

200HB generator. The native-2 data set was collected with a CCD detector at the A1 beamline of the Cornell High Energy Synchrotron Source (MacCHESS). The structure was determined by a combination of multiple isomorphous replacement (MIR) and molecular replacement (MR) methods. The position of the p53 core domain was determined by molecular replacement (MR) with the program X-PLOR (26), and the structure of the core domain from the p53-DNA complex (15). The correlation coefficient was 0.45 for 10 Å to 4 Å data. Heavy atom soaks were performed in well buffer lacking DTT, containing either 0.5 mM HgCl<sub>2</sub> or 0.5 mM lead acetate, for 4 and 24 hours, respectively. The MIR analysis with the CCP4 program suite (27) included anomalous scattering from the HgCl<sub>2</sub> derivative and had a mean figure of merit of 0.50 for 20.0 Å to 3.2 Å data. Initial maps were calculated by combining the MIR and the MR p53 phases with the program SIGMAA (27) and showed interpretable density for 53BP2. Successive rounds of model building, simulated annealing refinement with the program X-PLOR and phase combination allowed the complete interpretation of the 53BP2 structure. The model was then further refined by least squares refinement with the program TNT (28). In the crystals, the 53BP2 molecule has an overall temperature factor of 64.0 Å<sup>2</sup>, which is significantly higher than that of the p53 core domain (35.0 Å<sup>2</sup>). Consequently, the overall 53BP2 electron density is of lower quality than that of p53. The final model consists of residues 97 to 287 of

p53, residues 327 to 519 of 53BP2 and 275 water molecules. 53BP2 residues 291 to 326 from the NH<sub>2</sub>-terminus, residues 491 to 495 of the n-src loop of the SH3 domain, and residues 436 to 441 in the linker between the α helices of the fourth ankyrin repeat have poor electron density and are likely to be disordered in the crystals.

26. A. T. Brunger, *X-PLOR, a system for crystallography and NMR, Version 3.0 Manual* (Yale Univ. Press, New Haven, CT 1991).
27. Collaborative Computational Project, No. 4. *Acta Crystallogr.* **D50**, 760 (1994).
28. D. E. Tonrud, L. F. Ten Eyck, B. W. Matthews, *ibid.* **A43**, 489 (1987).
29. We thank S. Fields for providing us with the 53BP2 expression plasmid; the staff of the Cornell High Energy Synchrotron Source MacChess for help with data collection; S. Geromanos of the Sloan-Kettering Microchemistry Facility for NH<sub>2</sub>-terminal sequence analyses; W. Farley for help with the surface plasmon resonance experiments; M. Mayhew for helpful discussions; and R. Kenny for administrative assistance. Supported by the NIH (CA65698), the Pew Charitable Trusts, the Arnold and Mabel Beckman Foundation, the Dewitt Wallace Foundation, and the Samuel and May Rudin Foundation. Coordinates have been deposited with the Brookhaven Protein Data Bank (code 1YCS).

28 May 1996; accepted 5 September 1996

## U1-Mediated Exon Definition Interactions Between AT-AC and GT-AG Introns

Qiang Wu and Adrian R. Krainer\*

A minor class of metazoan introns has well-conserved splice sites with 5'-AU-AC-3' boundaries, compared to the 5'-GU-AG-3' boundaries and degenerate splice sites of conventional introns. Splicing of the AT-AC intron 2 of a sodium channel (SCN4A) precursor messenger RNA *in vitro* did not require inhibition of conventional splicing and required adenosine triphosphate, magnesium, and U12 small nuclear RNA (snRNA). When exon 3 was followed by the 5' splice site from the downstream conventional intron, splicing of intron 2 was greatly stimulated. This effect was U1 snRNA-dependent, unlike the basal AT-AC splicing reaction. Therefore, U1-mediated exon definition interactions can coordinate the activities of major and minor spliceosomes.

Precursor mRNA (pre-mRNA) splicing requires the stepwise assembly of the U1, U2, U4/U6:U5 small nuclear ribonucleoprotein (snRNP) particles, and numerous proteins on the substrate to form a spliceosome, which catalyzes two sequential trans-esterification reactions (1). Most introns have 5' and 3' splice sites that match degenerate consensus sequences, of which the 5'-GT and AG-3' intron ends are nearly invariant features. Many point mutations that result in a variety of human genetic diseases affect these dinucleotides and seriously impair gene expression (2). Mutation of G to A at the 5' end or of G to C at the 3' end is especially deleterious, but the double mutation restores splicing (3). Recently, a minor

class of natural pre-mRNA introns with 5'-AT and AC-3' boundaries emerged as an exception to the GU-AG rule (4, 5) (Table 1). The mouse cyclin-dependent kinase CDK5 and the human voltage-gated skeletal muscle and cardiac muscle sodium channel α subunits (SCN4A and SCN5A) also contain AT-AC introns (Table 1) (6).

Most AT-AC introns share highly conserved sequences at the 5' splice site, at the presumptive branch site, and a shorter sequence at the 3' splice site (Table 1). They lack the extensive polypyrimidine tract characteristic of most conventional 3' splice sites, and the distance between the presumptive branch site and the 3' splice site junction is unusually short. Intron 21 of human SCN4A and intron 25 of human SCN5A, which interrupt homologous positions of the coding sequence, also begin with AT and end with AC; however, the rest of their 5' splice site and putative

branch site sequences does not match the consensus, which may indicate that these two introns belong to a distinct subclass of AT-AC introns. There are no obvious structural relations or common expression patterns among the genes or gene families that contain AT-AC introns, and neither the length nor the position of AT-AC introns is conserved among unrelated genes. However, AT-AC introns and their positions within specific genes are conserved phylogenetically. Two members of the gene family of voltage-gated sodium channels, SCN4A and SCN5A, contain AT-AC introns at homologous positions, and it is likely that these introns will prove to be conserved in other sodium and perhaps calcium channel family members (7).

A pre-mRNA consisting of the SCN4A AT-AC intron 2, the two flanking exons, and nine nucleotides (nt) of the conventional intron 3 (Fig. 1) was transcribed *in vitro* (8). Optimal conditions for splicing of this substrate in HeLa cell nuclear extracts were established (9). The maximal splicing efficiency was 11% (Fig. 1A). SCN4A spliced less efficiently and with much slower kinetics than β-globin. The extent of splicing did not increase beyond 6 hours (10). The fidelity of splicing was verified by sequencing of the gel-purified mRNA after reverse transcriptase-polymerase chain reaction (RT-PCR) amplification (10, 11). In addition to the major AT-AC spliced product, a minor product accumulated that corresponds to conventional GT-AG splicing by means of cryptic 5' and 3' splice sites located in exons 2 and 3, respectively (Fig. 1B), as verified by sequencing (10, 11). RNA molecules corresponding to free exon 2 and lariat-exon 3 intermediates, and to released intron lariat, also appeared with appropriate kinetics (Fig. 1A). As expected, the lariat molecules had anomalously slow mobility on higher percentage gels (10, 12).

Recently, accurate *in vitro* splicing of the human P120 AT-AC intron was reported and could only be detected upon debilitation of U1 or U2 snRNPs with complementary 2'-O-methyl oligonucleotides (13). Because efficient accessibility of snRNPs requires incubation with adenosine triphosphate (ATP) (14), this approach precluded a study of the ATP requirement for AT-AC splicing. Because SCN4A AT-AC splicing can be detected in untreated extracts (Fig. 1A), we examined the requirement for ATP and magnesium and found that both are essential for AT-AC splicing (10).

To examine the role of major and minor snRNPs in SCN4A AT-AC splicing *in vitro*, we used oligonucleotide-directed ribonuclease H (RNase H) cleavage of the snRNAs (14), and the resulting extracts

Cold Spring Harbor Laboratory, Post Office Box 100, Cold Spring Harbor, NY 11724-2208, USA.

\*To whom correspondence should be addressed. E-mail: krainer@cshl.org

were used for splicing (15). We tested several oligonucleotides complementary to each of the five major snRNPs involved in conventional splicing, to the minor U11 and U12 snRNAs, and to the 7SK RNA (16). With the SCN4A pre-mRNA, the U12 and U2 oligonucleotides had very different effects (Fig. 2): the U12 oligonucleotide inhibited AT-AC intron splicing completely, whereas the U2 oligonucleotide was essentially neutral. The opposite effect was observed with the control  $\beta$ -globin pre-mRNA: U2 snRNA cleavage inhibited splicing, whereas U12 cleavage had no effect. Moreover, in the same reactions with SCN4A pre-mRNA, the smaller mRNA product generated by use of cryptic GT-AG splice sites disappeared upon cleavage of U2 snRNA, but was not affected by U12 cleavage. This result confirms that this aberrant splicing event is mediated by the conventional GT-AG pathway and serves as an internal control. Similar experiments with U4- and U6-specific oligonucleotides gave identical results to those obtained with U2 (10), indicating that the minor U12 snRNP is required for SCN4A AT-AC splicing, whereas the major U2 and U4/U6 snRNPs are not. The requirement of U12 snRNA for SCN4A AT-AC splicing is consistent with *in vivo* (17) and *in vitro* (13) data obtained with the AT-AC intron of the P120 gene.

The role of U11 and U5 snRNPs in the AT-AC splicing pathway could not be examined, because these snRNAs are resistant to RNase H (18). However, because most of U11 snRNA is complexed with U12 (19), it is likely to participate in AT-AC splicing, as postulated (5). Whereas multiple roles for the U11/U12 snRNP have been reported (5, 13, 17, 20), the role of the nucleoplasmic 7SK RNA has remained elusive. Like U6 snRNA, the 7SK RNA is transcribed by RNA polymerase III and has a

**Table 1.** Compilation of known AT-AC introns. P120, proliferating cell nucleolar antigen; CMP, cartilage matrix protein; SCN4A, voltage-gated skeletal muscle (type IV) sodium channel  $\alpha$  subunit; also conserved in the rat homolog (6); SCN5A, voltage-gated cardiac sodium channel  $\alpha$  subunit (6); CDK5, neuronal cyclin-dependent kinase (6); Rep-3, homolog of the bacterial DNA mismatch-repair protein MutS; Prospero, *Drosophila* homeodomain protein involved in neurogenesis. Intron sequences are shown in lowercase, flanking exon sequences in uppercase, and conserved sequence elements in bold. The dots indicate sequences that are either not shown or not known. The dashes denote 13 nt of the SCN4A intron 21, which are not shown. The question marks indicate that the corresponding sequences have not been determined or that the size of the intron is not yet known. The distance between the presumptive branch adenosine and the 3' splice junction varies from 9 to 12 nt in the examples shown, except in intron 21 of SCN4A, where it is predicted to be 21 nt. For conventional metazoan introns, the usual distance is 18 to 40 nt, although in some cases it can be 150 nt (1).

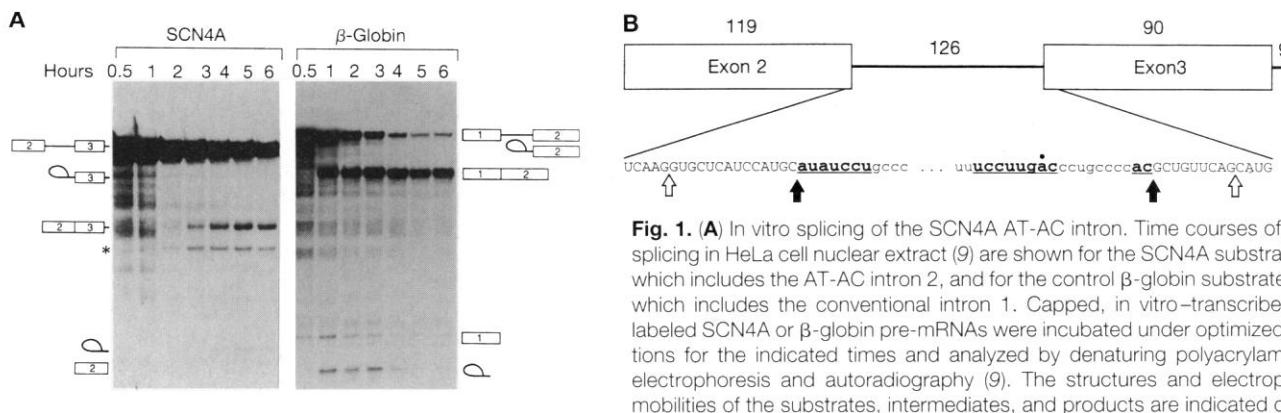
Species	Gene	Intron number	5' splice site	Presumptive branch site*	3' splice site	Size (nt)
Human	P120†	6	GG <b>at</b> atcctt. . .gttccttaac	aggccac	AT	99
Human	CMP†	7	GC <b>at</b> atcctt. . .tctccttaactctgagtcac	TCG	TG	644
Human	SCN4A	2	GC <b>at</b> atcctg. . .tttccttgac	cctgccccac	GC	126
Human	SCN5A	3	TC <b>at</b> atcc. . . ? . . .cccacgcac	GC	?	?
Human	SCN4A	21	AG <b>at</b> gagtat. . .tcaacctgac	----actatac	TT	800
Human	SCN5A	25	AG <b>at</b> acgt. . . ? . . .tctttgac	TT	?	?
Mouse	CDK5	9	CG <b>at</b> atcctc. . . ? . . .acatggacac	AC	440	440
Mouse	REP-3†	6	AG <b>at</b> atcctt. . .tttctttaat	cattactac	AT	?
Fly	Prospero†	2	CT <b>at</b> atcctt. . .aatccttgac	tcctttgac	TC	?
	AT-AC consensus		<b>at</b> atcct	<b>tccttrac</b>	<b>yac</b>	
	GT-AG consensus		AG gtaagt	ynytray	Y <sub>11</sub> nyag	GT

\*The underlined a is the branch nucleotide; in the case of AT-AC introns it has only been mapped so far for the human P120 intron 6 (13). †Previously compiled in (5); the corresponding AT-AC introns are also present in dog, mouse, and monkey P120, and in chicken CMP (5).

$\gamma$ -monomethylphosphate 5' cap (21). Regions of 7SK complementary to U4 and U11 snRNAs have been noted (16). We found that there is perfect complementarity between the sequence AAGGGUAAU, at position 186 of the 7SK RNA, and the AT-AC 5' splice site consensus; this complementarity is more extensive than the proposed pairing between U11 snRNA and the AT-AC 5' splice site (5). However, none of several 7SK-complementary oligonucleotides we tested resulted in specific inhibition of SCN4A splicing (10).

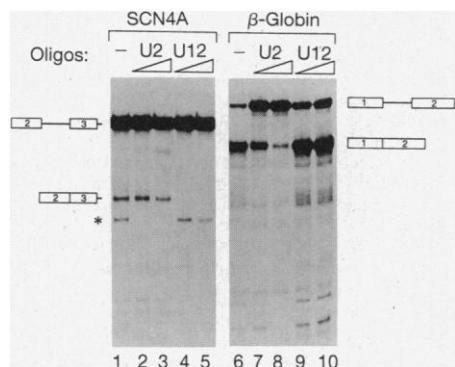
The U1 snRNP is involved in recognition of conventional 5' splice sites through base pairing (22), but has no significant

complementarity to AT-AC intron 5' splice sites (5). When the 5' end of U1 snRNA was cleaved to completion by RNase H and a complementary oligonucleotide, SCN4A AT-AC splicing was strongly but not completely inhibited (Fig. 3A). We reasoned that the requirement for intact U1 snRNA may be related to the presence of the short 5' segment of the conventional intron 3 at the 3' end of the pre-mRNA, which should be sufficient to bind U1 snRNP even in the absence of a downstream 3' splice site partner (23). To test this hypothesis, we constructed a second SCN4A template, in which the 5' splice site of the intron 3 segment was mutated to



**Fig. 1.** (A) *In vitro* splicing of the SCN4A AT-AC intron. Time courses of *in vitro* splicing in HeLa cell nuclear extract (9) are shown for the SCN4A substrate (left), which includes the AT-AC intron 2, and for the control  $\beta$ -globin substrate (right), which includes the conventional intron 1. Capped,  $^{32}$ P-labeled SCN4A or  $\beta$ -globin pre-mRNAs were incubated under optimized conditions for the indicated times and analyzed by denaturing polyacrylamide gel electrophoresis and autoradiography (9). The structures and electrophoretic mobilities of the substrates, intermediates, and products are indicated on each side. An aberrantly spliced SCN4A mRNA arising from use of conventional cryptic 5' and 3' splice sites is indicated by an asterisk. (B) Schematic structure of the SCN4A pre-mRNA substrate, showing the nucleotide sequence surrounding the exon-intron boundaries. The exon and intron sizes in nucleotides are indicated at the top. Exon sequences are in uppercase, intron sequences in lowercase; the consensus elements are bold and underlined, and the presumptive branch indicated by a dot. The solid arrows show the 5' and 3' splice AT-AC cleavage sites and the open arrows show the conventional cryptic sites. Translation of the aberrant mRNA would result in frameshifting and premature termination within exon 3.

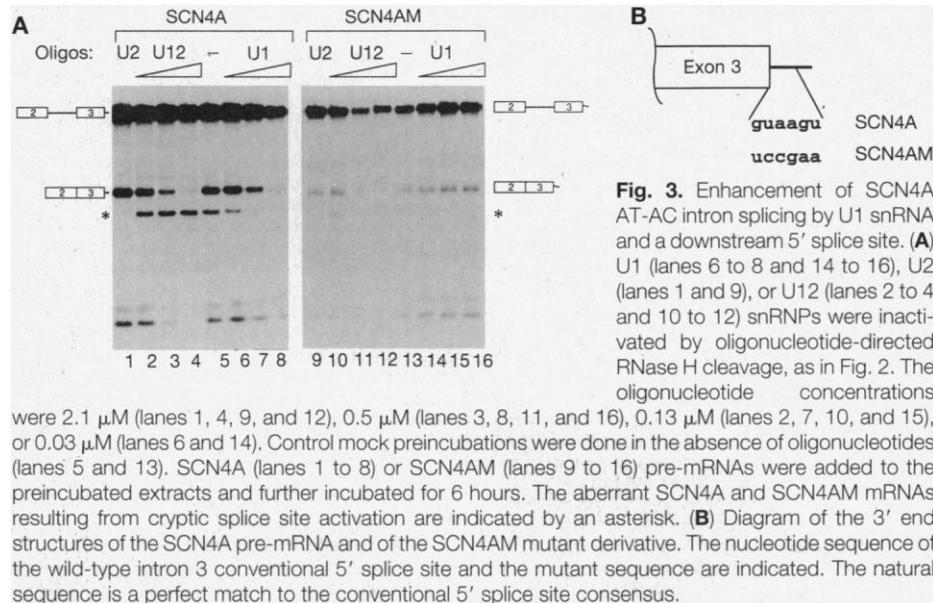
side. An aberrantly spliced SCN4A mRNA arising from use of conventional cryptic 5' and 3' splice sites is indicated by an asterisk. (B) Schematic structure of the SCN4A pre-mRNA substrate, showing the nucleotide sequence surrounding the exon-intron boundaries. The exon and intron sizes in nucleotides are indicated at the top. Exon sequences are in uppercase, intron sequences in lowercase; the consensus elements are bold and underlined, and the presumptive branch indicated by a dot. The solid arrows show the 5' and 3' splice AT-AC cleavage sites and the open arrows show the conventional cryptic sites. Translation of the aberrant mRNA would result in frameshifting and premature termination within exon 3.



**Fig. 2.** Requirement of U12 snRNA for SCN4A AT-AC intron splicing. U2 or U12 snRNPs were inactivated by cleavage of their snRNA moieties with RNase H in the presence of complementary oligonucleotides. The nuclear extract was preincubated for 15 min in the absence of oligonucleotides (lanes 1 and 6) or in the presence of a U2-specific oligonucleotide (lanes 2, 3, 7, and 8) or a U12-specific oligonucleotide (lanes 4, 5, 9, and 10). The oligonucleotide concentration was 4.2  $\mu$ M (lanes 2, 4, 7, and 9) or 8.4  $\mu$ M (lanes 3, 5, 8, and 10). SCN4A (lanes 1 to 5) or  $\beta$ -globin pre-mRNAs (lanes 6 to 10) were added to the preincubated extracts and further incubated for 6 or 4 hours, respectively. The aberrant mRNA resulting from use of cryptic splice sites is indicated by an asterisk.

an unrelated sequence (Fig. 3B) (8). The resulting pre-mRNA, SCN4AM, was spliced accurately but much less efficiently than the SCN4A pre-mRNA (Fig. 3A) (24). This AT-AC splicing reaction was inhibited by U12 cleavage but not by U2 cleavage, which is the same result obtained with the SCN4A pre-mRNA. Cleavage of U1 snRNA did not inhibit splicing of the SCN4AM pre-mRNA and instead resulted in a slight stimulation of this reaction. Conventional GT-AG splicing by way of the cryptic splice sites was also greatly reduced with the SCN4AM pre-mRNA compared with the SCN4A pre-mRNA, which indicates that this splicing event was also dependent on the downstream 5' splice site. Cryptic splicing of the SCN4AM pre-mRNA was further inhibited by U1 and U2 cleavage, which shows that this splicing event is mediated by the conventional pathway. With the SCN4A pre-mRNA, U1 cleavage inhibited cryptic splicing more severely than AT-AC splicing, presumably because this splicing event requires U1 snRNP to bind to the cryptic 5' splice site, as well as to the intron 3 5' splice site (23).

In short, the AT-AC splicing reaction was markedly stimulated by intact U1 snRNA (Fig. 3A), and this stimulation was entirely dependent on the presence of a wild-type downstream 5' splice site. We conclude that the basal AT-AC splicing reaction is U1-independent and can proceed by intron definition (25), but that this path-



way can be greatly stimulated by exon-definition interactions with U1 snRNP bound at a downstream conventional 5' splice site.

In the conventional splicing pathway, exon-intron boundaries are frequently specified in part through interactions between factors bound at adjacent introns, that is, across an exon, as proposed by the exon-definition model (25) and supported by considerable data (23). In this way, exons can serve as recognition units that allow the spliceosome to correctly define exon-intron boundaries. Protein-protein or protein-snRNP cooperative interactions across exons form functional bridges that promote splicing of adjacent introns. As a result, a downstream 5' splice site can stimulate removal of the upstream intron by binding to U1 snRNP (23). These interactions are strong enough that a 5' splice site can promote trans-splicing when present downstream of a 3' trans-splicing substrate (26). For the terminal introns, the 5' cap structure can stimulate splicing of the first intron (27), whereas U1-mediated interactions with the polyadenylation apparatus can enhance splicing of the last intron (28).

Basal levels of U1-independent AT-AC splicing can be detected in the absence of a downstream 5' splice site, indicating that interactions across the intron are sufficient to define an AT-AC intron, albeit inefficiently. Because the known AT-AC introns are always preceded and usually followed by conventional introns, the question arises whether components of the major and minor spliceosomes can functionally cooperate by establishing exon-definition interactions. Indeed, we have shown that AT-AC and GT-AG splicing reactions can take place under the same conditions, and that

the two types of spliceosomes can synergize. Thus, processing of pre-mRNAs containing both kinds of introns is likely to be a temporally and spatially integrated process in vivo. Our results suggest that U1 interacts cooperatively with U12, either directly or, more likely, through protein splicing factors. In the case of the major splicing pathway, interactions between U1 bound at a downstream 5' splice site and U2 bound at the upstream branch site are thought to be mediated by members of the SR protein family and by U2AF (29). The role of these or analogous factors in AT-AC intron splicing and exon definition interactions remains to be elucidated. The coexistence of AT-AC and GT-AG introns within single transcripts, which are processed by distinct but cooperating machineries, together with the phylogenetic conservation of the AT-AC introns in specific genes, suggest that the AT-AC introns may play important roles in the regulation of tissue-specific or developmental gene expression.

*Note added in proof:* After our manuscript was submitted, intron 2 of the mouse *Scn8a* sodium channel gene was shown to belong to the AT-AC class; the effects of spontaneous mutations on in vivo splicing suggested that intron definition interactions between AT-AC and GT-AG splice sites are unfavorable (30).

## REFERENCES AND NOTES

1. M. J. Moore, C. C. Query, P. A. Sharp, in *The RNA World*, R. F. Gesteland and J. F. Atkins, Eds. (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1993), pp. 303-358; T. W. Nilsen, *Cell* **78**, 1 (1994); H. D. Madhani and C. Guthrie, *Annu Rev. Genet.* **28**, 1 (1994).
2. M. Krawczak, J. Reiss, D. N. Cooper, *Hum. Genet.*

- 90, 41 (1992); K. Nakai and H. Sakamoto, *Gene* **141**, 171 (1994).
3. R. Parker and P. G. Siliciano, *Nature* **361**, 660 (1993); G. Chanfreau, P. Legrain, B. Dujon, A. Jacquier, *Nucleic Acids Res.* **22**, 1981 (1994); A. Deirdre, J. Scadden, C. W. Smith, *EMBO J.* **14**, 3236 (1995).
  4. I. J. Jackson, *Nucleic Acids Res.* **19**, 3795 (1991).
  5. S. L. Hall and R. A. Padgett, *J. Mol. Biol.* **139**, 357 (1994).
  6. I. Kiss *et al.*, *J. Biol. Chem.* **264**, 8126 (1989); R. N. Jenkins *et al.*, *ibid.* **265**, 19624 (1990); R. G. Larson *et al.*, *Cancer Commun.* **2**, 63 (1990); A. I. McClatchey *et al.*, *Hum. Mol. Genet.* **1**, 521 (1992); A. L. George Jr., G. S. Iyer, R. Kleinfield, R. G. Kallen, R. L. Barchi, *Genomics* **15**, 598 (1993); T. Ohshima *et al.*, *ibid.* **28**, 585 (1995); Q. Wang, Z. Li, J. Shen, M. T. Keating, *ibid.* **34**, 9 (1996).
  7. The published 5' splice site sequence of intron 2 of human CACNL1A1, the fibroblast voltage-gated L-type calcium channel, precisely matches the AT-AC consensus, whereas the reported sequences of two other calcium channels, CACNL1A2 and CACNL1A3, have GTATCC (rather than ATATCC) at the corresponding 5' splice site; however, all three calcium channel introns reportedly end with the conventional AG 3' splice site [N. M. Soldatov, *Genomics* **22**, 77 (1994); Y. Yamada *et al.*, *ibid.* **27**, 312 (1995); K. Hogan, R. G. Gregg, P. A. Powers, *ibid.* **31**, 392 (1996)]. The sodium and calcium channels are thought to derive from a common ancestral gene: they have considerable nucleotide and amino acid sequence homology, and the unusual intron interrupts a homologous position of the coding sequence in all five genes. Unless there are errors in some of the reported sequences, it will be interesting to determine whether the calcium channel transcripts are processed via the major pathway, the AT-AC pathway, or a hybrid pathway.
  8. A portion of the human SCN4A gene was amplified by PCR from human total genomic DNA (Promega) with primers containing restriction sites and matching exons 2 and 3, to generate a fragment comprising nt 886 to 1229 (numbering according to GenBank accession number L04216). This fragment was digested with Hind III and Xba I and subcloned into the corresponding sites of pSP64 (Promega) to generate the pSP64-SCN4A plasmid. For construction of pSP64-SCN4AM, a different downstream PCR primer containing the mutations and an Eco RI site was used to amplify a mutant fragment from the cloned wild-type template, which was then subcloned as a Hind III-Eco RI fragment in pSP64. All constructs were confirmed by sequence analysis. pSP64-SCN4A and pSP64-SCN4AM were linearized with Xba I or Eco RI, respectively, for use as templates for in vitro transcription with SP6 RNA polymerase. The transcripts contain short extensions at both ends, derived from the vector.
  9. Nuclear extract preparation and conditions for in vitro transcription and for  $\beta$ -globin splicing were as described [A. Mayeda and A. R. Krainer, *Cell* **68**, 365 (1992)]. The SCN4A in vitro splicing reaction was optimized by varying individual parameters. The optimum condition for SCN4A splicing was 60% (v/v) nuclear extract [giving final concentrations of 12% (v/v) glycerol, 12 mM Hepes-K<sup>+</sup> (pH 8.0), 60 mM KCl, 0.6 mM dithiothreitol, and 0.3 mM EDTA] plus an additional 32 mM Hepes-K<sup>+</sup> (pH 7.3), 3.5 mM MgCl<sub>2</sub>, 0.5 mM ATP, 20 mM creatine phosphate, 2.6% (w/v) polyvinyl alcohol, and 1.6 nM SCN4A pre-mRNA, incubated at 30°C for 6 hours. No splicing was detectable when ATP or MgCl<sub>2</sub> was omitted. RNA was recovered and analyzed on 4.5% denaturing polyacrylamide gels, followed by autoradiography. Splicing efficiency, defined as the molar ratio mRNA/(pre-mRNA + mRNA), was estimated by phosphor image analysis (Fujix, BAS2000).
  10. Q. Wu and A. R. Krainer, data not shown.
  11. To sequence the authentic and cryptic spliced mRNAs across the spliced junctions, we amplified each gel-purified RNA by RT-PCR with exon 2 and exon 3 primers (TCATCGTACTCAACAAGG and TACTCCACATTCTTGGAC). The amplified fragment, subcloned into PCR2.1 (Invitrogen), was sequenced with T7 Sequenase 2.0 (USB).
  12. P. J. Grabowski, R. A. Padgett, P. A. Sharp, *Cell* **37**, 415 (1984); B. Ruskin, A. R. Krainer, T. Maniatis, M. R. Green, *ibid.* **38**, 317 (1984).
  13. W.-Y. Tarn and J. A. Steitz, *ibid.* **84**, 801 (1996).
  14. A. Krämer, W. Keller, B. Appel, R. Lührmann, *ibid.* **38**, 299 (1984); A. R. Krainer and T. Maniatis, *ibid.* **42**, 725 (1985); D. L. Black, B. Chabot, J. A. Steitz, *ibid.*, p. 737; S. M. Berget and B. L. Robberson, *ibid.* **46**, 691 (1986); D. L. Black and J. A. Steitz, *ibid.*, p. 697.
  15. For RNase H inhibition experiments, the nuclear extract was preincubated for 15 min under splicing conditions in the presence or absence of the appropriate oligonucleotides. The oligonucleotides were complementary to U1 snRNA position 2 to 11, U2 snRNA position 1 to 15, or U12 snRNA position 11 to 24. snRNA cleavage is catalyzed by the endogenous RNase H (14), and exogenous RNase H had no additional effect (10). All snRNA cleavage and splicing inhibition experiments were carried out at least three times, with reproducible results.
  16. S. J. Baserga and J. A. Steitz, in *The RNA World*, R. F. Gesteland and J. F. Atkins, Eds. (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1993), pp. 359-381; D. A. Wassarman and J. A. Steitz, *Mol. Cell. Biol.* **11**, 3432 (1991).
  17. S. L. Hall and R. A. Padgett, *Science* **271**, 1716 (1996).
  18. K. A. Montzka and J. A. Steitz, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8885 (1988); D. L. Black and A. L. Pinto, *Mol. Cell. Biol.* **9**, 3350 (1989).
  19. K. M. Wasserman and J. A. Steitz, *Mol. Cell. Biol.* **12**, 1276 (1992).
  20. R. R. Gontarek, M. T. McNally, K. Beemon, *Genes Dev.* **7**, 1926 (1993).
  21. G. P. Shumyatsky, S. V. Tillib, D. A. Kramerov, *Nucleic Acids Res.* **18**, 6347 (1990).
  22. Y. Zhuang and A. M. Weiner, *Cell* **46**, 827 (1986); P. G. Siliciano and C. Guthrie, *Genes Dev.* **2**, 1258 (1988); B. Seraphin, L. Kretzner, M. Rosbash, *EMBO J.* **7**, 1533 (1988).
  23. B. L. Robberson, G. J. Cote, S. M. Berget, *Mol. Cell. Biol.* **10**, 84 (1990); H.-C. Kuo, F. H. Nasim, P. J. Grabowski, *Science* **251**, 1045 (1991); J. P. Kreivi, K. Zerivitz, G. Akusjärvi, *Nucleic Acids Res.* **19**, 6956 (1991).
  24. The SCN4AM mRNA was accurately spliced, as verified by RT-PCR sequencing. No spliced products were detected in the absence of ATP or magnesium. The SCN4AM cryptic spliced product was also confirmed by sequencing. Cleavage of U4 and U6 snRNAs had the same effect as U2 cleavage. An additional pre-mRNA, in which the downstream 5' splice site was deleted, had identical splicing and inhibition profiles as those of the SCN4AM substrate (10).
  25. S. M. Berget, *J. Biol. Chem.* **270**, 2411 (1995).
  26. M. D. Chiara and R. Reed, *Nature* **375**, 510 (1995).
  27. M. Ohno, H. Sakamoto, Y. Shimura, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5187 (1987); E. Izaurralde *et al.*, *Cell* **78**, 657 (1994).
  28. M. Niwa and S. M. Berget, *Genes Dev.* **5**, 2086 (1991); K. M. Wassarman and J. A. Steitz, *ibid.* **7**, 647 (1993); C. S. Lutz *et al.*, *ibid.* **10**, 325 (1996).
  29. B. E. Hoffman and P. J. Grabowski, *ibid.* **6**, 2554 (1992); J. Y. Wu and T. Maniatis, *Cell* **75**, 1061 (1993); D. Staknis and R. Reed, *Mol. Cell. Biol.* **14**, 7670 (1994).
  30. D. C. Kohrman, J. B. Harris, M. H. Meisler, *J. Biol. Chem.* **271**, 17576 (1996).
  31. We thank D. Horowitz for comments on the manuscript; A. Mayeda for sharing protocols and reagents; and M. Zhang, M. Murray, H.-X. Liu, and T.-L. Tseng for helpful discussions. Supported by grant GM42699 from NIH.

19 July 1996; accepted 3 September 1996

## TECHNICAL COMMENTS

### HIV-1 Evolution and Disease Progression

Steven M. Wolinsky *et al.* (1) studied the evolution of human immunodeficiency virus (HIV) in six infected patients with variable rates of disease progression and have represented their findings (1, 2) as inconsistent with our evolutionary hypothesis of HIV disease progression, also known as "antigenic diversity threshold theory" (3-6). Their account of our model is, however, incorrect. Consequently their interesting and important data do not falsify our theory. We proceed to give a short review of our theory and show that the findings of Wolinsky *et al.* (1) provide further evidence that the major assumptions of our evolutionary theory of HIV infection are valid.

The central tenet of our theory can be stated as follows: (i) virus load causes disease; (ii) immune responses reduce virus load; (iii) virus evolution during infection weakens the effect of the immune response to HIV and increases virus load. The key result is a dynamic threshold condition that specifies whether or not the HIV population in a given patient is controlled by immune responses. The precise location or value of this threshold, which can be breached by virus evolution leading to in-

creasing antigenic diversity, is likely to vary greatly among individual patients. In particular, weak immune responders should have a low antigenic diversity threshold and can therefore progress to disease rapidly and without significant antigenic variation (3-6). Admittedly, for simplicity our first papers presented analytic results for situations in which all patients had the same immune response, and all HIV strains had the same replication rate; these may have caused confusion. But even in these first papers, there is a discussion (and figures) for situations where diversity thresholds differ among infected individuals.

Wolinsky *et al.* (1) followed six HIV-1 infected patients longitudinally over up to 5 years after infection (Fig. 1). Two patients were rapid progressors and died within 36 and 42 months after infection. Genetic diversity (proviral DNA) was sampled in a region of the envelope protein, and cytotoxic T lymphocyte (CTL) precursor frequency was determined against the ENV, GAG, and POL proteins. The data were interpreted as rejecting our theory, apparently because high genetic diversity did not correlate with rate of CD4 cell loss and in particular