筑波大学

医学博士学位論文

DA 559 1987 C:4697.2

寄贈 平成 年 月 日 立井一明氏

Recognition of 5' and 3' splice site sequences in pre-mRNA studied with a filter binding technique.

（フィルター結合法を用いたメッセンジャーＲＮＡ前駆体上の５’及び３’スプライス部位配列の認識の機構の研究）

１９８７

筑波大学大学院博士課程医学研究科

立井　一明

92302879

The Journal of Biological Chemistry
© 1987 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 262, No. 24, Issue of August 25, pp. 11667-11674, 1987
Printed in U.S.A.

Recognition of 5′ and 3′ Splice Site Sequences in Pre-mRNA Studied with a Filter Binding Technique*

(Received for publication, November 19, 1986)

Kazuaki Tatei‡, Kazuta Takemura, Hideaki Tanaka§, Tomoh Masaki¶, and Yasumi Ohshima‖

From the Institute of Biological Sciences, University of Tsukuba, Sakura-mura, Ibaraki-ken 305, Japan, the §National Chemical Laboratory for Industry, Yatabemachi, Ibaraki-ken 305, Japan, and the ¶Institute of Basic Medicine, University of Tsukuba, Sakura-mura, Ibaraki-ken 305, Japan

A nuclear extract from HeLa cells was fractionated by DEAE-Sepharose chromatography. The obtained fractions were assayed for binding to an RNA transcript carrying a splice site sequence of 9–16 nucleotides by a filter binding technique. The U1 RNA-rich small nuclear ribonucleoprotein (snRNP) fractions, which showed binding activities for both 5′ and 3′ splice site RNAs, were studied for the sequence specificity of their binding. Results indicate that the U1-rich snRNP fraction can recognize both 5′ and 3′ splice site sequences. The U1 RNP, which was highly purified from the snRNP fractions, bound to at least some 5′ splice site sequences, but not to a consensus 3′ splice site sequence. Therefore, purified U1 RNP can directly recognize a 5′ splice site, but not a 3′ splice site. The binding activity for the 5′ splice sites was lost either by digestion with micrococcal nuclease or by digestion of the 5′ end of U1 RNA with RNase H and a complementary oligodeoxynucleotide, indicating the involvement of U1 RNA. Involvement of a protein moiety as well in this binding was suggested by the loss of binding activity upon heating at 60 °C. The binding activity to a 3′ splice site sequence was not sensitive to digestion by micrococcal nuclease and was removed by protein A-coupled anti-Sm antibody. This activity was found in sucrose gradient fractions of about 8 S.

Most eukaryotic pre-mRNA contain introns which can be removed by RNA splicing (1). Several reports suggest that the sequence around an exon-intron boundary (5′ splice site) and the sequence around an intron-exon boundary (3′ splice site) play an important role in this process (2–5). Development of an *in vitro* mRNA splicing system has led to the identification of a novel intermediate, known as lariat RNA, which contains a branched structure (6–8). Another signal sequence for mRNA splicing may reside around the branch point, although it is not highly conserved (8–10). Furthermore, in an *in vitro* splicing system, the mRNA splicing takes place in two steps. The first is cleavage at the 5′ splice site and lariat formation. The second is cleavage at the 3′ splice site followed by the ligation of two exons (6, 7, 11). The 5′ cleavage and lariat formation seem to be coupled, and even in this step both 5′ and 3′ splice site sequences are required (8, 12).

Several groups proposed that an snRNP[1] (small nuclear ribonucleoprotein complex, 13) contributes to mRNA splicing (14, 15). Evidence now supports the view that U1 RNP is involved in the recognition of the 5′ splice site (16–18). U2 RNP is also involved in mRNA splicing possibly by recognizing the branch point region (18, 19). Additionally, U4/U6 RNP, as well as other factors are associated with mRNA splicing (20–24). However, both the interaction among these factors and the role of each factor in the mRNA splicing remain unclear.

Using a filter binding technique, we previously found that a U1-rich snRNP fraction selectively bound to both 5′ and 3′ splice site sequences in single-stranded DNA or RNA (18). Evidence for the selective binding was 3-fold. First, alteration of conserved GT in the 5′ splice site and AG in the 3′ splice site of a single-stranded DNA reduced the binding. Second, an RNA transcript of about 80 nucleotides carrying a consensus sequence for a splice site showed much higher binding than did a control RNA. Third, the binding assay was done in the presence of a vast excess of denatured DNA or RNA from *Escherichia coli*. Recently, we used shorter RNA (32–37 nucleotides) to improve the specificity of binding and surveyed the binding activities more systematically (25). In this report, we demonstrate sequence-specific binding to the splice sites by a U1-rich snRNP fraction. We also present evidence for the nature of the factors that recognize the sequences, as well as some mechanisms for recognition.

EXPERIMENTAL PROCEDURES

Preparation of an RNA Carrying a Splice Site Sequence—Oligodeoxynucleotides were synthesized on a model 380A DNA synthesizer (Applied Biosystems, Inc.). They were purified with high pressure liquid chromatography using a C_{18} or C_8 reverse phase column (Nakarai Chemicals, Kyoto; Senshu Sciences, Tokyo). An oligodeoxynucleotide of 21 bases carrying a splice site sequence was annealed with a complementary oligonucleotide of 19 bases. The annealed DNA was inserted into pSP64 *Hin*dIII-*Acc*I region, and the RNA transcripts were prepared as described previously (25).

Preparation and Fractionation of a HeLa Cell Nuclear Extract—Preparation of a nuclear extract from HeLa cells, its fractionation, and the purification of U1 RNP were done as described previously (25, 26). The purified U1 RNP fraction was concentrated by a DEAE-Sepharose column (0.8 × 2 cm) for the binding assay.

RNA Binding Assay—The binding activity of an RNA transcript was assayed using the filter binding method as described previously (25) and 10 fmol of ^{32}P-labeled RNA in a 100-μl reaction volume. The reaction mixture was filtered through a Millipore GSWP filter using

* This research was aided by grants from the Ministry of Education, Science, and Culture of Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Dept. of Biology, Faculty of Science, Kyushu University, Hakozaki, Fukuoka-city 812, Japan.

‖ To whom correspondence should be addressed (present address): Dept. of Biology, Faculty of Science, Kyushu University, Hakozaki, Fukuoka-city 812, Japan.

[1] The abbreviations used are: snRNP, small nuclear ribonucleoprotein; RNP, ribonucleoprotein; SDS, sodium dodecyl sulfate; Pipes, piperazine-*N,N′*-bis(2-ethanesulfonic acid); EGTA, [ethylenebis(oxyethylenenitrilo]tetraacetic acid.

a 96-well microfiltration manifold (Ikeda Rika, Inc., Tokyo).

Preparation of an RNA Complementary to U1 RNA—An *Eco*RI fragment of about 130 base pairs with an internal sequence of a U1 RNA gene (27) was inserted into pSP65 plasmid in reverse orientation. It was linealized by *Hin*dIII and transcribed with SP6 RNA polymerase. Transcripts were purified on denaturing gel for RNase protection assay.

RNase Protection Assay—The method employed is based on that of Zinn *et al.* (28). A U1 RNP fraction was filtered as described above but without addition of any RNA. The retained RNA was eluted by incubating the filter with 0.2% SDS and 5 mM EDTA at 80 °C for 10 min. The eluted material was phenol-extracted and ethanol-precipitated. Aliquots were dissolved in 30 μl of hybridization buffer (80% formamide, 40 mM Pipes, pH 6.4, 400 mM NaCl, 1 mM EDTA) containing a labeled RNA probe complementary to U1 RNA. The samples were incubated at 85 °C for 10 min and then at 45 °C overnight. Three hundred and five μl of 10 mM Tris-HCl, pH 7.5, containing 5 mM EDTA, 300 mM NaCl, and 40 μg/ml RNase A was added, and the reaction mixture was incubated for 30 min at 30 °C. Fifty μg of proteinase K and 20 μl of 10% SDS were added. The mixtures were further incubated for 15 min at 30 °C, and the RNA was phenol-extracted. Aliquots of RNA samples were measured for radioactivity, separated on a 7 M urea, 10% polyacrylamide gel and followed by autoradiography.

RESULTS

Nucleotide Sequences of the RNA Transcripts—Table I shows the sequences of the RNA transcripts used herein. The 5455-RNA carries one of the 9-nucleotide consensus sequences, $^{C}_{A}$AG/GU$^{A}_{G}$AGU, for the 5′ splice site proposed by Mount (29). The consensus sequence CAG/GUAAGU chosen here is the most probable one and is sometimes found at the actual 5′ splice junctions in the pre-mRNAs such as mouse immunoglobulin for γ 2_b C_H chain and silkmoth chorion (29). The sequence outside the 9-nucleotide consensus in 5455-RNA was chosen either arbitrarily or for cloning to the *Hin*dIII and *Acc*I sites as described previously (25). RNAs numbered as 4142, 5657, 3940, or 5859 have a single base substitution in the strictly conserved GU or G immediately to the 5′ side of the GU. On the other hand, 1920-RNA carries a 16-nucleotide sequence representing the consensus $\binom{U}{C}_{11}$N AG/G for a 3′ splice site (29). We have prepared four RNAs (2122, 4950, 4748, and 4546) which carry a mutation in the 3′ consensus sequence of 1920. The remaining six RNAs carry a 14–16-nucleotide sequence around the splice site of the SV40 T/t antigen gene or human β-globin gene. There is accumulated information on the role of these splice sites in mRNA splicing (4, 6, 8, 30–33), which is available for comparison with our binding results. Specifically, SV40 T and t donors are interesting because they are alternatively used to produce

TABLE I

Sequences of the RNA transcripts used for the binding study

Sequences from pSP64 plasmid are shown by small letters, and inserted sequences by capital letters. An underline indicates a splice site or related sequence. A nucleotide shown by a hyphen (-) is the same as in the consensus RNA. * for a donor site indicates difference from the 9-nucleotide consensus shown. * for an acceptor indicates difference from the consensus of Mount (29). ΔG for a donor site shows free energy increase by presumed base-pairing with AUACψψACCUG at the 5′ end of U1 RNA as calculated according to Salser (44).

	RNA		Sequence	ΔG
Donor site (37 nucleotides)	5455	Consensus	gaauacaagcuUCGAAACAG/GUAAGUCUcgacucuag	−18.3 kcal
	4142	GU → GA (*)	----------------CAG/GAAAGU---------- (*)	−13.5 kcal
	5657	GU → CU (*)	----------------CAG/CUAAGU---------- (*)	−10.5 kcal
	3940	GU → GC (*)	----------------CAG/GCAAGU---------- (*)	−13.5 kcal
	5859	CAG → CAU (*)	----------------CAU/GUAAGU---------- (*)	−10.5 kcal
	2760	SV40 T-antigen	-------------AACUGAG/GUAUUUGC-------- (* **)	−10.8 kcal
	4344	SV40 T-antigen	------------GCUCUAAG/GUAAAUAU-------- (* *)	−13.8 kcal
	8061	Human β-globin IVS-1	------------CUGGGCAG/GUUGGU---------- (**)	−11.1 kcal
	8162	Human β-globin IVS-2	------------ACUUCAGG/GUGAGUCU-------- (** *)	−11.7 kcal
Acceptor site (37 nucleotides)	1920	Consensus	-----------CUUUCUUCUUCCAG/GAU--------	
	2122	AG → AU (*)	-----------CUUUCUUCUUCCAU/G---------- (*)	
	4950	AG → UG (*)	-----------CUUUCUUCUUCCUG/G---------- (*)	
	4748	CAGG → GAGC (* *)	-----------CUUUCUUCUUCGAG/C---------- (* *)	
	4546	11Pyr → 5Pyr	-----------GAAAAUUCUUCCAG/G---------- (*****)	
	7273	Human β-globin IVS-1	-----------UUUCCCACCCUUAG/G---------- (*)	
	7475	Human β-globin IVS-2	-----------UCUUCCUCCCACAG/C---------- (*)	
Control	2451	(35 nucleotides)	-----------UAGCAAGCUCCUAAU----------	

either T-mRNA or t-mRNA (30, 31).

U1-rich snRNP Fraction Can Correctly Recognize 5′ and 3′ Splice Site Sequences—Among the snRNP fractions obtained by fractionation of a HeLa cell nuclear extract on a DEAE-Sepharose column, the highest binding activities for RNA carrying a consensus 5′ or 3′ splice site sequence were found in the fractions with the highest U1 RNA concentrations (25). These fractions contain much lower amounts of U2, U4, U5, and U6 RNAs and, therefore, are called U1-rich snRNP fractions. Fig. 1 shows the binding of increasing amounts of these fractions to RNAs carrying various sequences related to a 5′ splice site. The highest binding activity was observed with 5455-RNA carrying the consensus sequence CAG/GUAAGU. Every RNA carrying a single base change from the consensus (that is, 5859, 3940, 4142, and 5657) displayed a drastically decreased activity in comparison with the consensus RNA. The two RNAs carrying human β-globin donor site sequences (8061, 8162) showed activities which were lower than that of the consensus, but higher than those of the mutated RNAs. In other experiments using different batches of the snRNP fraction, the SV40 t-antigen donor site RNA (4344) showed a high level of binding activity. This level was between those by the consensus RNA and β-globin first donor site RNA. In contrast, SV40 T-antigen donor site RNA showed activity as low as that found in RNAs carrying a mutation in U of the invariant GU (data not shown). Even though the absolute amount of bound RNA changed depending on the snRNP preparation, saturation level for 5455-RNA being 7–9 fmol, and the order of the binding activities was consistent.

Fig. 2 shows binding activities of the same U1-rich snRNP fraction to the RNAs which carry various sequences related to a 3′ splice site. Every RNA carrying an alteration in the conserved AG or pyrimidine stretch (2122, 4950, and 4546) exhibited significantly lower binding activity than the consensus RNA (1920). 4748-RNA carrying a CAGG → GAGC alteration showed even lower activity. In contrast, RNAs carrying human β-globin acceptor site sequences (7475 and 7273) showed higher binding activities than the consensus RNA.

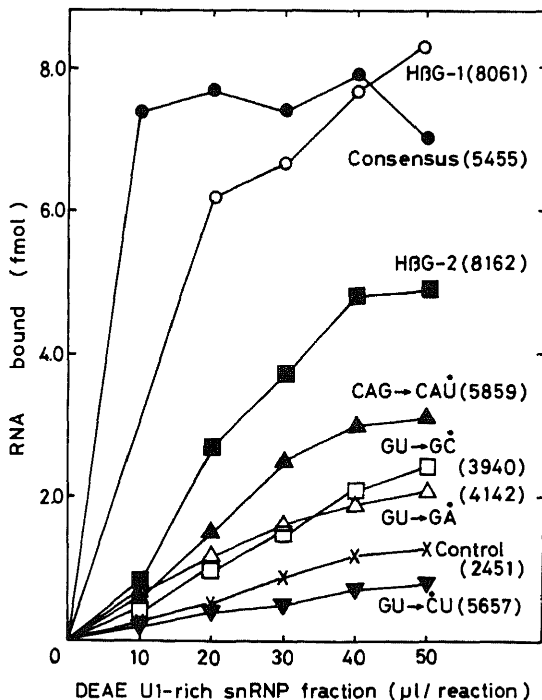

FIG. 1. **Binding of 5′ splice site RNAs to a U1-rich snRNP fraction obtained by DEAE-Sepharose chromatography.** Indicated volume of a U1-rich snRNP fraction (499 μg of protein/ml) was used for the RNA binding assay described under "Experimental Procedures." The sequence of the RNAs are described in Table I. *HβG,* human β-globin.

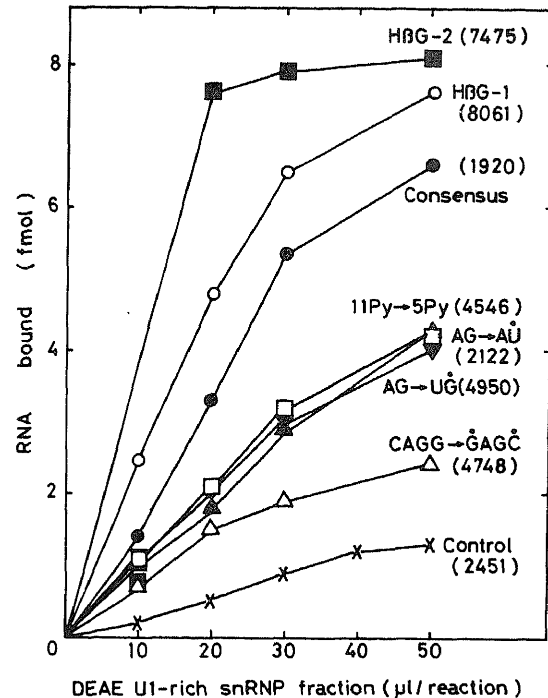

FIG. 2. **Binding of 3′ splice site RNAs to a U1-rich snRNP fraction.** The same U1-rich snRNP fraction as used in Fig. 1 was assayed for RNA binding activities as described under "Experimental Procedures." The RNAs used are described in Table I. *HβG,* human β-globin.

Properties of the Binding Activities—Fig. 3 shows the effect of pretreatment of a U1-rich snRNP fraction with micrococcal nuclease on the RNA binding activities. The binding activity for RNA carrying the consensus sequence for a 5′ splice site was drastically reduced, whereas that for a 3′ splice site was not. Degradation of the U1 RNA was monitored in the same digests: in the digest with 0.5 μg of the nuclease, no intact U1 RNA but a much smaller faint RNA band was detected by gel electrophoresis (data not shown). In the digest with 1 μg of the nuclease, no RNA band was detected. When the 5′ end of U1 RNA in the U1-rich snRNP fraction was digested with RNase H and a complementary oligodeoxynucleotide (Fig. 4), the binding activities for the three RNAs carrying a 5′ splice site were abolished. Addition of RNase H alone did not change binding to the consensus RNA (data not shown). As shown in Fig. 5, when a U1-rich snRNP fraction was pretreated with protein A-bound anti-Sm antibody, both binding activities were markedly reduced. To see whether these binding activities are associated with the same factor, a U1-rich snRNP fraction was further fractionated by sedimentation through a sucrose gradient containing 2 mM $MgCl_2$ (Fig. 6). The binding activity for a 5′ splice site RNA was rather broadly distributed around 12 S, whereas that for a 3′ splice site RNA was found around 8 S. When the same material was sedimented after treatment with micrococcal nuclease, little or no binding activity for the 5′ splice site RNA was found through the gradient, but the 8 S binding peak for the 3′ splice site RNA did not change significantly (data not shown). These results indicate that the factor binding to a 5′ splice site RNA is U1 RNP (see also next section), whereas that binding to a 3′

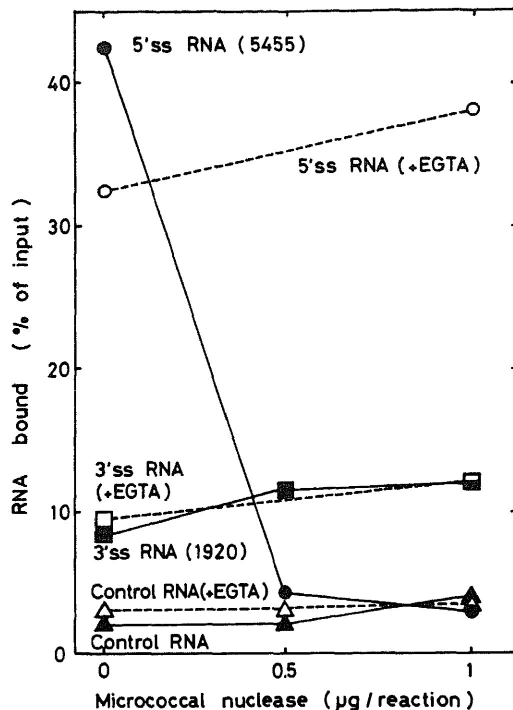

FIG. 3. **Effect of micrococcal nuclease digestion of a U1-rich snRNP fraction on the RNA binding activities.** A U1-rich snRNP fraction (28 μg of protein) was incubated with the indicated amount of micrococcal nuclease in a 100-μl solution containing 50 mM NH_4Cl and 3 mM $CaCl_2$ in the absence or presence of 9 mM EGTA at 30 °C for 1 h. EGTA was then added to the former reactions. The indicated RNA was added for binding assay as described under "Experimental Procedures." Each point shows the average of duplicate filtration results. *ss*, splice site.

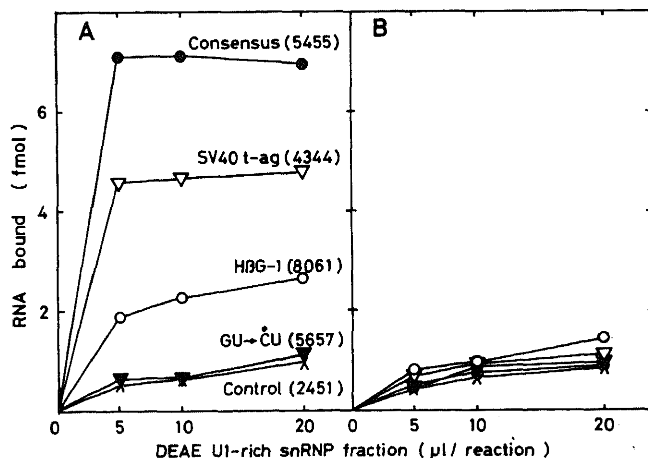

FIG. 4. **Effect of digestion of the 5′ end of U1 RNA on the binding activity for 5′ splice site RNAs.** A U1-rich snRNP fraction was incubated at 565 μg of protein/ml without (*A*) or with (*B*) 30 units/ml RNase H in the presence of 10 μg/ml of an oligodeoxynucleotide AGCTTCAGGTAAGTAAT (OS-11, Ref. 25) at 30 °C for 60 min. The reaction also contained 20 mM Tris-HCl, pH 7.5, 2 mM $MgCl_2$, 0.05% Nonidet P-40, 0.5 mM dithiothreitol, 50 mM NH_4Cl, and 5 μg/ml each chymostatin, pepstatin, and leupeptin. After digestion, the oligodeoxynucleotide was digested by 30 units/ml DNase I at 30 °C for 20 min. The RNA binding activities were assayed as described under "Experimental Procedures." *HβG*, human β-globin, *t-ag*, t-antigen.

splice site RNA is a different molecule or complex carrying an Sm antigen.

Purified U1 RNP Can Recognize a 5′ Splice Site Sequence—The U1 RNP was purified from U1-rich snRNP fractions by

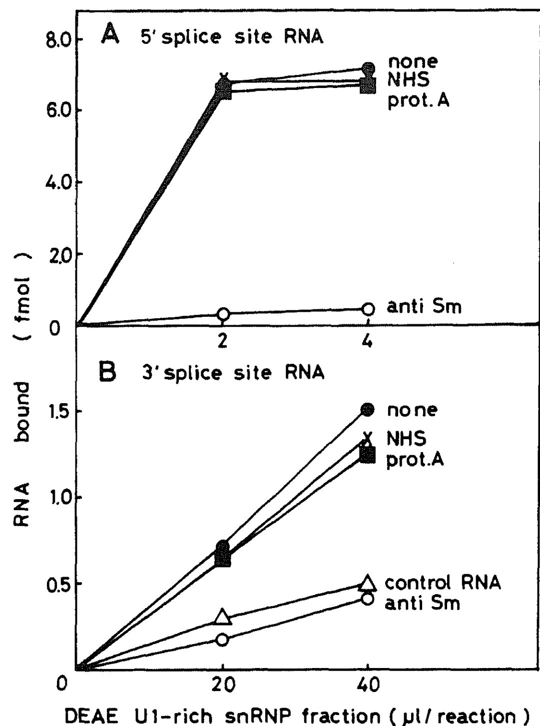

FIG. 5. **Binding of consensus splice site RNAs (*A*, 5455-RNA; *B*, 1920-RNA) to a U1-rich snRNP fraction after removal of Sm antigens.** Normal human serum (*NHS*, ×) or purified anti-Sm antibody from an SLE patient (○, 170 μg of protein) was reacted with 50 μl of protein A-Sepharose CL-4B in 300 μl containing 10 mM Tris-HCl, pH 7.8, 0.5 M NaCl, and 0.1% Nonidet P-40 overnight on ice. The protein A-Sepharose was washed three times by the same buffer, three times in a buffer containing 20 mM Tris-HCl, pH 7.8, 15 mM $MgCl_2$, and 0.05% Nonidet P-40, and then incubated with 80 μl of a U1-rich snRNP fraction (51 μg of protein) for 60 min on ice. Protein A-Sepharose was removed by centrifugation, and aliquots of the supernatant were assayed for the RNA binding activity. ● (*none*), without any treatment; ■ (*prot. A*), treated with protein A-Sepharose alone; Δ (*control RNA*), binding of 2451-RNA to nontreated snRNP fraction.

chromatography on heparin-agarose and blue-Sepharose according to the method of Kinlaw *et al.* (26). The binding activity for the consensus 5′ splice site RNA was co-purified with U1 RNP during the purification. However, the binding activity for the 3′ consensus splice site RNA was gradually lost (17, 25). Fig. 7 shows the proteins of the purified U1 RNP fraction. They include the nine polypeptides of molecular weights 68,000, 32,000 (*A*), 27,000 (*B*), 26,000 (*B′*), 18,000 (*C*), 12,000 (*D*), 10,000 doublets (*E,F*) and 9,000 (*G*), which were presumed to be U1 RNP components from their molecular weights and their immunoreactivity to anti-Sm or anti-RNP antibody (26, 34). The U1 RNP fraction also contained a protein of about M_r 50,000 and some other minor proteins whose relationship to U1 RNP is not clear. The purity of the U1 RNP was estimated to be around 80%.

Fig. 8 shows the binding activities of the purified U1 RNP fraction for the 5′ splice site sequences. That fraction showed significant binding to the RNAs carrying the consensus, SV40 t-antigen and human β-globin first donor site sequences. The order of binding activities for these RNAs was the same as for the crude U1 RNP fraction (see Fig. 1 and the text). In contrast, the binding activities for the other RNAs were not significant.

Involvement of a protein moiety in the binding of U1 RNP to a 5′ splice site was suggested previously (16). In order to test this, we studied the effect of preheating purified U1 RNP

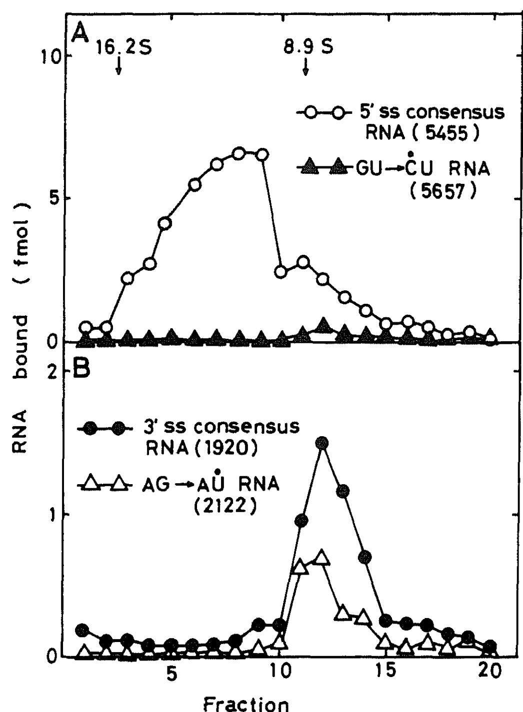

FIG. 6. **Sedimentation analysis of the binding factor for splice site RNAs.** 200 μl of a U1-rich snRNP fraction (847 μg of protein/ml) was centrifuged through 5 ml of a 5–20% sucrose gradient containing 20 mM Tris-HCl, pH 7.5, 50 mM KCl, and 2 mM $MgCl_2$ in a Beckman SW50.1 rotor at 39,000 rpm and 4 °C for 10 h. A portion of each fraction was used for the RNA binding assay described under "Experimental Procedures." *A*, binding to 5′ splice site RNAs; *B*, binding to 3′ splice site RNAs. *ss*, splice site.

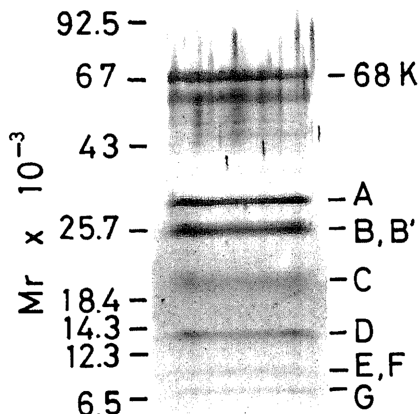

FIG. 7. **Analysis of proteins in a purified U1 RNP fraction.** A purified U1 RNP fraction (7 μg of protein) was electrophoresed in a 5–20% polyacrylamide gel containing 0.1% SDS and visualized by silver stain. The *letters* to the *right* of the *lane* refer to the U1 RNP proteins previously described (26, 34).

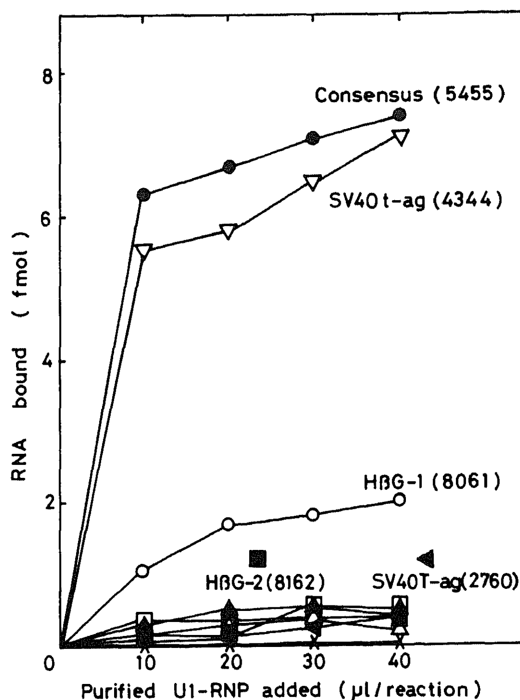

FIG. 8. **Binding of the purified U1 RNP fraction (91 μg of protein/ml) to 5′ splice site RNAs.** RNA binding activities were assayed as described under "Experimental Procedures." The *symbols* are indicated in this figure or in Fig. 1. *HβG*, human β-globin, *t-ag*, t-antigen; *T-ag*, T-antigen.

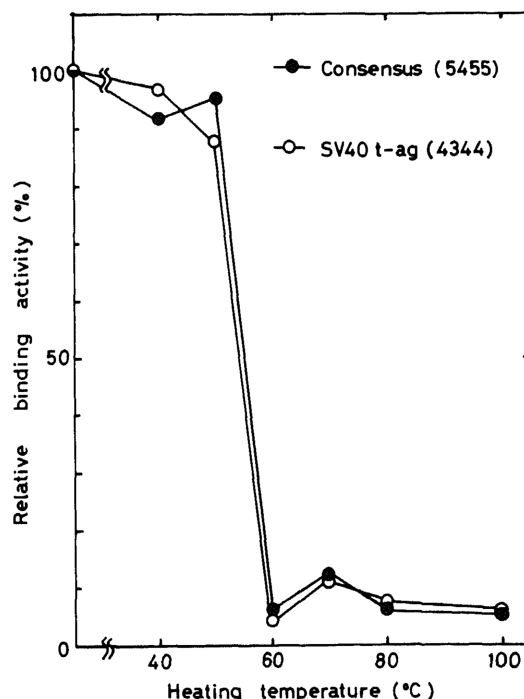

FIG. 9. **Effect of heating the purified U1 RNP fraction on the binding activity for 5′ splice site RNAs.** The same purified U1 RNP fraction (91 μg of protein/ml) was heated at the indicated temperature for 10 min and cooled on ice, and the binding activity for the 5′ splice site RNAs was assayed at 30 °C as described under "Experimental Procedures." *t-ag*, t-antigen.

on its binding activity for a 5′ splice site RNA (Fig. 9). Heating at 60 °C or a higher temperature greatly reduced that binding activity. Another experiment showed that the 50% inactivation temperature was 55–56 °C (data not shown). Since the filter binding assay requires a protein moiety, the loss of binding activity could be due to dissociation of all the protein components of U1 RNP from U1 RNA. However, Fig. 10*A* shows that this is not the case. In that experiment, the U1 RNP treated at various temperatures was filtered and the filter-bound U1 RNA was measured. Since the filter-bound U1 RNA was not reduced by heating, at least some protein component of U1 RNP was associated with U1 RNA after heating. This result also shows that at least a major part of

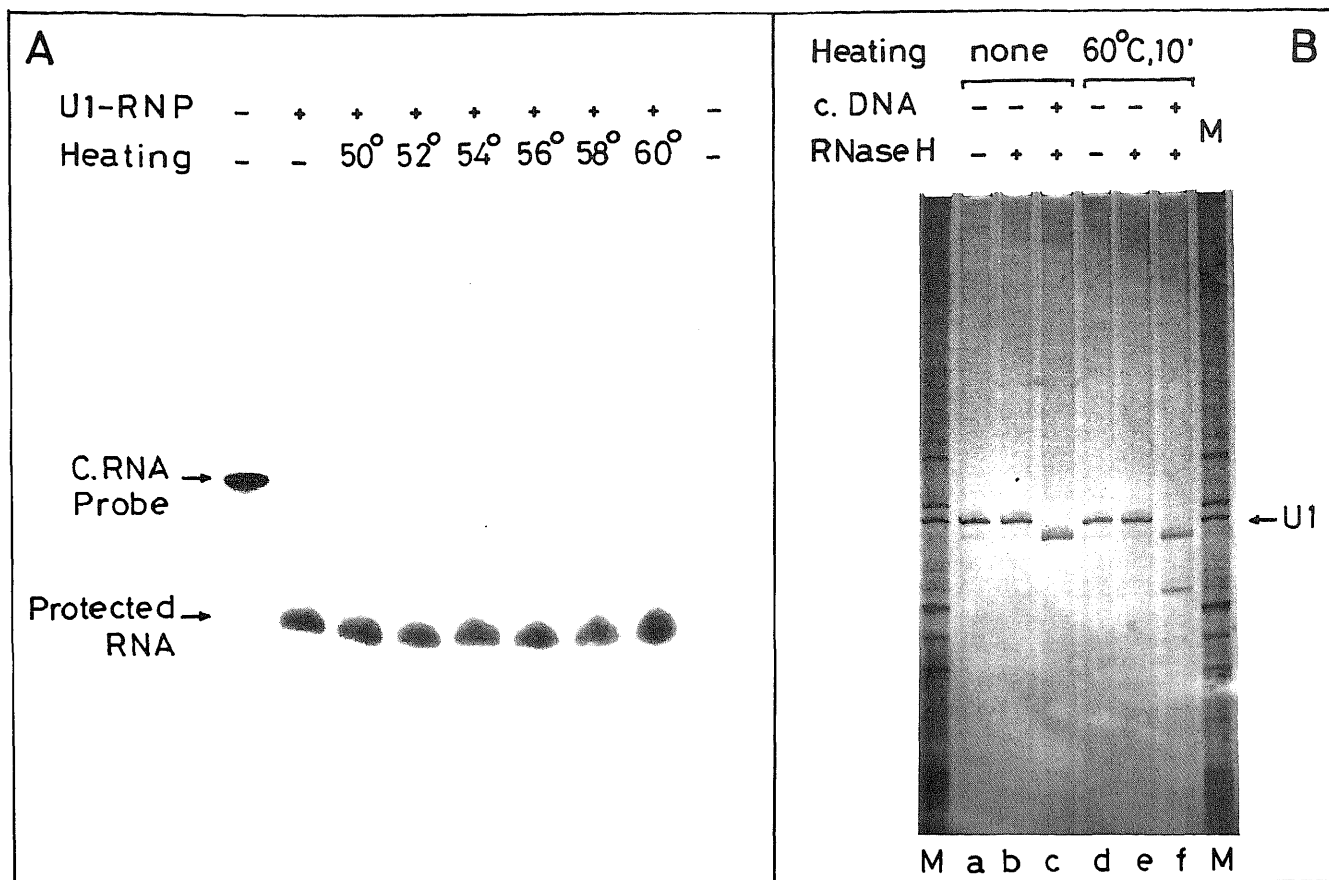

FIG. 10. *A*, retention of U1 RNA of the heated U1 RNP fraction on the filter. The purified U1 RNP fraction (7.6 ng of protein) heated at the indicated temperature was filtered through a Millipore filter, and then the retained RNA was extracted and analyzed by RNase protection assay described under "Experimental Procedures." *B*, accessibility of the 5′ end of U1 RNA of the U1 RNP by a complementary DNA. The purified U1 RNP (91 μg of protein/ml × 10 μl, *lanes a, b,* and *c*) or that heated at 60 °C for 10 min (*lanes d, e,* and *f*) was incubated in the presence or absence of 10 μg/ml of oligodeoxynucleotide OS-11 and 50 units/ml RNase H at 30 °C for 45 min. The RNA was extracted and analyzed on 7 M urea, 10% polyacrylamide gel followed by silver stain. *Lane M* indicates rat liver nuclear RNA as electrophoretic marker.

U1 RNA is intact. An experiment using RNase H and an oligodeoxynucleotide (Fig. 10*B*) indicates that the 5′ end of U1 RNA, which is essential for the binding activity, is intact after heat treatment. The *lower band* in *lane f* of Fig. 10*B* presumably resulted from cleavage in U1 RNA region 23–30 (UACCAUGA), which is partly homologous to the 5′ end of U1 and may be hybridizable with the oligonucleotide. These results suggest that heat inactivation is due to a change or loss of a protein component.

DISCUSSION

The Sequences and Mechanisms Involved in the Recognition of a 5′ Splice Site—The U1-rich snRNP fraction bound to the RNAs carrying different 5′ splice site sequences with different affinities (Fig. 1 and text). The RNAs with either the 9-nucleotide consensus sequence, the 14–16-nucleotide sequences for SV40 t, or the human β-globin donor sites showed higher binding than the RNAs with a mutation in one of the 9 nucleotides. These results indicate that this U1-rich snRNP fraction can recognize a 5′ splice site sequence. The invariant GU (positions +1, +2) and G at position −1 are clearly important for such recognition, since a mutation of any one of these bases reduced the binding greatly (Fig. 1). The 9 nucleotides of the SV40 and human β-globin donor sites differ from the consensus at −3, −2, +3, +4, or +5 position (Table I), suggesting that nucleotides at these positions may also be important. The 5–7 nucleotides outside the 9 are also natural in the RNAs carrying SV40 and β-globin donors, but their contribution to the recognition is not clear from our study. Our results indicate that the 9-nucleotide consensus sequence and the 14–16 nucleotides of the natural donor sites tested are sufficient to act as specific sequences for the selective binding detected here.

Digestion of the 5′ end of U1 RNA (Fig. 4) and binding by a purified U1 RNP fraction (Fig. 8) show that the U1 RNP itself is the binding factor, at least for the consensus, SV40 t, and β-globin first donor site sequences. This view is supported by the results with micrococcal nuclease treatment of both the U1-rich snRNP fraction (Fig. 3 and text on Fig. 6*A*) and the purified U1 RNP (data not shown), as well as by the experiment with anti-Sm antibody (Fig. 5*A*).

Although the binding factor for SV40 T and β-globin second donor sites, as well as for the mutated donor sites may be U1 RNP, we do not have any conclusive evidence for verification. However, there is some evidence indicating involvement of U1 RNP in the recognition of, or splicing with, the SV40 T donor site (30) and mutated or cryptic donor sites of human β-globin (35). Although Fradin *et al.* (30) did not detect the inhibition of t-mRNA splicing in *Xenopus* oocytes by antibodies to U1 or other snRNPs, our results strongly suggest

that U1 RNP recognizes t donor site. This apparent discrepancy may result because a lower level of U1 RNP is sufficient for the binding to t donor site but not for T donor site, since the former has a much higher affinity for U1 RNP. Our binding assay was done using 1×10^{-10} M substrate RNA and $10^{-8} \sim 10^{-7}$ M U1 RNP. Assuming that U1 RNP binds to a 5′ splice site RNA with one to one molar ratio, the dissociation constant of the binding should be in the range between 10^{-8} and 10^{-6} M. The concentration of U1 RNP in a HeLa cell nucleus is estimated to be in the order of 3×10^{-6} M assuming that it contains 10^6 particles of U1 RNP (36). Therefore, U1 RNP *in vivo* could bind substantially to a low affinity donor site such as SV40 T.

The U1 RNA contains a ACψψACCUG sequence near the 5′ end, which is complementary to the consensus donor sequence CAG/GUAAGU. This region in U1 RNP is hybridizable with a complementary oligonucleotide (37). When the 5′ end of U1 RNA in a nuclear extract was digested by RNase H, the splicing activity was lost (17, 19, 20). These results strongly suggest that the 5′ end region of U1 RNA is involved in the binding between U1 RNP and a 5′ splice site by base-pairing. Our results in Fig. 4 are consistent with this view. However, we suggest that at least some protein component in U1 RNP is also involved in the binding. This suggestion is based on the following two lines of results. First, preheating of purified U1 RNP at 60 °C abolished the binding activity (Fig. 9). This inactivation was due neither to dissociation of U1 RNA from the protein components nor to the loss of, or accessibility to, the 5′ of U1 RNA (Fig. 10). Second, the free energy increment (ΔG) from base-pairing between the 5′ end region of U1 RNA and 5′ splice site sequences is not in parallel with the binding affinities for U1 RNP (Table I and Fig. 1). The presumed heat-sensitive protein component(s) is likely to possess, or influence, sequence-specific binding ability. Such a protein component could wrap around base-paired RNA double helix, as in the case of λ phage cro repressor (38).

The Binding Activity for a 3′ Splice Site Sequence—Fig. 2 shows that the U1-rich snRNP fraction can recognize a 16-nucleotide acceptor site sequence. The invariant PyrAG and the pyrimidine cluster were shown to be important in this recognition. This binding was not reduced by pretreatment with micrococcal nuclease (Fig. 3), and the involved factor is different from U1 RNP (Fig. 6, text, and Ref. 25). We believe that this factor is involved in mRNA splicing although we have no direct evidence so far.

Two groups have recently reported a factor which can recognize a 3′ splice site in pre-mRNA. The 100-kDa protein described by Tazi *et al.* (39) co-migrates with U5 RNP and is precipitated by a monoclonal anti-Sm antibody only at a low Mg^{2+} concentration. The 70-kDa protein described by Gerke and Steitz (40) reacts with a monoclonal anti-Sm antibody and is not retained by a DEAE column at a high Mg^{2+} concentration. That snRNP could be U5 RNP (41). The factor we have detected reacts with polyclonal anti-Sm antibodies (Fig. 5) and is retained by a DEAE column at a high (15 mM) Mg^{2+} concentration. Therefore, our factor is similar to, and could be the same as, one of those reported by the two groups, although our factor shows a property apparently different from theirs.

We previously found binding activity for the consensus 3′ splice site RNA in the DEAE flow-through fraction as well (25). This binding activity showed specificity for a pyrimidine stretch but not for invariant PyrAG (data not shown), in contrast to the activity in this paper. Therefore, the factors responsible for these binding activities are likely to be different from each other.

Recently, the sequences outside the 16-nucleotide acceptor site sequence have been suggested to be important in the selection or utilization of an acceptor site (42). We have obtained a similar suggestion from computer analysis (43). The nature and role of an outside sequence in the recognition of a 3′ splice site remain to be elucidated.

Acknowledgments—We thank Dr. R. Mukai for supplying anti-Sm antibody, A. Mayeda and H. Kitayama for supplying materials and for discussion, and S. Ohshima for help in preparing the manuscript. We also thank Dr. S. Bannai for critical reading of the manuscript, and Prof. Y. Watanabe for encouragement.

REFERENCES

1. Breathnach, R., and Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383
2. Montell, C., Fisher, E. F., Caruthers, M. H., and Berk, A. J. (1982) *Nature* **295,** 380–384
3. Wieringa, B., Meyer, F., Reiser, J., and Weissmann, C. (1983) *Nature* **301,** 38–43
4. Treisman, R., Orkin, S. H., and Maniatis, T. (1983) *Nature* **302,** 591–596
5. Wieringa, B., Hofer, E., and Weissmann, C. (1984) *Cell* **37,** 915–925
6. Ruskin, B., Krainer, A. R., Maniatis, T., and Green, M. (1984) *Cell* **38,** 317–331
7. Padgett, R. A., Konarska, M. M., Grabowski, P. J., and Sharp, P. A. (1985) *Science* **225,** 898–903
8. Reed, R., and Maniatis, T. (1985) *Cell* **41,** 95–105
9. Ruskin, B., Greene, J. M., and Green, M. R. (1985) *Cell* **41,** 833–844
10. Padgett, R. A., Konarska, M. M., Aebi, M., Hornig, H., Weissmann, C., and Sharp, P. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 8349–8353
11. Grabowski, P. J., Seiler, S. R., and Sharp, P. A. (1985) *Cell* **42,** 345–353
12. Frendewey, D., and Keller, W. (1985) *Cell* **42,** 355–367
13. Busch, H., Reddy, R., Rothblum, L., and Choi, Y. C. (1982) *Annu. Rev. Biochem.* **51,** 617–653
14. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., and Steitz, J. A. (1980) *Nature* **283,** 220–224
15. Ohshima, Y., Itoh, M., Okada, N., and Miyata, T. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 4471–4474
16. Mount, S. M., Petterson, I., Hinterberger, M., Karmas, A., and Steitz, J. A. (1983) *Cell* **33,** 509–518
17. Kramer, A. R., Keller, W., Appel, B., and Luhrmann, R. (1984) *Cell* **38,** 299–307
18. Tatei, K., Takemura, K., Mayeda, A., Fujiwara, Y., Tanaka, H., Ishihama, A., and Ohshima, Y. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 6281–6285
19. Black, D. L., Chabot, B., and Steitz, J. A. (1985) *Cell* **42,** 737–750
20. Krainer, A. R., and Maniatis, T. (1985) *Cell* **42,** 725–736
21. Black, D. L., and Steitz, J. A. (1986) *Cell* **46,** 697–704
22. Perkins, K. K., Furneaux, H. M., and Hurwitz, J. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 887–891
23. Choi, Y. D., Grabowski, P. J., Sharp, P. A., and Dreyfuss, G. (1986) *Science* **231,** 1534–1539
24. Grabowski, P., and Sharp, P. A. (1986) *Science* **233,** 1294–1298
25. Mayeda, A., Tatei, K., Kitayama, H., Takemura, K., and Ohshima, Y. (1986) *Nucleic Acids Res.* **14,** 3045–3057
26. Kinlaw, C. S., Robberson, B. L., and Berget, S. M. (1983) *J. Biol. Chem.* **258,** 7181–7189
27. Watanabe-Nagasu, N., Itoh, Y., Tani, T., Okano, K., Koga, N., Okada, N., and Ohshima, Y. (1983) *Nucleic Acids Res.* **11,** 1791–1801
28. Zinn, K., DiMaio, D., and Maniatis, T. (1983) *Cell* **34,** 865–879
29. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472
30. Fradin, A., Jove, R., Hemenway, C., Keiser, H. D., Manley, J. L., and Prives, C. (1984) *Cell* **37,** 927–936
31. Noble, J. C. S., Prives, C., and Manley, J. L. (1986) *Nucleic Acids Res.* **14,** 1219–1235
32. Krainer, A. R., Maniatis, T., Ruskin, B., and Green, M. R. (1984) *Cell* **36,** 993–1005

33. Dobkin, C., and Bank, A. (1985) *J. Biol. Chem.* **260,** 16332–16337
34. Hinterberger, M., Pettersson, I., and Steitz, J. A. (1983) *J. Biol. Chem.* **258,** 2604–2613
35. Chabot, B., and Steitz, J. A. (1987) *Mol. Cell. Biol.,* in press
36. Weinberg, R. A., and Penman, S. (1968) *J. Mol. Biol.* **38,** 289–304
37. Rinke, J., Appel, B., Blöcker, H., Frank, R., and Lührmann, R. (1984) *Nucleic Acids Res.* **12,** 4111–4126
38. Anderson, W. F., Ohlendorf, D. H., Takeda, Y., and Matthews, B. W. (1981) *Nature* **290,** 754–758
39. Tazi, J., Alibert, C., Temsamani, J., Reveillaud, I., Cathala, G., Brunel, C., and Jeanteur, P. (1986) *Cell* **47,** 755–766
40. Gerke, V., and Steitz, J. A. (1986) *Cell* **47,** 973–984
41. Chabot, B., Black, D. L., LeMaster, D. M., and Steitz, J. A. (1985) *Science* **230,** 1344–1349
42. Reed, R., and Maniatis, T. (1986) *Cell* **46,** 681–690
43. Ohshima, Y., and Gotoh, Y. (1987) *J. Mol. Biol.* **195,** 247–259
44. Salser, W. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **42,** 985–1002