren. **Die defekte mRNA kann dann freigesetzt** werden, und das **Ribosom gleitet auf den Leserahmen über**, der durch die 10Sa RNA bereitgestellt wird.

Dieser Vorgang, den man auch als **RNA-Hüpfen** bezeichnen könnte, erinnert an ein 'frameshifting' oder dem ribosomalen Hüpfen, das von John Atkins 1990 beschrieben wurde (Atkins *et al.*, 1990). Der wesentliche Unterschied dabei ist, daß bei dem Peptid-tagging das Ribosom auf eine neue RNA überspringt (*trans*). Wenn das Ribosom, das sich im Leserahmen befindliche **Stopkodon der 10Sa RNA erreicht**, kann das, mit dem neuen Peptidschwanz versehene Polypeptid schließlich freigesetzt werden.

Dieser neue **Translationstrick** ist besonders eindrucksvoll, da er 2 Fliegen mit einer Klappe schlägt:

1. das Ribosom wird von der defekten mRNA befreit, und
2. der Peptidschwanz führt das möglicherweise zerstörerische Peptid dem Abbau zu.

Tsp ist ein ATP unabhängige Protease, die nur eine begrenzte Fähigkeit besitzt, stabil gefaltete Proteine zu spalten (Keiler *et al.*, 1996). Unstrukturierte Polypeptide mit dieser Schwanzsequenz sind vermutlich gute Substrate für dieses Enzym. Auch native Proteine wie z.B. das Penicillinbindeprotein 3 von *E. coli* oder das Photosystem D1 Protein von Cyanobakterien werden durch Tsp und dem cytoplasmatischen Homolog zu Tsp gespalten. Beide Enzyme haben nativ an ihrem Schwanzende eine Tsp Erkennungssequenz. Tsp wie auch 10Sa RNA haben **Homologe**, in sowohl **Gram-positiven** als auch **Gram-negativen** Bakterien (Anbudurai *et al.*, 1994).

Abb. 1:

Markierung fehlerhafte bakterieller Proteine und deren gezielter proteolytischer Abbau (Qualitätskontrollmechanismus auf Translationsebene)

Schritt 1:

Bei einer mRNA mit fehlendem Stopkodon kommt die Translation zum Stillstand und die nicht fertige Polypeptidkette bleibt an die tRNA gebunden. Die Alanin beladene 10Sa RNA bindet an das Ribosom und durch die Peptidyltransferase-Reaktion wird die Polypeptidkette an das Alanin, das an die 10Sa RNA gebunden ist, übertragen.

Schritt 2:

Dadurch kann die fehlerhafte mRNA freigesetzt werden.

Schritt 3:

Das Ribosom gleitet auf die 10Sa RNA über und translatiert einen kurzen ORF, der auf der 10Sa RNA enthalten ist und der, die Peptidsequenz ANDENENYALAA beinhaltet.

Schritt 4:

Die Translation endet beim Stopkodon TAG, wo es zur Freisetzung der Polypeptidkette kommt.

Schritt 5:

Die so am C-terminalen Schwanz markierten Polypeptide werden von der cytosolischen und die periplasmatischen (Tsp) Schwanz-spezifischen Protease erkannt und abgebaut.

Literatur:

- Atkins, J. F., Weiss, R. B. and Gestland, R. F. (1990) *Cell* **62**: 413.
Hara, H., Yamamoto, Y., Higashitana, A., Suzuki, H. and Nishimura, Y. (1991) *J. Bacteriol.* **173**: 4799.
Jentsch, S. (1996) When proteins receive deadly messages at birth. *Science* **271**: 955-956.
Keiler, K. C., Waller, P. R. H. and Sauer, R. T. (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**: 990-993.
Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K. and Inokuchi, H. (1994) A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **91**: 9223-9227.
Parsell, D. A., Silber, K. R. and Sauer, R. T. (1990) *Genes Dev.* **4**: 277.
Tu, G.-F., Reid, G. E., Zhang, J.-G., Moritz, R. L. and Simpson, R. J. (1995) C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide. *J. Biol. Chem.* **270**: 9322-9326.

Science 1996 Feb 16;271(5251):990-3
Science. 1996 Feb 16;271(5251):955-6.
Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Keiler KC, Waller PR, Sauer RT.

Variants of lambda repressor and cytochrome b562 translated from messenger RNAs without stop codons were modified by carboxyl terminal addition of an **ssrA-encoded peptide tag** and **subsequently degraded by carboxyl terminal-specific proteases** present in both the cytoplasm and periplasm of Escherichia coli. The tag appears to be added to the carboxyl terminus of the nascent polypeptide chain by cotranslational switching of the ribosome from the damaged messenger RNA to ssrA RNA.

tm-RNA Sequence from *Staphylococcus aureus*

. 10 . 20 . 30 . 40 . 50

GGGGACGUUCAUGGAUUCGACAGGGGUCCCCCGAGCUCAUUAAGCGUGUC
GGAGGGUUGUCUUCGUCAUCAACACACACAGUUUAUAAUAACU**ggcaau**
caacaauaauuucgagucgucUAAUUCGCACUCUGCAUCGCCUAAC
AGCAUUUCCUAUGUGCUGUUAACGCGAUUCAACCUUAAUAGGAUAUGCUA
AACACUGCCGUUUGAAGUCUGUUUAGAAGAAACUAAUCAAACUAGCAUC
AUGUUGGUUGUUUAUCACUUUCAUGAUGCGAAACCUAUCGAUAAACUAC
ACACGUAGAAAGAUGUGUAUCAGGACCUUUGGACGCGGGUUCAAAUCCCG
CCGUCUCCAcca - 3'

Predicted Proteolysis Tag: (A)GKSNNNFVAA*

Taxonomy: Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
Staphylococcaceae; *Staphylococcus aureus*; stains COL, NCTC8325, and
NCTC6571 NOTE:

Possible intraspecies variation; the equally reliable version 2 has the differing bases:
10U, 11C, 242G, 286U, 288U, 329C.

Sequencing: strain COL, nt 72741-73099 of contig 6206 from 6-Jul-00 release by
The Institute for Genomic Research.

Prokaryontische Introns (Belfort *et al.*, 1995 J.Bact. 177:3897)

Einleitung

Als wichtige **Unterscheidungsmerkmale** zwischen Pro- und Eukaryonten wurde bei den Eukaryonten der Besitz einer **Kernmembran** und der Anwesenheit von **Intron** und des **RNA-Spleißen** angesehen. Dieses Dogma wurde gründlich erschüttert, als man auch bei Prokaryonten Introns in bestimmten Genen entdeckte.

1984 fanden (Chu *et al.*, 1984) ein Intron, in dem Thymidylat Synthasegen des Bakteriophagen T4; in den folgenden Jahren wurden weitere Introns in bestimmten Genen bei Eubakterien und den *Archaea* entdeckt.

Das Dogma brach dann vollständig zusammen, als man bei den Bakterien auch **autokatalytische Spleiß ('splice')- Vorgänge** bei den **Gruppe I Intron** entdeckte. Ein weiterer Durchbruch war die Entdeckung eines **Kern-pre-mRNA ähnlichen Introns** in einem bakteriellen Plasmid bei *Agrobacterium rhizogenes* (Magrelli *et al.*, 1994).

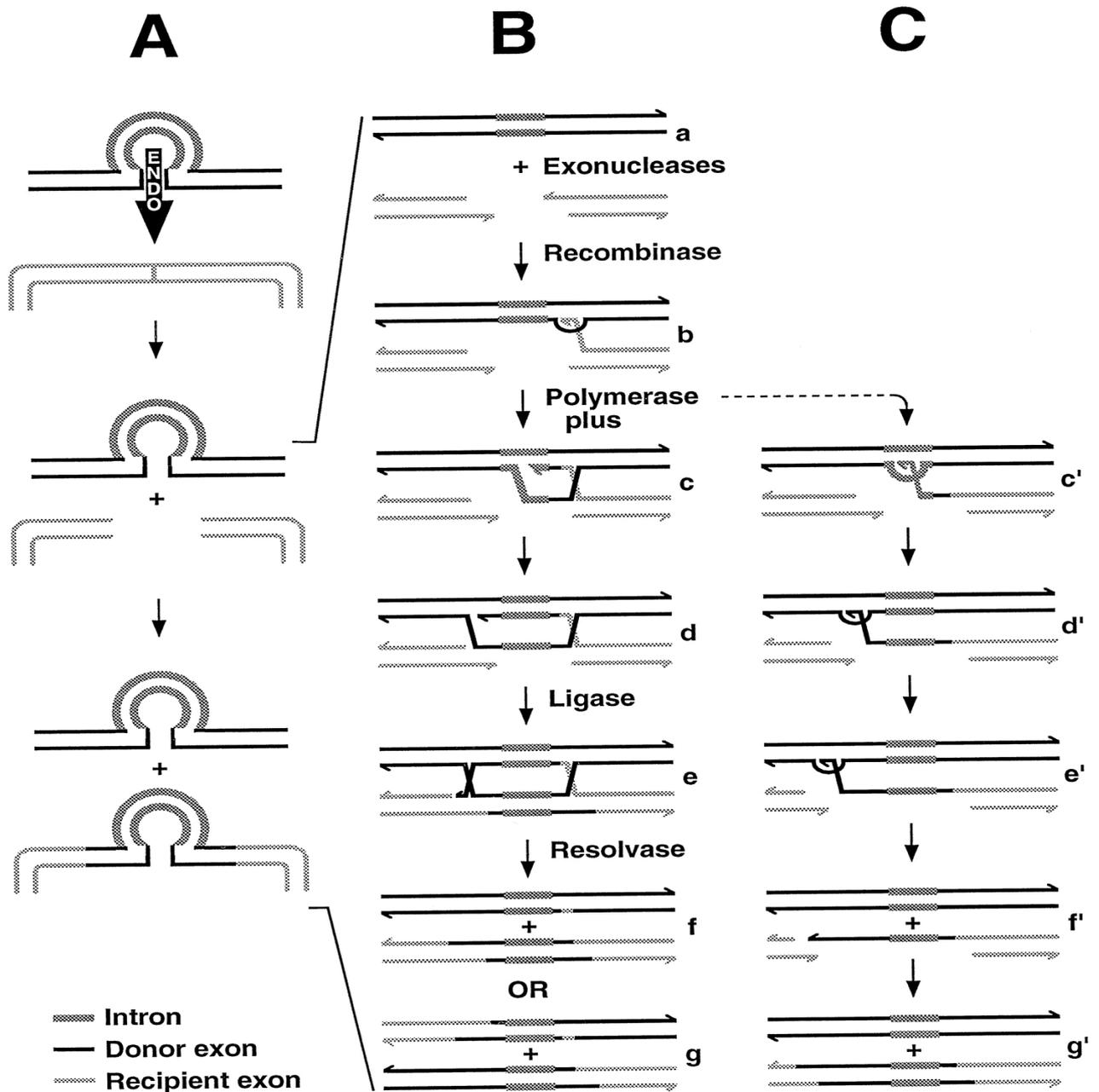
Im folgenden seien hier das Vorkommen und die charakteristischen Eigenschaften der verschiedenen Introngruppen bei den Prokaryonten beschrieben. Kurz zur **Nomenklatur**:

Introns sind **intervenierende Sequenzen (IVS, engl.: 'intervening sequences')** innerhalb eines Gens, die auf Transkriptionsebene durch einen Vorgang, der als **Spleißen** (engl.: 'splicing') bezeichnet wird, entfernt werden. Der Spleiß-Vorgang umfaßt einen **RNA-Bruch** an den beiden Enden des Introns und die **Religation** der **Exon-RNA-Moleküle**, die letztlich das reife Transkript darstellen.

Gruppe I Introns

Das erste Intron, das bei Prokaryonten entdeckt wurde, war ein Gruppe I Intron, das in dem Thymidylat-Synthasegen des *E. coli* Bakteriophagen T4 gefunden wurde (Chu *et al.*, 1984). In den folgenden Jahren wurden Gruppe I Introns auch in anderen T4-Coliphagen-Genen gefunden (Gott *et al.*, 1986), als auch in *B. subtilis* Phagen Genen (Bechhofer *et al.*, 1994; Goodrich-Blair *et al.*, 1990) und in tRNA- Genen von Cyanobakterien und Proteobakterien (Khusel *et al.*, 1990; Reinhold-Hurek and Shub, 1992; Xu *et al.*, 1990). Auch in den Genomen von **Chloroplasten** und **Mitochondrien**, welche endosymbiontische Abkömmlinge von Cyanobakterien und Proteobakterien repräsentieren, konnten Gruppe I Introns gefunden werden. Als Basis für die Einteilung in die Gruppe I Introns dient der gut untersuchte Spleiß-Mechanismus von *Tetrahymena thermophila* (Chech, 1993). Ein Charakteristikum der Gruppe I Introns ist der **autokatalytische Spleiß-Mechanismus** der *in vitro* in Abwesenheit von Proteinen erfolgen kann.

Abb. 1A: Gruppe I Introns (hier nicht gezeigt s. Lehrbücher der Biochemie)



DNA-mediated group I intron homing. A, overview of mobility pathway. Cleavage of the recipient by the endonuclease (ENDO) results in intron inheritance via gene conversion. B, the DSBR pathway. Subsequent to endonuclease cleavage the recipient allele undergoes exonucleolytic degradation and homologous sequence alignment with an intron-containing donor (a). A 3' - end of the recipient invades the donor, which serves as a template for repair synthesis (b). During DSBR (c-g), DNA synthesis through the intron results in expansion of a D-loop (c), which then serves as a substrate for repair synthesis of the non-invading strand (d). Holliday junctions formed during this process (e) are resolved to produce either non-crossover (f) or crossover (g) products. Activities involved in the T4 intron homing pathway are indicated, with

Polymerase plus signifying the requirements for polymerase accessory functions (3) ; presumably similar activities participate in homing events in other systems. C, the SDSA (BM) model (6, 7) . As DNA synthesis proceeds, the replication bubble migrates with the replicative end (c`-e`). The newly synthesized strand is released from the donor (f`) and serves as template for repair of the noninvading strand (f`-g`) to generate non-crossover products only (g`). Half-arrows indicate 3`-ends of DNA strands.

From:

Belfort et al. (1995) Procaryotic introns and inteins: a panoply of form and function. J. Bacteriol. 177: 3897-3903.

Schritt 1:

nukleophiler Angriff der 3' OH-Gruppe eines exogenen GTP (α - ^{32}P -GTP) auf die 5' Splice-Stelle;

Schritt 2:

nukleophiler Angriff der 3' OH-Gruppe, der an das 5' Exon mit anschließen der Ligation der beiden Exone und Freisetzung des Introns;

Schritt 3:

unter Umständen kommt es zur Transesterifikationsreaktion zwischen der 3' OH-Gruppe des Introns, wobei an einer internen Stelle es zur Zyklisierung kommen kann.

Die Exons werden durch offene Balken und die Introns durch eine fette Linie repräsentiert. Die Sekundärstruktur des Intron mit den 9 Haarnadelschleifen ist darunter gezeichnet, die Pfeile geben die Splice-Stelle an.

Wie in **Abb.1A** gezeigt, werden die Gruppe I Introns, ausgehend von einer Precursor-RNA durch eine Reihe Transesterifikations-Reaktionen gespleißt, die durch einen exogenen Guanosin-Kofaktor (GTP) initiiert werden. Die resultierenden Produkte dieses Spleiß-Vorgangs sind ligierte Exons und ein freies Intron, das an seinem 5' Ende das initiiierende Guanosin trägt. Die Anheftung des initiiierenden Guanosinrestes kann für den Nachweis der Gruppe I Introns herangezogen werden. Bakterielle RNA-Extrakte werden mit [α - ^{32}P]GTP inkubiert und dann auf radioaktiv markierte Introns getestet (Gott *et al.*, 1986).

Wie die Tabelle zeigt besitzen viele Bakteriophagen-Gene selbst-spleißende Gruppe I Introns.

Es scheint jedoch, daß unter *in vitro* Bedingungen, bestimmte Proteine, die RNA in unspezifischer Weise binden können, wie z.B. das *E. coli* ribosomale **Protein S12**, eine Chaperon ähnliche Aktivität besitzen, die die korrekte Faltung der RNA erleichtert. Gruppe I Introns in Pilz-Mitochondrien benötigen ein zusätzliches Protein für den Spleiß-Vorgang; einige Proteine werden von den Kerngenen kontrolliert andere (Maturase) sind im Intron selbst kodiert.

Sequenzvergleiche der verschiedenen Gruppe I Introns weisen keine ausgedehnte Sequenzhomologie auf. Die RNA besitzt lediglich eine **strukturelle Ähnlichkeit** insofern, als sie eine Reihe von kurzen und langen **Haarnadelschleifen** besitzt, die

in der Abb.1A als **P1-P9** bezeichnet werden. Sie stellen die einzige Grundlage für eine konservierte katalytisch aktive Tertiärstruktur dar. Hinzu kommt, daß viele Gruppe I Introns **ORFs** besitzen, die aus diesen konservierten Strukturen herausragen. Häufig kodieren diese ORFs für **Endonukleasen**, die eine Rolle bei der **Intronmobilität** spielen, auf die später noch eingegangen werden soll.

Gruppe II Introns

Die Gruppe II Introns wurden kürzlich in Cyanobakterien (*Calothrix* spp.) und Proteobakterien (*Azotobacter vinelandii*), den postulierten Vorläufern von Chloroplasten und Mitochondrien (Ferat and Michel, 1993) und natürlich auch in *E. coli* gefunden. Diese Entdeckungen waren vor allem aus evolutionärer Sicht besonders aufregend, da man bislang Gruppe II Introns nur in **Chloroplasten** und **Mitochondrien** von **Pilzen** und **Pflanzen** gefunden hat, und ihr Spleiß-Mechanismus dem der Kern-pre-mRNA Introns entspricht. Es war allgemein akzeptiert, daß die Gruppe II Introns als die **Vorfahren der Kern-pre-mRNA Introns** galten (Sharp, 1991).

Abb. 1B: Gruppe II Introns

Schritt 1:

nukleophiler Angriff der 2'OH-Gruppe des Adenosins nahe des 3' Endes des Introns an die 5'Spleiß-Stelle, wobei es zur Ausbildung einer 2'-5'-Bindung kommt;

Schritt 2: die 3'OH-Gruppe des 5'Exons reagiert mit der 3'Splice-Stelle, wobei es zur Ligation der beiden Exone unter Freisetzung des Intron-Lassos kommt. Die Sekundärstruktur des Introns mit den konservierten Basenpaaren I-VI und mit dem Adeninrest in Helix 6 ist darunter gezeigt.

Wie beim Kern-pre-mRNA Spleißen beginnt der Spleiß-Vorgang der Gruppe II Introns mit dem Angriff eines herausragenden 2' OH-Nukleotids auf die 5' Spleiß-Stelle. Das Resultat ist die Ausbildung einer 2'-5' Bindung und der damit einhergehenden Lassostruktur. Anschließend attackiert die endständige OH-Gruppe des 5' Exons die 3' Spleiß-Stelle. Das Resultat hier ist die Ligation der Exons und die Freisetzung des Intron Lassos.

Ähnlich wie die Gruppe I Introns besitzen auch die Gruppe II Introns eine konservierte Sekundärstruktur, die für das Spleißen essentiell ist. Diese Struktur besteht aus 6 konservierten Helices. Die Sequenzhomologien beschränken sich nur auf kurze Bereiche und auf einen herausragenden Adeninrest in Domäne 6, dem Verzweigungspunkt der Lassostruktur.

Obwohl einige Gruppe II Introns zum **autokatalytischen Spleißen befähigt** sind, tun sie das **nur unter nicht-physiologischen Bedingungen**. Die meisten benötigen **Protein-Cofaktoren**, einige davon sind im Intron selbst kodiert. Einige Intron ORFs kodieren für Maturasen, die den Spleiß-Vorgang erleichtern, oder aber auch als reverse Transkriptasen (RTs) bei der Mobilität der Gruppe II Introns eine Rolle spielen.

Eine Kern-pre-mRNA ähnliches Intron

Ein Ri-Plasmid-Gen aus *Agrobacterium rhizogenes* besitzt ein Intron in seiner nicht-translatierten Leitsequenz. Überraschenderweise besitzt dieses Intron die charakteristischen Eigenschaften eines pflanzlichen Kern-pre-mRNA Introns:

- a) das Intron besitzt am 5' Ende die charakteristischen Dinukleotide GT und am 3' Ende AG;
- b) der 3' Spleiß-Stelle geht eine AT reiche Region voraus;
- c) der AT Gehalt und die Homologie am 5' Exon/Intron-Schnittpunkt entspricht dem der pflanzlichen Konsensussequenz.

Wenn man dieses Gen in die Pflanze *Arabidopsis thaliana* transformiert, findet dort korrekte Transkription und Spleißen statt. Man kann daher annehmen, daß das **pflanzliche Spleißosom** die korrekte Prozessierung der bakteriellen pre-mRNA bewerkstelligen kann.

Introns in Archaea

Abb. 1C: Introns bei Archaea

Schritt 1:

endonukleolytische Spaltung am Schleifen-Helix-Motiv, das zur Ausbildung von 2' - 3' zyklischem Phosphat führt;

Schritt 2:

Exonligation mit möglicher Zirkularisierung der RNA.

Eine einzigartige Klasse von Introns, die keine Gemeinsamkeiten mit den bisher beschriebenen besitzt, hat man in den tRNA und rRNA-Genen von halophilen und hyperthermophilen *Archaea* gefunden. Diese Introns werden **nicht autokatalytisch prozessiert**, und sie besitzen nur ein halbwegs konserviertes Strukturmotiv im Intron/Exon-Grenzbereich (**Abb. 1C**). Dieses Motiv wird von einem **Enzym**, einer **RNA-Endoribonuklease** erkannt, das das Intron aus der Vorläufer-RNA herauschneidet. Als Ergebnis liegt dann ein 2'-3' zyklisches Phosphat und eine 5' Hydroxyl-Gruppe vor, was **charakteristisch für Kern-tRNA-Introns bei Eukaryonten** ist.

Eine bislang kaum erforschte Ligationsreaktion schafft eine 5'-3' Phosphordiesterbrücke, die die Exons verbindet. Verschiedene Archaea Introns besitzen ORFs, die für **Endonukleasen** kodieren. Diese Enzyme besitzen ein Sequenzmotiv.

LAGLIDADG

das auch bei den DNA-Endonukleasen, die bei eukaryontischen Gruppe I Introns kodiert werden, konserviert ist, und das auch bei Inteinien vorkommt.

Andere RNA Unterbrechungen

Die 23S und 16S rRNA Gene verschiedener Bakterienarten besitzen kleine (100-200 bp) IVSSs, welche aus der RNA herausgeschnitten werden, ohne daß es zu einer Exon Ligation kommt. Das Ergebnis sind fragmentierte rRNAs.

Es wurden in 23S rRNAs **auch IVS gefunden**, die **ORFs** besitzen, die 121-133 Aminosäuren kodieren (Ralph and McClelland, 1993). Diese IVSSs haben die

Fähigkeit, in ähnliche Haarnadelstrukturen zu falten und sie werden durch die **RNase III**, welche bevorzugt solche Duplexregionen erkennt, herausgeschnitten. Das gelegentliche Auftreten dieser Sequenzen innerhalb nah verwandter Organismen, ja sogar innerhalb ein und desselben Organismus und die Anwesenheit von ORFs in einigen, läßt vermuten, daß diese **IVS ebenfalls mobile Elemente** darstellen, die in rRNA Gene eingedrungen sind.

Ein weiterer Typ an IVS wurde im **Gen 60** des **Bakteriophagen T4** gefunden, das eine **Topoisomeraseuntereinheit** kodiert. Dieses **50 Nukleotid lange atypische IVS** wird weder gespleißt noch exzidiert. Stattdessen wird es während der Translation durch ein gigantisches '**frameshift**'-Ereignis übersprungen.

Für dieses **ribosomale Hüpfen** sind verschiedene Voraussetzungen notwendig:

1. ein *cis*- aktives 16 Aminosäure langes Segment des naszierenden Peptides,
2. die optimale Distanz zwischen Start- und Landekodon ist etwa 50 nts,
3. an der Startstelle befindet sich eine Haarnadelstruktur und
4. unmittelbar nach der Landestelle befindet sich ein Stopkodon.

Das ribosomale **Protein L9** scheint bei diesem Vorgang eine gewisse Funktion auszuüben (Herbst *et al.*, 1994).

Auch bei dem *E. coli* Gen das für den **trpR** kodiert, werden durch ein 'frameshift'-Ereignis 55nt übersprungen, was zur Ausbildung eines kleineren Produktes führt (Benhar and Engelberg, 1993).

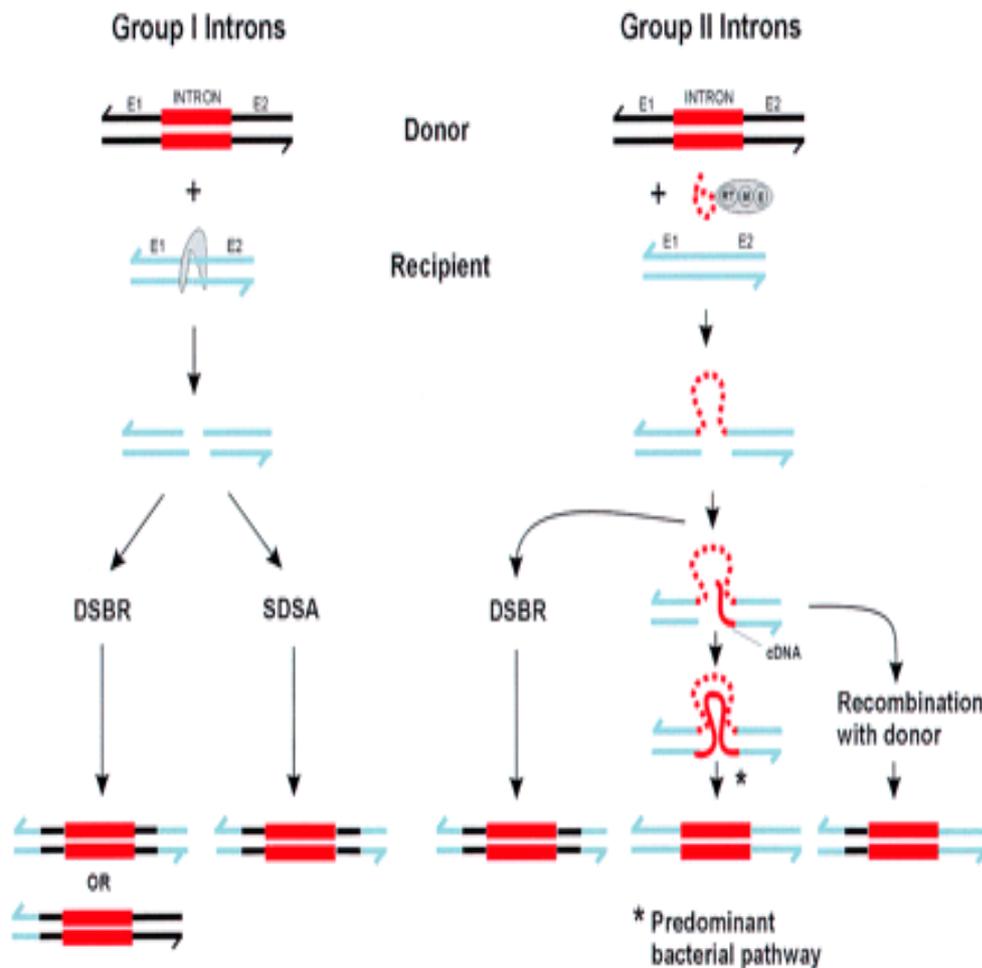
Literatur:

- Bechhofer, D. H., Hue, K. K. and Shub, D. A. (1994) An intron in the thymidylate synthase gene of *Bacillus* bacteriophage β 22: evidence for independent evolution of a gene, its group I intron, and the intron open reading frame. *Proc. Natl. Acad. Sci. USA* **91**: 11669-11673.
- Belfort, M., Reaban, M. E., Coetzee, T. and Dalgaard, J. Z. (1995) Procaryotic introns and inteins: a panoply of form and function. *J. Bacteriol.* **177**: 3897-3903.
- Benhar, I. and Engelberg, K., H. (1993) Framshifting in the expression of the *Escherichia coli* *trpR* gene occurs by the bypassing of a segment of its coding sequence. *Cell* **72**: 121-130.
- Chech, T. R. (1993) Structure and mechanism of the large catalytic RNAs: group I and group II introns and ribonuclease P. In *The RNA world*. Gesteland, R. F. and Atkins, J. F. (ed). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, pp. 239-269.
- Chu, F. K., Maley, G. F., Maley, F. and Belfort, M. (1984) Intervening sequences in the thymidylate synthase gene of bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **81**: 3049-3053.
- Ferat, J. and Michel, F. (1993) Group II self-splicing introns in bacteria. *Nature* **364**: 358-361.
- Goodrich-Blair, H., Scarlato, V., Gott, J. M., Xu, M. and Shub, D. A. (1990) A self-splicing group I intron in the DNA polymerase gene of *Bacillus subtilis* bacteriophage SPO1. *Cell* **63**: 417-424.
- Gott, J. M., Nichols, L. M., Gesteland, R. F. and Weiss, R. B. (1986) Multiple self-splicing introns in bacteriophage T4: evidence from autocatalytic GTP labeling of RNA *in vitro*. *Cell* **47**: 81-87.

- Herbst, K. L., Nichols, L. M., Gesteland, R. F. and Weiss, R. B. (1994) A mutation in ribosomal protein L9 affects ribosomal hopping during translation of gene 60 from bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **91**: 12525-12529.
- Khusel, M. G., Strickland, R. and Palmer, J. D. (1990) An ancient group I intron shared by eubacteria and chloroplasts. *Science* **250**: 1570-1573.
- Magrelli, A. K., Langenkemper, K., Dehio, C., Schell, J. and Spena, A. (1994) Splicing of the *rolA* transcript of *Agrobacterium rhizogenes* in *Arabidopsis*. *Science* **266**: 1986-1988.
- Ralph, D. and McClelland, M. (1993) Intervening sequences with conserved open reading frame in eubacterial 23S tRNA genes. *Proc. Natl. Acad. Sci. USA* **90**: 6864-6868.
- Reinhold-Hurek, B. and Shub, D. (1992) Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature* **357**: 173-176.
- Sharp, P. A. (1991) Five easy pieces. *Science* **254**: 663.
- Xu, M.-Q., Kathe, S. D., Goodrich-Blair, H., Nierzwicki-Bauer, S. A. and Shub, D. A. (1990) Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in *Cynobacteria*. *Science* **250**: 1566-1570.

Tabelle: Vorkommen von Introns in Bakterien und Organellen

| Gen | Vorkommen | Jahr |
|--------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------|-------------|
| Gruppe I | | |
| Thymidylat-Syntase | ΦT4 (<i>E. coli</i>) | 1984 |
| andere Gene | ΦT4 (<i>E. coli</i>) Nachweis durch α- ³² P-GTP-Markierung des Introns | 1986 |
| DNA-Polymerase | ΦSPO1 (<i>B. subtilis</i>) | 1990 |
| Thymidylat-Syntase | β22 (<i>B. subtilis</i>) | 1994 |
| t-RNA-Gene | Cyanobakterien / Chloroplasten | 1990 |
| t-RNA-Gene | viele Bakterien | 1992 |
| Gruppe II (Vorstufen der Kern Pre-mRNA-Introns) | | |
| rolA (spielt Rolle bei Haarwurzel-Infektion) | Ri plasmid A4 <i>Agrobacterium rhizogenum</i> (rolA wird auch in Arabidopsis gespleißt!) | 1994 |
| viele Gene | <i>E. coli</i> | 1994 |
| Archaea Introns | | |
| t und rRNA | <i>Halobacterium volcanii</i> (u.a. Hyperthermophile) (einige <i>Archaea</i> -Introns besitzen ORFs, Endoribonuclease) | 1985/91 |
| Andere RNA-Unterbrechungen | | |
| IVS in 23S / 16S-rRNA-Gene | viele Bakterien RNaseIII | |



Schematic of group I and II intron homing pathways. Solid black lines represent upstream (E1) and downstream (E2) exons of donor (intron-containing) alleles, and blue-gray lines indicate recipient (intronless) alleles. Group I or II introns are depicted by red boxes. Broken red lines indicate intron RNA, while solid red lines indicate cDNA synthesized by the reverse transcriptase (RT) function of group II intron proteins. The gray tooth-shaped structure represents a group I intron endonuclease. Group II intron proteins possess two functions in addition to RT, namely, maturase (M) and endonuclease (E) activities.

According to:

Edgell, D. R., M. Belfort, and D. A. Shub. 2000. Barriers to intron promiscuity in bacteria. *J Bacteriol* **182**:5281-9.

Inteine - Proteinspleißen

Einführung

to splice = spleißen, die Enden zweier Taue verflechten

Proteinspleißen ist ein neu entdeckter Vorgang, der auf Proteinebene das Äquivalent zum RNA-spleißen darstellt. Das Proteinspleißen erfolgt über ein verzweigtes Proteinintermediat. *In vitro* Untersuchungen lassen vermuten, daß die Reaktion **autokatalytisch** abläuft.

Die ausgeschnittenen **Inteinproteine** sind **ortsspezifische DNA-Endonukleasen**, die ihre eigene kodierende DNA-Sequenz in eine andere homologe Sequenz einbauen können, ein Vorgang, der im engl. als '**intein homing**' bezeichnet wird. Diese **post-translationale Prozessierung** umfaßt:

- a) die präzise Exzision einer intervenierenden Proteinsequenz aus dem Precursorprotein, gekoppelt mit
- b) einer Peptidbindung zwischen den flankierenden N- und C-terminalen Proteindomänen.

Das Resultat ist ein verknüpftes Proteinprodukt (**Abb. 1**). In Analogie zur RNA-spleißen-**Nomenklatur** wird der Vorgang als Proteinspleißen bezeichnet. Das **Intein** (interne Proteinsequenz) stellt eine Proteinsequenz dar, die im Leserahmen innerhalb eines Precursorproteins eingebettet ist und durch Proteinspleißen ausgeschnitten (exzidiert) wird.

Das **Extein** (bestehend aus den externen Proteinsequenzen), setzt sich aus den, das Intein flankierenden Polypeptidketten zusammen, die ebenfalls durch Spleißen verknüpft wurden.

Mit der Entdeckung von Inteinen wurde die Vielfalt der intervenierenden Sequenzen bereichert. Inteine kommen sowohl bei **Eukaryonten** (Kane et al. 1990), als auch bei **Archaea** (Perler et al. 1992; Xu et al. 1993) und **Eubakterien** (Davis et al. 1992 und 1994) vor.

Die Entdeckung des Proteinspleißens

Das erste Beispiel eines Proteinspleißens wurde bei dem *Saccharomyces cerevisiae* **VMA1** Gen entdeckt. Dieses Gen kodiert für die 69 kDa katalytische Untereinheit der vakuolen Protonen ATPase (V-ATPase). Ein durchgehender ORF läßt auf ein Protein von 119 kDa schließen, wobei die N- und C-terminale Aminosäuresequenz 75% Identität mit der homologen V-ATPase Untereinheit von *Neurospora crassa* aufweist.

Diese Homologie wird durch eine **454 Kodon** umfassende Sequenz unterbrochen. Diese Sequenz, als **Intein** bezeichnet, weist keine Homologie zu bekannten ATPase Untereinheiten auf. Schneidet man auf DNA Ebene die Intein kodierende Sequenz

heraus, so wird eine voll aktive V-ATPase Untereinheit in der korrekten Größe exprimiert. Dieses Ergebnis zeigt, daß das Intein in dem gespleißten Protein nicht enthalten ist.

Die Beteiligung eines RNA-spleißens konnte man durch Northern-blot Analysen ausschließen. Mutationen (Punkt- und Deletionsmutationen) innerhalb des Inteins führen zur Inaktivierung des Proteinspleißens; was für eine Beteiligung des Inteins am Spleißvorgang spricht.

Quantitative Untersuchungen ergaben, daß das **50 kDa große Intein** und das 69 kDa gespleißte V-ATPase Protein in **äquimolaren** Verhältnissen gebildet werden. Diese Ergebnisse stehen in Übereinstimmung mit der post-translationalen Prozessierung in 2 Proteine.

Tabelle: Vorkommen von Inteinen bei Bakterien und Eukaryoten

| Gen | Intein Vorkommen (Aminosäuren) | ortsspez. Endonuklease | |
|-------------------------------------------|-----------------------------------|-----------------------------------------------------------|--------|
| <i>TFP 1</i> Gen 69 kDa UE V-ATPase | 454 | <i>Saccharomyces c.</i> | + |
| <i>recA</i> | 440 | <i>Mycobacterium</i> | ? |
| <i>dna-pol</i> | 2 Inteine 538/390 | <i>Thermococcus litoralis</i> (Arch) <i>Pyrococcus</i> | + + |
| <i>dna-pol</i> | 537 | <i>Pseudomonas sp.</i> | ? |

Cooper / Stevens: TIBS **20**: 351-356 (1995)

Seit der ursprünglichen Entdeckung des Proteinspleißens (Kane et al 1990; Hirata et al 1990) bei Hefen wurden Inteine bei einer Reihe von Genen in Hefen, Bakterien und *Archaea* gefunden.

Die RecA Gene von *Mycobacterium tuberculosis* und *Mycobacterium leprae* beinhalten Intein kodierende Sequenzen, die durch Proteinspleißens entfernt wurden.

Gene, die für die DNA Polymerasen bei den extrem thermophilen *Archaea*-Arten *Thermococcus litoralis* und *Pyrococcus* kodieren, besitzen ebenfalls Intein kodierende Sequenzen, und diese Inteine werden auf Proteinebene ausgeschnitten. Diese DNA Polymerase Gene bei den *Archaea* zeichnen sich durch den Besitz **zweier Inteine** aus

Abb.1:

- a) Proteinspleißens des 119 kDa TFP 1 Precursor Proteins in ein 69 kDa V-ATPase Untereinheit (N-Extein-C-Extein) und eine 50 kDa große Endonuklease (Intein).

An den Nahtstellen zwischen Extein und Intein sind die Aminosäuresequenzen gezeigt. Die beiden Pfeilköpfe zeigen die Spaltstelle.

- b) Zusammenstellung der bekannten oder postulierten Inteine mit den hoch konservierten Aminosäureresten (fettgedruckt) an den Spleißstellen. Die Aminosäurenlänge der jeweiligen Inteine ist in Klammern angegeben. Die erste Aminosäure des assoziierten C-Exteins (C, S oder T) ist etwas abgesetzt gezeigt.

Welche Rolle spielen die Extein- bzw. Inteinsequenzen bei dem Spleißvorgang

Die N- und C-Exteine besitzen keine erkennbare Homologie, was darauf schließen läßt, daß sie für die Spleißreaktion entbehrlich sind. Verkürzt man bei der V-ATPase die Exteine bis auf 13 Aminosäurereste, so findet trotzdem noch ein Spleißen statt. Das *M. tuberculosis* RecA Intein hat man zusammen mit 8 N- und 32 C-Exteinkodons in das *lacZ* Gen inseriert und in *E. coli* transformiert. In *E. coli* wurde das Fusionsprotein nicht nur gebildet, sondern auch prozessiert. Dieser Versuchsansatz wurde auf eine 'in-frame' Insertion des *Saccharomyces cerevisiae* VMA Inteins (das keine Extein kodierende Sequenzen besaß) in das *S. cerevisiae* VAT2 Gen ausgedehnt. Auch dieses Intein wurde erfolgreich ausgeschnitten und die Exteine zu einer voll funktionierenden VAT2 zusammengefügt.

Diese Untersuchungen zeigen, daß alle Information, die für das Proteinspleißen notwendig sind, innerhalb des Inteins lokalisiert sind, also in *cis* vorliegen. Führt man innerhalb des Inteins 'in-frame' Deletionen ein, so findet kein Spleißen mehr statt. Die bekannten Inteine haben eine Größe von 360-537 Aminosäuren, die wenig Aminosäure-Sequenzhomologie aufweisen.

Am Extein-Grenzbereich treten häufig Thiol- oder Hydroxyaminosäuren auf und an der Intein-/C-Extein Grenze ist eine kurze Sequenz (His-Asn-Cys/Ser/Thr) konserviert (**Abb. 1**). Substitutionen von Cys, Ser oder Thr an den Resten der beiden Spleißbereiche verzögern oder inhibieren den Spleißvorgang. Der Asn-Rest am Intein-/C-Extein Grenzbereich ist für das Spleißen absolut notwendig, da sogar ein konservativer Austausch gegen Gln den Spleißvorgang verhindert.

Mechanismus des Proteinspleißens

Da das Proteinspleißen von Intein beinhaltenden Proteinen in heterologen Organismen erfolgt, kann man davon ausgehen, daß die Exzision des Inteins autokatalytisch erfolgt. Z.B die Expression des Hefegens *TFPI* in *E. coli* führt zu einer raschen Bildung des Inteins und des gespleißten reifen Proteins. Das Spleißen muß die Spaltung zweier Peptidbindungen und die Verknüpfung einer neuen Peptidbindung zwischen dem N- und C-Extein umfassen (**Abb.1**).

Die konservierte und essentielle Rolle von Asn am C-terminalen Ende des Inteins und die benachbarte Thiol-oder Hydroxygruppe von Cys bzw. Ser/Thr ist interessant.

Von **Asn** weiß man, daß es innerhalb einer Peptidkette durchaus reaktiv sein kann und z.B zur Spaltung der Peptidkette an seinem C-terminalen Ende führen kann (Perler et al.1992.; Clarke et al. 1992). Die Hydrolyse der Peptidbindung führt zur Ausbildung eines intermolekularen **Succinimidrings**, der durch die Reaktion der

Aminoseitenkette des Asn mit der benachbarten Carbonylgruppe zustande kommt. So eine **Zyklisierung des Intein-Asn** das zur Ausbildung eines **zyklischen Imids** am C-Terminus des Inteins führt, würde eine Spaltung zwischen Asn und der benachbarten C-terminalen Aminosäure bewirken (Cooper et al. 1993a und 1993b). Frühere Peptid- und Proteinstudien haben gezeigt, daß die Hydrolyse zyklischer Imide rasch erfolgt und häufig zu einer Mischung von **Asn- oder Asp-amiden** führt, wobei Aspartatamid charakteristisch für Succinimidbildung während der Reaktion ist (Clarke et al. 1992).

Zu einem Durchbruch des mechanistischen Verstehens führte die Entwicklung eines *in vitro* 'Spleiß assays' (Ku et al. 1993). Perler und Kollegen haben ein Fusionsgen konstruiert das ein Fusionsprotein kodiert. Das Fusionsprotein (**MIP**) besteht aus dem Maltose-Bindeprotein (**M**; N-Extein), dem *Pyrococcus* Polymerase Intein (**I**) und Paramyosin (**P**; C-Extein). Das Maltose-Bindeprotein erleichtert die Reinigung des MIP Precursorproteins. Spezifische Antikörper gegen M-, I- und P- Proteine dienen der Identifizierung der Precursorproteine und der gespleißten Produkte. Wurde das MIP Fusionsgen in *E. coli* exprimiert, fand eine Spleißreaktion bei 12-32° statt, aber sie war relativ langsam, was nicht überraschend war, da das Polymerase-Intein von einem extrem thermophilen Organismus stammt. Man konnte also bei niedriger Temperatur das Precursorprotein über Amylose Affinitäts-Chromatographie reinigen und bis zur Homogenität über MonoQ Chromatographie reinigen. Das gereinigte MIP Precursorprotein wurde dann für *in vitro* Spleißversuche herangezogen.

Dabei stellte sich heraus, daß das Spleißen von der Temperatur abhängig war und die Spleißeffizienz ab 37 bis 65°C kontinuierlich zunahm.

Eine sorgfältige Analyse aller Zwischenprodukte ergab, daß ein verzweigtes **Intermediärprodukt (MIP*)** entstand. Dieses verzweigte Protein erinnerte an die verzweigten RNA Intermediate, die bei den **Gruppe II Introns** beim mRNA Spleißen entstehen.

Das verzweigte Intermediärprotein war **alkalilabil** und eine Behandlung mit **6 M Guanidin-Hydrochlorid bei pH 9,0** führte zur Bildung von IP und M. Dieses Ergebnis weist darauf hin, daß das Intermediärprodukt aus dem N-Extein (M) besteht, das an IP über eine alkalilabile Verbindung verknüpft war. Bei dieser alkalilabilen Bindung konnte es sich entweder um eine Ester- oder eine Imidbindung handeln. Durch massenspektrometrische Methoden konnte man zeigen, daß am C-terminalen Ende des Imids ein Succinimidring entstanden ist.

Modelle für den Mechanismus des Proteinspleißens

Es wurden verschiedene Modelle für den Mechanismus des Proteinspleißens aufgestellt. Eines dieser Modelle von Perler und Mitarbeitern (Ku et al. 1994) ist in **Abb. 2** gezeigt.

Dieses Modell postuliert 2 Routen für die Entstehung des verzweigten Proteinintermediats:

- A) Bei dem einen Weg erfolgt ein nukleophiler Angriff der Hydroxylgruppe des konservierten Ser-Restes an der Intein-/C-Extein-Nahtstelle auf die Carbonylgruppe der N-Extein-/Intein Nahtstelle, was zu einem Serin-Acyl-Enzym Intermediat führt.

B) Der zweite Weg beinhaltet einen N-O 'acyl-shift' zwischen dem konservierten Ser-Rest an der N-Extein-/Intein Nahtstelle, gefolgt von einer Transesterifikation durch die Hydroxylgruppe des konservierten Ser-Restes an der Intein-/C-Extein-Nahtstelle.

Bei beiden Wegen kommt es zur Ausbildung eines verzweigten Proteinintermediates, bei dem der Carboxyterminus des N-Exteins mit dem Intein-/C-Extein Protein über eine Esterbrücke an die Hydroxylgruppe des konservierten Ser-Restes an der Intein-/C-Extein Nahtstelle gebunden wird (**Abb. 2**).

Das verzweigte Intermediat wird einer **Asn-Umlagerung** unterworfen, wobei es das Intein mit einem zyklischen Amid an seinem C-Terminus freisetzt. Die Esterbindung, die das N-Extein mit dem C-Extein verknüpft, wird einem O-N 'acyl-shift' unterworfen, wobei die Peptidbindung ausgebildet wird und damit das reife gespleißte Protein.

Genetische Mobilität von Inteinen

Eine der größten Überraschungen bei der Entdeckung des Proteinspleißens war die, daß Inteine **mobile genetische Elemente** darstellen. Häufig besitzen Inteine **ORFs**, die für **ortsspezifische Endonukleasen kodieren**.

Ein konserviertes Motiv dieser Endonukleasen ist das

LADLIDADG-Motiv.

Bei *S. cerevisiae* bezeichnete man diese ortsspezifischen Endonukleasen als '**homing endonuclease**' (**HO**).

Dieses Sequenzmotiv findet man auch bei den ORFs der **Gruppe I autospleißenden Introns**, die ebenfalls für deren genetische Mobilität verantwortlich sind. Die Endonukleasen vermitteln die genetische Mobilität von Gruppe I Introns in einem Vorgang der als '**intron homing**' bezeichnet wird. Das Intron wird in unidirektioneller Weise in einer Kopie des Gens, das dieses Intron besitzt, auf eine andere Kopie des Gens, das dieses Intron nicht besitzt (**Abb.3**), übertragen.

Bislang konnten bei 4 spleißenden Inteinen solche Endonukleasen nachgewiesen werden. Es handelt sich dabei um das

- a) VMA Intein bei *S. cerevisiae* ,
- b) das Polymeraseintein von *T. litoralis* Intein-1 und Intein-2 und
- c) dem Polymeraseintein von *Pyrococcus*.

Diese ortsspezifischen Endonukleasen führen nicht nur zur Spaltung der entsprechenden DNA-Region, sondern auch zur Integration der Inteinsequenz in dem entsprechenden Gen. Der Mechanismus des 'intron/intein homings' ist bei den Gruppe I Introns und bei den Inteinen der gleiche (**Abb. 3**).

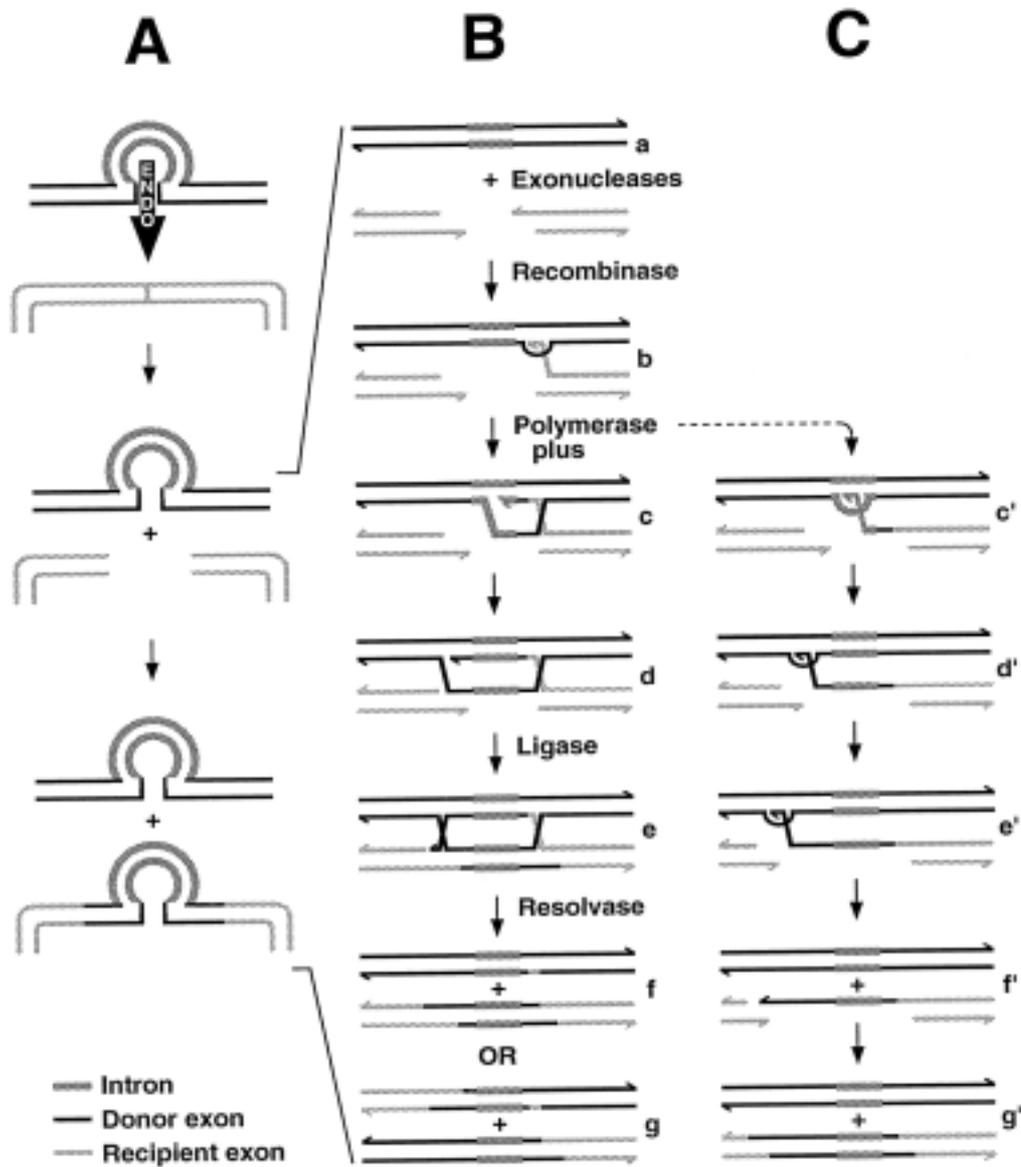


Abb. 3: DNA-mediated group I intron homing.

See:

Belfort et al. (1995) Prokaryotic introns and inteins: a panoply of form and function. J. Bacteriol. 177: 3897-3903.

Spleißen und genetische Mobilität von Gruppe I Introns und Proteinspleißen von Proteinen.

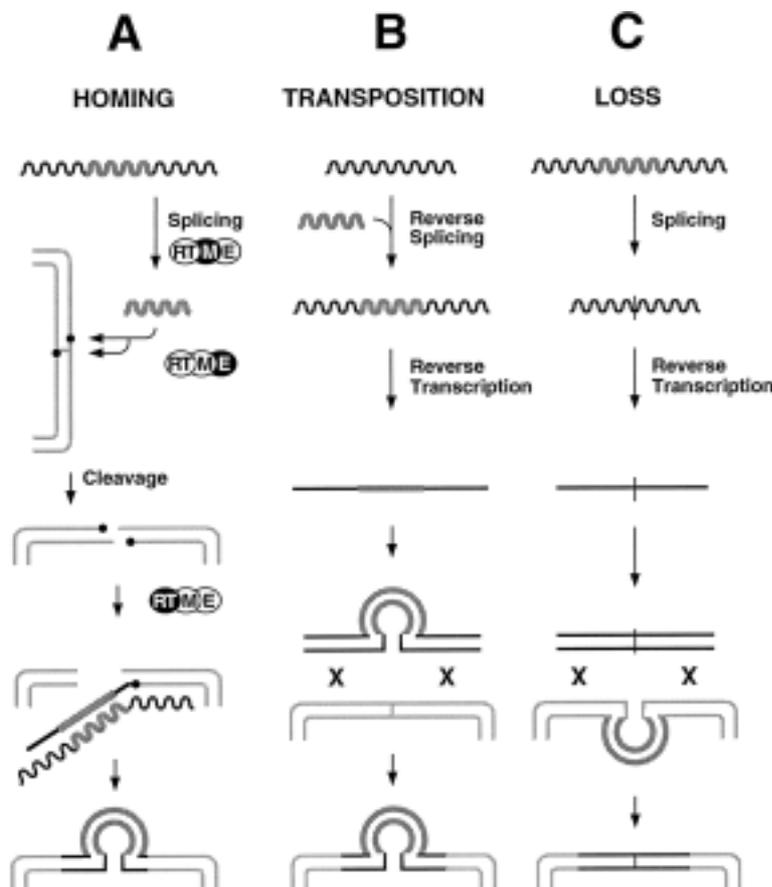
Inteine und Gruppe I Introns existieren beide als intervenierende Sequenzen, die entweder durch RNA-spleißen oder durch Proteinspleißen ausgeschnitten werden. Die Translation des Introns oder die Excision des Inteins führt zur Produktion einer

ortsspezifischen DNA Endonuklease, die nur jene DNA spalten kann, die keine intervenierende Sequenz besitzt (die DNA mit intervenierende Sequenz ist geschützt). Der resultierende Doppelstrangbruch innerhalb der I⁻ DNA wird mit Hilfe eines I⁺ Gens als Matrize repariert und es kommt damit zu einer Genkonversion von I⁻ zum I⁺ Allel. Dieser Vorgang wird als '**intron homing**' bezeichnet.

Wenngleich nur bei 4 Inteinen der Besitz einer ortsspezifischen Endonuklease gezeigt wurde, nimmt man an, daß vermutlich alle Inteine diese Nuklease kodieren, da alle bisher untersuchten Inteine dieses konservierte 12er Peptidmotiv für Endonukleasen besitzen.

Entfernt man dieses Endonukleasemotiv durch Deletion auf DNA Ebene geht zwar die Endonukleaseaktivität verloren, das Proteinspleißen wird dadurch jedoch nicht beeinträchtigt. Das zeigt, daß die Endonukleaseaktivität nichts mit dem Proteinspleißen zu tun hat, und daß beiden Prozessen unterschiedliche Mechanismen zugrunde liegen.

Von den 33 Mycobakterienarten, die man bislang untersucht hat, hat man nur bei 2, nämlich bei *M. tuberculosis* und *M. leprae* Inteine im *RecA* Gen gefunden. Die Inteine in diesen beiden Spezies waren unterschiedlich in ihrer Größe, 365 bzw. 440 Aminosäuren lang, und wiesen nur 33% Identität auf Aminosäureebene auf.



RNA-mediated mobility Events See: Belfort et al. (1995). Prokaryotic introns and inteins: a panoply of form and function. *J. Bacteriol.* 177: 3897-3903.

Ausblick

Es ist klar, daß das Proteinspleißen einen autokatalytischen Prozeß darstellt, die detaillierten Mechanismen der beteiligten Reaktionen sind jedoch noch nicht aufgeklärt. Die Entdeckung eines verzweigten Proteinintermediats ist für die Aufdeckung eines gespleißten Proteins wichtig.

Es stellt sich die Frage, wie weit verbreitet Inteine sind und ob alle Inteine genetisch mobil sind, oder zumindest von mobilen genetischen Elementen abstammen. In den 5 Jahren seit der Entdeckung des ersten Inteins (Kane et al. 1990) wurden 10 weitere Inteinsequenzen entdeckt. Aus dem bislang vorliegenden Sequenzdaten von Genomen verschiedener Mikroorganismen läßt sich nur sehr schwer auf die Anwesenheit von Inteinen schließen, da diese nur eine geringe Sequenzhomologie an ihren Randbereichen aufweisen.

Wenn man jedoch die Sequenz an den Nahtstellen von Intein/Extein und das Peptidmotiv für die Endonuklease heranzieht, kann man aus den Sequenzdaten auch auf die Anwesenheit neuer Inteine schließen. Aufgrund solcher Homologiestudien nimmt man an, daß bei *M. leprae* in dem *pps1* Gen ebenfalls ein Intein enthalten ist.

Abb. 4:

Das Proteinspleiß-Modell von Clarke:

In diesem Modell wird das konservierte Asn und die Imidbildung als das erste Ereignis angesehen. Es kommt zu einem nukleophilen Angriff des Amidstickstoffs des konservierten Asn auf die Peptidcarbonylgruppe der N-Extein-/Intein- Nahtstelle. Dieser nukleophile Angriff resultiert in der Ausbildung eines verzweigten Proteinintermediates, das von der Hydroxylgruppe des Ser an der N-Extein-Intein-Nahtstelle angegriffen wird. Das konservierte Asn innerhalb des aktivierten Precursors wandelt sich in eine zyklische Imidform um, gekoppelt mit einer Aminolyse, um das gespleißte Produkt freizusetzen. Die hypothetischen Protonenakzeptoren und Donatoren sind unbekannt.

Abb.: Phylogenetische Verteilung von Introns und Inteinen. Die Verteilung von Intron und Inteinen eines phylogenetischen Stammbaums innerhalb der 3 Königreiche Archaea, Bakterien und Eukaria ist gezeigt. Intervenierende Sequenzen (IVS) sind eingekreist und werden mit folgenden Abkürzungen bezeichnet: **GI** Gruppe I; **GII** Gruppe II; **Ar** Archaea; **In** Intein;

jene intervenierende Sequenzen, die in geschlossenen Kreisen angezeigt sind, sind vom Kerngenom kodiert, wogegen die anderen IVS in Prokaryonten oder Organellen vorkommen. Die Gruppe I Introns in Gram-positiven Bakterien sind auf Phagen genom kodiert.

Literatur:

Belfort, M., M. E. Reaban, T. Coetzee, and J. Z. Dalgaard. 1995.

Procaryotic introns and inteins: a panoply of form and function. *J. Bacteriol.* **177**:3897-3903.

Cooper, A. A., and T. H. Stevens. 1995. Protein splicing: self-splicing of genetically mobile elements at the protein level. *TIBS.* **20**:351-356.

Edgell, D. R., M. Belfort, and D. A. Shub. 2000. Barriers to intron promiscuity in bacteria. *J Bacteriol* **182**:5281-9.

Mueller, J. E., J. Clyman, Y.-J. Huang, M. M. Parker, and M. Belford. 1996. Intron mobility in phage T4 occurs in the context of recombination-dependent DNA replication by way of multiple pathways. *Genes Dev.* **10**:351-364.

Xu, M.-Q., and F. B. Perler. 1996. The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15**:5146- 5153.

Xu, M.-Q., M. W. Southworth, F. B. Mersha, L. J. Hornstra, and F. B. Perler. 1993. In vitro protein splicing of purified precursor and the identification of a branched intermediate. *Cell.* **75**:1371-1377.