BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **242,** 10–15 (1998)
ARTICLE NO. RC977553

Isolation and Characterization of EMS Induced Splicing Defective Point Mutations within the Intron of the *nrd*B Gene of Bacteriophage T4

Asad U. Khan,* Sunil K. Lal,†,[1] and Masood Ahmad*

**Department of Biochemistry, Faculty of Life Sciences, Aligarh Muslim University, Aligarh, India 202002; and †Virology Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi, India 110067*

Received September 15, 1997

The *nrdB* gene of bacteriophage T4 codes for the small subunit of ribonucleotide reductase and contains a 598-base-pair self-splicing intron which is closely related to other group I introns of T4 and eukaryotes. The screening, isolation, and mapping of 31 *nrd*B intron mutations were conducted by the strategic usage of the white halo phenotype exhibited by T4 mutants defective in dyhydrofolate reductase or thymidylate synthase. These intron mutations cluster towards the ends, mainly the 3′ end, and show a defect in self-splicing. These mutations map in regions of conserved structural elements, thus supporting secondary structure predictions. A distinct pattern of clustering is observed with the highest number of mutations mapping within three of the smaller regions (A, C, and D) of the nrdB intron and no mutations mapping in the largest (B) region. The highest density of mutations mapped in the smallest region (C) of the intron, containing only 96 bases, thus showing a distinct pattern of clustering within the catalytic core.

The auto-catalytic mechanism of group I introns has been well established [1-3]. Self-splicing of pre-mRNA *in vitro* occurs in the absence of proteins and external sources of energy; only magnesium and guanosine cofactors are necessary. The energy requirement is obviated by a *trans*-esterification mechanism in which the numbers of phosphodiester bonds formed and broken are the same. Self-splicing introns thus contain all the structural features necessary for the RNA catalysed reaction.

Bacteriophage T4 contains three self-splicing introns: the *td* intron (within the gene encoding dTMP synthase) [4], the *nrd*B intron (within the gene encoding the small subunit of ribonucleotide reductase) [5], and the *nrd*D intron (within the gene encoding an anaerobic ribonucleoside triphosphate reductase) [6]. These introns differ substantially both in sequence and in size: the *nrd*D and *td* introns respectively comprising 1033 and 1016 nucleotides, while the *nrd*B intron is only 598 nucleotides long. All three introns contain open reading frames (ORFs). While the intron ORFs are not homologous with one another and occur at different positions, all three ORFs are looped out of the RNA structure models, with only the 3′ ends of each of the ORFs extending into the secondary structure.

The white halo plaque phenotype exhibited by *frd* mutants (phage with a mutation in the dihydrofolate reductase gene) under special plating conditions [7] is an excellent system for the isolation of *nrd*B mutants which can be characterized further for intron mutants by using the *nrd*B intron subclone to rescue the intron mutation. Using this approach, we have mutagenized *frd* mutants (mutants with a mutation in the gene encoding dihydrofolate reductase) by *in vitro* ethylmethane sulphonate (EMS) and isolated 31 double mutants (*frd*1 *nrd*B) that map within the *nrd*B intron by marker rescue with a clone of the wild-type sequence of this intron.

Using subclones of the *nrd*B intron, we have further mapped the various point mutations into four regions of the *nrd*B intron. Interestingly, all these splicing defective mutants cluster into regions of the *nrd*B intron containing at least a portion of the conserved sequence (P, Q, R & S). The frequency of homologous recombination resulting in wild type phenotype has been used to predict the relative distances between various mutations.

[1] Corresponding author. Fax: 91-11-6162316. E-mail: sunillal@icgebnd.ernet.in.

Abbreviations used: EMS, ethlymethane sulphonate; *nrd*B, ribonucleotide reductase small subunit; *td*, thymidylate synthase; *frd*, dihydrofolate reductase; GPTG, glycerolphosphate tris gulcose media; *wt*, wild type.

0006-291X/98 $25.00

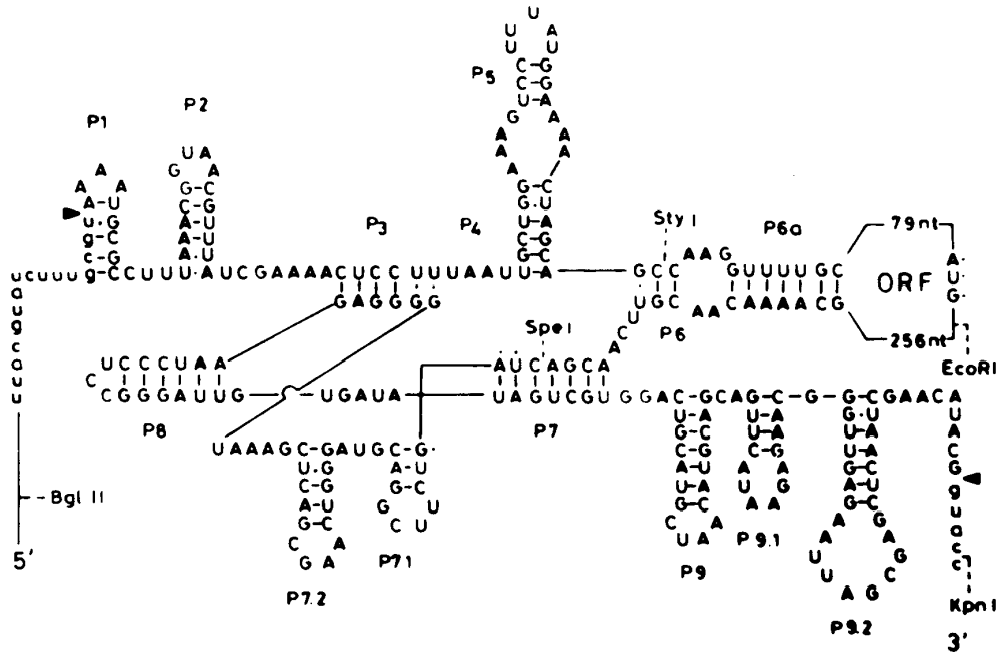

FIG. 1. A secondary structure model of the *nrd*B intron showing the various restriction enzyme sites used for subcloning the whole intron into the regions A, B, C and D. Arrowheads depict 5′ and 3′ cleavage sites of the intron. Flanking exon sequences are shown in lower case letters. pJBK1 = whole *nrd*B intron plus flanking exon sequences between the *Bgl*II and *Kpn*I sites (regions A+B+C+D); pJG106 = *Bgl*II --- *Eco*RI fragment (regions A+B); pJG108 = *Eco*RI --- *Kpn*I fragment (regions C+D); pJBS1 = *Bgl*II --- *Sty*I (region A), pJSR1 = *Sty*I --- *Eco*RI fragment (region B); pJSK7 = *Spe*I --- *Kpn*I fragment (region D). P1 through P9.2 are the various base pairing units contributing to the precise secondary structure of the self-splicing *nrd*B intron

MATERIALS AND METHODS

Bacterial strains, clones, and T4 phage. *E. coli* B, BB and OK305 strains were used as described in [8]. OK 305 is a derivative of *E. coli* B which has a defect in its pyrimidine metabolism. It requires uracil and has very low cytidine deaminase activity with no detectable deoxycitidine deaminase activity. The *nrd*B clones and subclones used are described in [9]. Bacteriophage T4Do and the *frd*1 mutant phage are described previously [10].

Mutagenesis. Mutagenesis of *frd*1 mutants was carried out by EMS as described previously [11].

Strategy for isolation and mapping of nrdB intron mutants. The mutagenized (*frd*1) phage were diluted appropriately and plated on LB plates with JM101 cells containing the pJSS10 plasmid (whole *nrd*B gene). For each one of the marker rescue plates, two well isolated plaques were picked into 1ml diluting fluid and three drops of chloroform were added to kill the cells. Serial dilutions were made and tested for halo phenotype. Appropriate control plates with *frd* and *wt* phage were plated for comparison with marker rescue plates. A minimum of 10^4 plaques were screened for each marker rescue. The recombination-based marker rescue technique is said to be positive if a mutant phage was infected into a cell which had a corresponding *wt* sequence for that mutation in the phage genome [12]. Some percent of the progeny phage would have picked up this *wt* sequence by homologous recombination with the insert in the plasmid and thus showed *wt* (halo$^+$) phenotype. The various subclones that divide the pJBK1 intron into various smaller regions are shown in Fig. 1. JG106 and JG108 divide the JBK1 region into two. The JG106 region is 442 bases and the JG108 region is 239 bases. JG106 is divided into regions A (JBS1) and B (JSR1). The A region is 158 bases and the B region is 284 bases. JG108 has a region which is covered by region D (JSK7). The region D is 143 bases. Thus JG108 has a (239−143=96) 96 base region for which there is no subclone (region C). Region A contains both the P and Q sequences, region C contains about half of the R sequence and region D contains the remaining half of the R sequence and the whole S sequence [9]. The scheme for isolation and mapping of *nrd*B mutants using the halo phenotype is shown in Fig. 2.

RNA extraction and dot blot hybridizations. Total RNA extraction was initiated after a 9-min infection to maximize the amount of *nrd*B mRNA isolated to maximize the amount of *nrd*B mRNA isolated. RNA isolation from T4-infected *E. coli* BB cells was performed as described in Lal and Hall [13]. RNA dot blot hybridizations (Fig. 3) were done using the protocol described in Belfort *et al.* [14].

RESULTS AND DISCUSSION

A total of 432 halo$^-$ plaques were isolated after EMS mutagenesis of the *frd*1 mutant of T4 phage. Since *frd*1 showed a halo$^+$ phenotype, the loss of this phenotype would result from a mutation in the genes *cd*, *nrd*A or *nrd*B. Our interest in this research was confined to mutations in the *nrd*B gene. Thus these 432 mutants were picked, purified and tested for marker rescue with pJSS10 (a clone containing the whole *nrd*B gene). 156 phage showed positive marker rescue thus indicating that the mutation was in the *nrd*B gene.

The pJBK1 clone contains the whole T4 *nrd*B intron along with 80 bases from the 3′ end of exon I and 4

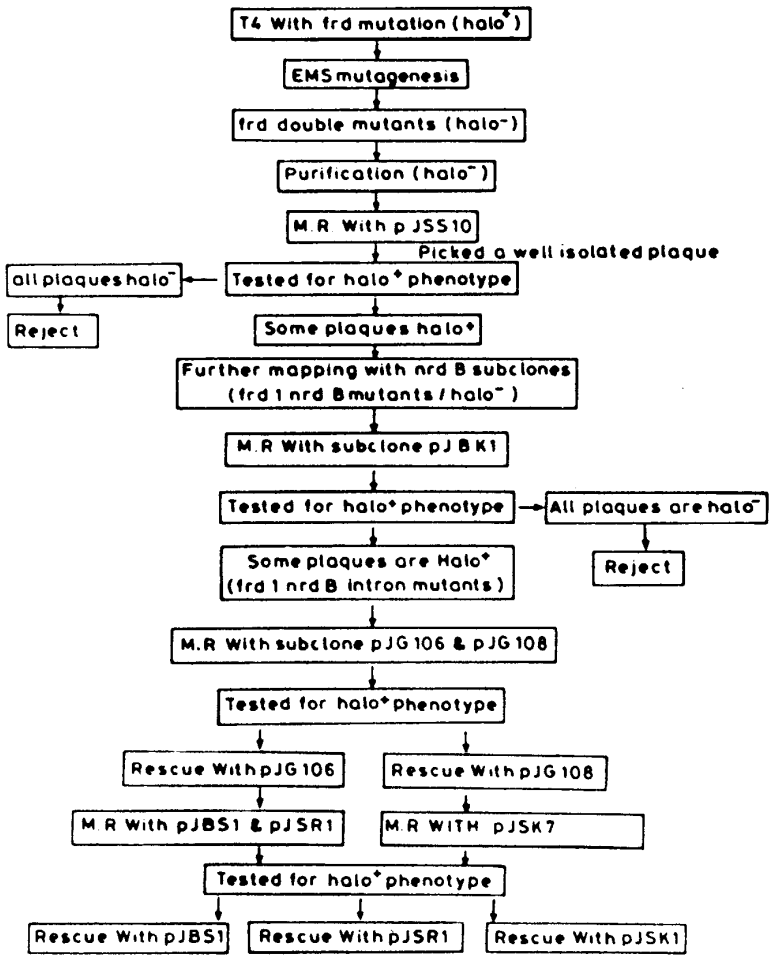

FIG. 2. Strategy for isolation and classification of the *nrd*B mutants into various intron regions by marker rescue using the white halo plaque phenotype. Halo$^+$ indicates that the halo phenotype was clearly visible at a minimum frequency of 5% of the plaques screened. Halo$^-$ indicates no plaques with halo phenotype (or less than 5%).

bases from the 5′ end of exon II (Fig. 1). The selected 156 phage were systematically crossed with the pJBK1 clone. Out of 156 phage, 31 exhibited the white halo phenotype with pJBK1, indicating that the mutation was within the *nrd*B intron. The frequency of the white halo phenotype plaques in the marker rescue picking varied among the different mutations. Most of the positive marker rescue picking from pJSS10 showed between 15-60% halos. Eight mutations (viz. *nrd*B11, *nrd*B49, *nrd*B57, *nrd*B112, *nrd*B342, *nrd*B445, *nrd*B717, *nrd*B779) showed below 10% halos. This suggests that these mutations may be situated towards one of the ends of the region covered by the JBK1 clone.

The various subclones that divide the pJBK1 intron into various smaller regions are shown in Fig. 1. JG106 and JG108 divide the JBK1 region into two parts. The JG106 region contains 442 bases and JG108 contains 239 bases. JG106 is divided into regions A (158 bases) and B (284 bases) by the plasmids pJBS1 and pJSR1, respectively. The JG108 region is divided into region C (96 bases) and D (143 bases) by the plasmid pJSK7. Each of the 31 mutants that mapped within or near the ends of the *nrd*B intron were tested for marker rescue with pJG106 and pJG108. Mutants that showed marker rescue with one of these clones was further tested for rescue with the pJBS1, pJSR1 or pJSK7 clones as shown in Table 1.

A total of 7 mutations mapped in the JG106 region which were further tested for marker rescue with pJBS1 and pJGR1. All seven mutations mapped in the JBS1 region. 24 mutations mapped in the JG108 region and were further tested for marker rescue with the pJSK7 clone (region D). Mutants that did not show positive marker rescue with subclone but had shown marker rescue with pJG108 mapped in region C. From a total of 24 mutants that mapped within the JG108 region, 13 mapped within the D region while the remaining 11 mapped in region C by default.

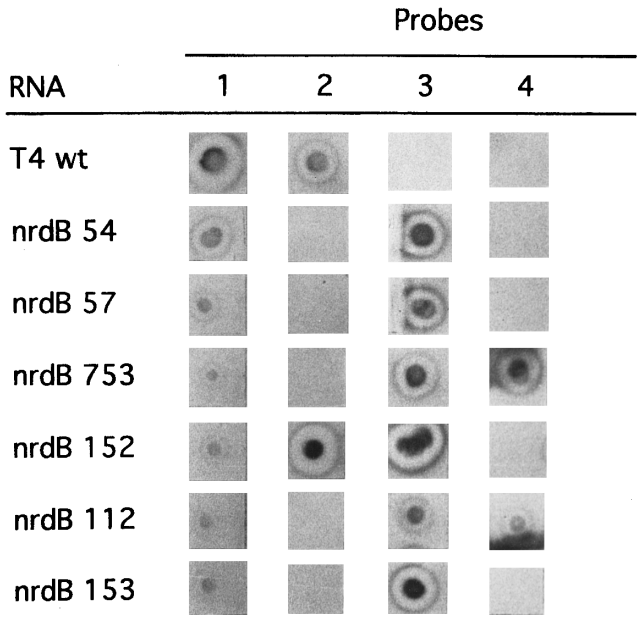

FIG. 3. Dot blot splicing assay for select splicing defective *nrd*B intron mutants. Probes 1, 2, 3 and 4 correspond to Exon II, Splice Junction, ExonI-Intron (5′ cutting) and Intron-ExonII (3′ cutting) oligonucleotides respectively.

The RNA from *E. coli* cells infected with splicing-defective *nrd*B intron mutants (Halo$^-$) phenotype as described above, was spotted on nitrocellulose paper. End-labelled DNA probes were designed such that their binding was dependent on the splicing of the *nrd*B intron [15]. Four DNA probes were used in this assay (Table 1 and Fig. 4). Appropriate *wt* RNA preparations were used to measure splicing proficiency. Mutants *nrd*B668, *nrd*B152, *nrd*B779, *nrd*B442, *nrd*B717 and *nrd*B730despite showing a Halo$^-$ phenotype showed no loss of spliced mRNA.

Twenty mutants (*nrd*B54, *nrd*B57, *nrd*B668, *nrd*B111, *nrd*B141, *nrd*B339, *nrd*B359, *nrd*B11, *nrd*B38, *nrd*B49, *nrd*B166, *nrd*B326, *nrd*B429, *nrd*B753, *nrd*B764, *nrd*B152, *nrd*B775, *nrd*B779, *nrd*B112 and *nrd*B153) showed a defect in 5′ cutting. Whereas 11 mutants (*nrd*B342, *nrd*B425, *nrd*B442, *nrd*B445, *nrd*B484, *nrd*B487, *nrd*B492, *nrd*B717, *nrd*B719, *nrd*B724 and *nrd*B730) showed no defect in 5′ cutting.

Fifteen mutations (*nrd*B54, *nrd*B57, *nrd*B668, *nrd*B111, *nrd*B141, *nrd*B399, *nrd*B359, *nrd*B38, *nrd*B49, *nrd*B166, *nrd*B326, *nrd*B429, *nrd*B152, *nrd*B775 and *nrd*B153) showed no defect in their 3′ cutting site. Whereas 16 mutations (*nrd*B11, *nrd*B753, *nrd*B764, *nrd*B779, *nrd*B112, *nrd*B342, *nrd*B425, *nrd*B442, *nrd*B445, *nrd*B484, *nrd*B487, *nrd*B492, *nrd*B717, *nrd*B719, *nrd*B724 and *nrd*B730) showed a defect in their 3′ cutting. Three mutants (*nrd*B668, *nrd*B152 and *nrd*B779) showed partial splicing defects. Four mutants (*nrd*B11, *nrd*B753,*nrd*B764 and *nrd*B112) showed a defect in their 5′ and 3′ cutting activity (Table 1)

The frequencey of halo+ phenotype recombinants was used to calculate the relative genetic distance between mutants that had point mutations mapping within the same marker rescue clone / subclone as shown in Fig. 3. For every genetic cross, approximately 5000 non-halo plaques were screened and the total genetic distance of that mutation from the ends of the clone / subclone and also the relative distance of mutants with each other (Table 2).

Different group I introns, including the *nrd*B intron,

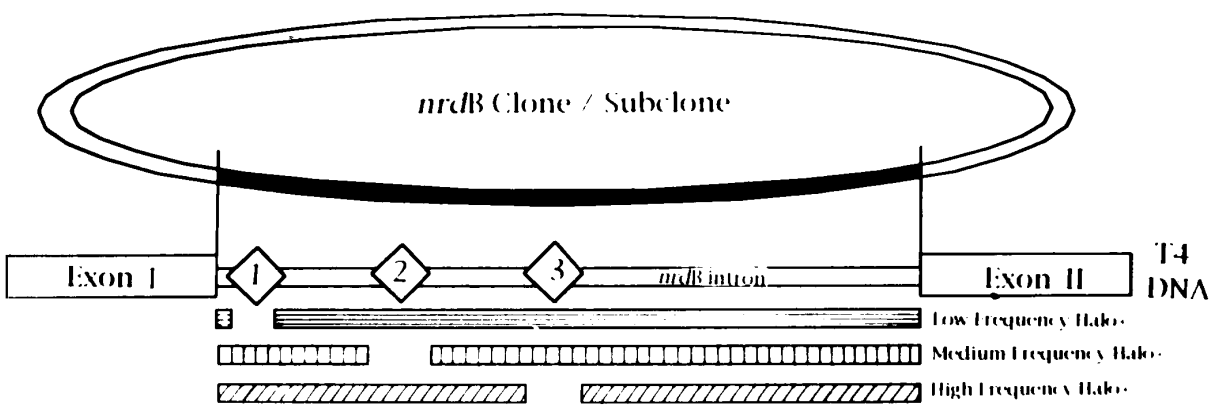

FIG. 4. A schematic view of the mechanism of marker rescue and its application in the prediction of genetic distance. The numbers 1, 2 and 3 within the rhombus show three different single point mutations within the *nrd*B intron for three independently isolated mutants. The plasmid contains the *wt* sequence for the *nrd*B clone being used for rescuing either one of the 1, 2 or 3 mutations (rhombus) within the *nrd*B intron of the T4 DNA. Depending upon the frequency of crossover, which is in direct relationship with the distance from the extremities of the *wt* insert, of relative genetic distance of that mutation can be estimated. Thus the mutation 3 will show a high degree of *wt* recombinants resulting in a high halo$^+$ frequency (diagonally shaded box). Mutation 2 will show a medium degree of *wt* recombinants showing halo$^+$ phenotype (vertically shaded box) and mutation 3 will show a low frequency of halo $^+$ phenotype (horizontally shaded box).

TABLE 1

Marker Rescue Mapping and Characterization of the Various *nrb*B Intron Mutants

Mutant number	Marker rescue assay with nrdB clones						Intron region	Conserved sequence	Mutations mapped	Dot-blot splicing assay				Splicing defect
	JBK1	JG106	JG108	JSK7	JBS1	JSR1				Ex II	Ex I - Ex II	Ex I - Int	Int - Ex II	
nrdB54	R	R	X	0	R	X	A(158)	P&Q	7	++	−	+++	−	5′
nrdB57	R	R	X	0	R	X				++	−	+++	−	5′
nrdB111	R	R	X	0	R	X				++	−	++	−	5′
nrdB141	R	R	X	0	R	X				++	−	+	−	5′
nrdB359	R	R	X	0	R	X				++	−	+	−	5′
nrdB399	R	R	X	0	R	X				++	−	+	−	5′
nrdB668	R	R	X	0	R	X				++	++	+++	−	PD(5′)
	0	0	0	0	0	0	B(284)	None	0					
nrdB11	R	X	R	X	0	0	C(96)	∼1/2R	11	++	−	+	++	5′&3′
nrdB38	R	X	R	X	0	0				++	−	+	−	5′
nrdB49	R	X	R	X	0	0				++	−	+	−	5′
nrdB152	R	X	R	X	0	0				++	+++	+++	−	PD(5′)
nrdB166	R	X	R	X	0	0				++	−	++	−	5′
nrdB326	R	X	R	X	0	0				++	−	++	−	5′
nrdB429	R	X	R	X	0	0				++	−	+++	−	5′
nrdB753	R	X	R	X	0	0				++	−	+++	+++	5′&3′
nrdB764	R	X	R	X	0	0				++	−	+++	+++	5′&3′
nrdB775	R	X	R	X	0	0				++	−	+++	−	5′
nrdB779	R	X	R	X	0	0				++	++	++	+++	PD(5′&3′)
nrdB112	R	X	R	R	0	0	D(143)	∼1/2 R&S	13	++	−	+++	++	5′&3′
nrdB153	R	X	R	R	0	0				++	−	+++	−	5′
nrdB342	R	X	R	R	0	0				++	−	−	++	3′
nrdB425	R	X	R	R	0	0				++	−	−	++	3′
nrdB442	R	X	R	R	0	0				++	+++	−	+++	PD(3′)
nrdB445	R	X	R	R	0	0				++	−	−	++	3′
nrdB484	R	X	R	R	0	0				++	−	−	+++	3′
nrdB487	R	X	R	R	0	0				++	−	−	+++	3′
nrdB492	R	X	R	R	0	0				++	−	−	++	3′
nrdB717	R	X	R	R	0	0				++	+++	−	++	PD(3′)
nrdB719	R	X	R	R	0	0				++	−	−	++	3′
nrdB724	R	X	R	R	0	0				++	−	−	++	3′
nrdB730	R	X	R	R	0	0				++	++	−	++	PD(3′)

Note. Mutants are listed from top to bottom in the various intron regions (A, B, C and D) they map in. Numbers within parentheses indicate the number of bases within the respective intron regions. R indicates positive marker rescue, X indicates negative marker rescue, and O indicates that the genetic crossing was not performed since the mutation showed no rescue with a larger subclone covering the same region of the intron. −, +, ++, +++ indicate the strength of radioactivity for the dot blot splicing assay where − indicates lowest or no signal and +++ indicates strongest signal. ExII (Exon II binding probe) was used as a positive control for the assay. ExI-ExII (Exon I-Exon II binding probe) was used for detection of splicing proficiency. The probes ExI-Int (ExonI-Intron) and Int-ExII (Intron-ExonII) further characterize the splicing defect and establish cutting defects at the 5′ and 3′ end of the intron respectively. PD indicates partial defect.

have relatively little sequence similarity, but all share a series of the short, conserved sequence elements P, Q, R and S with part of P/Q and R/S base pairing in conserved structure [16]. These conserved elements are known to be important, because they are sites of *cis*-acting splicing mutations. This is not to say that the less or non-conserved structures will be functionally irrelevant, they could serve to position the catalytic residues or to stabilize the overall structure of the intron, or they could provide binding sites for proteins that facilitate or regulate RNA splicing *in vivo* [16, 17, 18]. These sequence complementarity between pairs of elements may allow them to interact to form a complex higher order structure. This complex structure appears to be essential for splicing activity. The P and Q sequences are closer to the 5′ splice site of the intron whereas R and S sequences are closer to the 3′ splice site of the intron. The region A contains both P and Q sequences whereas region B (pJSR1) does not contain any of the conserved sequences. Region C contains about half of the R sequences that are less conserved than the region A. Region D contains the remaining half of the R sequences and the whole S sequences. In the present study we have obtained 7 mutations that mapped in the region pJG106 and all exhibiting positive marker rescue with pJBS1 while it is the biggest among four. Region D which contains a similar amount of conserved sequences to region A has 13 mutations

TABLE 2

Mutants Showing the Frequency of *wt* Recombinants from a Total of 5000 Plaques Screened for Every Genetic Cross

Mutant number	Intron region	Freq. of Halo+ by marker rescue		
		pJBK1	pJBS1	pJSK7
nrdB54	A	200	250	—
nrdB57	A	235	315	—
nrdB668	A	300	450	—
nrdB399	A	450	500	—
nrdB359	A	500	300	—
nrdB111	A	500	600	—
nrdB141	A	500	750	—
nrdB166	C	2000	—	—
nrdB326	C	2000	—	—
nrdB49	C	2450	—	—
nrdB753	C	2450	—	—
nrdB11	C	2500	—	—
nrdB779	C	2750	—	—
nrdB775	C	2850	—	—
nrdB38	C	3000	—	—
nrdB764	C	3150	—	—
nrdB152	C	3500	—	—
nrdB429	C	4200	—	—
nrdB342	D	225	—	900
nrdB717	D	250	—	250
nrdB112	D	300	—	900
nrdB724	D	300	—	200
nrdB153	D	500	—	1000
nrdB730	D	500	—	250
nrdB445	D	500	—	250
nrdB487	D	500	—	1250
nrdB719	D	550	—	1000
nrdB425	D	600	—	500
nrdB492	D	600	—	550
nrdB484	D	600	—	750
nrdB442	D	625	—	250

Note. Mutants are arranged in 5′ to 3′ order (see secondary structure model in Fig. 1) and increasing marker rescue frequency within each region. A dash indicates that the genetic crossing was not performed since the mutation showed no rescue with a larger subclone covering the same region of the intron.

mapped in it. Therefore the 3′ splice site of the *nrd*B intron is more proficient in splicing compared to the *td* intron which has shown 15 mutations within 167 nucleotides of the 5′ end and at least 12 mapped within 219 nucleotides of 3′ end of the intron [12]

Like many eukaryotic group I introns, the *nrd*B intervening sequence contains an open reading frame(ORF) [17]. The ORF sequences of other group I introns are not usually highly conserved and are looped out of secondary structure models for RNA folding [18]. Similarly for the *td* intron these genetic studies imply that the intron ORF is not required for formation of the active splicing conformation [12]. However our findings reveal that the in 3′ end of the *nrd*B intron ORF is essential for splicing since there are 11 mutants that mapped in the C region (between *Eco*RI and *Spe*I sites, see Fig. 1) which also contains the 3′ end of the intron ORF. These mapping results suggest that the conserved sequences in the *nrd*B intron, like the *td* and other group I introns, are important in pre-mRNA splicing.

ACKNOWLEDGMENTS

We thank Drs. D. H. Hall, D. Shub, and J. Gott for all their technical assistance and supplying various host strains and plasmid constructs. We are also grateful to Drs. K. K. Tewari and S. Jameel for their help in this project. This research was supported by the University Grants Commission fellowship to A.U.K. and internal funds from ICGEB.

REFERENCES

1. Cech, T. R. (1990) *Annu. Rev. Biochem.* **59,** 543–568.
2. Saldanha, R., Mohr, G., Belfort, M., and Lambowitz, A. M. (1993) *FASEB J.* **7,** 15–24.
3. Wittop Koning, T. H., and Grivell, L. A. (1986) *Trends Genetics* **2,** 51–55.
4. Chu, F. K., Maley, F., and Belfort, M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3049–3053.
5. Sjoberg, B.-M., Hahne, S., Mathews, C. Z., Mathews, C., Rand, K. N., and Gait, M. J. (1986) *EMBO J.* **5,** 2031–2036.
6. Gott, J. M., Shub, D. A., and Belfort, M. (1986) *Cell* **47,** 81–87.
7. Hall, D. H., Tessman, I., and Karlstrom, O. (1967) *Virology* **31,** 442–448.
8. Karlstrom, O. (1968) *J. Bacteriol.* **95,** 1069–1077.
9. Lal, S. K., and Hall, D. H. (1993) *Biochem. Biophys. Res. Commun.* **196,** 943–950.
10. Hall, D. H., and Tessman, I. (1966) *Virology* **29,** 339–345.
11. Drake, J. W., and Greening, E. D. (1970) *Proc. Natl. Acad. Sci. USA* **66,** 823–829.
12. Hall, D. H., Povinelli, C. M., Ehrenman, K., Pedersen-Lane, J., Chu, F., and Belfort, M. (1987) *Cell* **48,** 63–71.
13. Lal, S. K., and Hall, D. H. (1997) *Biophys. Biochem. Acta* **1350,** 89–97.
14. Belfort, M., Pedersen-Lane, J., West, D., Ehrenman, K., Maley, G., Chu, F., and Maley, F. (1985) *Cell* **41,** 375–382.
15. Kwon, H., Lal, S. K., and Hall, D. H. (1995) *Nuclosides and Nucleotides* **14,** 1811–1821.
16. Cech, T. R. (1988) *Gene* **73,** 259–271.
17. Chu, F. K., Maley, G. F., West, D. K., Belfort, M., and Maley, F. (1986) *Cell* **45,** 157–166.
18. Waring, R. B., and Davies, R. W. (1984) *Gene* **28,** 277–291.