

## Diskontinuierliche Gene und das Spleißen von RNA (Nobel-Vortrag)\*\*

Phillip A. Sharp\*

### Einleitung

Ende der siebziger Jahre war die physikalische Struktur eines Gens durch Arbeiten mit Bakterien bekannt. Die Sequenzen des Gens, der RNA und des Protein werden colinear organisiert und exprimiert. Molekulargenetiker vermuteten, daß Gene in eukaryotischen Organismen ein ähnliches Verhalten zeigen wie jene in prokaryotischen Organismen, und man hielt somit die bakterielle Genstruktur für universell. Daraus folgte, daß bei gleicher Genstruktur auch der Regulationsmechanismus ähnlich sein müßte. Was für Bakterien galt, würde also auch für Elefanten gelten.

Viele biochemische Aspekte des genetischen Materials und seiner Expression in Zellen mit Kernen deuteten jedoch darauf hin, daß die einfache Molekularbiologie der bakteriellen Genexpression nicht universell sein konnte. Erstens sind die von den Genen im Kern transkribierten RNAs von der Translationsmaschinerie im Cytoplasma physikalisch getrennt. Daher könnte das Kernkompartiment der Ort des selektiven RNA-Prozessierens und -Transportes sein. Zweitens variiert der DNA-Gehalt von eukaryotischen Keimzellen bei offenbar gleicher Anzahl von Genen beträchtlich zwischen den Organismen. Einige Organismen scheinen zehnmal mehr DNA zu haben, als für die Codierung aller Proteine nötig ist. Drittens vermutete man wegen des zuvor beschriebenen Phänomens der heterogenen nucleären RNA (hnRNA), daß lange RNAs von den diversen Kernsequenzen transkribiert würden<sup>[1]</sup>. Diese hnRNAs haben – verglichen mit cytoplasmatischen mRNAs – eine kurze Halbwertszeit und konnten daher Vorläufer der mRNAs sein. Des weiteren schienen sowohl die lange hnRNA als auch die kürzere mRNA modifizierte 5'- und 3'-Termini zu haben, nämlich eine 7mGppX-Kappe<sup>[2-4]</sup> bzw. Polyadenylreste<sup>[5-7]</sup>. Die Bedeutung dieser mRNAs wurde kontrovers diskutiert, da es nicht gelang, eine Vorläufer-Produkt-Beziehung zwischen nucleärer RNA-Population und cytoplasmatischer mRNA herzustellen.

Ob die Struktur von Genen in eukaryotischen Zellen der von Genen von Bakterien ähnelt, interessierte damals gar nicht so sehr. Man wollte den genauen biochemischen Weg zwischen einem Gen im Kern und seiner mRNA im Cytoplasma heraus-

finden. Es gab die Hypothese, daß dieser Vorgang mit der Initiation der Transkription durch die RNA-Polymerase II(B) beginnt und bis zur vollständigen Transkription über den Terminus des Gens hinaus verläuft. Die vom Gen transkribierte Vorläufer-RNA würde möglicherweise prozessiert und die mRNA selektiv in das Cytoplasma transportiert werden. Die Regulation der Genexpression, d.h. die Grundlage fast aller interessanten biologischen Abläufe, wie Krebsentstehung, die zelluläre Antwort auf Infektion und Entwicklungsvorgänge, würde primär auf Veränderungen der Geschwindigkeit oder der Effizienz der unterschiedlichen Schritte dieses Weges beruhen. Das Verständnis des Ablaufs der eukaryotischen Genexpression versprach neue Einsichten in den Mechanismus der Regulation und derjenigen biochemischen Ereignisse, die die verschiedenen biologischen und biomedizinischen Phänomene kontrollieren.

### Hintergrund

Das späte Stadium der Adenovirus-Infektion war das beste System, um den Weg vom Gen im Kern über nucleäre Vorläufer-RNA und cytoplasmatische mRNA zum Protein zu untersuchen. Wir hatten zuvor schon einige Kartierungen des viralen Genoms mit Restriktionsendonuklease-Schnittstellen durchgeführt (Abb. 1)<sup>[8]</sup> und Fragmente des Genoms benutzt, um die Positionen der cytoplasmatischen mRNAs, die während der frühen und der späten Stadien der Infektion gebildet werden, zu kartieren<sup>[9]</sup>. Wir hatten auch die Häufigkeiten (in Kopien pro Zelle) von nucleären und cytoplasmatischen RNAs bestimmt<sup>[10]</sup>.

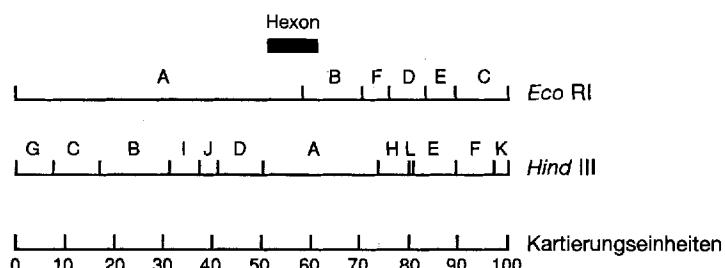


Abb. 1. Kartierung durch Schnittstellen der Restriktionsendonukleasen *Eco* RI (oben) und *Hind* III (Mitte). Den annähernd 35 000 Basenpaaren der DNA des Adenovirus 2 wurden 100 Kartierungseinheiten zugeordnet (unten). Die Position der Schnittstellen wurde durch vertikale Linien markiert, und die Fragmente wurden nach ihrer Länge mit Buchstaben gekennzeichnet. Die viralen r- und l-Strände werden nach rechts bzw. links transkribiert. Die Sequenzen des mRNA-Stückes, die das reichlich vorhandene Hexon(II)-Protein spezifizieren, werden durch den Balken über der Grenze der *Eco* RI-Fragmente A und B markiert.

\* Prof. P. A. Sharp

Department of Biology and the Center for Cancer Research  
Massachusetts Institute of Technology  
Cambridge, MA 02139 (USA)  
Telefax: Int. +617/253-3867

\*\* Copyright © The Nobel Foundation 1994. – Wir danken der Nobel-Stiftung, Stockholm, für die Genehmigung zum Druck einer deutschen Fassung des Vortrags.

und vermuteten, daß ausreichende Mengen von RNA aus beiden Kompartimenten gewonnen werden könnten, um ihre Strukturen direkt unter dem Elektronenmikroskop und durch die damals neu entwickelte Methode der RNA-DNA-Hybridisierung zu vergleichen. Diese frühen Untersuchungen zeigten, daß bestimmte virale RNA-Sequenzen auf den Kern beschränkt sind, was nahelegte, daß nur ganz bestimmte RNAs prozessiert und/oder in das Cytoplasma transportiert werden<sup>[11]</sup>. Schließlich wurde erkannt, daß lange nucleäre RNA während der späten Infektionsstadien vom Adenovirus-Genom transkribiert wird<sup>[12]</sup> und daß die stabilen cytoplasmatischen RNAs kürzer als die vorherrschenden nucleären RNAs sind. Somit wurde die Produktion von RNAs während der späten Stadien der Adenovirus-Infektion zum Musterbeispiel für das mit zellulären Genen assoziierte Phänomen der heterogenen nucleären RNA.

## Die Entdeckung des Spleißens von RNA

Zum Vergleich der Struktur einer cytoplasmatischen mRNA mit der einer nucleären Vorläufer-RNA<sup>[13]</sup> benötigte man gereinigte, spezifische, homogene mRNA. Die am häufigsten vorkommende mRNA, die das Hexon-Protein des Adenovirus codiert, wurde durch Gelelektrophorese von anderen viralen mRNAs getrennt und für elektronenmikroskopische Kartierungsuntersuchungen verwendet<sup>[14]</sup>. Ray White fand als Fellow in Stanford, daß RNA-DNA-Hybride in Lösungen mit hohen Formamidkonzentrationen stabiler waren als DNA-DNA-Duplexe<sup>[15]</sup>. Dieses physikalische Phänomen wurde von Davidson et al.<sup>[16]</sup> charakterisiert und bildete die Grundlage für die R-Schleifen-Methode, mit der ein Hybrid aus einem RNA- und einem DNA-Strang gebildet wird, wobei der andere DNA-Strang als leicht zu beobachtende Schleife verdrängt wird. Die Adenovirus-mRNA für das Hexon-Protein wurde durch die R-Schleifen-Methode dem *HindIII*-A-DNA-Fragment des Adenovirus 2 zugeordnet (siehe Abb. 1)<sup>[13]</sup>.

Die elektronenmikroskopische Analyse der R-Schleifen zwischen der Hexon-mRNA und dem *HindIII*-A-DNA-Fragment zeigte die Anwesenheit von RNA-Schwänzen am 5'- und am 3'-Ende des Hybrids (Abb. 2A, B und C). Der Einzelstrang-RNA-Schwanz am 3'-Ende des Hybrids war nicht überraschend, da die RNA bekanntermaßen nach der Transkription polyadenyliert wird. Der Einzelstrang-RNA-Schwanz am 5'-

Ende des Hybrids war jedoch unerwartet, hätte dieser doch aus den RNA-DNA-Hybriden durch Bildung von Duplex-DNA verdrängt werden können, ein Vorgang, der „branch migration“ genannt wird. Tatsächlich waren schon früher ähnliche 5'-RNA-Schwänze in R-Schleifen-Kartierungen von Adenovirus-mRNA gefunden worden und auf branch migration zurückgeführt worden<sup>[17, 18]</sup>. Gegen die branch migration sprach jedoch die Tatsache, daß die Längen der 5'-Schwänze mit etwa 170 Nucleotiden relativ einheitlich waren. Um die Verdrängung von RNA-Sequenzen durch konkurrierende DNA-Sequenzen zu vermeiden, benutzten wir einen denaturierten Einzelstrang des *HindIII*-Fragments (Abb. 2D, E und F), um ein RNA-DNA-Hybrid zu bilden. Überraschenderweise bildete der 5'-RNA-Schwanz wiederum kein Hybrid mit den benachbarten Sequenzen der Virus-DNA, so daß angenommen wurde, daß diese RNA-Sequenzen von anderen DNA-Sequenzen abstammten.

Könnten jene DNA-Sequenzen, die zur 5'-Schwanzsequenz, der Leader-RNA-Sequenz, transkribiert wurden, stromaufwärts von der Hexon-mRNA, vielleicht als Teil der langen nucleären DNA, lokalisiert sein? Um diese Frage zu beantworten, wurde ein Strang des *EcoRI*-A-DNA-Fragments mit der Hexon-RNA hybridisiert. Dieses DNA-Fragment enthielt alle linearen viralen Sequenzen, die von der Polymerase vor der Begegnung mit dem Rumpf der Hexon-mRNA hätten transkribiert werden können. Überraschenderweise hybridisierten die Leader-RNA-Sequenzen mit drei kurzen DNA-Sequenzen und bildeten drei unterschiedlich große Schleifen a, b und c aus intervenierender Einzelstrang-DNA (Abb. 2G und H). Die Längen dieser Schleifen zeigten die Positionen der Leadersequenzen L1, L2 und L3 bei etwa 16.9, 19.8 bzw. 26.9 Kartierungseinheiten auf dem Genom. Durch die Entfernung zwischen den drei Schleifen konnten die Längen der beiden inneren Leader-Sequenzen auf 80 und 110 Nucleotide geschätzt werden. Die Leader-Sequenz am 5'-Ende war sehr kurz, aber länger als 15 Nucleotide.

Wir schlügeln den Mechanismus des RNA-Spleißens vor, durch den die endgültige Hexon-mRNA gebildet würde<sup>[13]</sup>. Man wußte, daß der Kern von virusinfizierten Zellen lange RNAs enthielt, die von den viralen Sequenzen zwischen dem 5'-Leader L1 – bei 17 Kartierungseinheiten – und dem Rumpf der Hexon-mRNA – zwischen 51.7 und 61.3 Kartierungseinheiten – transkribiert wurden<sup>[19, 20]</sup>. Die lange nucleäre RNA enthielt wahrscheinlich Sequenzen für alle drei Leader-Sequenzen L1, L2 und L3 sowie für den Rumpf der mRNA. Man



*Professor P. A. Sharp wurde am 6. Juni 1944 in Kentucky geboren. Er erhielt sein Diplom in Chemie und Mathematik am Union College, Kentucky, und promovierte an der University of Illinois, Champaign-Urbana. Er war Postdoktorand bei Prof. N. Davidson am California Institute of Technology und bei Prof. J. Watson am Cold Spring Harbor Laboratory. 1974 wurde er von Prof. S. Luria ans Center for Cancer Research und das Department of Biology am Massachusetts Institute of Technology geholt. Von 1985 bis 1991 war er Direktor des Center for Cancer Research, 1991 übernahm er die Leitung des Department of Biology. Seine Forschungsaktivitäten konzentrieren sich auf die Molekularbiologie von Tumrviren und den Mechanismus des RNA-Spleißens, das er 1977 entdeckte. Seine Arbeiten legten die Grundlage zum Verständnis des Phänomens der diskontinuierlichen Gene in Säugetierzellen. Zahlreiche Preise und Auszeichnungen belohnten seine wissenschaftlichen Erfolge, die 1993 mit dem Nobel-Preis für Medizin gekrönt wurden.*

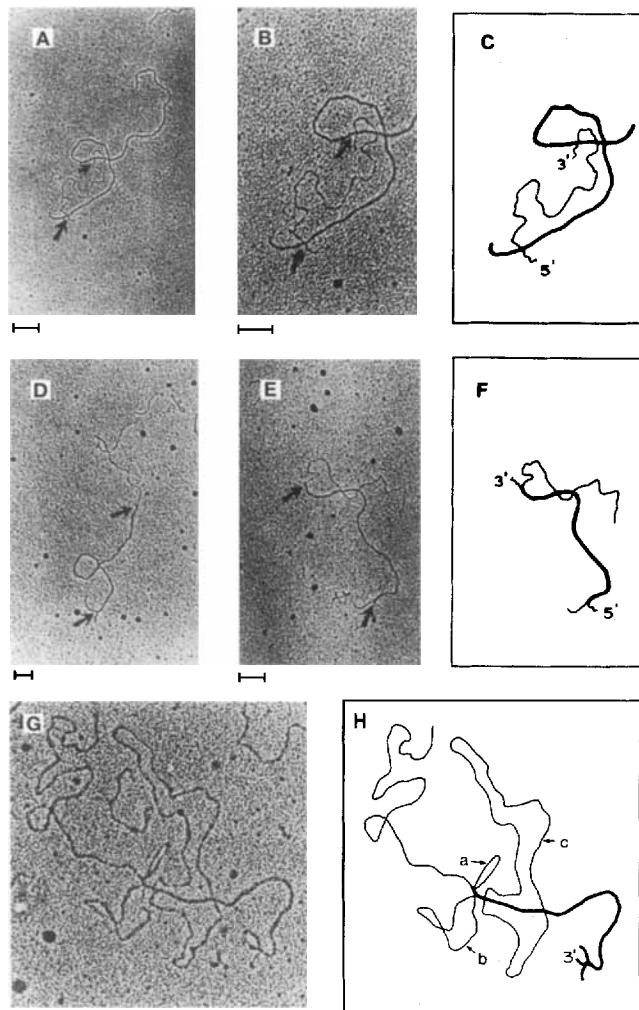


Abb. 2. Elektronenmikroskopische Aufnahmen und schematische Darstellungen von Hybriden der Hexon-mRNA und Fragmenten der Adenovirus-2-DNA [13]. A, B) Beispiele von R-Schleifen-Hybriden nach der Inkubation von Hexon-mRNA und Duplex-*Hind*III-A-Fragment-DNA. C) Schematische Darstellung von A und B. D, E) Zwei Beispiele von Hybriden aus Hexon-mRNA und dem einzelsträngigen *Hind*III-A-Fragment. F) Schematische Darstellung von E. Die einzelsträngige RNA am Ende der Hybridregion wird durch eine geschlängelte Linie dargestellt. In A, B, D und E sind die Positionen der RNA-Schwänze an den 5'- und 3'-Enden der Hybride durch Pfeile markiert. G) Beispiel eines Hybrids zwischen einzelsträngiger *Eco*RI-A-DNA und Hexon-RNA. H) Schematische Darstellung von G. Die Hybridregion ist durch die breite Linie hervorgehoben; die Schleifen a, b und c (einzelsträngige, nicht hybridisierte DNA) werden durch Hybridregionen verbunden, die durch Hybridisierung von „stromaufwärts“-DNA-Sequenzen mit dem 5'-Schwanz der Hexon-mRNA entstehen. Die Maßeinheit unter den Photos entspricht 0.1 µm.

vermutete, daß diese Sequenzen durch Excision der dazwischenliegenden Sequenzen und Ligation der flankierenden RNAs „zusammengeschweißt“ wurden, ein Vorgang, der RNA-Spleißen genannt wurde (Abb. 3). In der Tat kamen die nucleären RNAs relativ häufig vor und waren einfach durch Elektronenmikroskopie der RNA-DNA-Hybride zu beobachten<sup>[21]</sup>. Bei der Analyse der Struktur dieser RNA-DNA-Hybride wurden auch die Zwischenprodukte identifiziert, bei denen nur L1 und L2 oder L1, L2 und L3 durch Spleißen verbunden worden waren.

Die Hypothese des RNA-Spleißen klärte viele Widersprüche: Die längeren nucleären RNAs und die kürzeren cytoplasmatischen mRNAs können gemeinsame 5'-Kappen-Termini, 7mGpppX- und 3'-Polyadenosinreste enthalten<sup>[22]</sup>, da nur inne-

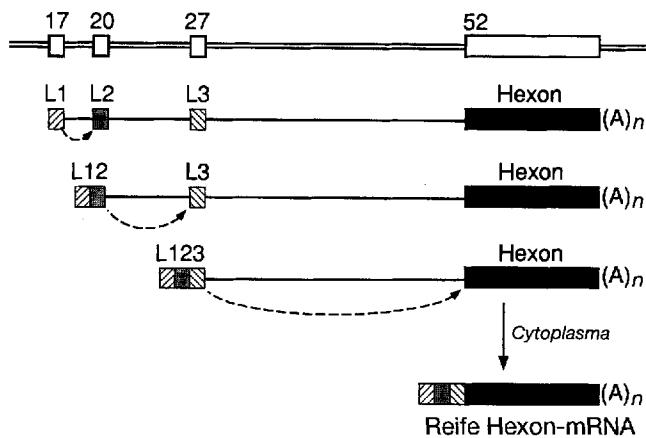


Abb. 3. Der RNA-Spleißmechanismus, der für die Synthese von mRNA des Hexon-Proteins vorgeschlagen wird. Eine lange nucleäre Vorläufer-RNA wird zwischen 16.9 Kartierungseinheiten und der Polyadenin-Stelle am Ende des Rumpfes der Hexon-mRNA (schwarzes Rechteck) transkribiert. Die Region des Adenovirus-2-Genoms, von dem die Vorläufer-RNA transkribiert wird, ist ganz oben in der Abbildung gezeigt. Die vier RNA-Segmente in der reifen cytoplasmatischen mRNA (unten) entstehen aus diesem Vorläufer durch Herausschneiden (Excision) von intervenierenden Sequenzen und Ligation (gestrichelte Pfeile).

re Sequenzen aus dem nucleären Vorläufer entfernt werden. Spezifisch für den Adenovirus war der Befund, daß viele virale mRNAs über eine gemeinsame 5'-terminale Sequenz zu verfügen schienen<sup>[23]</sup>. Diese Sequenz, ein langes T1-Oligonukleotid, enthält die Kappe und eine offenbar einzigartige Sequenz aus elf Resten. Nur wenn viele der späten viralen mRNAs durch Spleißen desselben Typs von Vorläufer-RNA entstünden, könnten sie eine gemeinsame Sequenz an ihren 5'-Termini enthalten. Noch wichtiger war, daß die Hypothese des RNA-Spleißen eine Erklärung für die hnRNA-Phänomene bei zellulären Genen lieferte. Heterogene nucleäre RNA, die von diversen zellulären Genen transkribiert wird, kann durch RNA-Spleißen zu kürzeren cytoplasmatischen mRNAs umgeformt werden. Wahrscheinlich enthalten die meisten zellulären Gene Sequenzen, die durch RNA-Spleißen entfernt werden, d. h. es handelt sich um diskontinuierliche Gene.

### Diskontinuierliche Gene: Intervenierende Sequenzen oder Introns in zellulären Genen

Kurz nach der Entdeckung des RNA-Spleißen und der diskontinuierlichen Gene im Adenovirus konnte gezeigt werden, daß auch viele zelluläre Gene intervenerende Sequenzen (= Introns) enthalten. So haben die Globin-Gene zwei intervenernde Sequenzen<sup>[24, 25]</sup>, das Ovalbumin-Gen ist in acht Sequenzen unterteilt<sup>[26]</sup>, und die Immunglobulin-Gene enthalten sowohl kurze als auch lange Introns<sup>[27]</sup>. Im Durchschnitt verfügen zelluläre Gene über etwa acht Introns, und das primäre Transkript ist typischerweise viermal länger als die reife mRNA. Wenig später erkannte man, daß es eine limitierte Anzahl von konservierten Sequenzen an jeder Introngrenze gibt<sup>[28]</sup>. Interessanterweise sind diese Consensus-Sequenzen bei Wirbeltier-, Pflanzen- und Hefezellen gleich<sup>[29]</sup>, was zu der Vermutung Anlaß gab, daß das Spleißen ein allgemein gültiger evolutionärer Vorgang ist. Die Introns der Hefen sind im allgemeinen kürzer und haben hochkonservierte Sequenzen an ihren Grenzen.

Der phylogenetische Vergleich von Sequenzen homologer Gene aus einer Vielzahl von Organismen zeigte, daß Intron-Sequenzen schneller verändert wurden als Exon-Sequenzen und daher wohl in der Regel nicht funktionell waren, zumindest in dem Sinn, daß sie keine langen spezifischen Sequenzen enthielten. Darüber hinaus veränderte sich die Länge von Introns in homologen Genen innerhalb evolutionärer Reihen signifikant, was auf einen geringen Selektionsdruck hindeutet. Schließlich wurde auch klar, daß spezifische Introns während der Evolution verloren gegangen sein konnten. Der für die Deletion von Introns verantwortliche Mechanismus ähnelt wahrscheinlich der Gen-Konversion mit einer cDNA-Kopie der mRNA oder einer partiell gespleißten intermediären RNA. Dieser Vorgang konnte für die Entfernung von Introns aus Hefe-Genen bewiesen werden<sup>[30]</sup> und wirft die Frage auf, warum Introns im Lauf der Evolution erhalten blieben.

## Mutationen in zellulären Genen

Viele Krankheiten des Menschen werden von Mutationen verursacht, die das korrekte RNA-Spleißen beeinträchtigen. Annähernd ein Viertel aller Thalassämie verursachenden Mutationen in den menschlichen Globin-Genen sind Mutationen in Sequenzen, die das Spleißen von RNA steuern. Die Konservierung von Information für akkurate Spleißen ist somit ein Selektionsdruck für genetische Systeme. Interessanterweise ist bei Globin-Genen die Zahl der Mutationen, die verändertes Spleißen verursachen, begrenzt<sup>[31]</sup>. Alle charakterisierten Mutationen ändern entweder die konservierte Sequenz der 5'- oder 3'-Spleißstelle an den Rändern des Introns oder sind Veränderungen innerhalb des Introns, die neue Consensus-Spleißstellen hervorbringen (Abb. 4). Das erstere ist einfach: Die Mutation der hochkonservierten GU- oder AG-Sequenzen an den Rändern des Introns inaktiviert das Spleißen an dieser Stelle, und dies führt häufig zur Aktivierung von nahegelegenen kryptischen Spleißstellen. Das zweite ist komplizierter: In diesem Fall aktiviert eine Mutation von CT nach GT innerhalb des Introns diese Stelle als 5'-Spleißstelle, und in der Folge wird eine stromaufwärts gelegene AG-Sequenz als 3'-Spleißstelle aktiviert. Dieses führt zur Bildung eines neuen, kurzen Exons aus der vormaligen Intron-Sequenz (Abb. 4). Daher können Mutati-

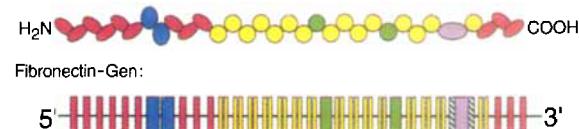
onen sowohl die Position von Spleißstellen als auch die Zahl der Exons verändern und so neue Proteine hervorbringen.

Fehler beim Spleißen von RNA komplexer Gene sind wahrscheinlich nicht ungewöhnlich. Gelegentlich können Exons übergangen werden, und in einigen Fällen kann zirkuläre RNA gebildet werden<sup>[32, 33]</sup>. Da vielen dieser RNAs noch keine spezifischen Funktionen zugeordnet worden sind, könnte ihre Bildung ein Grundrauschen des Systems sein. Interessanterweise wurde vor kurzem vorgeschlagen, daß in Zellen ein mRNA-Überwachungssystem RNAs beim Eintritt ins Cytoplasma zerstört, wenn sie einen offenen Leserahmen enthalten, der durch ein Translationsterminationssignal unterbrochen wird<sup>[34]</sup>. Dieses System würde wahrscheinlich die meisten RNAs mit Spleißfehlern beim Transport ins Cytoplasma abbauen.

## Warum gibt es diskontinuierliche Gene und RNA-Spleißen?

Es ist schwer zu erklären, warum Introns innerhalb der Gene der meisten, wenn nicht aller Eukaryoten während der Evolution erhalten blieben. Die Intron-Exon-Struktur der Gene war bei der Entwicklung neuer Gene im Lauf der Evolution sicherlich sehr wichtig. So ist das Gen für Fibronectin des Menschen aus drei Exon-Typen aufgebaut, wovon jeder einen spezifischen Typ von Proteinfaltungsdomänen codiert (Abb. 5)<sup>[35]</sup>. Die glei-

Fibronectin-Protein:



Fibronectin-Gen:

5' —————— 3'

### Protein-Subdomäne

- Typ I      ● - wird auch in Plasminogen-Aktivatoren im Gewebe gefunden
- Typ II     ● - wird auch in Blutgerinnungs-Proteinen gefunden
- Typ III    ● - wird auch in Zelloberflächen-Rezeptoren und anderen Proteinen der extrazellulären Matrix gefunden

Abb. 5. Fibronectin-Gen durch Exon-Duplikation. Fibronectin, ein großes Protein der extrazellulären Matrix, besteht hauptsächlich aus den drei Proteinsubdomänen I, II und III, die als schräge rote Ellipsoide, vertikale blaue Ellipsoide bzw. gelbe Kreise gezeigt werden. Jede dieser Domänen wird durch ein definiertes Exon-Muster aus einem oder zwei Exons codiert (vertikale Rechtecke). Diese Exon-Einheiten sind im Fibronectin-Gen und auch in anderen Genen, die diese Domänen enthalten, dupliziert. In zwei Fällen, gekennzeichnet durch die grünen Rechtecke, wurden die Introns, die die typischen Doppel-Exon-Strukturen der Subdomäne III trennen, entfernt. Diese Muster geben Anlaß zu der Vermutung, daß Teile dieser Gene von Abkömlingen einer gemeinsamen Vorläufer-Exon-Konfiguration gebildet wurden [36].

chen drei Proteinfaltungsdomänen und die korrespondierenden Exon-Strukturen werden auch in anderen Genen gefunden, von denen einige Zelloberflächen-Rezeptoren und Blutgerinnungs-Proteine codieren. So entstand das Fibronectin-Gen durch Tandem- und verstreute Duplikation von Exon-Einheiten über Brüche und Verknüpfungen innerhalb der Intron-Sequenzen. Da diese Exon-Einheiten durch präzises RNA-Spleißen verbunden werden und funktionelle Proteindomänen codieren, ist das resultierende Protein stabil und hat mehrere Funktionen. Ver-

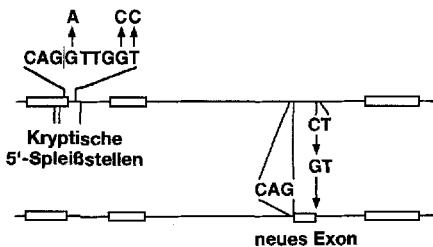


Abb. 4.  $\beta$ -Globin-Mutationen.  $\beta$ -Globin-Gene von Patienten mit Thalassämie haben Mutationen, die das RNA-Spleißen beeinflussen [31]. Die Pfeile markieren die vertauschten Einzelbasen in den zwei Intron-Strukturen des  $\beta$ -Globin-Gen. Drei dieser Mutationen verändern konservierte Sequenzen der 5'-Spleißstelle (links), und ein Teil der mRNA dieser Mutanten wird an kryptischen 5'-Spleißstellen prozessiert. Die vierte Mutation liegt innerhalb der Sequenz des zweiten Introns (rechts) und verursacht eine 5'-Spleißstelle an dieser Position. Dies führt zur Bildung eines neuen Exons mit den bezeichneten Sequenzen. Exon-Sequenzen sind als Rechtecke und Intron-Sequenzen als Linien dargestellt.

doppelung und Nutzung einer Exon-Einheit ist ein allgemeiner Mechanismus zur Bildung neuer Gene, durch die viele Zelloberflächen-Rezeptoren und andere Proteintypen bei Wirbeltieren codiert werden. Daher ist die Existenz von Exon-Einheiten, die funktionellen Proteindomänen entsprechen, entscheidend für die Evolution von komplexen Organismen gewesen.

Die Fähigkeit, auf der Stufe des RNA-Spleißens durch Wahl verschiedener Exon-Kombinationen Proteine mit verschiedenen Funktionen zu erzeugen, ist für die Entwicklung vieler Wirbeltiere unabdingbar. Etwa eines von zwanzig Genen wird in verschiedenen Zelltypen oder Entwicklungsstadien auf alternativen Wegen des Spleißens exprimiert. So werden nucleäre Vorläufer-RNAs des Fibronectin-Gens zu verschiedenen mRNAs gespleißt, die zwanzig Proteine codieren. Diese Proteine haben infolge ihrer Strukturunterschiede auch leicht unterschiedliche Funktionen. Auch tauchen einige Exons in mRNAs aus Leberzellen nicht auf, die in mRNAs anderer Zelltypen auftreten (Abb. 6). Das von der Leber sekretierte Fibronectin zirkuliert

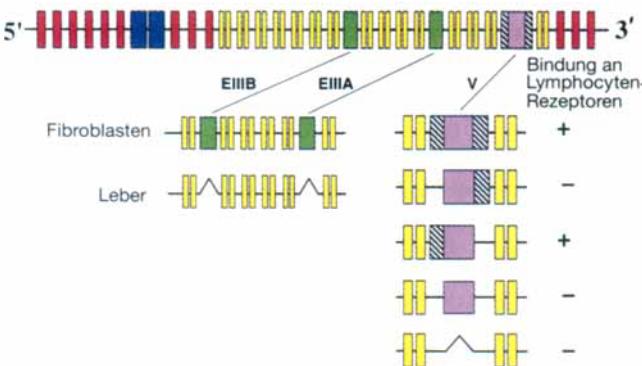


Abb. 6. Alternatives Spleißen von Fibronectin-codierender RNA. Die Exons, die als EIIIB und EIIIA bezeichnet sind, werden zur reifen RNA gespleißt, die in vielen Zelltypen einschließlich Fibroblasten, synthetisiert wird. Die mRNA aus Leberzellen enthält diese zwei Exons jedoch nicht, sondern sie werden bei der Synthese der mRNA ausgelassen. Des Weiteren sind fünf Variationen des Spleißmusters der V-Region gezeigt. Alle diese Spleißvariationen werden in den meisten Zelltypen gefunden. Die Fibronectin-Proteine, die von der ersten und dritten mRNA codiert werden, binden unterschiedlich an Rezeptoren von Lymphozyten. Insgesamt werden mindestens zwanzig verschiedene Proteine von einem einzigen Fibronectin-Gen synthetisiert.

früher im Blutkreislauf als andere Fibronectin-Typen, weil es weniger Proteindomänen enthält, die durch Zelloberflächen-Rezeptoren erkannt werden<sup>[36]</sup>. Das von anderen Zellen sekretierte Fibronectin wird typischerweise in der intrazellulären Matrix fester Gewebe gefunden. Ähnlich unterschiedliche Zelladhäsion findet man auch nach alternativem Spleißen in einem anderen Teil der Fibronectin-mRNA. Durch alternatives Spleißen kann also dasselbe Satz von Gensequenzen für unterschiedliche Funktionen genutzt werden.

Alternatives Spleißen von Vorläufer-mRNA in verschiedenen Zelltypen muß auch Unterschiede bei den Faktoren, die den Spleißvorgang regulieren, widerspiegeln. Diese Faktoren konnten nur in wenigen Fällen, vor allem durch die Analyse von Mutanten mit einem Defekt in der Regulation, identifiziert werden. Die Entwicklung des Geschlechts-Dimorphismus bei der Fruchtfliege *Drosophila* wird auf der Ebene des alternativen RNA-Spleißens durch die Expression einer Gen-Kaskade reguliert. Schon in einem frühen Stadium wird das weibliche oder

männliche Spleißmuster durch ein Ablesen des Verhältnisses von Geschlechtschromosomen zu autosomalen Chromosomen festgelegt. Danach wird der Entwicklungsvorgang von einem autoregulatorischen Prozeß kontrolliert, der die weiblichen oder männlichen Spleißmuster in vielen, wenn nicht gar allen Zellen aufrechterhält<sup>[37]</sup>. Interessanterweise gehören die meisten der Proteine, die durch diese geschlechtsregulierenden Gene codiert werden, zu einer Familie von Proteinen, die eine oder mehrere RNA-Erkennungssubdomänen und die auffällige, sich wiederholende Aminosäuresequenz Arg-Ser gemeinsam haben<sup>[38]</sup>. Wahrscheinlich wechselwirken diese Proteine über die Arg-Ser-Reste direkt mit anderen Proteinen, was für die Bildung des Spleißapparats wichtig ist.

### Gab es diskontinuierliche Gene in Progenoten?

Die diskontinuierliche Genstruktur muß sehr alt sein und schon vor der Divergenz der Organismen, aus der Pflanzen, Hefe und Menschen hervorgingen, bestanden haben<sup>[29]</sup>. Erstens existiert im Kern aller Eukaryoten – wie im folgenden beschrieben wird – ein komplexer Spleißapparat, an dem mehr als 50–100 Proteine und fünf kleine RNAs beteiligt sind. Dieser Spleißapparat ist hochkonserviert und kann nicht spontan in den verschiedenen Zelllinien entstanden sein. Zweitens sind Introns in Positionen innerhalb homologer Gene von Pflanzen- und Tierzelllinien konserviert. Dies beweist, daß diskontinuierliche Gene und RNA-Spleißen in gemeinsamen Vorläufer-Organismen, den Progenoten, schon vor einer Milliarde Jahren vorhanden waren. Manche Wissenschaftler postulieren, daß Gene sich ursprünglich als Exons herausbildeten und daß die Progenoten, aus denen sich die heutige prokaryotischen und eukaryotischen Organismen entwickelten, eine diskontinuierliche Genstruktur besaßen<sup>[39–42]</sup>. Man stellt sich vor, daß diese ursprünglichen Exon-Sequenzen stabile Proteinfaltungsdomänen codieren. Durch den Zusammenbau mehrerer Exon-Sequenzen durch RNA-Spleißen entsteht dann ein Protein mit stabilen Faltungsdomänen, das mit hoher Wahrscheinlichkeit entweder strukturell oder katalytisch wirksam wäre. Wenn Gene tatsächlich so entstanden sind, dann dürfte die Position der Introns in bezug auf die Sekundärstruktur der Proteine nicht zufällig sein. Beweise für diese Hypothese wurden in der Exon-Intron-Struktur von Entwicklungsgeschichtlich alten Proteinen gesucht, die für den Energiehaushalt wichtig sind. Beispielsweise ist die Struktur und die Funktion des Enzyms Pyruvat-Kinase in Bakterien, Hefen und Hühnern konserviert (Abb. 7)<sup>[43]</sup>. Das Gen in Bakterien und Hefen ist nicht durch Introns unterbrochen, enthält aber in Hühnern zehn Introns, davon neun innerhalb der Codierungssequenz. Die Positionen dieser neun Introns sind – wenn man die Proteinstruktur betrachtet – nicht zufällig. Erstens befinden sich die ersten drei Introns im N-terminalen Teil des Proteins zwischen sich wiederholenden Strukturmotiven eines  $\alpha$ -Helix- $\beta$ -Faltblatts. Zweitens hat die Pyruvat-Kinase am Mononucleotid-Bindungsort eine gemeinsame Intron-Position mit einem anderen alten metabolischen Enzym, der Alkohol-Dehydrogenase. In der Pyruvat-Kinase besetzt das achte Intron eine Position, die der eines Introns in der Dehydrogenase ähnlich ist. Da die Bildung dieser zwei Enzyme aus einer gemeinsamen Proteinsubdomäne in die Zeit vor der entwicklungs-



Abb. 7. Die Struktur der Pyruvat-Kinase und die Position der Introns [44] (mit freundlicher Genehmigung, Copyright © 1993, Cell Press, Cambridge, MA, USA). Links wird die Pyruvat-Kinase schematisch als ein Protein mit Sekundär- und Tertiärstruktur gezeigt. Rechts ist die an den Intron-Positionen aufgeweitete Tertiärstruktur dargestellt. Die ersten drei Introns liegen zwischen repetitiven  $\alpha$ -Helix- (Zylinder) und  $\beta$ -Faltblatt-Strukturen (Pfeile). Das Intron 8 liegt in annähernd äquivalenten Positionen in der Mononucleotidbindungsfurche der Pyruvat-Kinase und der Alkohol-Dehydrogenase aus Mais.

geschichtlichen Divergenz von prokaryotischen und eukaryotischen Organismen zurückreicht, lassen diese Ergebnisse vermuten, daß die Vorläufer-Organismen von prokaryotischen und eukaryotischen Organismen eine diskontinuierliche Gen-Struktur gehabt haben, die eher für eine heutige eukaryotische Zelle typisch ist (siehe Lit. [44] für eine weiterführende Diskussion).

### Ist das Gen ein Exon?

Zuerst wurde das Gen als Vererbungseinheit auf einem Chromosom genetisch definiert. Die chemische Definition eines Gens wurde jedoch schwieriger, als die Komplexität der genetischen Information und ihrer Modifikationen entdeckt wurden. Die Existenz des alternativen Exon-Spleißens, durch den die Information einer Exon-Einheit wahlweise nur in bestimmten Zellen exprimiert werden kann, würde bedeuten, daß ein Exon einem Gen und damit dem Minimum an Information entspricht, das als diskrete Einheit exprimiert wird. Dieses Konzept wird beim *trans*-Spleißvorgang besonders relevant<sup>[45]</sup>. In diesem Fall werden Exons, die von verschiedenen Chromosomorten und in vielen Fällen sogar von verschiedenen Chromosomen transkribiert worden sind, durch RNA-Spleißen verbunden. Das *trans*-Spleißen von Exons und Introns wurde bei parasitisch lebenden Trypanosomen<sup>[46]</sup> und dem Plattwurm *C. elegans*<sup>[47]</sup> gefunden und auch für einige menschliche Gene diskutiert<sup>[48]</sup>. Im Fall des *trans*-Spleißens können sowohl die Vererbungseinheit als auch

der Ort auf einem Chromosom einem einzigen Exon entsprechen.

Es ist jedoch unwahrscheinlich, daß die derzeitige Arbeitsdefinition eines Gens als die lineare Aneinanderreihung von Exons, die durch RNA-Spleißen verbunden werden, in naher Zukunft radikal verändert wird. Das ist insofern sinnvoll, als daß die Existenz vieler Vorgänge, die zusammen RNA-Editing genannt wird, die biochemische Definition eines Gens weiter kompliziert<sup>[49]</sup>. Da die früheste Einheit der genetischen Information möglicherweise als Exon evolvierte, könnte sich das allgemeine Konzept von Exons als Geneinheiten gegenüber anderen Vorschlägen durchsetzen.

### Das Spleißen von nucleären Vorläufer-RNAs zu mRNAs

Die Entwicklung einer Reaktion, in der lösliche Zellkomponenten die Vorläufer-RNAs (prä-mRNAs) zu mRNAs umsetzen, war entscheidend für unser Verständnis des Spleißvorgangs<sup>[50]</sup>. Mit hochradioaktiven prä-mRNA-Substraten wurde so eine biochemische Analyse des Spleißens möglich<sup>[51, 52]</sup>. Es überrascht nicht, daß kinetische RNA-Zwischenprodukte der Spleißreaktion bald identifiziert wurden<sup>[53, 54]</sup>. Sie haben eine Lasso-Struktur, in der das 5'-Nucleotid des Introns durch eine 2',5'-Phosphodiesterbindung an ein Adenosinrest innerhalb des Introns geknüpft ist<sup>[55]</sup>. Da die Adenosineinheit durch 3',5'-

und 2',5'-Phosphodiesterbrückenkovalent gebunden ist, liegt hier eine RNA-Verzweigung vor. Die Existenz solcher Verzweigungen wurde durch Nuclease-Verdauungsuntersuchungen der gesamten nucleären RNA aus menschlichen Zellen gezeigt<sup>[56]</sup>. Die Bildung der Verzweigung geht mit der Spaltung an der 5'-Spleißstelle einher. Die entstandene Lasso-RNA wandert typischerweise bei der Elektrophorese durch ein engporiges Polyacrylamidgel langsamer als die prä-mRNA.

Das Spleißen der prä-mRNA erfolgt in zwei Schritten (Abb. 8, links). Wie schon erwähnt, wird im ersten Schritt an der 5'-Spleißstelle gespalten und gleichzeitig die Verzweigung gebildet. Zu diesem Zeitpunkt hat die 5'-Exon-RNA eine 3'-Hydroxygruppe, und die intermediäre Lasso-RNA enthält das Intron und das 3'-Exon. Im zweiten Schritt wird die RNA an der 3'-Spleißstelle gespalten und gleichzeitig die beiden Exons verbunden. Das Intron wird als Lasso-RNA freigesetzt und ist *in vitro* recht stabil. In *vivo* werden Intron-RNAs hingegen fast immer schnell abgebaut.

Die Tatsache, daß die beiden Bruchstücke nach dem ersten Schritt sehr effizient zu den Endprodukten umgewandelt wurden, weist darauf hin, daß diese RNAs in einem Komplex gebunden bleiben. Dieser Komplex wurde aufgrund seines Sedimentationskoeffizienten von 60 s in einem Glycerolgradienten identifiziert und Spleißosom oder Spleißkörper genannt<sup>[57, 58]</sup>. In früheren Arbeiten hatte man schon die Bedeutung von kleinen nucleären Ribonucleotid-Protein(snRNP)-Partikeln für den Spleißvorgang diskutiert; das Spleißosom enthält die kleinen nucleären RNAs (snRNAs) U 2, U 4, U 5 und U 6 sowie unter bestimmten Umständen auch U 1<sup>[59]</sup>. Ähnlich wie ein Ribosom besteht das Spleißosom also aus einer Substrat-RNA und einer Anzahl von stabilen zellulären RNA-Protein-Komponenten.

### Gruppe I, Gruppe II und das Spleißosom

Ein Vergleich der zwei Spleißosom-Spleißschritte mit den RNA-katalysierten Selbst-Spleißreaktionen der Introns der Gruppen I und II zeigt einige frappierende Ähnlichkeiten (Abb. 8). In allen drei Fällen ist der erste Schritt eine Spaltung an der 5'-Spleißstelle. Bei Introns der Gruppe I wird für die Spaltung ein Guanosin-Cofaktor benötigt, der eine Bindungsstelle in den katalytischen Intron-Sequenzen spezifisch besetzt<sup>[60]</sup>. Die 3'-Hydroxygruppe dieses Guanosins wird aktiviert und verdrängt in einer Umesterung die 3'-Hydroxygruppe des 5'-Exons. Die selbst-spleißenden Introns der Gruppe II spalten an der 5'-Spleißstelle durch Aktivierung der 2'-OH-Gruppe an der Verzweigungsstelle, wodurch eine Lasso-RNA entsteht<sup>[61, 62]</sup>, die der von Spleißosomen gebildeten recht ähnlich ist. Im zweiten Schritt aller drei Vorgänge werden durch eine Reaktion an der 3'-Spleißstelle die Exons verbunden und die Introns verdrängt. Der einfachste Mechanismus für den zweiten Schritt des Spleißens von Introns der Gruppen I und II ist eine einfache Umesterung. Dies konnte für die Reaktionen von Introns der Gruppe I gezeigt werden.

Wegen der Ähnlichkeiten zwischen dem spleißosomalen Vorgang und dem Selbst-Spleißen der Introns ist anzunehmen, daß diese Prozesse evolutionsgeschichtlich verwandt sind. Dies gilt insbesondere für Reaktionen der Introns der Gruppe II und

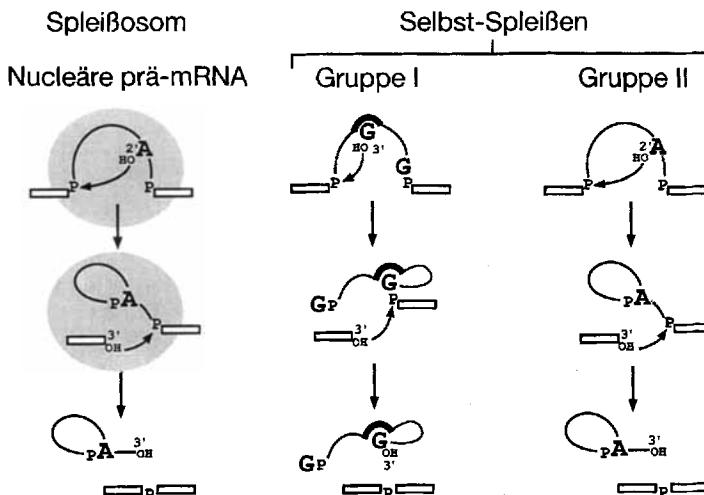


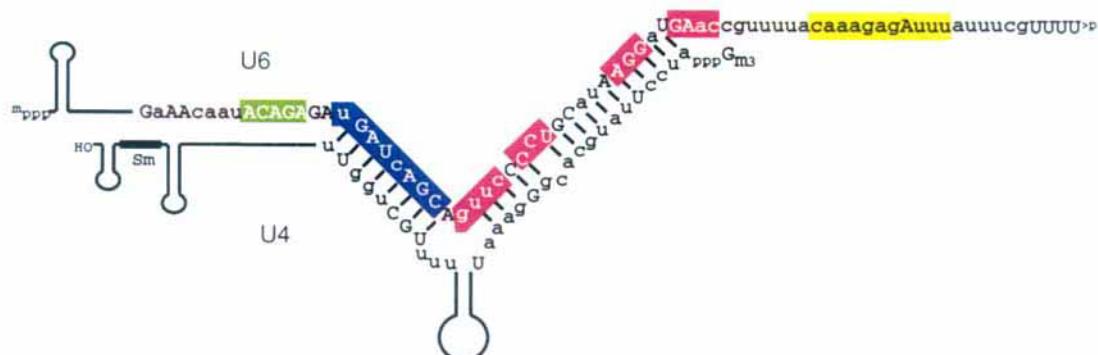
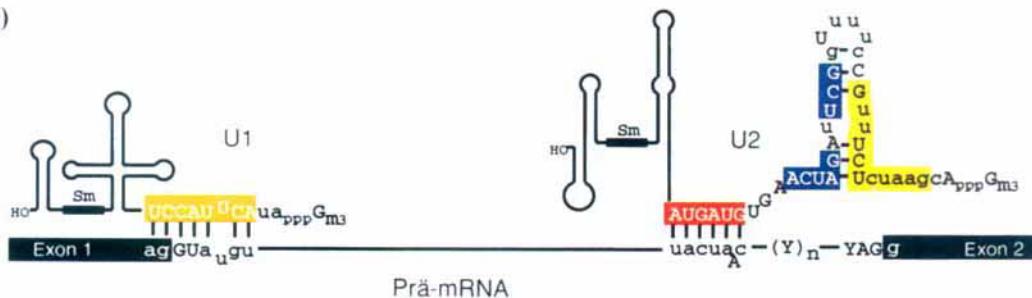
Abb. 8. Vergleich der Mechanismen des nucleären prä-mRNA-Spleißens und des Selbst-Spleißen. Die erste Spalte zeigt den mRNA-Vorläufer-Spleißmechanismus. Die schraffierte Kreise stellen einen Multikomponentenkomplex, das Spleißosom, dar, der die Spleißreaktion katalysiert. Die zweite Spalte zeigt den Spleißmechanismus der selbst-spleißenden Introns der Gruppe I. Dieser Vorgang wird von einer Guanosin-Bindungsstelle (dunkle Halbkreise) innerhalb des Introns katalysiert und nutzt einen Guanosin(G)-Cofaktor im ersten Schritt. Die dritte Spalte zeigt den Spleißmechanismus der selbst-spleißenden Introns der Gruppe II. Dieser Prozeß wird ebenfalls von RNA-Strukturen (dunkle Halbkreise) innerhalb des Introns katalysiert, nutzt aber anstatt eines Cofaktors einen Adenosinrest(A) im Intron, um eine Lasso-RNA zu bilden. Jeder der drei Mechanismen beinhaltet zwei Schritte: die Reaktion an der 5'-Spleißstelle und die anschließende Reaktion an der 3'-Spleißstelle.

der Spleißosomen. Die snRNAs im Spleißosom kann man sich als ein zerstückeltes Intron der Gruppe II vorstellen, wobei diese RNAs die katalytischen Stellen für beide Reaktionen sind. Die Proteine in den Spleißosomen könnten für die Erkennung der Vorläufer-RNA nötig sein, indem sie die snRNAs zu katalytischen Strukturen organisieren und die Komponenten umordnen, so daß die Reaktionskette leicht beendet werden kann. Früh in der Evolution waren vielleicht alle Introns selbst-spleißend. Dann könnten sich im Laufe der Zeit *trans*-reagierende Komponenten entwickelt haben, die Introns durch Sequenzerkennung an den Spleißstellen spleißen, wodurch die *cis*-Sequenzen innerhalb jedes Introns verkümmerten<sup>[63]</sup>. *Trans*-reagierende Segmente von Introns der Gruppe II, die das Spleißen von Introns mit partiell katalytischer Struktur ergänzen, sind nachgewiesen worden<sup>[63]</sup>.

Einige Jahre nach der Entdeckung des Lasso-Zwischenproduktes beim *cis*-Spleißen konnte auch beim *trans*-Spleißen ein verzweigtes Zwischenprodukt nachgewiesen werden. Die Chemie des *trans*-Spleißens ist demnach der des *cis*-Spleißens sehr ähnlich<sup>[64]</sup>. Es ist daher sehr wahrscheinlich, daß die *cis*- und *trans*-Spleißvorgänge Variationen eines grundlegenden Mechanismus sind. Trypanosom-Parasiten, die mRNA ausschließlich durch *trans*-Spleißen einer kurzen Leader-RNA synthetisieren, exprimieren offenbar weder U 1- noch U 5-snRNA, aber dafür U 2-, U 4- und U 6-snRNAs<sup>[65]</sup>.

Daß die Sequenzkomplementarität zwischen dem 5'-Ende der U1-snRNA und den Consensus-Sequenzen an der 5'-Spleißstelle eine entscheidende Rolle für das Spleißen spielt<sup>[66, 67]</sup>, konnte durch eine Reihe von Experimenten, einschließlich der Inhibition des Spleißens *in vitro* durch Zugabe von Antiseren, die für U1-snRNP spezifisch sind, bestätigt werden<sup>[68]</sup>. Nach

a)



b)

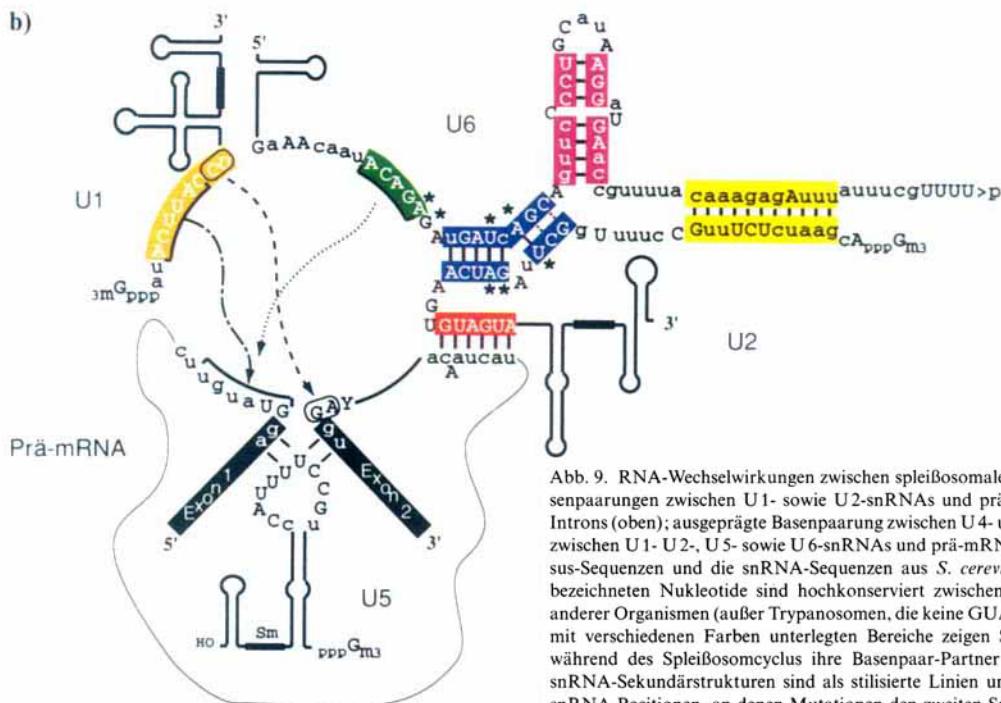


Abb. 9. RNA-Wechselwirkungen zwischen spleißosomalen snRNAs und prä-mRNA-Substraten. a) Basenpaarungen zwischen U1- sowie U2-snRNAs und prä-mRNA an der linken bzw. rechten Seite des Introns (oben); ausgeprägte Basenpaarung zwischen U4- und U6-snRNAs (unten). b) Wechselwirkungen zwischen U1- U2-, U5- sowie U6-snRNAs und prä-mRNA. In (a) und (b) sind die prä-mRNA-Consensus-Sequenzen und die snRNA-Sequenzen aus *S. cerevisiae* wiedergegeben; die mit Großbuchstaben bezeichneten Nukleotide sind hochkonserviert zwischen *S. cerevisiae* und den bekannten Sequenzen anderer Organismen (außer Trypanosomen, die keine GUAGUA-Sequenz in der U2-snRNA haben). Die mit verschiedenen Farben unterlegte Bereiche zeigen Sequenzen in den U2- und U6-snRNAs, die während des Spleißosomzyklus ihre Basenpaar-Partner wechseln. Interne, in (a) und (b) identische snRNA-Sekundärstrukturen sind als stilisierte Linien und Schleifen gezeigt. Sternchen bezeichnen die snRNA-Positionen, an denen Mutationen den zweiten Spleißschritt spezifisch blockieren.

der Entdeckung der Lasso-RNA wurde eine Consensus-Sequenz in der Region neben der Verzweigungsstelle gefunden. Diese Consensus-Sequenz ist zu einer konservierten inneren Sequenz in der U2-snRNA komplementär, und durch Mutationsanalysen konnte gezeigt werden, daß diese Wechselwirkung auch für das Spleißen wichtig ist<sup>[69]</sup>. Schon früh in der Spleißeaktion erkennen U1- und U2-snRNAs Consensus-Sequenzen an der 5'-Spleiß- und an der Verzweigungsstelle (Abb. 9a). Zu diesem Zeitpunkt sind U4- und U6-snRNAs durch eine ausgedehnte komplementäre Region aneinander gebunden<sup>[70]</sup>. Die

snRNA U6 unterscheidet sich dadurch von den anderen snRNAs, daß sie nicht durch core-Peptide, die von Sm-Lupus-Antiseren erkannt werden, gebunden wird; außerdem wird sie von einer anderen Polymerase transkribiert und ist am 5'-Terminus anders modifiziert. Die U6-snRNA ist zudem das am höchsten konservierte Mitglied der snRNA-Familie und könnte eine besondere katalytische Rolle spielen. Das U4/U6-snRNP bindet das U5-snRNP und der entstandene spezifische tri-snRNP-Komplex bildet mit anderen Komponenten wahrscheinlich das Spleißosom<sup>[71]</sup>.

Nach der Bildung des Spleißosoms erfolgen wichtige Umorganisationen der snRNAs (Abb. 9b). Genetische und biochemische Untersuchungen zeigen, daß die U 6- mit der U 2-snRNA durch eine ausgedehnte komplementäre Region im Spleißosom gepaart sind<sup>172, 731</sup>. Zur Paarung der U 2- und U 6-Strukturen muß die U4-snRNA von der U 6-snRNA abdissoziieren. Tatsächlich kann die U 4-snRNP nach dieser Umwandlung vom aktiven Spleißosom abgegeben werden. Die komplementären Sequenzen zwischen der U 1-snRNA und der 5'-Spleißstelle dissoziieren wahrscheinlich auch vor dem ersten Spleißschritt und die 5'-Spleißstelle paart sich mit einer anderen Region der U 6-snRNA<sup>174, 751</sup>. Man glaubt, daß die U 5-snRNA wichtig für die Erkennung der Exon-Sequenzen ist, die den Spleißstellen direkt benachbart liegen. Mutationen in diesen benachbarten Sequenzen haben starke Auswirkungen auf die Spleißeffizienz, und diese Effekte können durch komplementäre Veränderungen in den U 5-snRNAs verringert werden<sup>1761</sup>. Es ist möglich, daß es zu weiteren Umorganisationen in den von den snRNAs gebildeten Sekundärstrukturen zwischen dem ersten und dem zweiten Spleißschritt kommt. Einige Mutationen in U 2 und U 6 (Positionen mit \* gekennzeichnet), inaktivieren nur den zweiten Schritt<sup>172, 731</sup>. Betrachtet man das Netzwerk der komplementären Sequenzen zwischen snRNAs und der prä-mRNA im Spleißosom genauer, findet sich eine Verdichtung von snRNA-Strukturen – möglicherweise eine Tertiärstruktur – nahe der Verzweigungstelle und den Sequenzen der 5'-Spleißstellen. Diese Anordnung von snRNAs könnte das katalytische Zentrum für den ersten Schritt sein. Weitere Umorganisationen führen möglicherweise zu einer weiteren katalytischen Stelle für den zweiten Schritt.

## Die Bildung des Spleißosoms

Einige stabile Komplexe, die snRNAs und prä-mRNA enthalten, wurden teilweise charakterisiert und ein Spleißosomzyklus entworfen (Abb. 10)<sup>1771</sup>. Der Komplex CC bildet sich an der prä-mRNA durch Erkennung der Sequenz der 5'-Spleißstelle durch U 1-snRNP und der Sequenzen der Verzweigungs- und der 3'-Spleißstelle<sup>178–801</sup>. Für Säugetiersysteme ist eine Pyrimidinsequenz nahe der 3'-Spleiß- und der Verzweigungstelle besonders wichtig. Die darauffolgende Bindung von U 2-snRNP an die prä-RNA führt zur Bildung des sehr stabilen Komplexes A und hängt entscheidend vom Protein U 2 AF ab<sup>1811</sup>, das die Polypyrimidinsequenz durch prototypische RNA-bindende Domänen erkennt und Wechselwirkungen mit den anderen Komponenten des Spleißens durch Arg-Ser-Wiederholungssequenz vermittelt. Interessanterweise enthält das – auch für die stabile U 2-snRNP-Bindung nötige – U 1-snRNP ein gebundenes Protein mit einer Arg-Ser-Wiederholungssequenz, das U 1-70-kd-Polypeptid. Ob U 2 AF und das U 1-70-kd-Polypeptid direkt über das Intron kommunizieren, ist nicht klar.

Das B1-Spleißosom bildet sich, wenn der U 4/U 6/U 5-snRNA-Komplex an den Komplex A bindet. Wahrscheinlich dissoziiert im Komplex B1 die U 4-snRNA von der U 6-snRNA und auch die U 1-snRNA von der 5'-Spleißstelle. Diese Ereignisse definieren den Komplex B 2, welcher der Vorläufer des Komplexes C 1 ist. Der erste Schritt des Spleißens findet nun im Komplex C 1 statt. Für die Entstehung des Komplexes C 2 sind wahr-

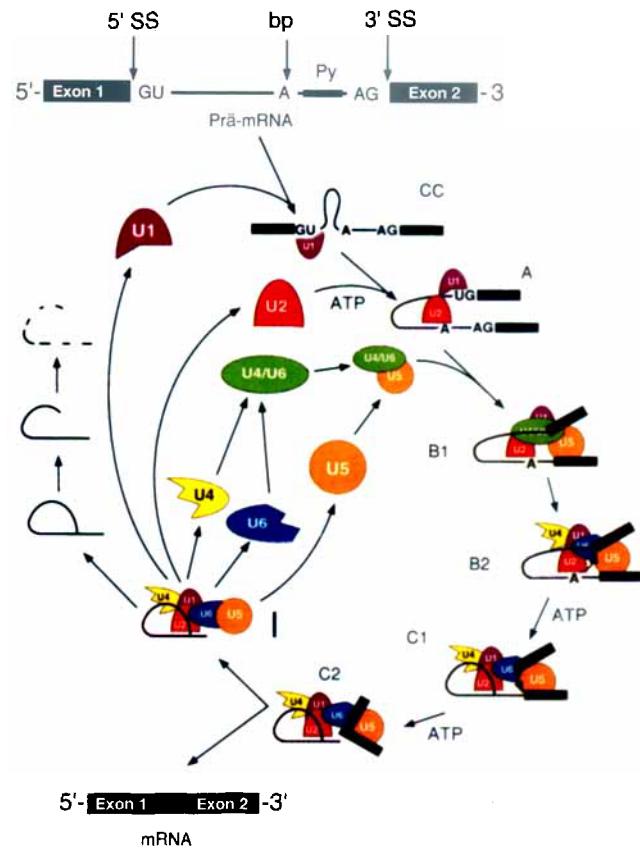


Abb. 10. Schematische Darstellung des Spleißosomzyklus unter Berücksichtigung der Rolle von kleinen nucleären Ribonucleoprotein(snRNP)-Partikeln bei prä-mRNA-Spleißen. Prä-mRNA (oben), die zwei durch ein Intron getrennte Exons enthält, tritt in die snRNP-Spleiß-Komplexe ein und verläßt sie als mRNA (unten) und als herausgeschnittenes Lasso-Intron (links). Andere, nicht-snRNP-Faktoren werden für die Spleißosombildung benötigt, wurden aber der Einfachheit halber weggelassen. CC, A, B1, B2, C1, C2 und I sind Komplexe innerhalb des Spleißzyklus, die biochemisch und/oder genetisch unterschieden werden konnten. 5'-SS, 3'-SS, bs und Py bezeichnen 5'- und 3'-Spleißstellen, Verzweigungsstellen bzw. Polypyrimidinsequenzen und U 1, U 2, U 4, U 5 sowie U 6 die einzelnen snRNPs.

scheinlich weitere Umorganisationen notwendig, um den zweiten Spleißschritt zu katalysieren. Der Komplex C 2 dissoziiert in die frisch verbundenen Exons und den snRNP-Intron-Komplex I. In der Zelle wandert dann das gespleißte Produkt ins Cytoplasma. Komplex I wird zurückgehalten und umgesetzt: Die snRNPs werden für weiteres Spleißen in vivo wiederverwertet, die Lasso-Intron-RNA wird abgebaut. Da alle snRNAs sehr stabil sind, durchlaufen sie wahrscheinlich viele Spleißzyklen.

Man sollte erwarten, daß am Zusammenbau und an der Umorganisation der Komplexe, die so kompliziert sind wie die des Spleißosomzyklus, einige Proteine beteiligt sind. Durch ausführliche genetische Analyse wurden mehrere Genprodukte aus Hefe identifiziert, die für das Spleißen in vivo oder in vitro bedeutsam sind<sup>182, 831</sup>. Temperaturempfindliche Mutationen in Precursor-RNA-processing(PR P)-Genen verursachen eine Anreicherung von ungespleißter prä-mRNA im Kern und/oder erzeugen Extrakte, die für das Spleißen in vitro ungeeignet sind (Abb. 11). Man schätzt, daß mindestens 100 Gene Produkte codieren, die für das DNA-Spleißen von Bedeutung sind; das entspricht etwa 2% des gesamten Hefe-Genoms. Der Spleißapparat ist also eine bedeutende Komponente des Kerns. Die Aminosäuresequenzen, die für einige der PR P-Gene vorhergesagt wurden, haben vermutlich spezifische Funktionen, wie RNA-Bindung

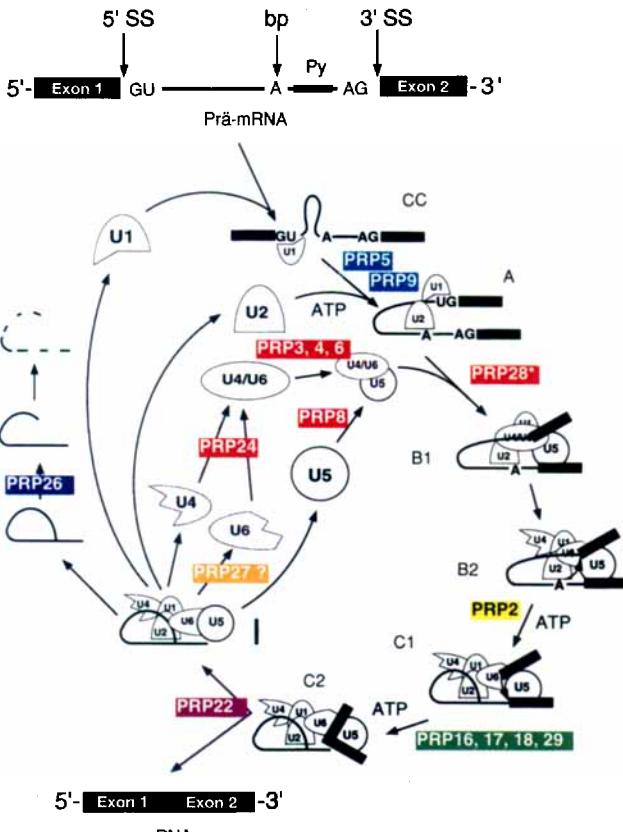


Abb. 11. Übergänge im Spleißosomzyklus, für die ein PRP-Protein benötigt wird. Die erforderliche PRP-Mutante ist an den Pfeilen angegeben, die die Übergänge im Zyklus *in vitro* zeigen.

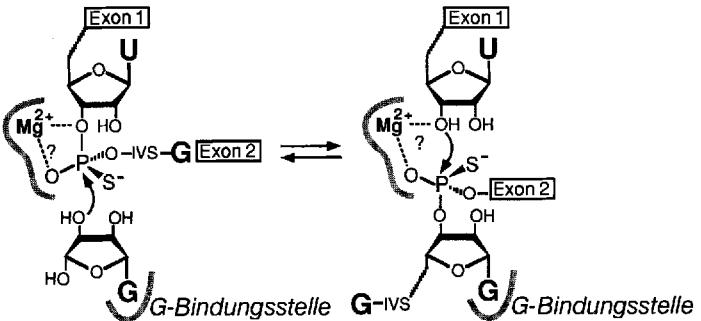
und RNA-Helicase- sowie Protein-Protein-Wechselwirkungen. Es ist eine herausfordernde Aufgabe, diese hypothetischen Funktionen den Vorgängen in Spleißosomen zuzuordnen.

Die Schritte im Spleißosomzyklus, für die besondere PRP-Proteine benötigt werden (Abb. 11), stimmen mit dem Zyklus überein, der durch kinetische und biochemische Methoden bestimmt wurde. Die meisten Übergänge zwischen spezifischen Formen des Spleißosoms benötigen ein oder mehrere spezifische Proteine. Zudem spleißen einige PRP-Mutanten nicht, da sie snRNPs nicht wieder zusammenbauen können, um weiteres Spleißen zu ermöglichen. Es beweisen also sowohl genetische als auch biochemische Untersuchungen, daß der Spleißosomzyklus für die Excision von Introns aus diskontinuierlichen Genen verantwortlich ist.

## Die Chemie des Spleißosoms

Beide Schritte im Spleißosomzyklus sind Reaktionen zwischen Hydroxygruppen und einer Phosphodiesterbindung. Die Produkte und Substrate beider Schritte enthalten die gleiche Zahl von kovalenten Bindungen, weshalb jeder Schritt mit einer einzigen Umesterung erreicht werden kann. Das konnte für die selbstspleißenenden Introns der Gruppe I durch Analyse der Stereochemie der beiden Reaktionen gezeigt werden (Abb. 12)<sup>[84–86]</sup>. Phosphorothioate in Diesterbindungen der RNA sind chiral, sie sind je nach Position des Schwefelatoms entweder  $R_p$ - oder  $S_p$ -konfiguriert. Die Konfiguration kann über die Empfindlichkeit

gegenüber bestimmten Nucleasen festgelegt werden. Wenn eine chirale Phosphorothioatgruppe nur an einer Umesterung beteiligt ist, wird ihre Konfiguration verändert, z.B. von  $R_p$  zu  $S_p$ . So können die Umesterung während eines Prozesses ausgezählt und potentielle kurzlebige Zwischenprodukte entdeckt werden.



Erster Schritt :  $R_p$  aktiv

Zweiter Schritt :  $S_p$  aktiv

Abb. 12. Der stereochemische Verlauf des ersten und zweiten Schritts der selbst-spleißenenden Reaktion von Introns der Gruppe I. Im ersten Schritt ist das Phosphorothioat der 5'-Spleißstelle des Exons 1 in der  $R_p$ -Konfiguration, im zweiten Schritt das Phosphorothioat der 3'-Spleißstelle des Exons 2 in der  $S_p$ -Konfiguration gezeigt. Diese Konformere inhibieren den ersten bzw. zweiten Schritt nicht<sup>[84–86]</sup>. Die Wechselwirkungen zwischen dem Sauerstoffatom des Phosphates und einem  $Mg^{2+}$ -Ion sind hypothetisch und sollen erklären, warum ein Phosphorothioat mit einem Schwefelatom an dieser Stelle nicht reaktiv ist. IVS = Intervenierende Sequenz.

Wie bereits angedeutet, sind  $R_p$ - und  $S_p$ -Phosphorothioate in vielen katalytischen Stellen nicht gleich aktiv. Man schreibt dies der Wechselwirkung zwischen einem Sauerstoffatom des Phosphates und einem Metall-Ion zu; die entsprechende Wechselwirkung eines Schwefelatoms ist weniger günstig. Diese Zusammenhänge wurden für die selbst-spleißenenden Introns der Gruppe I und auch für das Spleißosom untersucht.

Die zwei Spleißschritte durch Introns der Gruppe I erfolgen wahrscheinlich als Hin- und Rückreaktion an einer einzigen katalytischen Stelle (Abb. 12). Im ersten Schritt wird der Guanosin-Cofaktor an die G-Bindungsstelle gebunden und seine 3'-OH-Gruppe für die Umesterung aktiviert. In dieser Reaktion ist das  $R_p$ -konfigurierte Phosphorothioat aktiv und nicht das  $S_p$ -konfigurierte. Das deutet darauf hin, daß das Sauerstoffatom in dieser Reaktion über ein Metall-Ion gebunden wird. Übereinstimmend mit nur einer Umesterung im ersten Schritt wird das  $R_p$ -konfigurierte Zentrum in ein  $S_p$ -konfiguriertes überführt<sup>[84]</sup>. Im zweiten Schritt besetzt die Guanosingruppe an der 3'-Spleißstelle die G-Bindungsstelle und das 3'-OH des 5'-Exons wird für die Umesterung aktiviert. Entsprechend dem Prinzip der mikroskopischen Reversibilität ist das  $S_p$ -konfigurierte Phosphorothioat in dieser Reaktion aktiv<sup>[86]</sup>, weshalb das  $R_p$ -konfigurierte Produkt gebildet wird. Aus der stereochemischen Analyse der Reaktionen der Introns der Gruppe I folgt daher, daß nur ein katalytisches Zentrum für beide Schritte existiert.

Der stereochemische Verlauf der beiden Schritte des Spleißosomprozesses wurde durch Synthese von Substrat-RNAs, die entweder ein  $R_p$ - oder ein  $S_p$ -Phosphorothioat an der 5'- oder der 3'-Spleißstelle enthielten, untersucht<sup>[87]</sup>. Das jeweilige Stereoisomer wurde durch eine Kombination von chemischer Synthese, oligo-geprimeter Transkription und Ligation der

RNAs mit überbrückenden Oligodesoxynucleotiden und T4-DNA-Ligase eingebaut<sup>[88]</sup>. Das Spleißen der Phosphorothioat-enthaltenden RNA-Substrate wurde nach verschiedenen Inkubationszeiten in Kernextrakten geprüft. Die Analyse der Substrat- und Produkt-RNAs ergab, daß a) beide Spleißschritte jeweils nur eine Umesterung mit Inversion am Chiralityzentrum des aktiven Phosphorothioates beinhalten und b) in beiden Schritten das *S<sub>P</sub>*-Diastereomer und nicht das *R<sub>P</sub>*-Diastereomer das aktive Phosphorothioat ist<sup>[87]</sup>.

Ein Vergleich des stereochemischen Verlaufs und der wesentlichen Komponenten der beiden Schritte des Selbst-spleißens von Introns der Gruppe I mit den Reaktionen an Spleißosomen ist aufschlußreich (Abb. 13): Wie schon diskutiert, sind die beiden

entsprechen. Im ersten Schritt wird eine 2'-Hydroxygruppe eines Adenosinrestes, im zweiten Schritt eine 3'-Hydroxygruppe am 5'-Exon aktiviert. Des weiteren ist der konservierte Guanosinrest durch eine 5'-Bindung im ersten Schritt und durch eine 3'-Bindung im zweiten Schritt an das aktivierte Phosphat gebunden. Der Spleißosomprozeß und wahrscheinlich auch der selbst-spleißende Prozeß von Introns der Gruppe II müssen also an zwei getrennten katalytischen Zentren ablaufen.

Das für den zweiten Schritt verantwortliche katalytische Zentrum des Spleißosoms ähnelt wahrscheinlich dem der selbst-spleißenden Introns der Gruppe I. Diese zwei Zentren haben ähnliche Eigenschaften: Beide katalysieren die Umesterung, wobei a) ein *S<sub>P</sub>*- und nicht ein *R<sub>P</sub>*-Phosphorothioat im aktiven Zentrum liegt, b) das aktivierte Phosphat an die 3'-Position eines konservierten Guanosinrestes gebunden ist und c) eine 3'-Hydroxygruppe am 5'-Exon aktiviert wird. Diese Ähnlichkeiten sind wahrscheinlich die Folge gemeinsamer chemischer Strukturelemente in den beiden katalytischen Zentren. Möglicherweise spiegelt diese gemeinsame Mikrostruktur eine ähnliche RNA-Tertiärstruktur der katalytischen Zentren und einen gemeinsamen Entwicklungsgeschichtlichen Ursprung wider.

### Kernstruktur und RNA-Spleißen

Das Prozessieren von Introns erfolgt sowohl aus naszierender RNA, die durch Polymerase verlängert wird, als auch aus post-transkriptioneller Vorläufer-RNA<sup>[89]</sup>. Der Spleißapparat muß also in der Nähe des Gens und in den Regionen zwischen dem Gen und den Kernporen gelegen sein. Nach Beendigung des Spleißens wird die mRNA durch eine der etwa 4000 Kernporen einer typischen Säugetierzelle ins Cytoplasma transportiert<sup>[90]</sup>. Der Mechanismus, durch den nucleäre RNAs vom Transkriptions- und Spleißort zu den Poren transportiert werden, ist nicht geklärt.

Die subnukleäre Lokalisation einiger Proteine und RNAs, die für das RNA-Spleißen wichtig sind, wurde licht- und elektronenmikroskopisch untersucht. Das Auffälligste ist die Konzentration der snRNPs U2, U4/6 und U5 in 20–50 gefleckten Strukturen und auch in 1–5 Zentren, die „coiled bodies“ genannt werden (Abb. 14)<sup>[91, 92]</sup>. Das U1-snRNP ist zwar auch in solchen Zentren und Flecken konzentriert, darüber hinaus jedoch gleichmäßiger über den Kern verteilt. Die gefleckten Strukturen entsprechen Regionen, die von Elektronenmikroskopikern als granuläre Interchromatinhaufen und Perichromatinfibrillennetzwerk beschrieben werden. Durch hochauflösende In-situ-Hybridisierungsmethoden und Pulsmarkierung wurde neu synthetisierte RNA in den gefleckten Regionen und auch in kurvilinearen Anordnungen gefunden, die sich vom Gen zur Kernperipherie hin erstrecken<sup>[93, 94]</sup>. Je weiter entfernt die nucleäre Vorläufer-RNA entlang dieser Anordnungen, die von seinem Ursprungsgen ausgehen, lokalisiert ist, desto niedriger ist die Konzentration der Intron-Sequenzen im Vergleich zu der der Exon-Sequenzen. Das deutet darauf hin, daß die Vorläufer-RNA durch die gefleckten Strukturen hindurchläuft, während sie prozessiert und transportiert wird<sup>[94]</sup>. In diesen Regionen sind wahrscheinlich aktive Spleißosome konzentriert.

Neben den snRNPs sind auch einige für das Spleißen wichtige Proteine in den gefleckten Strukturen konzentriert. Dazu gehö-

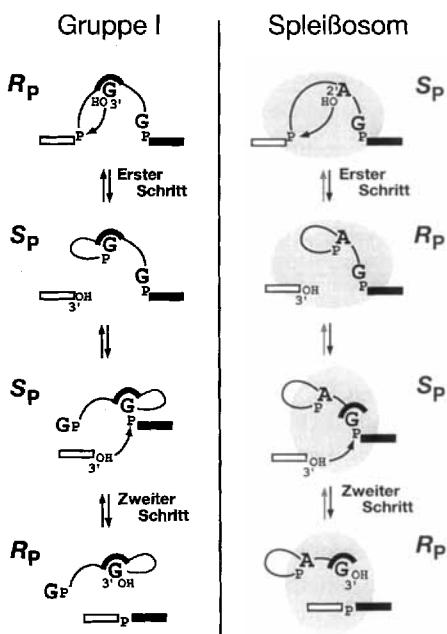


Abb. 13. Chemische Mechanismen und Konfigurationsangaben zu den beiden Schritten des Selbst-Spleißens der Introns der Gruppe I (links) und der des Spleißosomalen prä-mRNA-Spleißens (rechts). Unterschiedlich sind im ersten Schritt unter anderem die Phosphorothioatdiastereomerpräferenzen, die Nucleophile (2'-OH bzw. 3'-OH) sowie die Platzierung (5' bzw. 3') des konservierten Guanosinrestes (G) bezüglich des Phosphatrestes an der 5'-Spleißstelle. Ähnlich sind im zweiten Schritt die Präferenz für das *S<sub>P</sub>*-Phosphorothioatdiastereomer, der Einsatz einer 3'-OH-Gruppe als Nucleophil und die Bindung eines konservierten Guanosins (G) über ein 3'-Sauerstoffatom an das Phosphat an der 3'-Spleißstelle. *R<sub>P</sub>* und *S<sub>P</sub>* kennzeichnen für jeden Schritt das eingesetzte Phosphorothioatdiastereomer, das nicht signifikant inhibitorisch ist, sowie das resultierende Produktdiastereomer. Die gestrichelten Pfeile deuten auf die potentielle, aber bislang nicht beobachtete Reversibilität eines jeden Schrittes des nucleären prä-mRNA-Spleißens hin. Die dunklen Halbkreise, die im Spleißosom für den zweiten Schritt gezeigt werden, bedeuten eine hypothetische Stelle, die der der Introns der Gruppe I ähnelt.

Schritte der Gruppe I Hin- und Rückreaktionen an einem einzigen katalytischen Zentrum. Dies ist sicherlich nicht der Fall bei den Reaktionen an Spleißosomen. Hier ist in beiden Schritten das *S<sub>P</sub>*-Phosphorothioat aktiv, und daher ist der zweite Schritt nicht die Rückreaktion des ersten. Diese Ergebnisse sind in Einklang mit der Vorstellung, daß das Spleißosom zwei katalytische Zentren für die beiden Schritte bereitstellt. Die Bestandteile dieser Zentren können teilweise überlappen, die Zentren müssen getrennt sein. Das würde auch der unterschiedlichen chemischen Natur der Reaktionskomponenten der beiden Schritte

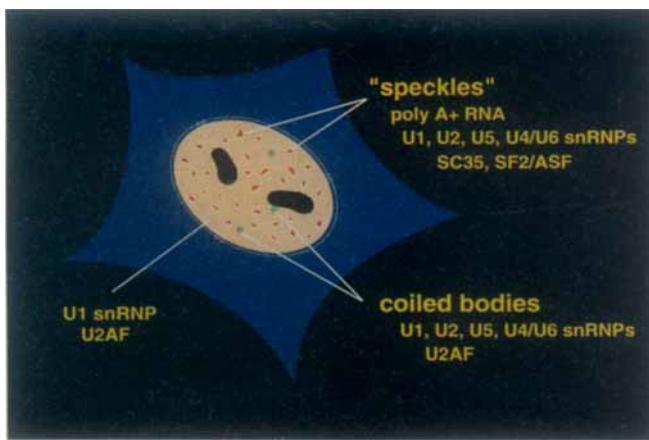


Abb. 14. Subnukleäre Lokalisation der Spleißfaktoren im Kern. Es wurden spezifische Antiseren und In-situ-Hybridisierungsmethoden eingesetzt, um Komponenten des Spleißosoms und Spleißfaktoren im Kern von Säugetierzellen zu lokalisieren. Der Umriss der Zelle wird ohne cytoplasmatische Struktur gezeigt. In der Kernmembran sind drei Kernporen dargestellt. Die großen dunklen Strukturen sind Nukleoli, die Stellen der ribosomalen RNA-Synthese. Einige Komponenten der Flecken („speckles“), d.h. der Regionen, die prä-mRNA enthalten, sind aufgelistet. Die „coiled bodies“ unterscheiden sich davon und scheinen keine prä-mRNAs zu enthalten [93]. Die rosafarbene Schraffierung stellt die nukleäre Verteilung von U1-snRNP und U2AF dar.

ren die Arg/Ser-Proteine SF2/ASF<sup>[95, 96]</sup> und SC35<sup>[97]</sup>. Überraschenderweise sind diese Proteine – wie auch die snRNPs in aktiven Spleißosomen – wahrscheinlich an eine durch ihre Funktion definierte Kernmatrix gebunden. Diese Matrix ist diejenige Struktur, die nach Extraktion fast aller Chromatin-Proteine und DNA durch sequenzielle Behandlungen mit DNase I und einer Salzlösung hoher Konzentration im Kern verbleibt<sup>[96]</sup>. Es handelt sich um eine fibrilläre Matrix, die gegenüber RNase empfindlich ist, was auf eine große Bedeutung der RNA für die Struktur hinweist. Es konnte gezeigt werden, daß Spleißosome mit der Kernmatrix assoziiert sind, da pulsmarkierte endogene Vorläufer-RNA, die als Teil der Matrix extrahiert wurde, prozessiert werden konnte<sup>[98]</sup>.

Aktive Spleißosome könnten an die Kernmatrix über Wechselwirkungen mit Proteinen der Arg/Ser-Familie gebunden sein<sup>[99, 100]</sup>. Eine Reihe von monoklonalen Antikörpern wurde gegen Komponenten der Kernmatrix aus menschlichen Zellen hergestellt. Drei dieser Antikörper färben die Matrix spezifisch und finden sich im wesentlichen bei den snRNPs in den gefleckten Strukturen. Überraschenderweise können mit diesen drei Antikörpern spezifisch Spleißosomkomplexe, die Exon-Sequenzen enthalten, d.h. eine Fraktion der Vorläufer-RNA, die Lasso-Zwischenprodukt-RNA, das assoziierte 5'-Exon und auch die gespleißten Exon-Produkt-RNAs immunpräzipitiert werden<sup>[101]</sup>. Die Antikörper präzipitieren nicht den Intron-enthaltenden Spleißosomkomplex I, obwohl er fast alle snRNPs und assoziierte Proteinkomponenten enthält. Diese bemerkenswerten Ergebnisse deuten darauf hin, daß die Matrixkomponenten an Exon-RNA-Sequenzen binden, die spezifisch zu Spleißkomplexen zusammengesetzt sind.

Die von den monoklonalen Antikörpern gegen die Kernmatrix erkannten Proteine sind höchstwahrscheinlich Arg/Ser-Protein oder mit solchen assoziiert. Vorläufer-RNAs werden von Arg/Ser-freien Kernextrakten nicht immunpräzipitiert. Werden zusätzlich gereinigte Arg/Ser-Proteine hinzugegeben, wird die

spezifische Immunfluoreszenz von subnukleären Strukturen blockiert. Diese Ergebnisse bestätigen ein Modell, nach dem die Arg/Ser-Proteine mit der Kernmatrix assoziiert sind oder sie bilden<sup>[101]</sup>. Neu synthetisierte RNA-Sequenzen werden von den Arg/Ser-Proteinen, wie SC35, ASF, U1-70-kd und U2AF, erkannt und sind an die Matrix möglicherweise über Wechselwirkungen mit den Arg/Ser-Sequenzen gebunden. An diesen Strukturen würden sich vollständige Spleißosome bilden, die die in der Nähe der Matrix gelegenen Spleißstellen umfassen, und das Spleißen durchführen. Die Exon-Produkt-RNAs würden an die Matrix gebunden bleiben und sich durch einen bisher unbekannten Mechanismus zu den Kernporen bewegen.

### Schlussfolgerung

Die Entdeckung diskontinuierlicher Gene und des Spleißen von RNA war entscheidend für das Verständnis der Biologie eukaryotischer Organismen. Das Spleißen von RNA ist wichtig für die Genregulation, die eine zentrale Stellung bei allen biologischen Phänomenen einnimmt. Dies gilt besonders dann, wenn Vorläufer-RNAs über unterschiedliche Wege prozessiert werden, um mRNAs zu bilden, die verschiedene Proteine codieren. Der Mechanismus des Spleißen durch das Spleißosom ist wahrscheinlich dem selbst-spleißenden Prozeß der Introns der Gruppen I und II verwandt. Der Spleißprozeß der Spleißosome ist entwicklungsgeschichtlich alt, vielleicht so alt wie die Translation durch Ribosomen. Man kann sich vorstellen, daß die eukaryotische Zelle aus zwei Kompartimenten besteht, dem Kern, wo die Spleißosome durch RNA-Katalyse RNA-Vorläufer prozessieren, und dem Cytoplasma, wo die Ribosomen durch RNA-Katalyse mRNAs translatieren. Die subnukleäre Lokalisation von Spleißosom-verwandten Komponenten deutet darauf hin, daß der Kern in Kompartimenten organisiert ist. Mit weiteren Untersuchungen zum RNA-Spleißen und -Transport wird die Natur dieser spezifischen Organisation sicher aufgeklärt werden können.

*Ich danke meinen Mitarbeitern über die Jahre (in alphabetischer Reihenfolge): A. S. Baldwin, S. M. Berget, A. J. Berk, K. Berkner, B. Blencowe, M. A. Brown, S. Buratowski, C. Carr, R. W. Carthew, C. Cepko, D. Chang, D. Chasman, L. A. Chodosh, G. Chu, R. G. Clerc, J. D. Crispino, D. J. Donoghue, A. Z. Fire, D. E. Fisher, S. J. Flint, M. Garcia-Blanco, A. Gil, S. Gilbert, P. J. Grabowski, H. Handa, U. Hansen, S. Hardy, S. Harper, T. Harrison, M. Horowitz, P. S. Jat, R. Kaufman, J. Kim, R. Kingston, J. Kjems, T. Kobayashi, M. M. Konarska, T. Kristie, A. I. Lamond, F. Laski, J. LeBowitz, K. LeClair, F. Lee, I. Lemischka, A. MacMillan, R. Marciniak, P. McCaul, R. Meyers, C. Moore, M. Moore, M. Morton, M. Murata, R. Padgett, J. Parvin, J. L. Pomerantz, C. Query, M. E. Samuels, J. Sedivy, S. Seiler, S. Shykind, H. Singh, H. Skolnik-David, M. Timmers, A. Virtanen, J. Weinberger und Q. Zhou.*

Eingegangen am 12. Januar 1994 [A 46]  
Übersetzt von Frau Dr. Christiane Koszka, Berlin

- [1] J. E. Darnell, Jr., *Harvey Lectures* **1975**, *69*, 1–47.
- [2] Y. Furuchi, M. Morgan, S. Muthukrishnan, A. J. Shatkin, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 362–366.
- [3] C. M. Wei, B. Moos, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 3014–3018.
- [4] F. Rottman, A. Shatkin, R. P. Perry, *Cell* **1974**, *3*, 197–199.
- [5] M. Edmonds, M. H. Vaughan, Jr., H. Nakazoto, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1336–1340.

- [6] S. Y. Lee, J. Mendecki, G. Brawerman, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1331–1335.
- [7] J. E. Darnell, Jr., R. Wall, R. J. Tushinski, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1321–1325.
- [8] U. Pettersson, C. Mulder, H. Delius, P. A. Sharp, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 200–204.
- [9] P. A. Sharp, P. H. Gallimore, S. J. Flint, *Cold Spring Harbor Symp. Quant. Biol.* **1974**, *34*, 457–474.
- [10] S. J. Flint, P. H. Gallimore, P. A. Sharp, *J. Mol. Biol.* **1975**, *96*, 47–68.
- [11] S. J. Flint, P. A. Sharp, *J. Mol. Biol.* **1976**, *106*, 749–771.
- [12] S. Bachenheimer, J. E. Darnell, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4445–4449.
- [13] S. M. Berget, C. Moore, P. A. Sharp, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 3171–3175.
- [14] J. Lewis, J. F. Atkins, C. Anderson, P. R. Baum, R. F. Gesteland, *Proc. Natl. Acad. Sci. USA* **1975**, 4445–4449.
- [15] M. Thomas, R. L. White, R. W. Davis, *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 2294–2298.
- [16] J. Casey, N. Davidson, *Nucleic Acid Res.* **1977**, *4*, 1539–1552.
- [17] H. Westphal, J. Meyer, J. Maizel, *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 2069–2071.
- [18] I. T. Chow, J. M. Roberts, J. B. Lewis, T. R. Broker, *Cell* **1977**, *11*, 819–836.
- [19] S. Goldberg, J. Weber, J. E. Darnell, Jr., *Cell* **1977**, *10*, 617–622.
- [20] J. Weber, W. Jelinek, J. E. Darnell, Jr., *Cell* **1977**, *10*, 611–616.
- [21] S. M. Berget, P. A. Sharp, *J. Mol. Biol.* **1979**, *129*, 547–565.
- [22] R. P. Perry, D. E. Kelley, *Cell* **1976**, *8*, 433–442.
- [23] R. E. Gelinas, R. J. Roberts, *Cell* **1977**, *11*, 533–544.
- [24] A. J. Jeffreys, R. A. Flavell, *Cell* **1977**, *12*, 1097–1108.
- [25] S. M. Tilghman, D. C. Tiermeier, J. G. Seidman, B. M. Peterlin, M. Sullivan, J. V. Maizel, P. Leder, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 725–729.
- [26] R. Breathnach, J.-L. Manel, P. Chambon, *Nature* **1977**, *270*, 314–319.
- [27] S. Tonegawa, A. M. Maxam, R. Tizard, O. Bernard, W. Gilbert, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 1485–1489.
- [28] R. Breathnach, P. Chambon, *Annu. Rev. Biochem.* **1981**, *50*, 349–384.
- [29] R. A. Padgett, P. J. Grabowski, M. M. Konarska, S. Seiler, P. A. Sharp, *Annu. Rev. Biochem.* **1986**, *55*, 1119–1150.
- [30] G. R. Fink, *Cell* **1987**, *49*, 5–6.
- [31] R. Treisman, S. Orkin, T. Maniatis, *Nature* **1983**, *302*, 591–596.
- [32] J. M. Nigro, K. R. Cho, E. R. Pearson, S. E. Kern, J. M. Ruppert, J. D. Oliver, D. W. Kinzler, B. Vogelstein, *Cell* **1991**, *64*, 607–613.
- [33] C. Cocquerelle, P. Daubersies, M.-A. Majerus, J.-P. Kerckaert, B. Bailleul, *EMBO J.* **1992**, *11*, 1095–1098.
- [34] R. Pulok, P. Anderson, *Genes Dev.* **1993**, *7*, 1885–1897.
- [35] R. S. Patel, F. Odermatt, J. F. Schwarzbauer, R. O. Hyynes, *EMBO J.* **1986**, *6*, 2565–2572.
- [36] R. O. Hyynes, *Fibronectin*, Springer, New York, **1989**.
- [37] B. S. Baker, *Nature* **1989**, *340*, 521–524.
- [38] M. Tian, T. Maniatis, *Science* **1992**, *256*, 237–240.
- [39] W. Gilbert, *Nature* **1978**, *271*, 501.
- [40] J. E. Darnell, Jr., *Science* **1978**, *202*, 1257.
- [41] W. F. Doolittle, *Nature* **1978**, *272*, 581.
- [42] C. C. F. Blake, *Nature* **1978**, *273*, 267.
- [43] *Intervening Sequences in Evolution and Development* (Hrsg.: E. M. Stone, R. J. Schwartz), Oxford University Press, New York, **1990**.
- [44] N. Lonberg, W. Gilbert, *Cell* **1985**, *40*, 81–90.
- [45] N. Agabian, *Cell* **1990**, *61*, 1157–1160.
- [46] P. Borst, *Ann. Rev. Biochem.* **1986**, *55*, 701–732.
- [47] M. Krause, D. Hirsh, *Cell* **1987**, *49*, 753–761.
- [48] M. Vellard, A. Sureau, J. Soret, C. Martinovic, B. Perbal, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 2511–2515.
- [49] R. Benne, J. van den Burg, J. P. J. Branhoff, P. Sloff, J. H. van Boom, M. C. Tramp, *Cell* **1986**, *46*, 819–826.
- [50] R. A. Padgett, S. F. Hardy, P. A. Sharp, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 5230–5234.
- [51] M. R. Green, T. Maniatis, D. A. Melton, *Cell* **1983**, *32*, 681–694.
- [52] N. Hernandez, W. Keller, *Cell* **1983**, *35*, 89–99.
- [53] P. J. Grabowski, R. A. Padgett, P. A. Sharp, *Cell* **1984**, *37*, 415–427.
- [54] A. R. Krainer, T. Maniatis, B. Ruskin, M. R. Green, *Cell* **1984**, *36*, 993–1005.
- [55] M. M. Konarska, P. J. Grabowski, R. A. Padgett, P. A. Sharp, *Nature* **1985**, *313*, 552–557.
- [56] J. C. Wallace, M. Edmonds, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 950–954.
- [57] P. J. Grabowski, S. R. Seiler, P. A. Sharp, *Cell* **1985**, *42*, 345–353.
- [58] F. Brody, J. Abelson, *Science* **1985**, *228*, 963–967.
- [59] J. A. Steitz, D. L. Black, V. Gerke, K. A. Parker, A. Krämer, D. Frendeway, W. Keller, in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles* (Hrsg.: M. L. Birnstiel), S. 115–154, Springer, New York, **1988**.
- [60] T. R. Cech, *Cell* **1985**, *43*, 713–716.
- [61] C. L. Peebles, P. S. Perlman, K. L. Mecklenburg, M. L. Petrillo, J. H. Tabor, K. A. Jarrell, H.-L. Cheng, *Cell* **1986**, *44*, 213–223.
- [62] R. van der Veen, A. C. Arnberg, G. van der Horst, L. Bonen, H. F. Tabak, I. A. Grivell, *Cell* **1986**, *44*, 225–234.
- [63] P. A. Sharp, *Science* **1991**, *254*, 663.
- [64] G. J. Hannon, P. A. Maroney, J. A. Denker, T. W. Nilsen, *Cell* **1990**, *61*, 1247–1255.
- [65] K. R. Watkins, J. M. Dungan, N. Agabian, *Cell* **1993**, im Druck.
- [66] M. R. Lerner, J. A. Boyle, S. M. Mount, S. L. Wohin, J. A. Steitz, *Nature* **1980**, *283*, 220–224.
- [67] J. Rogers, R. Wall, *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 1877–1879.
- [68] R. A. Padgett, S. M. Mount, J. A. Steitz, P. A. Sharp, *Cell* **1983**, *35*, 101–107.
- [69] R. Parker, P. G. Siliciano, C. Guthrie, *Cell* **1987**, *49*, 229–239.
- [70] C. Guthrie, B. Patterson, *Annu. Rev. Genet.* **1988**, *22*, 387–419.
- [71] M. M. Konarska, P. A. Sharp, *Cell* **1987**, *49*, 763–774.
- [72] H. D. Madhani, C. Cuthrie, *Cell* **1992**, *71*, 803–817.
- [73] D. S. McPhee, J. Abelson, *Cell* **1992**, *71*, 819–831.
- [74] H. Sawa, J. Abelson, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11269–11273.
- [75] D. A. Wassarman, J. A. Steitz, *Science* **1992**, *257*, 1918–1925.
- [76] A. Newman, C. Norman, *Cell* **1992**, *68*, 743–754.
- [77] M. J. Moore, C. C. Query, P. A. Sharp, in *The RNA World* (Hrsg.: R. F. Gesteland, J. F. Atkins), Cold Spring Harbor Laboratory Press, **1993**, S. 303–357.
- [78] B. Seraphin, M. Rosbash, *EMBO J.* **1991**, *10*, 1209–1216.
- [79] S. Michaud, R. Reed, *Genes Dev.* **1991**, *5*, 2534–2546.
- [80] S. F. Jamison, M. A. Garcia-Blanco, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5482–5486.
- [81] B. Ruskin, P. D. Zamore, M. R. Green, *Cell* **1988**, *52*, 207–219.
- [82] S. W. Ruby, J. Abelson, *Trends Genet.* **1988**, *7*, 79–85.
- [83] C. Guthrie, *Science* **1991**, *253*, 157–163.
- [84] J. A. McSwigan, T. R. Cech, *Science* **1989**, *244*, 679–683.
- [85] J. Rajagopal, J. A. Doudna, J. W. Szostak, *Science* **1989**, *244*, 692–694.
- [86] E.-R. Suh, R. B. Waring, *Nucleic Acids Res.* **1992**, *20*, 6303–6309.
- [87] M. J. Moore, P. A. Sharp, *Nature* **1993**, *365*, 364–368.
- [88] M. J. Moore, P. A. Sharp, *Science* **1992**, *256*, 992–997.
- [89] A. L. Beyer, Y. N. Osheim, *Genes Dev.* **1988**, *2*, 754–765.
- [90] H. McHlia, B. Dancholt, U. Skoglund, *Cell* **1992**, *69*, 605–613.
- [91] D. L. Spector, *Annu. Rev. Cell. Biol.* **1993**, *9*, 265–315.
- [92] U. Nyman, H. Hallman, G. Hadlochzy, I. Pettersson, G. Sharp, N. R. Rinzel, *J. Cell. Biol.* **1986**, *102*, 137–144.
- [93] K. C. Carter, D. Bowman, W. Carrington, K. Fogarty, J. A. McNeil, F. S. Fay, J. B. Lawrence, *Science* **1993**, *259*, 1330–1334.
- [94] Y. Xing, C. V. Johnson, P. R. Dobner, J. B. Lawrence, *Science* **1993**, *259*, 1326–1330.
- [95] H. Ge, J. L. Manley, *Cell* **1990**, *62*, 25–34.
- [96] A. R. Krainer, G. C. Conway, D. Kozak, *Cell* **1990**, *62*, 35–42.
- [97] X.-D. Fu, T. Maniatis, *Nature* **1990**, *343*, 437–441.
- [98] S. Zeitlin, R. C. Wilson, A. Efstratiadis, *J. Cell. Biol.* **1989**, *108*, 765–777.
- [99] M. B. Roth, C. Murphy, J. G. Gall, *J. Cell. Biol.* **1990**, *111*, 2217–2223.
- [100] M. B. Roth, A. M. Zahler, J. A. Stolk, *J. Cell. Biol.* **1991**, *115*, 587–596.
- [101] B. J. Blencowe, J. A. Nickerson, R. Issner, S. Penman, P. A. Sharp, unveröffentlicht.