STUDIES OF NOVEL METHODS OF GENE TRANSFER AND GENE REGULATION IN EUKARYOTIC CELLS

A Thesis
Submitted in Partial Fulfillment
of the Requirements for the
Master of Science Degree
in Chemistry

Steven G. Sogo
Division of Chemistry and Chemical Engineering
California Institute of Technology
Pasadena, CA 91125

June 23, 1988

TABLE OF CONTENTS

Specific Interactions Between Enhancer Elements and Two Cellular Factors

Keith D. Harshman, Steve Sogo and Carl S. Parker

K. D. Harshman and S. Sogo are graduate research assistants and C. S. Parker is an assistant professor of chemical biology in the Division of Chemistry of the California Institute of Technology, Pasadena, CA 91125.

Contribution #7369

Two factors isolated from calf thymus nuclei have been identified that bind specifically and independently to the SV40 enhancer. One partially purified enhancer binding protein (EP1) binds to a sequence of approximately 20 base pairs on the origin-proximal side of the SV40 enhancer. The other enhancer binding protein (EP2) binds to an approximately 20 base pair region containing the core consensus sequence on the origin-distal side of the SV40 enhancer. EP2 has been purified from calf thymus to apparent homogeneity by DNA affinity chromatography. The affinity resin was prepared by coupling a concatamerized synthetic oligonucleotide containing the core consensus sequence to Sepharose. Several distinct polypeptides that were isolated by affinity chromatography were further purified by preparative SDS-gel electrophoresis. After elution from the gel and renaturation, all of the protein species were shown to bind to the same core sequence as does native EP2. Both EP1 and EP2 will also bind specifically to a number of other enhancer elements. A comparison of the EP2 sites shows that they contain sequences with a high degree of homology to the core element GTGGAAAG. A comparison of EP1 sites reveals a potential conserved element with the sequence AGTpyACpy. Similar activities have also been identified in HeLa cell preparations, with the addition of a third binding activity not previously seen in the bovine thymus extracts.

Enhancers are DNA sequences that can potentiate transcription from an adjacent promoter at relatively great distances and in an orientation-independent fashion (for review see *1-4*). A large number of DNA and RNA viruses have been found to contain enhancer elements including SV40 (*5-7*), polyoma (*8*), BK (*9*), RSV (*10*) and mouse mammary tumor virus (*11*). Additionally, enhancer elements have been shown to be associated with cellular genes such as the rat chymotrypsin, insulin (*12*), and albumin (*13*), as well as the mouse histocompatibility genes (*14*) and immunoglobulin genes (*15-17*). One conclusion that emerges from these studies is that enhancer elements can activate transcription of a variety of different genes in various contexts.

The mechanism(s) by which enhancers stimulate transcription from a promoter in *cis* is unknown. Indeed, only recently has it been shown that an enhancer can function by stimulating transcription in a direct way (*18, 19*). These experiments revealed that the number of RNA polymerase molecules actively engaged in transcription of a human β-globin gene was greatly increased when the enhancer was adjacent to the gene. Other experiments have been reported which address various aspects of the proposed models for enhancer function (*20-22*). Unfortunately, it has not been definitively shown that any model is correct.

To determine the molecular mechanism of enhancer element function it is essential that proteins which recognize these elements be identified and purified to homogeneity. The presence of such cellular proteins has been inferred by *in vivo* competition studies (*23, 24*). In addition, genomic footprinting experiments (*25*) and footprints using crude nuclear extracts (*26*) have also supported the idea that *trans*-acting factors directly interact with enhancer sequences. The proteins responsible for the apparent binding, however, have yet to be purified and biochemically characterized. Furthermore, it is likely that enhancer binding proteins will be rare,

thus an abundant source of material in which enhancer binding proteins have been identified would be desirable.

With these goals in mind we have identified two enhancer binding proteins from calf thymus nuclear lysates. One of these proteins, EP1 (enhancer binding protein 1) has been partially purified through several chromatographic steps. A second activity, EP2 (enhancer binding protein 2) has been extensively purified using DNA-affinity chromatography (*27, 28*) and found to consist of several protein species. EP1 and EP2 binding proteins are chromatographically distinct and bind independently to separate domains of the SV40 enhancer. These proteins also recognize and bind to other enhancer sequences of both a viral and cellular origin. Preparations of highly purified enhancer binding proteins will now permit a detailed biochemical characterization of the mechanism(s) by which these enhancer binding factors activate transcription.

Identification and Partial Purification of Enhancer Binding Proteins

Calf thymus was selected as the initial starting material for the identification and purification of enhancer binding proteins because of its relative low cost and abundance. It was not, however, intrinsically obvious that calf thymus would be a good source for enhancer binding proteins. An extract was prepared from freshly obtained calf thymus following the protocol described in the legend to Fig. 1. A direct analysis of the crude extract failed to reveal the presence of proteins that could specifically bind to the SV40 enhancer as determined by DNAse I footprinting experiments. The extract was applied to a DEAE-cellulose column and the flow through applied directly to a heparin-Sepharose column and the adsorbed proteins eluted with a linear salt gradient. This step resulted in a net 3-fold increase in the concentration of bound proteins. At this point it is possible to see (by DNase I footprinting) specific interactions in two regions of the SV40 enhancer (Fig. 1 and 2, panels A). These two regions are approximately 20 base pairs in size and are

positioned over the 5′ and 3′ ends of the SV40 72 base pair repeat. The binding activity that elutes first from the heparin-Sepharose column was designated EP2. This protein binds to the distal end of the SV40 72 base pair repeat [relative to the 21 base pair repeats and the origin of replication (see Fig. 3, panel C)]; included in the binding site is the core consensus sequence (*29*). The other binding activity elutes at a higher salt concentration from the heparin-Sepharose resin and was designated EP1. EP1 binds to a 20 base pair region of the SV40 enhancer proximal to the 21 base pair repeats (Fig. 3, panel C). EP1 forms a significantly weaker footprint at this stage in the purification than does EP2. This is probably because the EP1 activity is present at lower concentration in the extract than EP2. EP1 was further purified by phosphocellulose and DNA cellulose chromatography as shown in Fig. 1. EP2 was purified to apparent homogeneity by AcA44 chromatography and sequence-specific DNA-affinity chromatography as shown in Fig. 2 and described in detail below. A flow diagram describing the purification of both EP1 and EP2 is shown in Fig. 4.

Detailed Analysis of EP1 and EP2 Interactions with the SV40 Enhancer

Detailed footprint analysis of EP1 binding to the SV40 enhancer has been performed using both DNase I (*30*) and DMS (dimethyl sulfate) (*31*). A DNase I footprint of EP1 on both DNA strands of the SV40 enhancer is shown in Fig. 1 (panel D). Protection is observed on the top strand (the coding strand for early gene expression) between residues 113 and 127. Binding is also observed on the complementary bottom strand (the non-coding strand for early gene expression) between residues 110 and 130.

Dimethyl sulfate was used to determine both minor groove adenine contacts and major groove guanine contacts between EP1 and its target sequence. Residues protected from methylation are absent or diminished in intensity compared to control reacitons in Fig. 1 (panel D) and indicate a close contact by EP1. Similarly, some bases exhibit a greatly enhanced reactivity toward DMS alkylation in the

presence of bound protein, also indicating a close interaction between the protein and DNA. Those purine residues whose reactivity to DMS are modified by the binding of EP1 are shown alongside the data in Fig. 1 (panel D). Guanine residue 113 of the top strand shows an enhanced DMS reactivity while guanine 117 of the top strand and guanines 112 and 115 of the bottom strand show a decreased DMS reactivity. In addition, a minor groove interaction is observed on adenine 110 of the bottom strand. An anlaysis of EP1 binding activity has been conducted on the Rous sarcoma viral and polyoma viral enhancers as well as the immunoglobulin heavy chain enhancer (results not shown). A comparison of all these sites bound by the partially purified factor reveals a potential conserved element with the sequence AGTpyAGpy.

The SV40 element bound by EP1 has been implicated in enhancer function by recent *in vivo* studies as well as *in vitro* footprinting studies using HeLa cell crude extracts (*26, 32, 33*). The footprinting results obtained with partially purified calf thymus extract differ somewhat from those obtained with the HeLa cell crude extract. The HeLa cell crude extract footprints are relatively more complex and show a smaller protected region in the EP1 binding domain. Although it is difficult to make direct comparisons between the calf thymus activities and those observed with the crude extracts from HeLa cells, a very good correlation can be observed once the HeLa cell EP1 activity is further purified (see below).

EP2 binding to the SV40 enhancer has also been examined in detail. DNase I protection on the top strand as a result of EP2 binding lies between residues 160 and 179 (Fig. 3, panel B). On the bottom strand protection is observed between residues 160 and 175 (Fig. 3, panel A). It is interesting to note that the core sequence is situated asymmetrically to the distal side of the DNase I footprint. DMS footprints on the EP2 sequence reveal two enhanced guanine cleavages at 156 and 157 and one reduced guanine cleavage at 155 of the bottom strand. Mutations of these residues

(156 and 157) have little effect, however, on enhancer function *in vivo* (*29*). EP2 interactions determined by DMS on the top strand consist of two strongly suppressed methylations at guanines 172 and 173 and a weakly enhanced guanine reaction at residue 168 (Fig. 3, panel B). The two guanine residues at positions 172 and 173 are critical for *in vivo* enhancer activity as shown by Herr and Clark (*34*).

A third reagent, MPE (methidium propyl-EDTA) was used to determine more precisely the boundaries of the EP2 recognition sequence. MPE's relatively small size allows it to bind and cleave DNA nearer to the sites of protein-DNA interaction than DNase I (*35, 36*). MPE will also cleave DNA with less sequence-specificity than DNase I, so one obtains a very uniform cleavage ladder, allowing a more precise determination of the boundaries of protein binding. MPE footprints with partially purified EP2 were performed on both strands of the SV40 enhancer and are shown in Fig. 3 (panel A and B). A small region of protection is observed over nine residues on the bottom strand between positions 167 and 175, inclusive. On the top strand, an eight base pair region of protection is observed between residues 165 and 172. As expected, this small protected region includes the core consensus sequence (*29*).

EP2 was also observed to protect a region outside of the 72 base pair repeats, between bases 263 and 291. The protected sequences can be seen in the column profiles of Fig. 2 as well as in Fig. 5, 6 and 7. Contained within this region of DNase I protection are two sequences homologous to the core sequence. The relative positions of all the EP2 binding sites determined in our footprint assays are shown in Fig. 3 (panel C). These EP2 sites correspond reasonably well to those observed with HeLa cell crude extracts (*26*). A systematic mutagenesis screen through the SV40 enhancer, by Chambon and colleagues has shown that all of the EP2 sites are important for enhancer activity *in vivo* (*32*).

A Double-Point Mutation in the Core Sequence Prevents EP2 Binding.

Herr and Clarke (*34*) have shown that a double point mutation in the core consensus sequence (dpm 26) greatly reduces enhancer function *in vivo*. If EP2 is a protein that in some way mediates enhancer function, then one might predict that such a double point mutation would reduce EP2 binding. Shown in Fig. 5 is a titration of partially purified EP2 on the wild-type SV40 enhancer (left) and on the double point mutation (right). It is evident that EP2 is unable to bind to the sequence containing the point mutations at levels which fully protect the wild-type sequence. It is noteworthy (as shown in Fig. 5) that the two alterations in dpm 26 correspond to guanine residues 172 and 173 (wild-type GTGGAAAG to dpm 26 GTCCAAAG). These residues were shown by DMS footprints to make close contact to EP2 (Fig. 3, panel B). Guanine 172 and 173 are also included within the MPE footprint (Fig. 3, panel A and B). Evidence showing the *in vivo* importance of guanine 172 for SV40 enhancer function has been described by Weiher *et al.* (*29*). In this case, mutation of guanine 172 to adenine completely abolished enhancer activity. As mentioned in the previous section, an EP2 binding site exists outside of the 72 base pair repeat, this site serves as an internal control for EP2 binding activity for the dpm 26 template used in the experiment shown in Fig. 5.

Purification of EP2 to Apparent Homogeniety

The EP2 binding activity was purified to near homogeneity using a combination of gel filtration and affinity chromatography. Gel filtration was performed on an AcA44 resin. The chromatographic profile of EP2 activity on this resin is shown in Fig. 2 (panel B). The activity profile was bimodal with a small amount of EP2 activity eluting in the void volume of the column while greater than 80% of the EP2 activity was well included within the column. The included activity has an apparent molecular weight between 50 and 30 kD (fractions 32 through 43). These fractions were pooled and concentrated on Biorex 70 prior to DNA-affinity

chromatography. (The low molecular weight DNase activity that elutes between fractions 44 through 68 was discarded.)

The bimodal elution of EP2 binding activity can be explained in a number of ways. Partial proteolysis of a large intact protein may result in the accumulation of active EP2 proteins possessing a molecular weight of 50 kD and below. The activity that eluted in the void volume may be the unproteolyzed native protein or an active proteolysis product of an even larger protein. Alternatively, the two activities may represent two different gene products (intact or degraded) which bind to the core sequence. It is also possible that the higher molecular weight activity is actually an aggregate of the smaller species, in which case its position in the elution profile would not correctly reflect its size. The majority of the binding activity, however, clearly eluted in the 50-20 kD range, and these fractions were used to complete the purification using DNA-affinity chromatography.

A DNA-affinity resin was constructed by concatamerizing a synthetic oligonucleotide containing the core consensus sequence and coupling the concatamers to cyanogen bromide-activated Sepharose (27). The active fractions from the Biorex 70 concentration step (Fig. 4) were slowly applied to the column. Next, the column was washed extensively with buffer at a salt concentration high enough to elute the majority of non-specifically bound proteins from the resin, while the EP2-site specific binding protein(s) remain adsorbed. The tightly bound proteins were eluted with a salt gradient from 0.27 to 1.5 M KCl. To complete the purification, it was necessary to repeat the affinity chromatography step to remove minor contaminating proteins. DNase I footprint assays of the fractions eluted from the first passage through the affinity resin are shown in Fig. 2 (panel C). The binding activity elutes between approximately 0.35 M KCl and 1.1 M KCl. DNase I footprint assays of fractions from the second passage are shown in Fig. 6 (panel A). Analysis of the protein composition of the active fractions by SDS polyacrylamide gel

electrophoresis is shown in Fig. 6 (panel B). Several polypeptides are observed, including a major species at 39 kD, and two other species with molecular weights of 43 kD and 41 kD. Additionally, an interesting series of evenly spaced polypeptides ranging in molecular weight from 38 to 34 kD are observed.

To identify which of these polypeptides is responsible for the specific binding activity observed we performed a preparative SDS protein gel purification step. The protein bands were visualized by KCl staining, individually excised and electrophoretically eluted from each gel slice. The eluted proteins were further denatured with guanidium-HCl and renatured by slow dialysis (*37*; see Fig. 7 legend for complete details). Each eluted and renatured protein was subsequently analyzed on an SDS protein gel to measure recovery and purity (Fig. 6, panel B; note that the 70 kD band present in each lane is BSA added after electroelution to aid in protein recovery during the renaturation procedure). In this particular experiment two bands from the 34-38 kD cluster as well as the 43 kD band were not efficiently recovered (Fig. 6B). The binding activity from each of the isolated polypeptides is shown in Fig. 7 (panel A). Footprinting reactions were conducted with 5 µl and 40 µl from each fraction. A footprint was observed with each of those polypeptides recovered in good yield from the gel. In other experiments, the 43 kD polypeptide has been shown to bind to the core sequence in a fashion identical to the other polypeptides (data not shown).

The relationship between the bands exhibiting binding activity is unclear. It is possible that these proteins are proteolysis products of a single, high molecular weight protein (as mentioned earlier). A more interesting possibility is that these different protein species represent modified forms of the same protein. Such putative modifications might be important for their overall activity independent from DNA-binding. It is also possible that several distinct gene products exist, each with its own *in vivo* function, which bind to the core enhancer sequence. Monoclonal

antibody preparation against the EP2 polypeptides should allow these possibilities to be resolved. A summary of the purification of EP1 and EP2 is given in Table 1.

EP2 Binds to Cellular Enhancer Sequences

Enhancer sequences have been identified in mammalian immunoglobulin genes. In the case of the μ heavy chain gene, the enhancer is found in the intervening sequences between the constant and joining regions (*15, 17*). Recently, evidence has been presented suggesting that factors present in a number of cell lines will bind to several sites on the heavy chain enhancer element both *in vivo* and *in vitro*. The sequence-specific binding observed *in vivo* was only observed in lymphoid cell lines (*25*). In contrast, specific binding was observed on the same enhancer element *in vitro* with extracts derived from both lymphoid and non-lymphoid cell lines (38).

Partially purified EP2 from calf thymus was able to bind specifically to three portions of the μ enhancer as shown in Fig. 8. All three sites contain one or two elements of close homology to the SV40 core consensus sequence. Although a heterologous tissue sourse, calf thymus glands are composed primarily of thymocytes. Therefore, T-cells probably contain at least the two general enhancer binding proteins, EP1 and EP2. It is perhaps surprising that those sites in Fig. 8 were not revealed in the *in vivo* footprints with B-cells (*25*). Recent data presented by Sigh and Baltimore (*38*) indicates that B-cell extracts might contain an EP2 like activity. However, competition studies show that the μ heavy chain enhancer (identical to the one used in our experiments) did not compete for the binding of this factor. Purification of this B-cell activity, along with more detailed binding studies will be required to resolve this issue.

Enhancer Binding Proteins from HeLa Cells

Our primary reason for studying enhancer binding proteins is to determine their role in transcriptional activation. Calf thymus nuclear extracts, however, have

not been found to be active for *in vitro* transcription studies. A large number of possible factors can be involved to explain this experimental observation, but the extracts as currently prepared will not permit the kind of experimentation necessary to determine the activities of the enhancer binding proteins. HeLa cells have already proven to be a good source for active RNA polymerase II transcription factors. Because the SV40 early and late promoters are actively transcribed in HeLa cell nuclear extracts (*39*), we examined HeLa cell nuclear extracts for the presence of SV40 enhancer binding proteins. As with the calf thymus nuclear extracts, DNase I footprinting was used as the assay to determine the presence of these proteins.

The presence of enhancer binding proteins in HeLa cell extracts has been previously inferred by both *in vivo* and *in vitro* transcription as well as with footprinting data using crude extracts (*26, 32, 33*). HeLa cell nuclear extracts prepared as described by Parker and Topol (*40*) were chromatographed on DEAE-cellulose followed by heparin-Sepharose. Shown in Fig. 9 (panel A) is the DNase I footprinting pattern of the heparin-Sepharose column fractions on the 72 bp repeat of the SV40 enhancer. Clearly visible are footprints over the previously described EP1 and EP2 sites. Also visible is a third protected region not previously seen with the calf thymus nuclear extracts. This binding site includes the octamer sequence that has been found in important control regions of the human histone H2b, immunoglobulin and U2 small nuclear RNA genes (*41-43*). The protein(s) responsible for the binding observed to the octamer sequence from HeLa cells is designated EP3.

A more detailed comparison between the HeLa cell and calf thymus EP2 activities is shown in Fig. 9 (panel B). The DNase I footprinting pattern produced by the HeLa cell EP2 is smaller in size than the calf thymus EP2. The human EP2 used in the experiment shown in Fig. 9 (panel B) had been further purified after the heparin-agarose step by core consensus sequence DNA affinity chromatography.

Although the human EP2 footprint is smaller than the calf thymus EP2 footprint, it is centered over the core consensus sequence. The functional significance of the difference in the footprint dimensions, if any, is unknown and may simply reflect a difference in size of the two binding activities. HeLa cell derived EP1 after heparin-Sepharose chromatography appears to footprint identically to EP1 from calf thymus.

Prospectus

With the development of sequence-specific affinity chromatography, biochemists are in a position to purify very rare protein factors that bind to critical control regions of the genes that they are studying. The purification of these proteins will be essential for detailed analysis of their mechanism of action. In the current case, sequence-specific affinity chromatography has allowed the homogeneous purification of a set of polypeptides, designated EP2, which bind to the core enhancer element. Homogeneous preparations of this binding activity will allow the determination of a partial amino acid sequence and preparation of antibodies against the protein(s); both endeavors are currently being undertaken in this laboratory. The antibodies will be useful in studying the localization and potential modifications that might occur to the sequence-specific binding proteins. The partial amino acid sequence will also facilitate cloning the gene for the factor by allowing synthesis of appropriate oligonucleotide probes to screen a recombinant DNA library. Additionally, sequence-specific affinity chromatography is being used to purify both the human EP1 and EP3 binding activities. It is hoped that homogeneous preparations of these enhancer binding proteins, when used with an *in vitro* transcription system, will allow the determination of the molecular mechanism(s) by which these proteins bring about transcriptional enhancement.

REFERENCES

1. P. Chambon, A. Dierich, M.P. Gaub, S. B. Jakowlew, J. Jongstra, A. Krust, J. P. Le Pennec, P. Ondet, T. Ruedelhuber, *Recent Prog. Horm. Res.* **40**, 1-42 (1984).
2. G. Khoury and P. Gruss, *Cell* **33**, 313-314 (1983).
3. *Eukaryotic Transcription - The Role of cis and trans Acting Elements in Initiation*, Y. Gluzman, Ed. (Cold Spring Harbor Laboratory, 1985).
4. B. Wasylyk, in *Maximizing Gene Expression*, W. Reznikoff and L. Gold, Eds. (Butterworth, London, 1986).
5. J. Banerji, S. Rusconi, W. Schaffner, *Cell* **27**, 299-308 (1981).
6. P. Moreau, B. Ritten, B. Wasylyk, R. Everett, M. P. Gaub, P. Chambon, *Nucl. Acids Res.* **9**, 6047-6068 (1981).
7. M. Fromm and P. Berg, *J. Molec. Appl. Genet.* **1**, 457-481, (1982).
8. P. Herbomel, B. Bonlachot, M. Yaniv, *Cell* **39**, 653-662 (1984).
9. N. Rosenthal, M. Kress, P. Gruss, G. Khoury, *Science* **222**, 749-755 (1983).
10. P. A. Luciw, J. M. Bishop, H. E. Varmus, M. R. Capecchi, *Cell* **32**, 705-716 (1983).
11. V. L. Chandler, B. A. Maler, K. R. Yamamoto, *Cell* **33**, 489-499 (1983).
12. M. D. Walker, T. Edlund, A. M. Boulet, W. J. Rutter, *Nature* **306**, 557-561 (1983).
13. M. O. Ott, L. Sperling, P. Herbomel, M. Yaniv, M. C. Weiss, *EMBO J.* **3**, 2505-2510 (1984).
14. S. D. Gilles, V. Folsoun, S. Tonagawa, *Nature*, **310**, 594-597 (1984).
15. J. Banerji, L. Olson, W. Schaffner, *Cell* **33**, 729-740 (1983).
16. C. Queen and D. Baltimore, *Cell* **33**, 741-748 (1983).
17. S. D. Gilles, S. L. Morrison, V. T. Oi, S. Tonegawa, *Cell* **33**, 717-728 (1983).
18. R. Treisman and T. Maniatis, *Nature* **315**, 72-75 (1985).

19. F. Weber and W. Schaffner, *Nature* **315**, 75-77 (1985).

20. S. E. Plon and J. C. Wang, *Cell* **45**, 575-580 (1986).

21. A. J. Courey, S. E. Plon, J. C. Wang, *Cell* **45**, 567-575 (1986).

22. X.-F. Wang and K. Calame, *Cell* **47**, 241-247 (1986).

23. H. R. Schöler and P. Gruss, *Cell* **36**, 403-411 (1984).

24. M. Mercola, J. Goverman, C. Mirell, K. Calame, *Science* **227**, 266-270 (1985).

25. A. Ephrussi, G. Church, S. Tonegawa, W. Gilbert, *Science* **227**, 134-140 (1985).

26. A. G. Wildeman, M. Zenke, C. Schatz, M. Wintzerith, T. Grundström, H. Matthes, K. Takahashi, P. Chambon, *Molec. Cell. Biol.* **6**, 2098-2105 (1986).

27. G. Wiederrecht, D, J. Shuey, W. A. Kibbe, C. S. Parker, *Cell* **48** (1987).

28. J. T. Kadonaga and R. Tjian, *Proc. Natl. Acad. Sci. USA* **83**, 5889-5893 (1986).

29. H. Weiher, M. König, P. Gruss, *Science*, **219**, 626-631 (1983).

30. P. Galas and A. Schmitz, *Nucl. Acids Res.* **5**, 3157-3170 (1978).

31. U. Siebenlist and W. Gilbert, *Proc. Natl. Acad. Sci. USA* **77**, 122-126 (1980).

32. M. Zenke, T. Grundström, H. Matthes, M. Wintzerith, C. Schatz, A. Wildeman, P. Chambon, *EMBO J.* **5**, 387-397 (1986).

33. I. Davidson, C. Fromental, P. Augereau, A. Wildeman, M. Zenke, P. Chambon, *Nature* **323**, 544-548 (1986).

34. W. Herr and J. Clarke, *Cell* **45**, 461-470 (1986).

35. M. Van Dyke and P. Dervan, *Nucl. Acids Res.* **11**, 5555-5567 (1983).

36. M. Sawadogo and R. Roeder, *Cell* **43**, 165-175 (1985).

37. D. Hager and K. Burgess, *Anal. Biochem.* **109**, 76-86 (1980).

38. R. Sen and D. Baltimore, *Cell* **46**, 705-716 (1986).

39. W. S. Dynan and R. Tjian, *Nature* **316**, 774-778 (1985).

40. C. S. Parker and J. Topol, *Cell* **36**, 357-369 (1984).

41. H. L. Sive, R. G. Roeder, *Proc. Natl. Acad. Sci. USA* **83**, 6382-6386 (1986).

42. N. F. Landolfi, J. D. Capra, P. W. Tucker, *Nature* **323**, 548-551 (1986).

43. L. M. Standt, H. Singh, R. Sen, T. Wirth, P. Sharp, D. Baltimore, *Nature* **323**, 640-643 (1986).

44. M. M. Bradford, *Anal. Biochem.* **72**, 248-254 (1976).

45. B. Alberts, G. Herrick, *Meth. Enzymol.* **21**, 198-213 (1971).

46. B. L. Davidson, T. Leighton, J. C. Rabinowitz, *J. Biol. Chem.* **254**, 9220-9226 (1979).

47. A. Maxam, W. Gilbert, *Meth. Enzymol.* **65**, 499-560 (1980).

48. *Molecular Biology of Tumor Viruses*, 2nd Edition, Revised: *DNA Tumor Viruses*, J. Tooze, Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1981).

49. U. Laemmli, *Nature* **227**, 680-685 (1970).

50. We thank Winship Herr for the gift of the dpm 26 and 1 x 72 SV40 clones. We also are indebted to the members of the Parker lab for their valuable comments on this manuscript. We also are grateful to Phoebe Ray for her expert preparation of the manuscript. CSP is a Rita Allen Foundation Scholar. This research was supported in part by NIH Biomedical Research Support Grant RR07003.

FIGURE LEGENDS

Fig. 1. Purification and binding analysis of EP1. All steps are performed at 4°C. The calf thymus (1 Kg) tissue was blended at low speed in 3 l of A buffer [15 mM KCl, 10 mM Hepes (pH 7.6), 2 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF] for 1 minute, followed by 30 seconds at medium speed and 30 seconds at high speed. The nuclei were then collected by centrifugation at 10,000 rpm for 10 minutes in a Beckman JA10 rotor (Beckman J221 centrifuge). The nuclei pellets were resuspended in 2 l of A buffer plus 1 mg/ml 1,10-phenanthroline and 1 mg/ml benzamidine hydrochloride (Sigma) by blending at low speed for 10 seconds. The nuclei were lysed by the addition of 310 ml (1/10 volume) of 4 M ammonium sulfate (pH 7.9). At this stage diisopropyl-fluorophosphate (Sigma) was added to 0.4 mM. The DNA was sedimented from the lysate by centrifugation at 35,000 rpm for 90 minutes in a Beckman 45 Ti rotor. The clear supernatant was collected and the protein precipitated by the addition of 0.25 gm ammonium sulfate per ml of supernatant. After mixing for 30 minutes, the precipitated proteins were collected by centrifugation at 15,000 rpm for 15 minutes in a Beckman JA10 rotor. The supernatant was discarded and the pellet was resuspended in 100 ml C buffer [10% glycerol, 50 mM Hepes (pH 7.6), 40 mM KCl, 0.1 mM EDTA, 0.1 mM DTT]. The protein suspension was subsequently dialyzed against C buffer for 8 hours. The standard chromatography buffer was HGED, which consists of 25 mM Hepes (pH 7.6), 10% (v/v) glycerol, 0.1 mM EDTA, 1 mM dithiothreitol. KCl was added and its molarity indicated as follows: 0.1 HGKED is 0.1 M KCl in the otherwise standard buffer. The protein concentrations were determined by the Bradford procedure with BSA as the standard (*44*). DEAE-cellulose (Whatman DE52), phosphocellulose (Whatman P11) and Biorex 70 (Bio-Rad), resins were all prepared according to the manufacturer's recommendations. Double-stranded DNA cellulose was prepared following the method of Alberts and Herrick (*45*). Heparin-Sepharose was prepared

following the procedure of Davidson *et al.* (*46*). (A) A 300 ml DE52 column was packed and equilibrated with 0.1 M HGKED. 150 ml of calf thymus nuclear extract was loaded onto the column and washed with 900 ml 0.1 M HGKED. The proteins which flowed through the DE52 column were passed directly onto a 90 ml heparin-Sepharose column equilibrated with 0.1 M HGKED. Following a 300 ml 0.1 M HGKED wash the heparin-Sepharose bound proteins were eluted with a 5-column volume linear salt gradient from 0.1 M to 0.8 M HGKED. Protein containing fractions were dialyzed and assayed (40 µl) by DNase I footprinting (see legend to Fig. 5 for DNase I footprinting procedure). The fragment used in the footprint analysis was 1 x 72 (Fig. 5) cut and 5′ end-labeled at the BamHI with the secondary cleavage at EcoRI. The boundaries of the EP1 footprint are marked; the fractions containing EP1 binding activity are bracketed. Fractions 24 and 25 were pooled and dialyzed against 0.1 M HGKED. Lane S is a minus protein control. (B) This material was loaded onto a 10 ml P11 column equilibrated with 0.1 M HGKED. The column was washed with 30 mls 0.1 M HGKED and bound protein eluted with a 5-column volume linear salt gradient from 0.1 to 1.0 M HGKED. The fractions were dialyzed and assayed (30 µl) using the 1 x 72 fragment. (C) Fractions 19 through 24 from P11 chromatography were dialyzed and loaded onto a 5 ml dsDNA cellulose column. The column was washed and bound proteins eluted with a salt gradient from .06 to 0.8 M HGKED. Fractions 12 through 16 were dialyzed and assayed (20 µl). Active fractions were frozen in small aliquots at -80°C. Lane S is a minus protein control. (D) DNase I and DMS footprints with EP1. Plasmid pXho (see Fig. 3, panel C) was 5′ or 3′ end-labeled at Xho and re-cut with the KpnI. DNase I footprinting experiments were performed with 20 µl of the DNA cellulose chromatography fractions following the procedure described in the legend to Fig. 5. Lanes 1 (Top and Bottom strands) are DNase I controls where no protein was added; lanes 2 (Top and Bottom strands) show the protection patterns resulting from the

addition of partially purified EP1. The bases included in the footprint and their coordinates are shown. DMS footprints were performed as follows: binding reactions, conducted for 5 minutes at 22°C in a total volume of 25 µl, were 5 mM $MgCl_2$. DMS was added to a final concentration of 5 mM. The mixture was incubated at 22°C for 5 minutes. The DMS reaction was terminated by addition of 6 µl stop solution (1.5 M sodium acetate (pH 7.0), 1.0 M mercaptoethanol, 100 µg/ml tRNA), followed by the addition of 50 µl of sarkosyl solution [1% sarkosyl (w/v), 0.1 M Tris (pH 7.8)], 0.1 M NaCl, and 0.01 M EDTA). After phenol and ether extractions, the DNA was precipitated with ethanol. The DNA pellet was resuspended in 20 µl TE (10 mM Tris (pH 7.8), 1 mM EDTA). Base cleavage was accomplished by addition of NaOH to a final concentration of 0.15 M and heating at 90°C for 20 minutes. The DNA was precipitated with ethanol, dried and electrophoresed according to the procedures of Maxam and Gilbert (*47*). Lanes 3 (Top and Bottom strands) are negative controls showing the DMS cleavage pattern in the absence of added protein. Lanes 4 (Top and Bottom strands) show the effect of 20 µl partially purified EP1. Bands whose reactivity to DMS are altered by EP1 are indicated with arrows. The effected bases are circled in the accompanying sequences and marked as to whether their reactivity toward DMS has been increased (arrow up) or decreased (arrow down). Lane G shows the chemical cleavages specific for guanine residues. Lane R shows the chemical cleavages specific for purine residues. Numbering scheme is acording to Tooze (*48*).

Fig. 2. Purification of EP2. (A) Heparin-Sepharose chromatography was performed as described in the legend to Fig. 1. The binding site as well as those fractions containing EP2 binding activity are shown by brackets. Fractions 21, 22 and 23 were pooled and dialyzed against 0.15 M HGKED. The protein was concentrated by applying the sample to a 30 ml Biorex 70 column and eluting the protein with 0.7 M

HGKED (see Fig. 4). (B) The protein sample was then loaded onto a 500 ml AcA44 column equilibrated with 0.7 M HGKED. The protein was eluted with the same buffer and the fractions assayed for activity by DNase I footprinting. The active fractions - 32 through 44 - were pooled, dialyzed against 0.15 M HGKED and concentrated as described above on a 2.0 ml Biorex 70 column. (C) The concentrated protein sample was dialyzed against 0.25 M HGKED and slowly loaded onto a 0.8 ml EP2 oligonucleotide Sepharose column equilibrated with the same buffer. The column was prepared according to the procedure of Wiederrecht *et al. (27)*. After loading, the resin was washed with 50 column volumes of 0.27 M HGKED. The bound protein was eluted from the resin with a 5 column volume linear salt gradient from 0.27 to 1.5 M HGKED. EP2 binding activity eluted in fractions 12 through 26. These fractions were pooled, dialyzed, and subjected to one additional round of affinity chromatography. The active fractions were pooled, dialyzed against C buffer, and stored at -80°C. An aliquot of the pooled material was further purified by preparative SDS gel electrophoresis (see text and Fig. 7). The results of the second pass over the affinity column are shown in Fig. 6.

Fig. 3. Analysis of EP2 binding to the SV40 72 base pair repeat. All footprinting reactions were conducted with 20 µl purified EP2. Fragments used for footprinting reactions are described in the legend to Fig. 2, as are the DMS footprinting reaction conditions. DNase I footprinting reaction conditions are described in the legend to Fig. 5. MPE·Fe(II) footprinting reactions were conducted as follows: reactions were carried out in 10 mM Tris (pH 7.4) and 50 mM NaCl in a final reaction volume of 30 µl. Each reaction contained 20 µl purified EP2, 2 ng end-labeled DNA fragment (12,000-15,000 cpm) and 750 ng freshly denatured salmon sperm DNA. After a 10 minute binding reaction at 22°C, 1.5 µl 100 µM MPE·Fe(II) was added and incubated at 22°C for 3 minutes. Cleavage was initiated by addition of 1 µl 40

mM dithiothreitol and continued for 10 minutes at 22°C. Cleavage was stopped by the addition of 1 µl 0.6 M EDTA (final concentration 20 mM). After butanol extraction, 100 µl of termination buffer [1% sarkosyl (w/v), 0.1 M Tris (pH 7.8), 0.1 M NaCl, 0.01 M EDTA, 25 µg/ml salmon sperm DNA, and 100 µg/ml proteinase K] was added and the samples incubated at 37°C for 15 minutes followed by 2 minutes at 90°C. The reactions were extracted with phenol and the DNA fragments ethanol precipitated and washed. Gel electrophoresis was performed according to the procedures of Maxam and Gilbert (*47*). (A) Bottom strand analysis of EP2 binding activity. MPE-lane 2, DNase-lane 1, and DMS-lane 3 are the no protein control reactions. MPE-lane 1, DNase-lane 2, and DMS-lane 4 show the result of purified EP2 binding on the cleavage patterns of these three reagents. Lanes R and G show the chemical cleavages at purine and guanine residues, respectively. (B) Top strand analysis. MPE-lane 1, DNase-lane 1, and DMS-lane 3 are the no protein control reactions. MPE-lane 2, DNase-lane 2, and DMS-lane 4 show the effect of EP2 binding on the respective reagents' cleavage pattern. (C) Number and relative position of EP2 (E2) and EP1 (E1) binding sites in the SV40 early promoter region.

Fig. 4. Fractionation scheme of the calf thymus nuclear extract.

Fig. 5. EP2 footprint titration on wild-type (1 x 72) and mutant (dpm 26) core elements. Increasing amounts of partially purified EP2 were added to DNase I footprinting reactions. The general DNase I footprinting reaction is as follows: reactions, in a final total volume of 50 µl, contained 1-3 ng of labeled fragment, 25 µg/ml freshly denatured salmon sperm DNA, 5 mM $MgCl_2$ and the indicated volume of EP2 (or EP1). After a 20 minute 0°C binding reaction, DNase I (Cooper) was added to 50 µg/ml. After 30 seconds at 0°C, the DNase I digestion was terminated by the addition of 100 µl of a termination buffer. The samples were incubated at 37°C for 15

minutes, followed by a 90°C incubation for 2 minutes. The samples were phenol extracted, ethanol precipitated, washed, and dried under vacuum. Samples were electrophoresed according to the procedures of Maxam and Gilbert (*47*). In the titration, both templates were 5′ end-labeled at BamHI and then cut at EcoRI. Lanes 1 and 4 are no protein controls; lanes 2 and 5 contain 5 µl added EP2; lane 6, 10 µl EP2; lanes 3 and 7, 20 µl EP2. R and G correspond to the chemical cleavage reactions specific for purines and guanine residues, respectively. Reactions performed with wild-type (1 x 72) or mutant (dpm26) core containing templates are as indicated. Equal amounts of end-labeled template was added to both sets of reactions. Footprints are bracketed and are shown with map coordinates. Also shown is a diagram of the construct with the locations of the EP2 binding sites, the wild-type core sequence and the changes introduced into the mutant by point mutagenesis.

Fig. 6. Sequence-specific oligonucleotide chromatography of EP2. Chromatography was performed as described in Fig. 2 (panel C). (A) Footprint analysis of fractions eluting from the second cycle of affinity chromatography. Numbers correspond to the fraction being assayed; lane S is no protein control; lane i contains protein sample from column input. Reactions were performed with 20 µl of the indicated column fractions. (B) Polypeptide composition of samples assayed in panel (A). Lanes M contain protein gel marker with polypeptides of molecular weights in kD as indicated. Samples contained 20 µl of the indicated column fraction. Approximate molecular weight of bands eluting from the column are indicated at right. SDS 8% polyacrylamide gels were run according to the method of Laemmli (*49*).

Fig. 7. Renaturation of EP2 binding activity. Those fractions containing EP2 binding activity after the affinity chromatography steps were combined and dialyzed

against C buffer. Protein was precipitated and pelleted by the addition of 3 volumes of acetone, incubated for 30 minutes in a dry ice-ethanol bath, followed by centrifugation at 12,000 xg for 30 minutes at 0°C. The pellet was first vacuum dried and then resuspended in an SDS protein gel loading buffer. The sample was applied to a 10% SDS polyacrylamide gel and electrophoresed at 250 volts for 3 hours at 4°C. The gel was then soaked for 5-10 minutes in 0.1 M KCl. At the appearance of the protein-SDS-potassium precipitate, the bands were excised from the gel and placed into small dialysis bags along with 700 µl SDS gel running buffer. The protein was electroeluted from the gel slice for 3 hours at 250 volts. The bags were then dialyzed against 0.05 HGKED for 3 hours to remove glycine. The protein was precipitated as before and washed twice with 80% acetone to remove residual SDS. The pellet was vacuum dried and resuspended in 300 µl C buffer containing 6 M guanidine hydrochloride. The protein sample was dialyzed overnight against C buffer to remove the guanidine hydrochloride. (A) Footprint analysis of protein eluted from bands excised from preparative SDS gel. Lane 1, no protein control; lane 2, sample (20 µl) of pooled fractions from second cycle of affinity chromatography; lanes 3 and 4, 5 µl and 40 µl, respectively, of sample obtained from the cluster of polypeptide bands at 36 kD; lanes 5 and 6, 5 µl and 40 µl of sample obtained from 39 kD polypeptide; lane 7 and 8, 5 µl and 40 µl of sample obtained from 41 kD polypeptide band; R and G correspond to chemical cleavage reactions specific for purines and guanine residues, respectively. Footprints are bracketed and sequences contained within the footprint along with map coordinates are shown. (B) Polypeptide analysis of gel eluted protein samples. 8% polyacrylamide SDS gel was run using 40 µl of samples obtained from excised gel slices described in (A). Lane M, protein gel molecular weight markers; lane 1, sample used in lanes 3 and 4 of footprint analysis; lanes 2, sample used in lanes 5 and 6 of footprint analysis; lane 3, sample used in

lanes 7 and 8 of footprint analysis. The molecular weights of the markers and samples are indicated.

Fig. 8. Analysis of EP2 binding to immunoglobulin heavy chain enhancer. Lane 1, DNase I footprinting control with no added protein; lane 2, footprinting reaction in presence of 20 µl purified EP2. R and G are the chemical sequencing reactions specific for purine and guanosine residues, respectively. The footprints are bracketed. Sequences contained within the footprints are displayed with those bases showing homology to the core sequence (*29*) circled. The numbering system is that used by Ephrussi *et al.* (*25*).

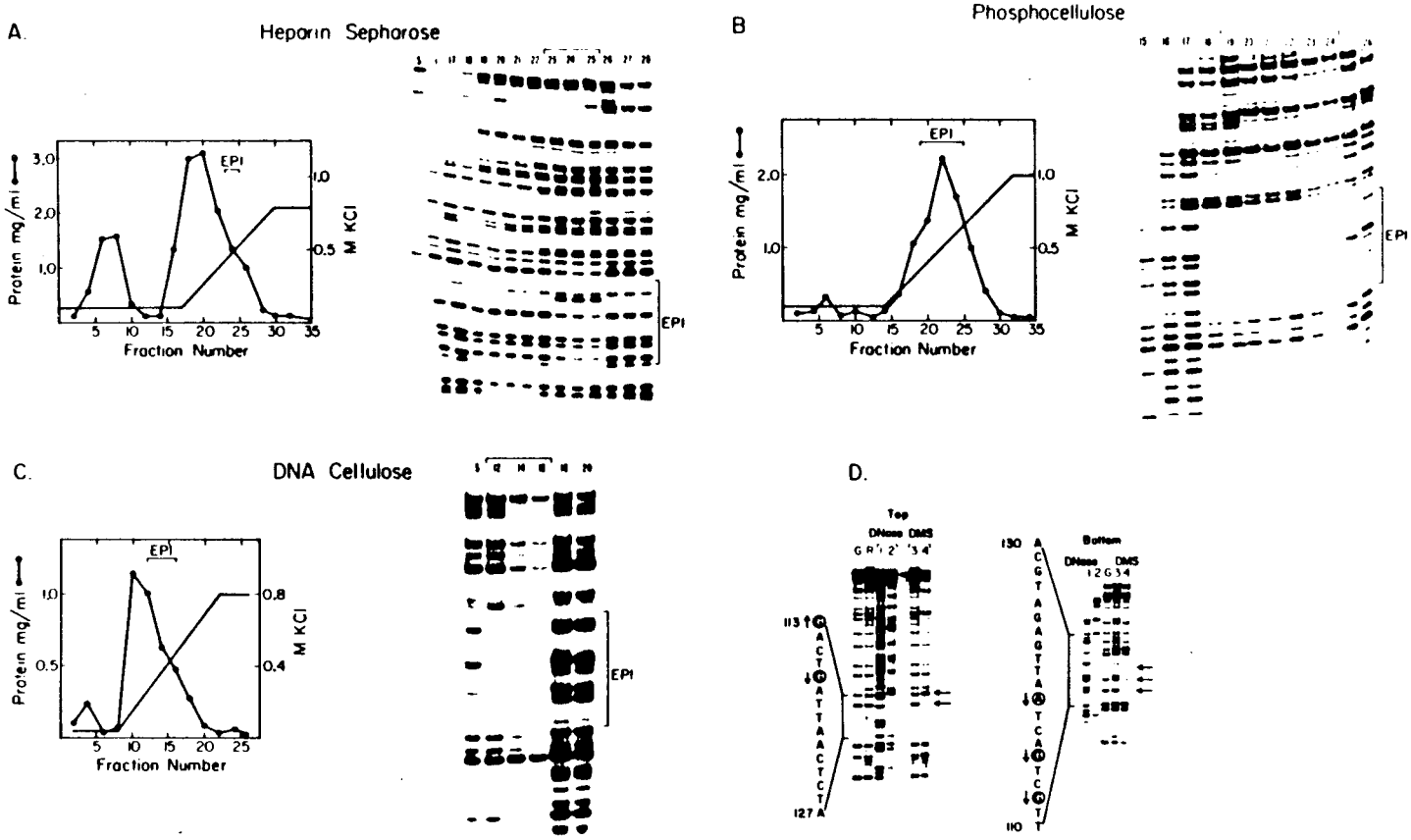
Fig. 9. (A) Heparin-Sepharose chromatography of HeLa cell enhancer binding proteins. Nuclear extract from HeLa cells was prepared according to the method of Parker and Topol (*40*). Extract was applied to a DEAE-cellulose column and the protein which flowed through applied directly to a heparin-Sepharose column and eluted according to the procedure described in the legend to Fig. 1. Samples (40 µl) from the collected fractions were analyzed for enhancer binding activity on the 1 x 72 fragment by DNase I footprinting. Footprints are bracketed and the bound sequences along with map coordinates are shown. (B) Comparison of HeLa cell and calf thymus affinity purified EP2. Lane 1, no protein control; lane 2, calf thymus EP2 after one cycle over affinity column; lane 3, HeLa cell EP2 after one cycle over affinity column (sizing step was excluded from HeLa cell purification procedure). R and G are chemical sequencing reactions specific for purine and guanine residues, respectively.

Table 1. Purification summary of EP1 and EP2

Fraction	Volume (ml)	Protein concentration (mg/ml)	Total protein (mg)	Units[1]	Specific activity (units/mg)
EP-1					
Extract	100	20	2000	ND	ND
Heparin Agarose (pool)	30	1.5	45	375	8.3
Phosphocellulose (pool)	12	1.8	22	480	22
DNA Cellulose (pool)	2.5	0.5	1.25	167	133
EP-2					
Extract	100	20	2000	ND	ND
Heparin Agarose (pool)	40	2.6	104	1000	9.6
Biorex 70	11	9	99	725	7.5
EP2 Affinity Column 1x	1.5	0.4	0.6	160	270
EP2 Affinity Column 2x	1.5	0.030	0.045	100	2220

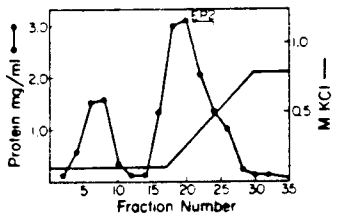
[1]One unit equals 1 fmole of binding site protected.

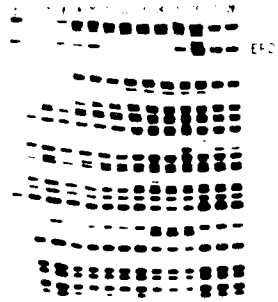
EPI Chromatography



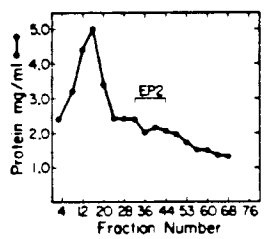
EP2 Chromatography

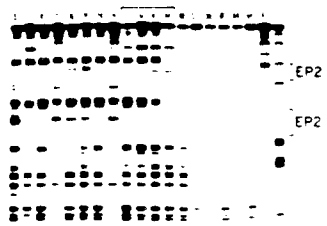
A. Heparin Sepharose



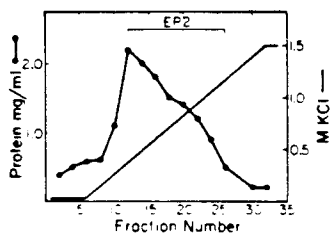


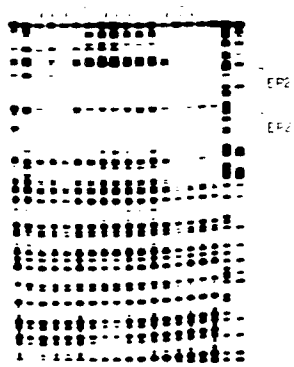
B. AcA 44

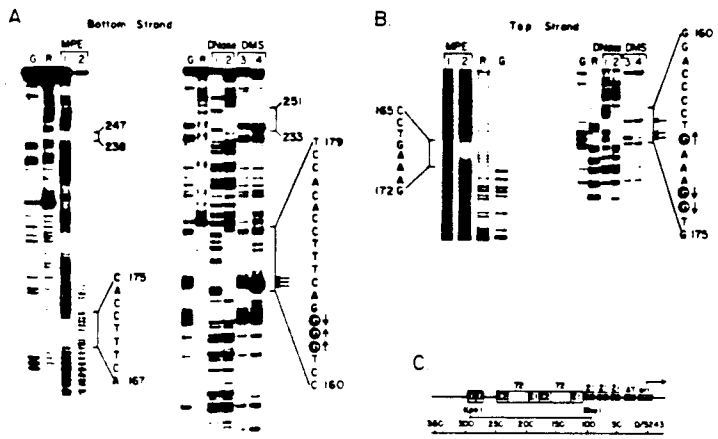


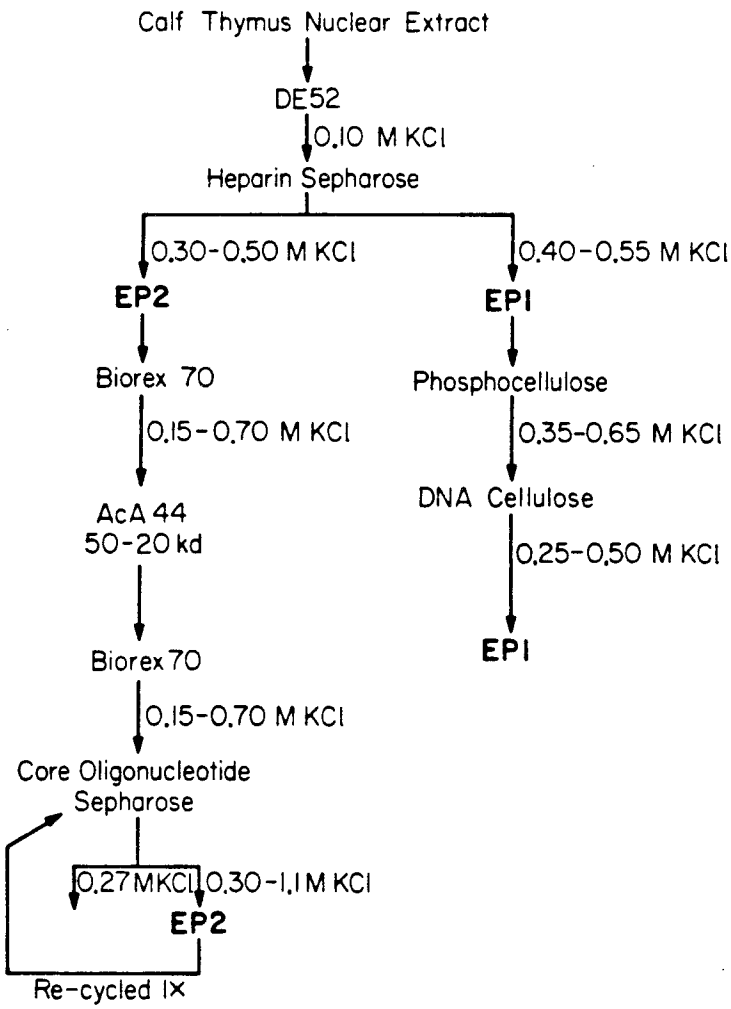


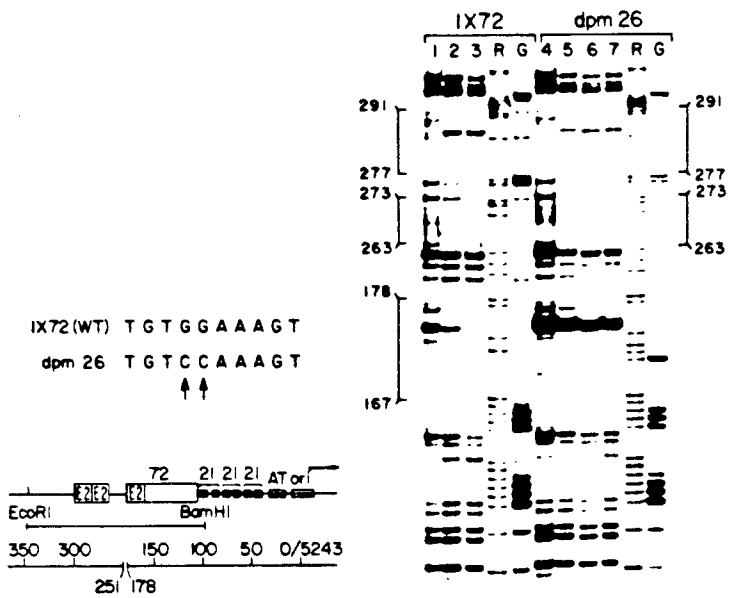
C. Oligonucleotide Sepharose



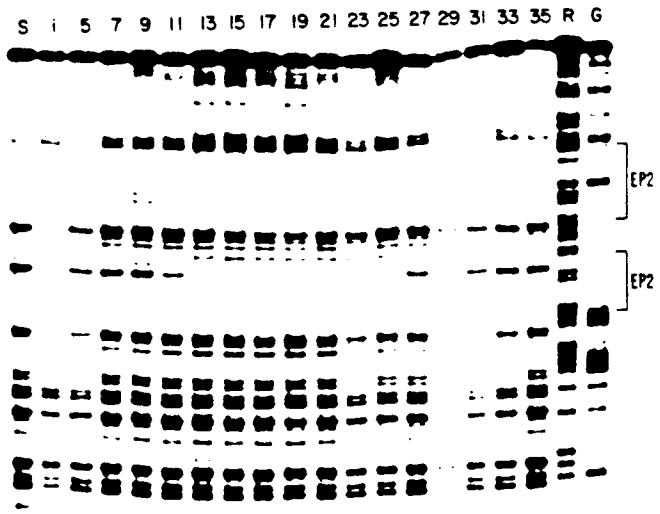



A
Bottom Strand
MPE
DNase DMS
247
238
251
233
179
175
167
160
B
Top Strand
MPE
DNase DMS
165
172
160
175
C

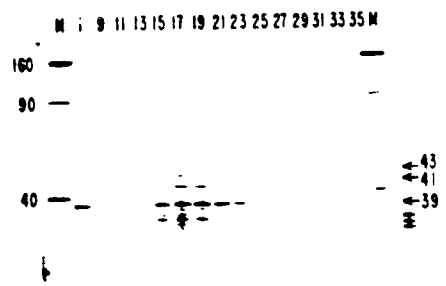

Calf Thymus Nuclear Extract
DE52
0.10 M KCl
Heparin Sepharose
0.30-0.50 M KCl
0.40-0.55 M KCl
EP2
EP1
Biorex 70
Phosphocellulose
0.15-0.70 M KCl
0.35-0.65 M KCl
AcA 44
50-20 kd
DNA Cellulose
0.25-0.50 M KCl
EP1
Biorex 70
0.15-0.70 M KCl
Core Oligonucleotide
Sepharose
0.27 M KCl
0.30-1.1 M KCl
EP2
Re-cycled 1x


IX72
dpm 26
1 2 3 R G 4 5 6 7 R G
291
277
273
263
178
167
IX72(WT) T G T G G A A A G T
dpm 26 T G T C C A A A G T
72
21 21 21
AT or
EcoRI
BamHI
350 300 150 100 50 0/5243
251 178

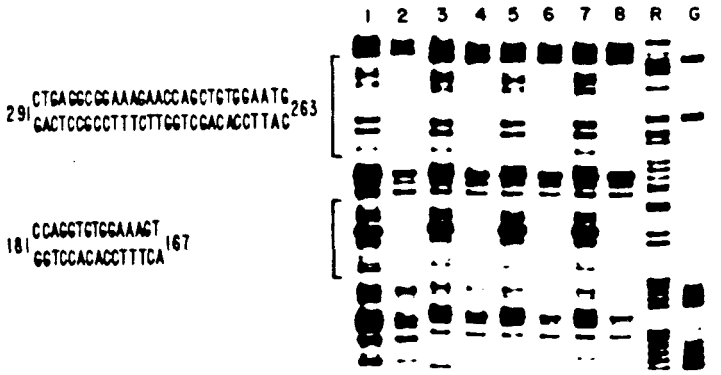
A.


S 1 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 R G
EP2
EP2

B.

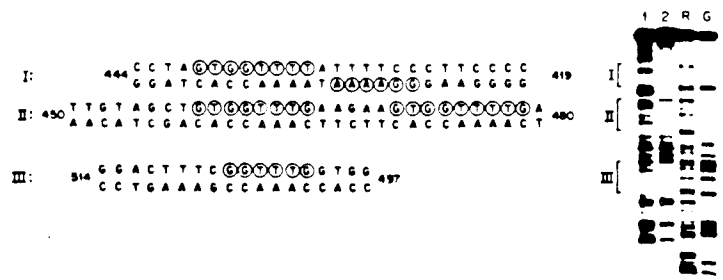

M 1 9 11 13 15 17 19 21 23 25 27 29 31 33 35 M
160
90
40
43
41
39

A.

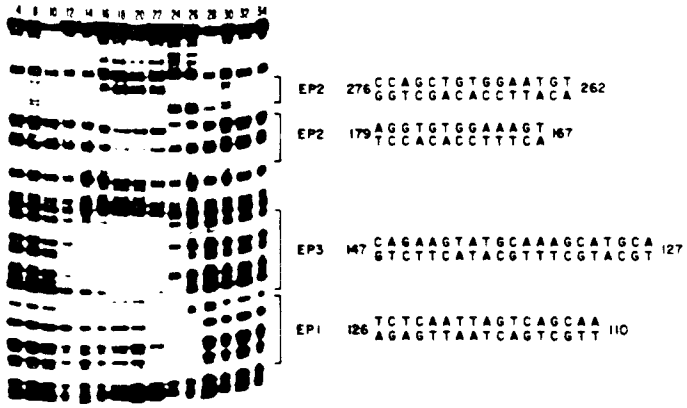

1 2 3 4 5 6 7 8 R G
291 CTGAGCCGGAAAGAACCAGCTGTGGAATG 263
GACTCCGCCTTTCTTGGTCGACACCTTAC
181 CCAGGTGTGGAAAGT 167
GGTCCACACCTTTCA

B.

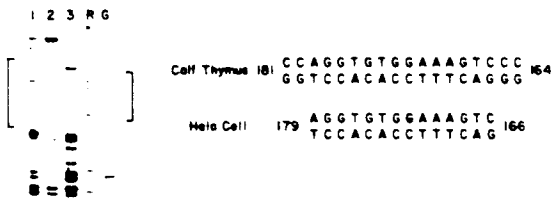

M 1 2 3
160
90
40
41
39
36


1 2 R G
I:
444
419
II:
450
480
III:
514
497

A.


EP2 276 CCAGCTGTGGAATGT 262
GGTCGACACCTTACA
EP2 179 AGGTGTGGAAAGT 167
TCCACACCTTTCA
EP3 147 CAGAAGTATGCAAAGCATGCA 127
GTCTTCATACGTTTCGTACGT
EP1 126 TCTCAATTAGTCAGCAA 110
AGAGTTAATCAGTCGTT

B.


1 2 3 R G
Cell Thymus 181 CCAGGTGTGGAAAGTCCC 164
GGTCCACACCTTTCAGGG
Hela Cell 179 AGGTGTGGAAAGTC 166
TCCACACCTTTCAG

TRANSPLANTATION OF THE BACTERIOPHAGE LAMBDA RECEPTOR

TO MAMMALIAN CELLS

USE OF RECONSTITUTED SENDAI VIRUS ENVELOPES TO TRANSPLANT THE PHAGE LAMBDA RECEPTOR TO MAMMALIAN CELLS

Introduction:

Bacteriophage lambda is one of the best understood viruses in nature. Its genome consists of 48kb of linear duplex DNA. During infection, the phage injects its DNA through the membrane of the host bacteria. Once the DNA is inside the bacterium, it may undergo one of two fates. The first of these is the so-called lytic cycle, in which multiple phage progeny are produced within the bacterium and are eventually released by lysing the bacterial cell. The second course of infection is known as lysogeny, in which the phage genome integrates into the host cell chromosome and remains in a quiescent state, replicating passively along with the cell's own genome.

The receptor for phage lambda has been identified as *lamB*, an outer membrane protein which is normally used by bacteria for the transport of maltose. Phage lambda binds to *lamB* and presumably injects its DNA through this protein during infection. *LamB* can be isolated from bacteria without losing its receptor activity. If *lamB* is reconstituted into phospholipid vesicles, phage lambda will bind to these vesicles and inject its DNA into the interior of the vesicles (1).

Phage lambda has become one of the most commonly used cloning vectors in modern molecular biology. The central portion of the phage genome can be replaced by foreign DNA without impairing the phage's ability to replicate in bacterial hosts. Recombinant phage are routinely produced by ligating desired DNA fragments of suitable size into lambda phage vectors. When these phage replicate inside their host cells, the inserted DNA sequence is also replicated, resulting in an efficient means for

amplifying the DNA of interest. Desired clones are normally isolated by nucleic acid hybridization procedures.

Much use has been made of phage "libraries", which are constructed by taking the entire genome of a given organism, cutting it into suitably sized pieces, inserting these pieces into lambda phage, and amplifying the recombinant phage in a suitable host. These libraries contain inserted sequences spanning the entire genome of the organism, and can be used to isolate any desired gene sequence if an appropriate probe is available.

The aim of the project described below is to combine the wealth of knowledge about both the molecular biology of phage lambda and its use as a cloning vector to produce a novel method of DNA transfer in eukaryotic cells. The idea is to implant the lambda receptor into the membrane of eukaryotic cells, then allow phage lambda to bind to these cells and inject recombinant DNA into the cells. This method may lead to a new method for introducing desired genes into eukaryotic cells, as outlined in figure 1.

The strategy for accomplishing the desired transplantation of the lambda receptor consists of three basic parts. First is reconstitution of *lamB* into phospholipid vesicles. Second is the fusion of these vesicles with the target cells. Third is the characterization of phage lambda interactions with these target cells.

RECONSTITUTION OF LAMBDA RECEPTOR IN PHOSPHOLIPID VESICLES

The *E. coli* receptor for phage lambda is *lamB*, a trimeric outer membrane maltose porin (2). *LamB* can be extracted in fairly pure form from bacterial membranes

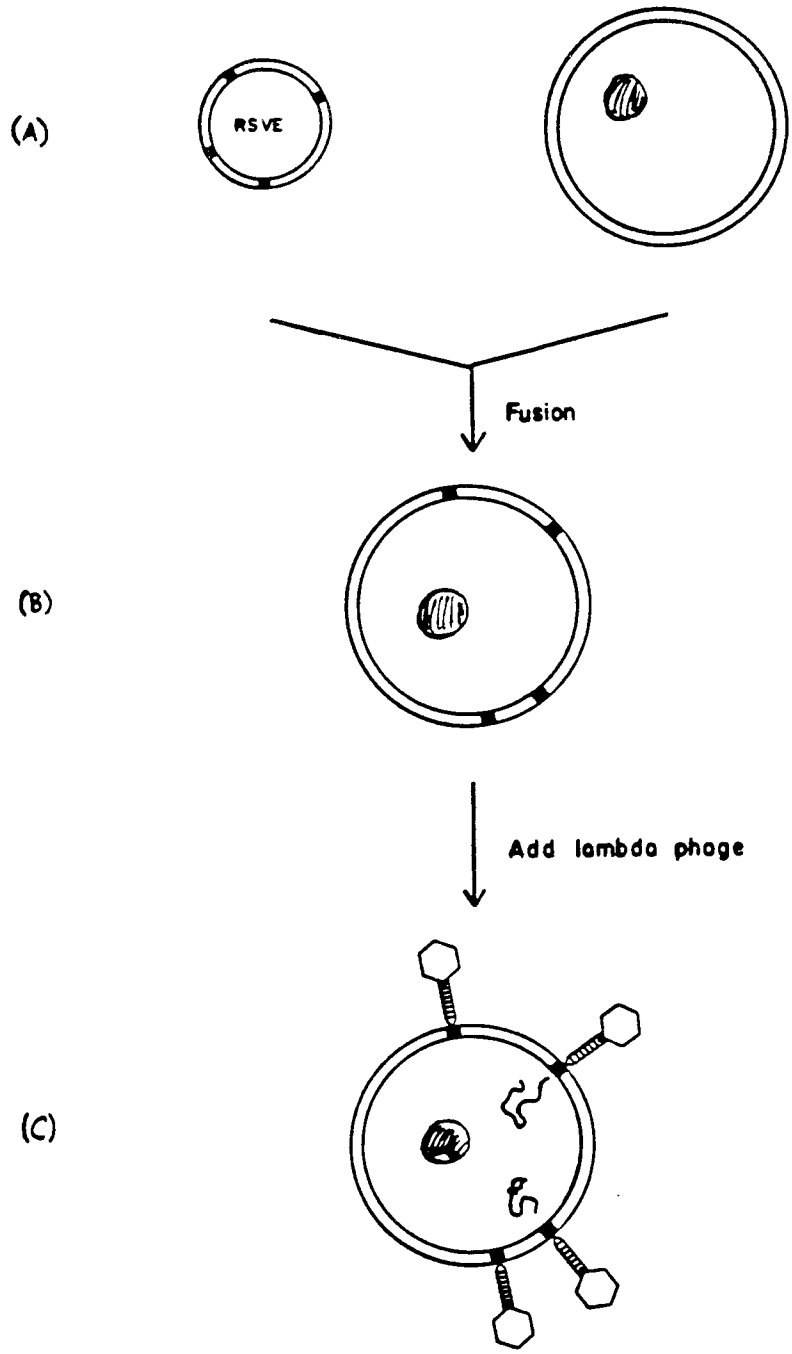


Figure 1: Schematic diagram of the use of lambda phage mediated gene transfer following transplantation of the lambda receptor. In (A), receptor-bearing vesicles (RSVE), are fused with target cells. This results in cells which carry the lambda receptor on their surfaces (B). Addition of recombinant phage results in efficient transfer of genes to the cellular interior (C).

by detergent solubilization. To induce synthesis of *lamB*, the bacteria are grown in the presence of maltose.

Purified *lamB* can be reconstituted in artificial phospholipid membranes using techniques pioneered by Racker and associates in the early 1970's (3). In this procedure, phospholipids (such as phosphatidyl choline, phosphatidyl ethanolamine, and phosphatidyl serine) and the protein to be incorporated are co-solubilized in an aqueous solution of detergent. This mixture produces mixed micelles containing both protein and phospholipid. If the mixture is put into dialysis, which gradually removes detergent molecules, the phospholipids will spontaneously form closed bilayer structures known as liposomes. Integral membrane proteins, such as *lamB*, insert themselves into these bilayer membranes as the detergent is removed. The incorporation of membrane proteins into liposomes is a very efficient process, and often results in biological activity that is very similar to that seen in natural membranes.

Incorporation of *lamB* into liposomes by the detergent dialysis technique was previously demonstrated by Roessner et al. (1). Using electron microscopy, these investigators found that phage lambda bound to *lamB*-containing liposomes and appeared to inject at least part of its genome into the vesicle interior. These finding suggest that *lamB* alone can bind phage lambda and mediate injection of DNA across a membrane. Incorporation of *lamB* in a eukaryotic cell membrane might make it possible for phage lambda to inject DNA directly into the cytoplasm of the cell.

FUSION OF *lamB*-CONTAINING LIPOSOMES WITH EUKARYOTIC CELLS

In order to incorporate the lambda receptor into the plasma membrane of a eukaryotic cell, it is necessary to fuse *lamB*-containing liposomes with the membrane of the target cell. Membrane fusion is a poorly understood process, but there are a number

of chemical and biological agents that are known to promote fusion of bilayer membranes. The most effective chemical fusogens are dehydrating agents such as polyethyleneglycol which appear to work by removing water from the interface between bilayers, thus facilitating tight apposition of the membranes to be fused. Known biological fusogens include a number of proteins, the best characterized being the proteins present on the exterior of enveloped viruses.

In this project, fusion has been accomplished using Sendai virus envelope proteins. This is a fairly well characterized system that has been used by many researchers for incorporation of exogenous agents into eukaryotic cells (reviewed in reference 4). In a project similar to this one, reconstituted Sendai virus envelopes (RSVE, see below) were used to implant the Epstein-Barr virus receptor into receptor-negative cells (5). Delivery of the EBV receptor to these cells conferred susceptibility to EBV infection.

Molecular Biology of Sendai Virus

Sendai is an enveloped murine virus belonging to the paramyxovirus family (6). It is composed of a nucleocapsid containing 16kb of single-stranded RNA surrounded by a highly proteinaceous lipid bilayer called the envelope. The envelope contains two viral glycoproteins which are present in multiple copies on each viral particle. The envelope proteins are responsible for binding of the virus to target cells and the fusion of the viral envelope with the plasma membrane of cells. The attachment and fusion functions appear to reside in separate proteins, attachment being mediated by the hemagglutinin/neuraminidase protein (HN), while fusion is mediated by the fusion protein (F). All other viral proteins, including the nucleoprotein (NP), matrix protein (M), and viral polymerase (P), are enclosed within the envelope.

When Sendai virus infects a cell, it first binds to the plasma membrane of the target cell using a sialic acid receptor (7). The viral envelope the fuses with the plasma membrane, dumping the nucleocapsid into the cytoplasm of the cell. Once the viral RNA has entered the cell, it replicates itself and produces viral proteins under the control of virally encoded enzymes. Mature virus particles are formed by budding through the plasma membrane of the infected cell.

Growth of Sendai Virus

Sendai virus is most commonly grown in fertilized chicken eggs (8). Ten day old embryonated eggs are innoculated with 0.1 to 0.2 ml of a 1:1000 dilution of stock virus (see below). This dilution is carried out using sterile PBS, and antibiotics are added to the innoculum to prevent bacterial contamination of the eggs. After innoculation, the eggs are incubated at 37^{o} for 72 hours. At this time, the eggs are chilled to 4^{o} for approximately 5 hours, which kills the embryos.

To harvest the virus from the eggs, the shell and the membrane covering the airspace are removed. A ten ml pipette is used to suck out the allantoic fluid (5 to 10 ml per egg), care being taken not to rupture the yolk. The allantoic fluid is collected in a beaker set on ice. An aliquot of this fluid is flash frozen and kept at -70^{o} to be used as the virus stock for the next prep's innoculum.

The allantoic fluid is processed first by a low speed centrifugation step which removes blood cells and other aggregated materials. The virus is then pelleted by centrifuging at 40,000 x g for 45 minutes. The virus is taken up in PBS and pelleted again as above. Virus at this stage is fairly pure, but further purification can be obtained if needed by sedimentation in a 15-60% sucrose gradient.

The yield of virus was fairly reproducible and ranged from 0.5 to 1 mg of virus protein from each egg. Two dozen eggs were routinely used in each virus prep. Quantities of virus were measured using a modified Lowry protein assay (9). It has been estimated that 1 mg of virus protein corresponds to 10^{12} virus particles (10).

Reconstituted Sendai Virus Envelopes (RSVE)

When Sendai virus is placed in a solution of non-ionic detergent, such as Triton X-100 or octyl glucoside, the envelope of the virus is disrupted. This results in solubilization of the lipids and glycoproteins which compose the envelope. The nucleocapsid remains intact and in not solubilized. This selective solubilization makes it possible to easily separate the envelope components from the nucleocapsid by high speed centrifugation. Centrifugation at 100,000 x g for 1 hour pellets the nucleocapsid, resulting in a clear supernatant containing the envelope components in detergent solution.

If this solution of envelope components is subjected to dialysis, membrane bilayers spontaneously reassemble, producing vesicles. The envelope glycoproteins, HN and F, insert themselves into these vesicles, resulting in structures which resemble the original viral envelopes. These reconstituted Sendai virus envelopes (RSVE) are capable of binding to and fusing with cell membranes in a manner similar to native Sendai virus (11-13). The construction of RSVE is diagrammed in figure 2.

In this project, the detergent octyl glucoside was used as the disrupting agent. This detergent has a high critical micellar concentration (20 mM) which makes it very easy to remove during dialysis. Studies have shown that no residual detergent remains in RSVE made from octyl glucoside, in contrast to RSVE made from Triton X-100 (13).

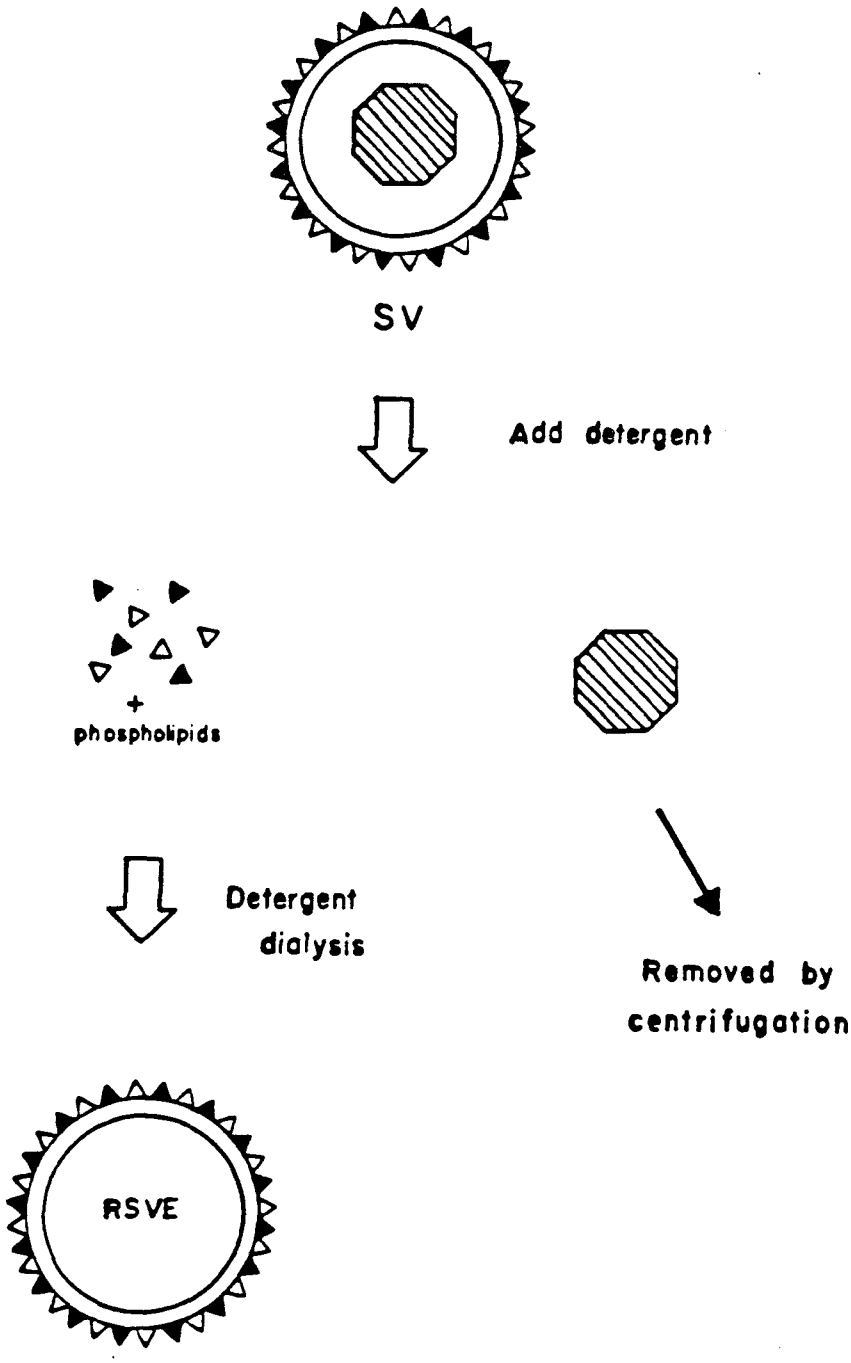


Figure 2: Construction of RSVE. Intact Sendai virus (SV) is disrupted with a non-ionic detergent which selectively solubilizes the envelope glycoproteins (represented by closed and open triangles). The insoluble nucleocapsid is removed by centrifugation, leaving a clear supernatant containing the envelope lipids and glycoproteins. Upon removal of the detergent by dialysis, structures similar to the original viral envelope form spontaneously (RSVE).

When using octyl glucoside to make RSVE, the rate of removal of detergent during dialysis is important. Accordingly, the slow dialysis procedure outlined by Harmsen et al. (13) was used in this project. It has been suggested that the rate of formation of the bilayer structures must be slow enough to allow the proteins to orient themselves properly in the membrane in order for the reconstitution process to be successful.

Characterization of RSVE properties

The RSVE made by octyl glucoside dialysis were examined under a Phillips 200 electron microscope. They appeared as approximately spherical particles about 50nm in diameter. It was difficult to determine fine details of protein orientation using this instrument, but Harmsen et al. (13), using a similar procedure, found that nearly all of the glycoprotein spikes were oriented toward the outside of the vesicles. The reason for this asymmetric reassembly has not yet been determined, but steric packing forces have been suggested.

To investigate the ability of RSVE to bind and fuse with target cells, the membrane of the RSVE was labeled with a fluorescent probe. Octadecyl rhodamine B (R_{18}, see figure 3) is a commercially available compound which is useful for these purposes. It contains the highly fluorescent rhodamine moiety coupled to an 18 carbon aliphatic chain which makes it substantially lipophilic.

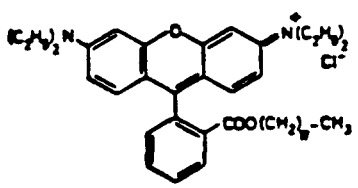


Figure 3: Structure of Octadecyl Rhodamine B (R_{18}). The rhodamine moiety is strongly fluorescent, and the eighteen carbon tail permits stable insertion into a lipid bilayer.

If an ethanolic solution of R_{18} is injected into a suspension of membranes, the probe will spontaneously insert into the membranes. If R_{18} is incorporated into a lipid bilayer at a concentration greater than 1 mole percent relative to the bilayer lipids, self-quenching of the rhodamine fluorescence is observed. The degree of quenching increases linearly with increasing R_{18} concentration up to 8 mole percent.

Hoekstra et al. have exploited the properties of this dye to create a simple assay for measuring membrane fusion kinetics (14). R_{18} is incorporated at a self-quenching concentration into one of the membranes involved in the fusion reaction. When fusion takes place with an unlabeled membrane, the R_{18} is free to diffuse into the new membrane, resulting in a relief of quenching. By monitoring the rate of fluorescence increase, the rate of membrane fusion can thus be determined.

Addition of a detergent such as Triton X-100 to R_{18} labeled membranes results in complete solubilization of the membranes. This effectively dilutes the probe into a very large volume, resulting in complete dequenching. The amount of fluorescence seen after addition of Triton X-100 is taken as the maximum fluorescence that can be observed. In the fusion assay, the rate and extent of fusion is expressed as a percentage of this "infinite dilution" value.

The R_{18} fusion assay was used to characterize the fusogenic properties of the RSVE made in this project. R_{18} was incorporated into RSVE so that the observed fluorescence was about 25 to 35 percent of that seen after addition of Triton X-100. This degree of quenching corresponds to approximately 6 to 7 mole percent R_{18} in the RSVE membrane (14). Fusion was measured using erythrocyte ghosts as the target membrane. Ghosts were prepared by hypotonic lysis using standard procedures (15).

In a typical fusion reaction, 2-3 ug of RSVE protein was placed in a fluorimeter cuvette in a final volume of 1 ml. The cuvette was warmed to 37°, and fusion was

initiated by the addition of 50 ug of ghost protein (equivalent to 10^8 cells). Fluorescence was monitored using an SLM 4800 spectrofluorimeter with excitation at 535 nm and emission at 585 nm.

A typical fusion assay fluorescence profile is shown in figure 4. In general, a burst of fusion is observed within the first few minutes followed by a slower rate of fluorescence increase. In this particular experiment, the burst of fluorescence increase corresponds to about 15% of the maximal fluorescence seen after addition of Triton X-100. Since fusion of a single RSVE with a ghost results in sufficient dilution of the R_{18} probe to result in complete dequenching, the percentage of maximal fluorescence evolved can be directly correlated to the percentage of RSVE that have fused with target membranes.

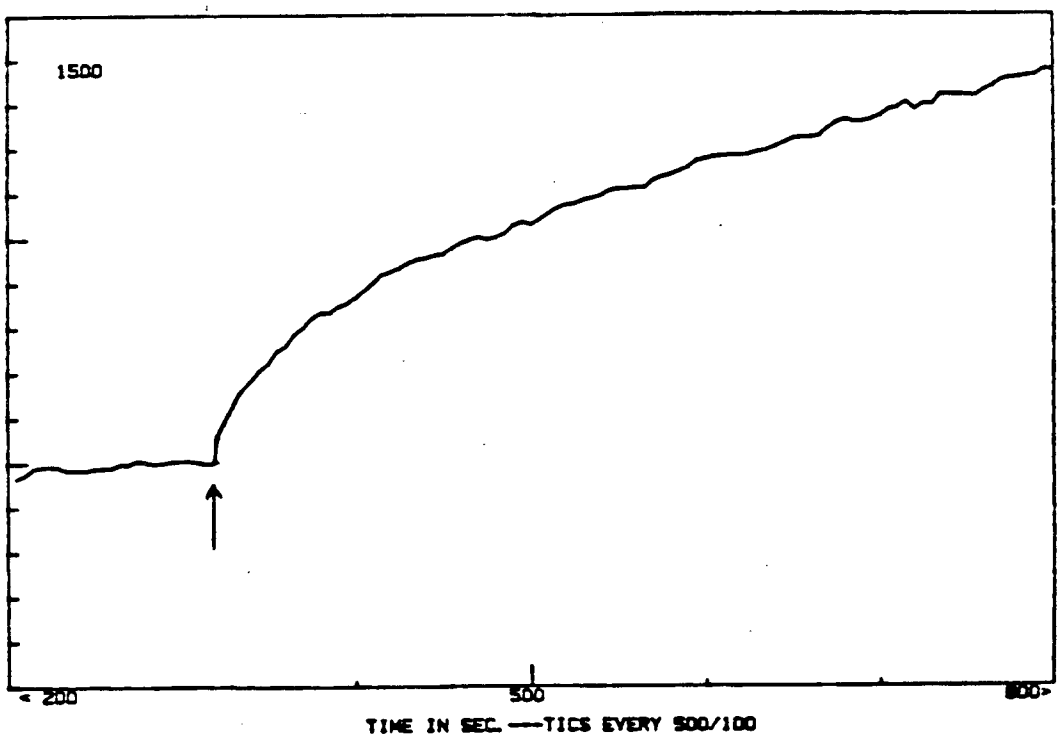


Figure 4: Typical fusion assay using R_{18}-labeled RSVE. The Y-axis represents fluorescence intensity, the X-axis time. Addition of erythrocyte ghosts (at the point marked by the arrow), results in an increase in rhodamine fluorescence as labeled RSVE fuse with the ghost membranes.

Binding of RSVE to target cells can also be easily measured using R_{18} labeled RSVE. In a typical binding assay, cells and R_{18} labeled RSVE are mixed in a buffered 1 ml solution. After allowing adequate time for binding to proceed (generally a few minutes), the mixture is centrifuged in an eppendorf microcentrifuge (at about 10,000 x g) for 5 minutes. This pellets the cells, but will not pellet RSVE that remain unassociated with the cells. The supernatant is removed from the cell pellet, and the relative ratios of fluorescence in both pellet and supernatant are measured after addition of Triton X-100. This ratio provides a measure of the percentage of RSVE which have bound to the cells. Typical experiments using 2 ug RSVE protein and 50 ug ghost protein showed that about 65% of the RSVE were capable of binding to ghosts.

Incorporation of LamB into RSVE

The method for incorporating exogenous proteins into RSVE is very simple. The protein of interest is generally added to the detergent solubilized Sendai virus envelope components prior to dialysis. As the detergent is dialyzed out, soluble proteins are trapped within the interior of the RSVE while membrane proteins insert into the vesicle bilayer (5).

To create RSVE containing the lambda receptor, *lamB* was isolated from *E. coli pop154*, a derivative of *E. coli* K-12 which expresses the lambda receptor from *Shigella sonnnei* (1). 200ug of *lamB* was added to an octyl glucoside solution of viral envelope components (2 mg of viral envelope protein). The mixture was then dialyzed as usual.

Characterization of LamB-containing RSVE

To determine whether *lamB* incorporated into RSVE maintained its ability to bind phage lambda, a competition experiment was designed as schematized in figure 5. Presence of functional receptors in the RSVE was determined by measuring the ability of such RSVE to inhibit phage binding to receptor-bearing bacteria. Addition of *lamB*-bearing RSVE to a mixture of bacteria and phage resulted in a substantial decrease in binding of phage to the bacteria, demonstrating the presence of functional receptors on the RSVE. Using an estimate of 6,000 receptors per bacterium (16), these experiments suggest that there are 5-10 functional receptors per RSVE.

To determine whether incorporation of *lamB* in the RSVE membrane had any effect on the activities of the viral glycoproteins, binding and fusion assays were performed as described above. No significant differences in activity were found in the *lamB*-containing RSVE compared to normal RSVE.

Fusion of lamB-containing RSVE with erythrocyte ghosts

Fusion of *lamB*-containing RSVE with erythrocyte ghosts was performed under conditions similar to those employed for assaying RSVE fusion activity. 100ug of RSVE (not R_{18} labeled) were added to ghosts (1 mg protein) in PBS, and the mixture was incubated at 37^{o} for 45-60 minutes. Based on results from fusion assays done previously, it is estimated that 30-40% of the RSVE would have fused with the ghosts within this time period. Non-fused RSVE were then removed by repeated centrifugation and washing with PBS.

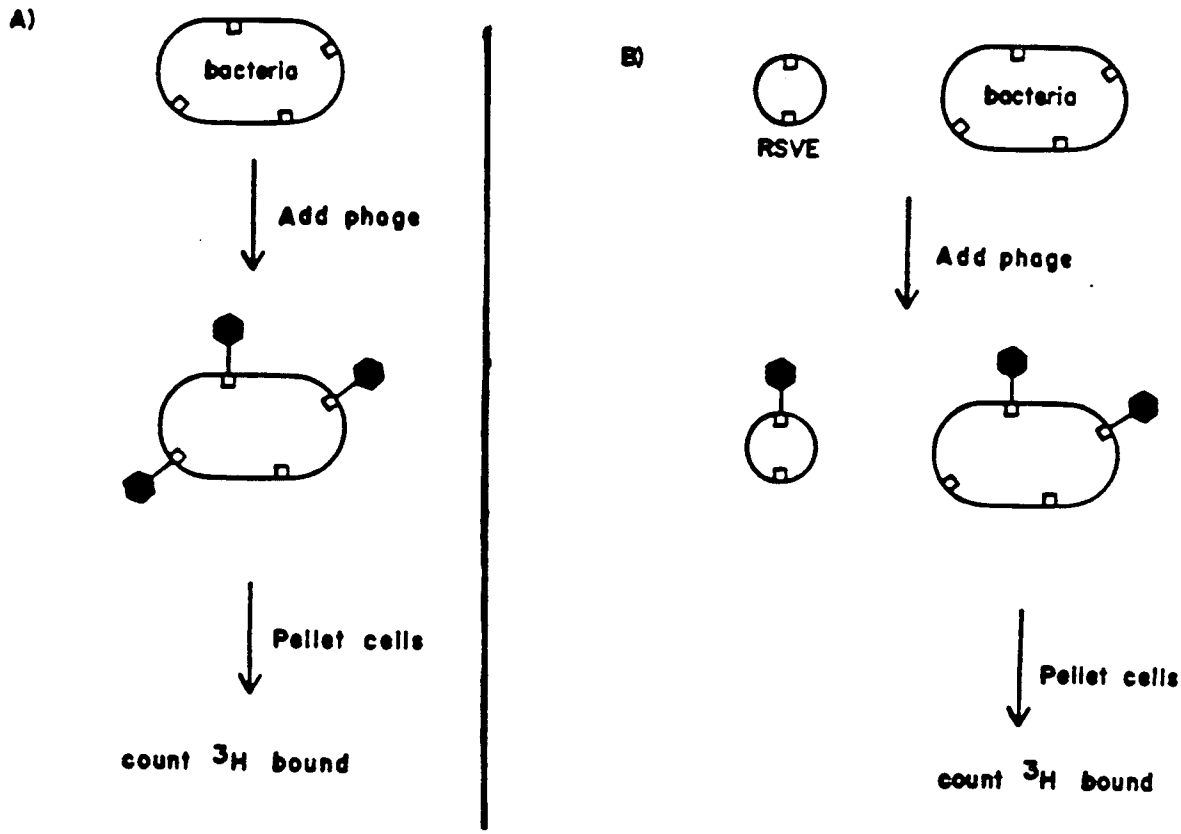


Figure 5: Competition experiment designed to measure phage binding activity of *lamB*-containing RSVE. In panel (A), ^{3}H labeled phage are allowed to bind to receptor-expressing bacteria in the absence of added RSVE. After spinning down the bacteria, the amount of ^{3}H associated with the pellet is counted to quantify the number of phage bound. In panel (B), *lamB*-containing RSVE are added to the binding mixture. These RSVE compete with the bacteria for the binding of labeled phage. When the bacteria are pelleted, the RSVE with their bound phage remain in the supernatant, resulting in fewer counts are associated with the cell pellet. Quantitation of the amounts bound under varying quantities of added RSVE allows an estimation of the number of receptors present in the RSVE.

Characterization of LamB incorporated into ghost membranes.

As a first attempt to determine whether *lamB* was functionally incorporated into the ghost membranes by fusion with *lamB*-containing RSVE, phage lambda interactions with *lamB*-bearing ghosts were visualized using electron microscopy. Phage lambda was added to a preparation of *lamB*-bearing ghosts (10^{10} pfu per 400ug ghost protein), and this mixture was incubated at room temperature for 30 minutes. EM grids were prepared from aliquots of this mixture and stained with 1% uranyl acetate. The grids were then observed in a Phillips 200 electron microscope at magnifications of 20,000 to 70,000x.

Examination of the electron micrographs suggested that lambda phage does interact with *lamB*-containing ghosts. However, this evidence is purely qualitative and does not provide solid evidence that phage lambda is interacting specifically with its receptor in these ghosts.

In order to gain more quantitative information on whether functional *lamB* is present in the membranes of these ghosts, binding studies using ^{3}H labeled phage were employed. Phage binding was measured by incubating cells with ^{3}H labeled phage at room temperature for 10 minutes. The cells were then pelleted in an eppendorf microfuge and washed once with PBS. The cells pellet from this washing was counted in a liquid scintillation counter to determine the quantity of phage which remained bound to the cells.

Results from these experiments are shown in table 1. From this data it appears that functional receptors are present in the *lamB*-containing ghosts. "Functional" in this case refers to the ability of the receptor to bind phage. No experiments so far have shown that this binding is followed by injection of phage DNA into the cellular interior.

Lambda mediated gene transfer into eukaryotic cells

The next step in this project is the demonstration that phage lambda is capable of injecting its DNA into eukaryotic cells into which the lambda receptor has been implanted. The ultimate demonstration of this would be the expression of a selectable marker in the eukaryotic cells following lambda-mediated DNA injection. Efforts in this direction are currently being made by others associated with this project.

Table 1: Test of Receptor Binding Activity in Erythrocyte Ghosts.

Cell Type	^{3}H cpm bound
Bacteria	950 ± 44
Normal Ghosts	106 ± 17
Ghosts with transplanted *lamB*	301 ± 21

To test the ability of phage lambda to bind to transplanted receptors in ghost membranes, 4 x 10^8 pfu of phage lambda (2,000 cpm) were added to 3 x 10^8 cells of the types shown above. The cells were incubated with phage for 10 minutes at room temperature, then spun out, washed with buffer, and spun out again. The pellet was then counted in a liquid scintillation counter to determine the quantity of phage bound to the cells. Each cell type was tested in triplicate.

REFERENCES

1. Roessner, C.A., D.A. Struck and G.M. Ihler (1983) *J. Biol Chem.* **258**, 643-648.

2. Schwartz, M. (1980) In: *Virus Receptors, Part 1: Bacterial Viruses* (L.L. Randall and L. Philipson eds.) pp.59-94; Chapman and Hall with Methuen, Inc., New York.

3. Reviewed in: Racker, E. (1976) *A New Look at Mechanisms in Bioenergetics*, pp. 107-126.

4. Loyter, A., and D.J. Volsky (1982) *Cell Surf. Rev.* **8**, 215-266.

5. Volsky, D.J., I.M. Shapiro, and G. Klein (1980) *Proc. Natl. Acad. Sci USA* **77**, 5453-5457.

6. Kingsbury, D.W. (1977) In: *The Molecular Biology of Animal Viruses* (D.P. Nayak ed.) Marcel Dekker Inc., New York.

7. Markwell, M.A.K., A. Portner and A.L. Schwartz (1985) *Proc. Natl. Acad. Sci. USA* **82**, 978-982.

8. Hoskins, J.M. (1967) *Virological Procedures*; Butterworths, London.

9. Markwell, M.A.K., S.M. Haas, L.L. Bieber and N.E. Tolbert (1978) *Anal. Biochem.* **87**, 206-210.

10. Hoekstra, D., and K. Klappe (1986) *J. Virol.* **58**, 87-95.

11. Hosaka, Y., and Y.K. Shimizu (1972) *Virology* **49**, 627-639.

12. Volsky, D.J., and A. Loyter (1978) *FEBS Lett.* **92**, 190-194.

13. Harmsen, M.C., J. Wilschut, G. Scherphof, C. Hulstaert, and D. Hoekstra (1985) *Eur. J. Biochem.* **149**, 591-599.

14. Hoekstra, D., T. de Boer, K. Klappe and J. Wilschut (1984) *Biochemistry* **23**, 5675-5681

15. Steck, T.L. and J.A. Kant (1974) *Methods Enzymol.* **31**, 172-180.

16. Schwartz, M. (1975) *J. Mol. Biol.* **99**, 185-201.