CONTROL OF CELLULAR GENE EXPRESSION BY VIRAL

REGULATORY PROTEINS

By

BARBARA PANNING, B. SC.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
CONTROL OF CELLULAR GENE EXPRESSION BY VIRAL PROTEINS
DOCTOR OF PHILOSOPHY (1994)  McMaster UNIVERSITY
(Medical Sciences)  Hamilton, Ontario

TITLE: Control of cellular gene expression by viral regulatory proteins.

AUTHOR: Barbara Panning, B. Sc. (McMaster University)

SUPERVISOR: Dr. J. R. Smiley

NUMBER OF PAGES: xiv, 299
Abstract

Herpes simplex virus type 1 and adenovirus type 5 are nuclear DNA viruses that encode a number of regulatory polypeptides that serve to facilitate expression of the viral genome at the expense of host gene expression. These viruses depend on cellular transcriptional apparatus, and viral regulatory proteins function primarily to ensure that host factors engage the viral template and that viral genes are expressed in the correct spatial and temporal sequence. The results presented in this thesis demonstrate that regulatory proteins encoded by herpes simplex virus and adenovirus modulate the expression of two classes of cellular genes: the globin genes and Alu repetitive sequences. Herpes simplex virus gene products were required to stimulate the expression of human α-globin, rabbit β-globin genes and Alu elements introduced into cells as part of the herpes genome. In addition, infection with herpes simplex virus or adenovirus stimulated expression of host α-globin and Alu sequences and de novo synthesized viral gene products were required for induction of these cellular genes. Viral induction of α-globin and Alu expression are both unusual instances of activated gene expression: viral infection allows the α-globin gene to escape its erythroid-restricted transcription pattern, and the viral activation of RNA polymerase III transcription of Alu sequences is the first instance of high level class III transcription of these repetitive DNA elements in vivo. The identification of the herpes simplex virus and adenovirus gene products which mediate the transcriptional activation of these host sequences has contributed to the understanding of the mechanisms which regulate expression of α-globin genes and Alu repetitive elements and also of the mechanisms by which viral regulatory proteins modulate gene expression.
Acknowledgements

There is a huge number of people I would like to thank. First and foremost is Jim Smiley, my thesis supervisor. He combined patient guidance and intellectual generosity in a (hopefully successful) effort to teach the two most important aspects of science: thoughtful evaluation and fun. Jim also makes great tabouleh. What more could anyone ask from a supervisor?

David Johnson and John Capone comprise the remainder of my supervisory committee. They have provided valuable suggestions, criticism and support and have always been generous with their time and ideas. I would also like to thank them in advance for not reading this thesis too closely. I think we've all had enough of it.

The faculty of the Cancer Research Group, William Rawls, Frank Graham, Silvia Bachetti, Phil Branton, and more recently Lori Frappier and Al Edwards, have provided an excellent intellectual environment in which to do and learn science. I would also like to thank the National Cancer Institute of Canada for giving me a scholarship and my one chance to be a cover girl.

The students and staff of the CRG helped to make my many years at McMaster enjoyable ones. The Smileys (over several years and in no particular order): Craig Smibert (aka Big/old Craig), Joanne Duncan, Carol Lavery, Tom Hupel, Phil Kibler, Peter Cheung, Frank Jones, Karen Koop and Margaret Howes, were a constant source of ideas (even some scientific ones), help and amusement. It was from these people that I learned to make dry ice cannons (endless hours of exciting, destructive amusement); that sharing a bay with someone messier than you generally makes your bench look great (thanks Craig); that sharing musical tastes with a bay-mate makes depressing results look good (thanks Frank - except for those YES and DEAD tapes); that starting over in a new lab can actually be fun - if you have e-mail (thanks Pete), and that technicians really are god-like (thanks Carol) and generally pregnant (just kidding Joanne). Non-Smileys who have provided moral, intellectual or athletic support include Ian York, Craig Brunetti (aka Young/little Craig), Jean Barwell, Sally Hemmings, Hashem Salloukh, Llloyd Hutchison, Ramsi Haddad, Steve Orlicky, Kim Goldsmith and Geoff Bovett.

Finally, I would like to thank my family, especially my parents, for their patience, support and understanding during my McMaster years. I could never have done it without them.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive Note</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iv</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>vi</td>
</tr>
<tr>
<td>Preface</td>
<td>xiv</td>
</tr>
</tbody>
</table>

1. **Introduction**  
   1.1 Regulation of eukaryotic gene expression  
   1.1.i Regulation of transcription in eukaryotes  
   1.1.ii *Cis*-acting sequences and *trans*-acting factors  
   1.1.iii Chromatin structure  
   1.1.iv RNA processing  
   1.1.v Viruses as model systems for regulated gene expression  
   1.2 Regulation of gene expression by HSV-1 proteins  
   1.2.i HSV-1 structure and genome organization  
   1.2.ii Temporal regulation of HSV-1 gene expression  
   1.2.iii Vmw65 and IE gene expression  
   1.2.iv IE gene products  
   1.2.v Inhibition of gene expression by viral products  
   1.3 Regulation of Ad gene expression  
   1.3.i Ad5 structure and genome organization  
   1.3.ii Temporal regulation of Ad5 gene expression  
   1.3.iii E1a gene products and viral gene expression  
   1.3.iv E1 proteins and cellular gene expression  
   1.3.v Regulation of gene expression by E1b and E4 products  
   1.4 Statement of purpose  

2. **Materials and Methods**  
   2.1 Manipulation of plasmids and phage  
   2.1.i Bacterial culture  
   2.1.ii Transformation of bacteria  
   2.1.iii DNA manipulations and cloning  
   2.1.iv Purification of DNA fragments  
   2.1.v Polymerase chain reaction procedure  
   2.1.vi Large scale isolation of plasmid DNA  
   2.1.vii Large scale isolation of M13 phage DNA  
   2.1.viii DNA sequencing  

v
2. **Materials and Methods (continued)**

2.2 Cells and Virus

2.2.i Tissue culture

2.2.ii Virus culture

2.2.iii Construction of recombinant viruses

2.2.iv Screening of recombinant viruses by Southern blot

2.3 Analysis of gene expression

2.3.i Isolation of RNA

2.3.ii Primer extension analysis

2.3.iii S1 nuclease analysis

2.3.iv Nuclear run-on assays

2.3.v Sequencing of RNA using reverse transcriptase

3. **Regulation of cellular genes transduced by herpes simplex virus**

Journal of Virology 63:1929-1937

3.2 Segue

55

4. **Activation of RNA polymerase III transcription of human Alu elements by herpes simplex virus**

in press, Virology

4.1 Segue

59

4.2 Segue

101

5. **Activation of RNA polymerase III transcription of human Alu repetitive elements by adenovirus type 5: requirement for the E1b 58-kilodalton protein and the products of E1 open reading frames 3 and 6**

Molecular and Cellular Biology 13:3231-3244

5.1.i Appendix

5.2 Segue

107

108

6. **Expression of multiple subfamilies of Alu elements is activated by herpes simplex virus and adenovirus infection**

prepared for submission, Molecular and Cellular Biology

6.2 Segue

112

113

159
7. Illegitimate expression of human α-globin, but not β-globin, is activated by herpes simplex virus and adenovirus type 5.

7.1 Forward
7.2 Introduction
7.3 Methods and Materials
7.4 Results
7.5 Discussion
7.6 References

8. Discussion
8.1 Alu repetitive elements
8.1.i Virus-induced expression of Alu repetitive elements
8.1.ii Future experiments
8.2 α-globin
8.2.i Expression of human α-globin
8.2.ii Future work with α-globin
8.3 Concluding remarks

9. References
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Structure of gC-beta</td>
<td>J. Virol 63:1930</td>
</tr>
<tr>
<td>3.2</td>
<td>Southern blot analysis of gC-beta DNA</td>
<td>J. Virol 63:1931</td>
</tr>
<tr>
<td>3.3</td>
<td>Expression of gC RNAs in strain gC-beta</td>
<td>J. Virol 63:1931</td>
</tr>
<tr>
<td>3.4</td>
<td>Time course of β-globin RNA expression during infection with gC-beta</td>
<td>J. Virol 63:1931</td>
</tr>
<tr>
<td>3.5</td>
<td>Effects of inhibitors on β-globin expression</td>
<td>J. Virol 63:1932</td>
</tr>
<tr>
<td>3.6</td>
<td>Structure of strain tk-alpha</td>
<td>J. Virol 63:1932</td>
</tr>
<tr>
<td>3.7</td>
<td>Southern blot analysis of tk-alpha DNA</td>
<td>J. Virol 63:1933</td>
</tr>
<tr>
<td>3.8</td>
<td>Time course of α-globin and Alu expression during infection with tk-alpha</td>
<td>J. Virol 63:1933</td>
</tr>
<tr>
<td>3.9</td>
<td>Effects of inhibitors on expression of α-globin and Alu RNAs</td>
<td>J. Virol 63:1934</td>
</tr>
<tr>
<td>3.10</td>
<td>Location of the 3' end of the AluI transcript</td>
<td>J. Virol 63:1935</td>
</tr>
<tr>
<td>4.1</td>
<td>Primer extension analysis of Alu RNAs produced during HSV infection</td>
<td>68</td>
</tr>
<tr>
<td>4.2</td>
<td>Requirement for viral gene expression and protein synthesis</td>
<td>70</td>
</tr>
<tr>
<td>4.3</td>
<td>Requirements for induction of Alu RNAs</td>
<td>72</td>
</tr>
<tr>
<td>4.4</td>
<td>Role of IE gene products in activation of Alu expression</td>
<td>74</td>
</tr>
<tr>
<td>4.5</td>
<td>Analysis of additional ICP27 mutants</td>
<td>76</td>
</tr>
<tr>
<td>4.6</td>
<td>Nuclear run-on analysis of Alu transcription in infected HeLa cells</td>
<td>78</td>
</tr>
<tr>
<td>4.7</td>
<td>HSV induces Alu expression in multiple cell lines</td>
<td>80</td>
</tr>
<tr>
<td>4.8</td>
<td>Effect of infection with AdICP0 and AdICP4 recombinants on levels of Alu expression</td>
<td>102</td>
</tr>
<tr>
<td>4.9</td>
<td>IE RNA levels from AdICP0 and Ad ICP4</td>
<td>104</td>
</tr>
<tr>
<td>5.1</td>
<td>Induction of Alu expression upon Ad5 infection</td>
<td>Mol. Cell. Biol. 13:3232</td>
</tr>
<tr>
<td>5.2</td>
<td>Table listing mutant Ad5 strains and their sources</td>
<td>Mol. Cell. Biol. 13:3233</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.5</td>
<td>Effects of high multiplicity infection with dl312</td>
<td>107</td>
</tr>
<tr>
<td>5.6</td>
<td>Effect of inhibition of DNA replication with ara C</td>
<td>123</td>
</tr>
<tr>
<td>5.7</td>
<td>Role of early gene products in activation of Alu expression</td>
<td>125</td>
</tr>
<tr>
<td>5.8</td>
<td>Alu transcription following infection with E1b and E4 mutants</td>
<td>127</td>
</tr>
<tr>
<td>5.9</td>
<td>Role of individual E1b gene products</td>
<td>129</td>
</tr>
<tr>
<td>5.10</td>
<td>Effects of deletion of various E4 gene products on Alu activation</td>
<td>129</td>
</tr>
<tr>
<td>5.11</td>
<td>Effect of inhibiting protein synthesis on Alu RNA levels</td>
<td>130</td>
</tr>
<tr>
<td>5.12</td>
<td>Table listing levels of induction of steady state amounts of Alu RNA</td>
<td>132</td>
</tr>
<tr>
<td>5.13</td>
<td>Effects of inhibition of DNA replication on viral and host protein synthesis</td>
<td>134</td>
</tr>
<tr>
<td>5.14</td>
<td>Role of early gene products in transcription of VA RNA</td>
<td>136</td>
</tr>
<tr>
<td>5.15</td>
<td>Effect of inhibition of DNA replication with hydroxyurea</td>
<td>138</td>
</tr>
<tr>
<td>6.1</td>
<td>Relationship between Alu subfamilies identified by Jurka and coworkers</td>
<td>140</td>
</tr>
<tr>
<td>6.2</td>
<td>Competition between Alu oligonucleotide primers</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>Analysis of Alu RNA using subfamily specific primers</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Table comparing quantities of different classes of Alu RNA in mock infected HeLa cells</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>Table showing levels of induction of steady state RNA from various Alu subfamilies</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>Structure of F-Alu2</td>
<td></td>
</tr>
<tr>
<td>6.7</td>
<td>Analysis of Alu expression from F-Alu2</td>
<td></td>
</tr>
<tr>
<td>6.8</td>
<td>Comparison of Sb and S family consensus sequences</td>
<td></td>
</tr>
<tr>
<td>6.9</td>
<td>Sequencing of Alu RNA using the primer Alu2</td>
<td></td>
</tr>
<tr>
<td>6.10</td>
<td>Sequencing of Alu RNA using the Sb primer</td>
<td></td>
</tr>
<tr>
<td>6.11</td>
<td>Levels of Alu RNA from various subfamilies in different cell types</td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>7.1</td>
<td>Analysis of globin expression in HeLa cells</td>
<td>170</td>
</tr>
<tr>
<td>7.2</td>
<td>Time course of α-globin induction</td>
<td>172</td>
</tr>
<tr>
<td>7.3</td>
<td>Activation of α-globin expression requires viral IE and early gene expression</td>
<td>174</td>
</tr>
<tr>
<td>7.4</td>
<td>Role of IE gene products in activation of α-globin expression</td>
<td>176</td>
</tr>
<tr>
<td>7.5</td>
<td>Analysis of α-globin expression induced by Ad/HSV recombinants</td>
<td>178</td>
</tr>
<tr>
<td>7.6</td>
<td>Expression of viral RNA from AdICP4 and 0pro-0</td>
<td>180</td>
</tr>
<tr>
<td>7.7</td>
<td>Effect of Ad infection on β-globin expression</td>
<td>182</td>
</tr>
</tbody>
</table>
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>S</td>
<td>sedimentation coefficient</td>
</tr>
<tr>
<td>pol I</td>
<td>RNA polymerase I</td>
</tr>
<tr>
<td>pol II</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>pol III</td>
<td>RNA polymerase III</td>
</tr>
<tr>
<td>TFII</td>
<td>class II transcription factor</td>
</tr>
<tr>
<td>TFIII</td>
<td>class III transcription factor</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA binding protein</td>
</tr>
<tr>
<td>TAF</td>
<td>TBP associated factor</td>
</tr>
<tr>
<td>A</td>
<td>adenine residue</td>
</tr>
<tr>
<td>T</td>
<td>thymidine residue</td>
</tr>
<tr>
<td>C</td>
<td>cytidine residue</td>
</tr>
<tr>
<td>G</td>
<td>guanine residue</td>
</tr>
<tr>
<td>R</td>
<td>purine residue</td>
</tr>
<tr>
<td>Y</td>
<td>pyrimidine residue</td>
</tr>
<tr>
<td>W</td>
<td>A or T</td>
</tr>
<tr>
<td>N</td>
<td>any nucleotide residue</td>
</tr>
<tr>
<td>MOI</td>
<td>multiplicity of infection</td>
</tr>
<tr>
<td>PFU</td>
<td>plaque forming unit</td>
</tr>
<tr>
<td>CPE</td>
<td>cytopathic effect</td>
</tr>
<tr>
<td>HSV-1</td>
<td>herpes simplex virus type 1</td>
</tr>
<tr>
<td>PAA\textsuperscript F5</td>
<td>herpes simplex virus strain KOS PAA\textsuperscript F5</td>
</tr>
<tr>
<td>F-strain</td>
<td>herpes simplex virus strain F</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr virus</td>
</tr>
<tr>
<td>tk</td>
<td>thymidine kinase</td>
</tr>
<tr>
<td>tk-</td>
<td>thymidine kinase deficient</td>
</tr>
<tr>
<td>gC</td>
<td>glycoprotein C</td>
</tr>
<tr>
<td>gD</td>
<td>glycoprotein D</td>
</tr>
<tr>
<td>gI</td>
<td>glycoprotein I</td>
</tr>
<tr>
<td>V\textsubscript mw</td>
<td>virion molecular weight</td>
</tr>
<tr>
<td>ICP</td>
<td>infected cell protein</td>
</tr>
<tr>
<td>VHS</td>
<td>virion host shutoff</td>
</tr>
<tr>
<td>Ad5</td>
<td>adenovirus type 5</td>
</tr>
<tr>
<td>E\textsubscript 1a</td>
<td>early region 1a</td>
</tr>
<tr>
<td>E\textsubscript 1b</td>
<td>early region 1b</td>
</tr>
<tr>
<td>E\textsubscript 2a</td>
<td>early region 2a</td>
</tr>
<tr>
<td>E4</td>
<td>early region 4</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>IE</td>
<td>immediate early</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pairs</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>E. Coli</td>
<td><em>Escherichia Coli</em></td>
</tr>
<tr>
<td>HB101</td>
<td><em>E. Coli</em> strain HB101</td>
</tr>
<tr>
<td>JM109</td>
<td><em>E. Coli</em> strain JM109</td>
</tr>
<tr>
<td>MV1190</td>
<td><em>E. Coli</em> strain MV1190</td>
</tr>
<tr>
<td>Dh5a</td>
<td><em>E. Coli</em> strain Dh5a</td>
</tr>
<tr>
<td>HeLa</td>
<td>transformed human cervical carcinoma cells</td>
</tr>
<tr>
<td>293</td>
<td>Ad5 E1 transformed human embryonic kidney cells</td>
</tr>
<tr>
<td>MEL</td>
<td>mouse erythroleukemia cells</td>
</tr>
<tr>
<td>Vero</td>
<td>transformed african green monkey kidney cells</td>
</tr>
<tr>
<td>BHK21</td>
<td>transformed baby hamster kidney cells</td>
</tr>
<tr>
<td>°C</td>
<td>degrees celcius</td>
</tr>
<tr>
<td>min</td>
<td>minutes</td>
</tr>
<tr>
<td>ml</td>
<td>millilitre</td>
</tr>
<tr>
<td>μl</td>
<td>microlitre</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>U</td>
<td>arbitrary units</td>
</tr>
<tr>
<td>ddH2O</td>
<td>double distilled, deionized water</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulphoxide</td>
</tr>
<tr>
<td>CaCl2</td>
<td>calcium chloride</td>
</tr>
<tr>
<td>RbCl2</td>
<td>rubidium chloride</td>
</tr>
<tr>
<td>MgCl2</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propane-sulphonic acid</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaOH</td>
<td>sodium hydroxide</td>
</tr>
<tr>
<td>MgSO4</td>
<td>magnesium sulphate</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloric acid</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>tris(hydroxymethyl)aminomethane buffered with HCl</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>KCl</td>
<td>potassium chloride</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenosine nucleotide triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>deoxycytidine nucleotide triphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>deoxythymidine nucleotide triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>deoxyguanidine nucleotide triphosphate</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine nucleotide triphosphate</td>
</tr>
<tr>
<td>CTP</td>
<td>cytidine nucleotide triphosphate</td>
</tr>
<tr>
<td>GTP</td>
<td>guanidine nucleotide triphosphate</td>
</tr>
<tr>
<td>UTP</td>
<td>uridine nucleotide triphosphate</td>
</tr>
<tr>
<td>α-32p</td>
<td>radioactive phosphate at the alpha position</td>
</tr>
<tr>
<td>γ-32p</td>
<td>radioactive phosphate at the gamma position</td>
</tr>
<tr>
<td>Ci</td>
<td>Curie</td>
</tr>
<tr>
<td>cpm</td>
<td>counts per minute</td>
</tr>
<tr>
<td>rpm</td>
<td>rotations per minute</td>
</tr>
<tr>
<td>X g</td>
<td>times gravity</td>
</tr>
</tbody>
</table>
Preface

Several of the chapters in this thesis consist of paper which have been published or submitted for publication. Authorship on these papers is shared by my thesis supervisor, Dr. J. R. Smiley, and myself. It is difficult to determine relative contributions to a body of work that has evolved over six years. In the production of each paper, I carried out all the experiments, organized and prepared the data for publication, and wrote first drafts. There was invariably supervisory revision and the degree of resemblance between first and final drafts increased with time.
Chapter 1: Introduction
1.1 Regulation of eukaryotic gene expression

1.1.1 Regulation of transcription in eukaryotes

The expression of the genome of all organisms is precisely regulated to coordinate the processes of cellular proliferation and differentiation and thus ensure that the correct developmental program is followed. Each cell consists of an assemblage of macromolecules, which are ordered in the precise spatial and temporal sequence such that appropriate alterations in cellular physiology are achieved in response to environmental stimuli. A recurring theme in cellular biology is that metabolic pathways are spatially organized within the cells by compartmentalization, often in membrane bound organelles or on an intracellular matrix. RNA biogenesis, an essential component in the series of events obligatory for gene expression, involves several steps: transcription of the DNA template, processing of the nascent RNA and transport to the appropriate cellular compartment. Each of these steps involves several enzymes and cofactors, and much evidence indicates that RNA metabolism is regulated by the production and spatial organization of the relevant biological macromolecules. Thus there are a variety of mechanisms by which RNA production may be regulated to ensure correct spatial and temporal expression of the genomes of the smallest viruses to the most complex eukaryotes.

Transcription, the first step in RNA biogenesis, requires that the enzymes and cofactors involved in RNA synthesis engage the DNA template and initiate transcription at the correct location in response to
cues which indicate timing and rate of activity. Eukaryotic transcription is carried out by three RNA polymerases which differ in biochemical composition, substrate specificity, and type of RNA produced (Jacob et al., 1970; Kedinge et al., 1970; Lindell et al., 1970; Roeder and Rutter, 1969; Stirpe and Fiume, 1967). Class I genes are transcribed by RNA polymerase I (pol I) and encode 18S and 28S RNAs, RNA components of ribosomes (reviewed in (Sollner-Webb and Tower, 1986). Pol II transcribed class II genes include hnRNAs, the precursor of mRNA, and snRNAs (reviewed in Chambon, 1975; Sawadogo and Sentenac, 1990). A complex assemblage of snRNA-protein particles (snRNPs) is required for the processing of hnRNA into translatable mRNA (Rogers and Wall, 1980; Yang et al., 1981; reviewed in Guthrie and Patterson, 1988) Finally, class III genes encode a variety of abundant, small cytoplasmic RNA molecules, which generally contribute to ribonucleoprotein complexes required for translation (reviewed in Geiduschek and Tocchini, 1988). These include tRNA, 7SL RNA and 5S RNA, another ribosomal constituent. Each gene carries signals which determine which class of polymerase it will engage, in addition to those which modulate the precise time and level of expression.

1.1.ii Cis-acting sequences and trans-acting factors

Correctly regulated transcription is determined by the combination and arrangement of DNA sequence motifs that bind regulatory factors which in turn determine class of polymerase which engages the template and the timing of gene expression. Thus, each class of genes is defined by a different set of regulatory sequences, though some motifs may be shared
between classes of genes (reviewed in Sollner-Webb and Tower, 1986; Chambon, 1975; Geiduschek and Tocchini, 1988). As a result some class II-like genes are transcribed by pol III: the U6 snRNA 7SL RNA and 7SK RNA genes are the best characterized examples (Brow and Guthrie, 1990; Lobo et al., 1990; Murphy et al., 1989; Murphy et al., 1986; Ullu and Weiner, 1985). Promoters, the sequence motifs that bind RNA polymerase and its accessory factors and direct RNA synthesis to begin at a specific site, are located close to the site of transcription initiation (Dynan and Tjian, 1985; McKnight and Tjian, 1986; McKnight and Kingsbury, 1982). Additional regulation is conferred by other cis-acting sequences, including silencers and enhancers which bind transcriptional activators or repressors to modulate basal transcription, usually in a tissue specific or cell-cycle dependent fashion (Banerji et al., 1983; Banerji et al., 1981; Brand et al., 1985; Gillies et al., 1983; Grosschedl and Birnstiel, 1980; reviewed in Atchison, 1988). The precise mechanisms by which regulatory molecules which interact with DNA at sites distant from the gene they modulate probably involves protein-protein interaction between factors acting at the promoter and enhancer. Therefore specific DNA-protein and protein-protein interactions are the basic regulatory components of RNA biosynthesis.

General transcription factors (TF) are polymerase-specific multiprotein complexes that direct RNA polymerases to initiate basal transcription at the correct location on the DNA template (Cozzarelli et al., 1983; Lewis and Burgess, 1980; Lin and Riggs, 1975; Matsui et al., 1980; Segall et al., 1980; Wilkinson and Sollner, 1982; reviewed in Brown, 1984).
The TATA-binding protein (TBP) is an essential component of the cellular transcriptional machinery: it is a subunit of SL1, TFII D and TFIIIB, TF absolutely required for the transcription by pol I, II and III, respectively (Comai et al., 1992; Cormack and Struhl, 1992; Kassavetis et al., 1992; Lobo et al., 1992; Schultz et al., 1992; Taggart et al., 1992; White et al., 1992; White and Jackson, 1992; reviewed in Rigby, 1993). TBP shows some specificity for the hexanucleotide sequence TATAWA present 25 to 35 bp upstream of sites of initiation of transcription of a number of class II and III genes (Parker and Topol, 1984; Sawadogo and Roeder, 1985; Workman and Roeder, 1987), however the requirement for this sequence is not absolute. Basal transcription can be activated by the binding of sequence-specific transactivators to enhancers or upstream regulatory elements (reviewed in Johnson and McKnight, 1989). Transactivators are generally modular, consisting of separable DNA-binding and activation domains. The precise mechanism by which transactivation domains function remains unclear since they show little primary amino acid sequence conservation (though there are some common motifs, such as regions rich in glutamine, proline or acidic residues) and structural analysis suggests that transactivation domains do not generally have a discernible, ordered secondary structure (Ptashne, 1988). Transactivating factors are thought to interact with the general TF through protein-protein interactions involving the transactivation domains and facilitated by a third set of factors known as coactivator molecules (Mitchell and Tjian, 1989). TBP may be the focal point of such interactions: it has the ability to bind a number of proteins,
including TBP associated factors (TAF), other general TF, transactivators and transcriptional inhibitors (Hernandez, 1993).

1.1.iii Chromatin structure

The bulk of protein which interacts with DNA functions to condense and package DNA into chromatin, and sequence specific DNA binding proteins must compete with these packaging proteins (Felsenfeld, 1992). The primary structural units of chromatin are nucleosomes, which consist of approximately 140 bp of DNA wrapped around an octameric histone protein core (Bryan et al., 1979; Lutter, 1979; Prunell et al., 1979; Simpson and Kunzler, 1979). The histone H1 packages nucleosomes into a 30 nm fibre, a helical structure composed six nucleosomes per turn (Allan et al., 1980; Butler, 1984; Thoma et al., 1979; reviewed in Pederson et al., 1986). The 30 nm fibre is organized into looped domains, by attachment to an insoluble protein scaffold which anchors DNA in interphase nuclei and provides the structural basis for metaphase chromosomes (Adolph et al., 1977; Benyajati and Worcel, 1976; Cook and Brazell, 1976; Cook and Brazell, 1978; Earnshaw et al., 1985; Earnshaw and Heck, 1985; Earnshaw and Laemmli, 1983; Igo and Zachau, 1978; Lebkowski and Laemmli, 1982; Lebkowski and Laemmli, 1982; Paulson and Laemmli, 1977). Organization of DNA into nucleosomes and supercoiled loops is thought to be important in the regulation of gene expression since they determine the accessibility of the template by altering DNA topology (Gellert, 1981; Kornberg and Lorch, 1992; Weintraub and Groudine, 1976). Loops may also represent fundamental organizational units of DNA, since discrete segments
containing a transcription unit (or coordinately regulated group of
transcription units) and the cis-acting elements required to direct
expression are often delineated by nuclear matrix attachment sites (Gasser
and Laemmli, 1986; Mirkovitch et al., 1984). In addition, the nuclear
matrix may provide a basic framework or skeleton for the intranuclear
localization of replication foci, transcription factories and gates that engage
RNA processing and transport apparatus (Berezney and Coffey, 1975;
Comings, 1968; Haaf and Schmid, 1991; Xing et al., 1993; Jimenez-Garcia
and Spector, 1993).

Transcriptional activation is often associated with the alterations in
chromatin structure (Weintraub and Groudine, 1976). Histones are subject
to a number of post-translational modifications such as acetylation (De
Lange et al., 1969; Liew et al., 1970; Ogawa et al., 1969), methylation (De
Lange et al., 1969; Ogawa et al., 1969), poly(ADP)ribosylation (Mullins et al.,
1977; Stone et al., 1977), phosphorylation (Balhorn et al., 1972; Gurley et al.,
1973; Lake et al., 1972; Louie and Dixon, 1973) and ubiquitination (Goldknopf
and Busch, 1977; Goldknopf et al., 1977), which alter their structure,
stability and the transcriptional status of the regions of DNA they are
complexed in (reviewed in Bradbury, 1992). Non-histone chromosomal
proteins, such as the high mobility group proteins, probably also play an
important role in determining the transcriptional status of regions of DNA
(Elgin et al., 1974; Elgin and Weintraub, 1975; Paul and Gilmour, 1968;
Shea and Kleinsmith, 1973). Another group of genes, such as PRP20/PIM,
SPT/SIN and SNF/SWI genes in yeast, and EMB-5 and RCC1, their
homologues in multicellular organisms, also modulate chromatin
structure: these families of proteins are required to maintain the transcriptionally active interphase chromatin (Fleischmann et al., 1991; Forrester et al., 1992; Winston and Carlson, 1992). A second interesting gene is XIST, which encodes a functional RNA that colocalizes with the Barr body and is required for X-inactivation, suggesting that in some instances RNA molecules may also play an important role in determining chromatin structure (Brockdorff et al., 1992; Brown et al., 1992; Kay et al., 1993). Gene products that cause more local alterations in the chromatin structure of higher eukaryotes, such as those associated with developmental activation of tissue specific gene expression, have yet to be identified, though there is some evidence that TF and transactivators play a role in this process (Svaren and Horz, 1993; Workman and Roeder, 1987).

1.1.iv RNA processing

The final steps in RNA biogenesis involve the processing of primary transcripts and the transport of functional RNA molecules to the correct intracellular location. The three classes of RNA undergo different forms of post-transcriptional modification (reviewed in Deutscher, 1984; Goessens, 1984; Sharp, 1985). The best characterized example of RNA processing is the extensive alteration of hnRNA to produce translatable mRNA. The majority of newly synthesized class II RNA is processed at 5' and 3' ends, by capping (Contreras and Fiers, 1981; Gidoni et al., 1981; Hagenbuchle and Schibler, 1981; Shatkin, 1976) and polyadenylation (Acheson, 1978; Logan et al., 1980; Proudfoot and Brownlee, 1976; Setzer et al., 1980), and internally, by splicing (Berget et al., 1977; Chow et al., 1977; Kessig, 1977). The same
primary transcript can be processed to yield different mRNAs by the use of alternate splicing and polyadenylation signals (Alt et al., 1980; Early et al., 1980; Maki et al., 1981; Nevins and Darnell, 1978). There is evidence that these processes are regulated, though the mechanisms which determine differential splicing and polyadenylation are not well understood (Andreadis et al., 1987; Leff et al., 1986). The small ribonucleoproteins which catalyze splicing differ in protein and snRNA composition between cell types (Forbes et al., 1984; Pederson, 1983), suggesting that post-transcriptional modification of class II RNA is determined by the type of the splicing machinery mRNA precursors encounter. In addition, RNA and RNA processing apparatus are physically associated with nuclear subdomains (Spector, 1993; Xing et al., 1993), which provide the structural framework for the coordinated nuclear transport and processing of RNA molecules and may provide an additional level of regulation.

1.1.v Viruses as model systems for regulated gene expression

A number of animal viruses have provided invaluable models for the study of the mechanisms which regulate eukaryotic gene expression. Viral genomes are significantly less complex than those of their hosts, and viruses are obligately dependent on cellular factors required for the execution of their genetic programs. Herpes simplex virus type 1 (HSV-1) and adenovirus type 5 (Ad5) have played especially important roles in the development of our understanding of eukaryotic gene regulation. Both these viruses encode factors that interfere with cellular regulatory systems and allow these viruses to usurp host transcriptional apparatus and
preferentially express viral genes. Viral regulatory proteins modulate gene expression at the level of transcription and RNA processing, providing tools for investigation of the cellular mechanisms that regulate both these processes.

HSV-1 and Ad5 are both nuclear DNA viruses (Horowitz, 1990; Roizman and Sears, 1990) which encode regulatory proteins have no apparent sequence homology and are not functionally interchangeable (Spessot et. al., 1989; Zhu et al., 1989), yet they share a number of similarities in the manner in which they appropriate host transcriptional apparatus to facilitate viral replication (reviewed in Horowitz, 1990; Roizman and Sears, 1990). The majority of HSV-1 and Ad5 promoters are composed of cis-regulatory motifs that bind host transactivators and TF and are indistinguishable from cellular promoters. The physical interaction of Ad5 and HSV-1 transactivating proteins with elements of the cellular transcription machinery results in the organization of active transcription complexes on viral genomes, often at the expense of host gene expression. Infection with both these viruses results in changes in nuclear organization: discrete intranuclear compartments, which are sites of viral transcription and DNA synthesis, are formed and the distribution of RNA processing apparatus is altered. The onset of viral DNA synthesis marks the passage between stages of viral gene expression for both these viruses, mimicking a simple developmental system. DNA replication appears to induce a cis-acting modification of viral templates that facilitates expression of a group of genes which had previously been silent (or expressed at very low levels). This progression is modulated by viral
proteins which regulate RNA processing or transport, in addition to transcriptional activators. Thus HSV-1 and Ad5 provide model systems to study a variety of aspects of eukaryotic gene expression.

1.2 Regulation of gene expression by HSV-1 proteins

1.2.i HSV-1 structure and genome organization

Herpes simplex virions consist of linear DNA molecules of approximately 152 kb (Becker et al., 1968), surrounded by an icosadeltahedral capsid (reviewed in Roizman and Sears, 1990). Nucleocapsids are contained within a lipid/glycoprotein envelope, and the region between the capsid and envelope, the tegument, contains an amorphous mass of polypeptides (Roizman and Furlong, 1974). The HSV-1 genome consists of two covalently linked DNA fragments, each flanked by an inverted repeat (Sheldrick and Berthelot, 1975; Wadsworth et al., 1975). The two components freely isomerize with respect to each other, creating a population of virus consisting of equimolar portions of the four possible isotopes (Delius and Clements, 1976; Hayward et al., 1975). HSV-1 encodes more than 70 genes (McGeoch et al., 1988), which can be grouped into three broad classes based on their temporal expression patterns (Honess and Roizman, 1974).

1.2.ii Temporal regulation of HSV-1 gene expression

The kinetic class of HSV-1 genes is defined by the time of production of the gene product and the requirement for viral protein synthesis or DNA replication (Fenwick and Roizman, 1977; Honess and Roizman, 1973;
Honess and Roizman, 1974; Honess and Roizman, 1975; Mackem and Roizman, 1981; Mackem and Roizman, 1982a). The first class of genes to be transcribed during lytic infection, the immediate early (IE) genes, are expressed in the absence of de novo viral protein synthesis (Anderson et al., 1980; Clements et al., 1979; Mackem and Roizman, 1980; Watson et al., 1979). The IE gene products play crucial roles in facilitating the expression of the remaining two classes of gene, the early and late genes (reviewed in Everett, 1987a). Early genes encode factors required for DNA replication and are maximally expressed prior to the onset of viral DNA replication (Honess and Roizman, 1974; Roizman and Sears, 1990). The two subclasses of late genes, the leaky late and true late genes, show varying degrees of dependence on viral DNA replication for high level expression and encode structural proteins required for virion assembly (Holland et al., 1980; Holland et al., 1979; Roizman and Sears, 1990; Wagner, 1984).

Transcription of each of the approximately 70 genes of HSV-1 is driven by individual pol II promoters and viral mRNAs are capped and polyadenylated (Bachenheimer and Roizman, 1972; Costanzo et al., 1977; Silverstein et al., 1973; Silverstein et al., 1976; Stringer et al., 1977). Temporal regulation during infection is determined by promoter sequences, and control occurs primarily at the level of transcription (Godowski and Knipe, 1986; Homa et al., 1988; Post et al., 1981; Silver and Roizman, 1985; Weinheimer and McKnight, 1987). HSV-1 gene expression is determined by two classes of cis-regulatory motifs: those corresponding to the binding sites of cellular transcription factors and sequences required for the binding and activity of regulatory proteins encoded by the virus.
(reviewed in Everett, 1987a; Roizman and Sears, 1990). Analysis of a number of HSV-1 promoters, heterologous eukaryotic promoters and artificial promoters, has demonstrated that temporal class is determined by promoter complexity. The IE promoters consist of TATA-like sequences, binding sites for ubiquitous cellular transcription factors, and IE specific enhancer-like elements (Jones and Tjian, 1985; Jones et al., 1985; Kristie and Roizman, 1984; Mackem and Roizman, 1982; Mackem and Roizman, 1982; Preston et al., 1984). Unlike IE promoters, early and late promoters lack virus specific or class specific regulatory sequences, leading to the suggestion that IE proteins stimulate viral gene expression by altering the activity of the host transcription apparatus (Coen et al., 1986; Eisenberg et al., 1985; El Kareh et al., 1985; Everett, 1984; Homa et al., 1988; Homa et al., 1986; Johnson and Everett, 1986a; Johnson and Everett, 1986b). Early, but not late, promoters contain binding sites for cellular transactivators and though these motifs required for binding by IE proteins they do not appear to play a major role in determining early kinetics of expression. Instead, the TATA-box/cap site region is essential for induction by IE gene products and in many instances determines kinetic class, suggesting that the TATA binding factor TFIID plays a important role in transactivation by IE proteins and temporal regulation of viral gene expression (Homa et al., 1988; Homa et al., 1986; Johnson and Everett, 1986a; Kibler et al., 1991). Thus HSV-1 gene expression is regulated by the interaction of viral regulatory proteins with the host transcriptional machinery. The mechanisms by which these viral polypeptides alter the specificity of the
cellular transcriptional apparatus has contributed greatly to our understanding of eukaryotic gene expression.

1.2.iii Vmw65 and IE gene expression

Though the precise mechanisms by which HSV-1 trans-acting polypeptides modulate gene expression are unclear, one common feature is that they interact with the basal cellular pol II transcription machinery. During lytic infection, the virion protein Vmw65 (also known as VP16, ICP25, or α-TIF) initiates the cascade of HSV-1 gene expression by stimulating transcription of IE genes (Campbell et al., 1984; Post et al., 1981). Vmw65 is a nuclear phosphoprotein that induces IE gene expression through an IE gene-specific cis-regulatory sequence, TAATGARAT (Bzik and Preston, 1986; Campbell et al., 1984; Kristie and Roizman, 1984; Mackem and Roizman, 1982a; Mackem and Roizman, 1982b; Mackem and Roizman, 1982c; Preston et al., 1984). Unlike other transactivators Vmw65 does not directly bind DNA (Marsden et al., 1987; McKnight et al., 1987), instead it associates with Oct-1, a member of the POU-domain class of cellular transcription factors, which recognizes the consensus octamer motif included as part of the TAATGARAT element (O'Hare and Goding, 1988; Preston et al., 1988). The GARAT sequence is not required for Oct-1 binding, but is required for assembly of Vmw65 into a multi-protein-DNA complex, that includes at least one additional cellular factor (Gerster and Roeder, 1988; Katan et al., 1990; Stern and Herr, 1991; Xiao and Capone, 1990). Vmw65 contains an acidic carboxy-terminal domain that is not required for complex assembly, but is essential for transcriptional
activation and is one of the most potent transactivating domains identified to date (Cousens et al., 1989; Sadowski et al., 1988). This region has been shown to interact with TFIID and TFIIB (Ingles et al., 1991; Roberts et al., 1993; Stringer et al., 1990; Walker et al., 1993). These factors, along with TFIIA, are the first to assemble on the template, forming a transcription preinitiation complex (Buratowski et al., 1989; Lin and Green, 1991; Van Dyke et al., 1988). Other components of the general transcription machinery, including TFIIE, TFIIH, TFIIJ and RNA pol II, then assemble on the DNA in an ordered fashion to initiate transcription (Reinberg et al., 1987; Reinberg and Roeder, 1987; Samuels et al., 1982; Samuels and Sharp, 1986; Sawadogo and Roeder, 1985; Sopta et al., 1989; Zheng et al., 1990; Zheng et al., 1987). Therefore, Vmw65 appears to stimulate IE gene expression by interacting with the basal transcription apparatus to facilitate transcription initiation.

It is interesting that the transactivating domain of Vmw65 is not essential for viral growth in tissue culture, suggesting that IE gene expression is not totally dependent on this activity (Greaves and O'Hare, 1989). In contrast, a viral strain containing a deletion of all Vmw65 coding sequences must be propagated on a complementing cell line (Weinheimer et al., 1992). Vmw65(+) virions containing genomes of the Vmw65(-) genotype show abnormal gene expression late in infection, suggesting that de novo synthesized Vmw65 is crucial for normal HSV-1 temporal expression and implying that the portion of the protein which does not contain the transactivation domain also regulates viral gene expression (Weinheimer et al., 1992).
1.2.iv IE gene products

Four of the five IE gene products are crucial for the correct levels of and timing of expression of early and late genes (Costa et al., 1985; DeLuca et al., 1985; Everett, 1984; Everett, 1986; Gelman and Silverstein, 1985; O'Hare and Hayward, 1985a; O'Hare and Hayward, 1985b; Preston, 1979; Quinlan and Knipe, 1985; Sacks et al., 1985; Sacks and Schaffer, 1987; Sears et al., 1985; Stow and Stow, 1986; Watson and Clements, 1980). The mechanisms by which these nuclear phosphoproteins, designated as infected cell proteins (ICP) (Roizman and Sears, 1990), stimulate early and late transcription and mediate the transition between these stages are complex. This complexity is manifested at the level of individual IE proteins, since these polypeptides are generally multifunctional, and in the cross-regulatory interactions between IE proteins, since the activity of several IE gene products is required for correct temporal expression of most early and late genes.

ICP4 is essential for viral growth (Preston, 1979; Watson and Clements, 1980), functioning as a transcriptional activator of early and late genes and a repressor of its own expression (DeLuca et al., 1985; DeLuca and Schaffer, 1985; Everett, 1984; Everett, 1986; Gelman and Silverstein, 1985; Gelman and Silverstein, 1987; O'Hare and Hayward, 1985a; Quinlan and Knipe, 1985). This 175 kDa protein binds DNA nonspecifically, but shows some preference for sites related to the bipartite sequence ATCGTNNNYCGRC (Faber and Wilcox, 1986; Pizer et al., 1991) [a second form of this consensus sequence is ATCGTNNNNCGG (Muller,
ICP4 binds DNA as a 350 kDa homodimer, and DNA binding is crucial for its transcriptional activation and repression functions (DeLuca and Schaffer, 1987; Paterson and Everett, 1988a; Paterson and Everett, 1988b; Shepard et al., 1989). DNA binding by partially purified ICP4 or isolated DNA binding domain causes a bend in DNA very close to the binding site, implying that ICP4-induced alterations in DNA conformation may play an important role in modulating transcriptional repression and activation (Everett et al., 1992). A high affinity ICP4 binding site spanning the site of ICP4 transcription initiation is essential for autoregulation, and it has been suggested that ICP4 binding represses transcription by a combination of steric hindrance at the site of initiation and by affecting the function of upstream activators (Gu et al., 1993; Roberts et al., 1988). In contrast, there is no good evidence for the existence of cis-acting sequences which are critical for ICP4 inducibility (Imbalzano et al., 1990; Shepard and DeLuca, 1991; Smiley et al., 1992), implicating non-specific DNA binding as important in transcriptional activation. Since low affinity ICP4 binding sites are predicted to occur at high frequency in the HSV-1 genome (DiDonato et al., 1991), it seems probable that ICP4 activity is modulated by the positioning of high and low affinity binding sites relative to sites of transcription initiation and other cis-regulatory elements.

Mutational analysis of ICP4, which has demonstrated that ICP4 is composed of a number of discrete functional domains (DeLuca and Schaffer, 1988; Everett et al., 1990; Paterson and Everett, 1988a; Paterson and Everett, 1988b; Shepard and DeLuca, 1991; Shepard et al., 1989). ICP4 molecules purified from infected cells carry post-translational
modifications including poly(ADP)-ribosylation and guanylation at least four GTP-binding sites (Blaho et al., 1992; Blaho and Roizman, 1991), suggesting that this protein is subject to considerable post-translational regulation. Since ICP4 stimulates transcription from promoters in which the only recognizable regulatory motif is the TATA sequence (Homa et al., 1988; Homa et al., 1986; Johnson and Everett, 1986a; Kibler et al., 1991) it seems probable that ICP4 operates through the basal transcriptional machinery acting at the TATA box and that some of the post-translational modifications of will be important in regulating this interaction, perhaps even in a temporally controlled fashion. It is therefore compelling that ICP4 is able to form a tripartite complex with recombinant TBP and TFIIB, and that its ability to form this complex correlates with its ability to bind DNA and activate transcription (DiDonato and Muller, 1989; Smith et al., 1993).

A second essential IE gene product, ICP27, mediates the transition between the early and late stages of infection (McCarthy et al., 1989; Rice et al., 1989; Sacks et al., 1985). ICP27 is a 68 kDa protein which binds zinc ions and interacts with single stranded DNA (Hardwicke et al., 1989; Vaughan et al., 1992). Viral strains which fail to express this gene product exhibit a reduction in levels of viral DNA replication, fail to repress early gene expression late in infection, and show dramatically reduced levels of late gene expression (McCarthy et al., 1989; Rice et al., 1989; Sacks et al., 1985). Transfection experiments have demonstrated that ICP27, in cooperation with other viral IE gene products, can repress expression of some IE and early genes and stimulate expression of some late genes and that these two
functions are mutationally distinct (Everett, 1988; Sekulovich et al., 1988; Su and Knipe, 1989). Activator mutants were trans-dominant in transfection assays, indicating that ICP27 exerts its effects by interacting with a host protein present in limited quantities (Smith et al., 1991). Detailed analysis of finer ICP27 mutations in virus confirmed that ICP27 carries out two mutationally separable functions that facilitate the early to late transition: one which stimulates viral DNA replication and the expression of leaky late genes, and a second which activates expression of true late genes, independent of DNA replication (Rice et al., 1993; Rice et al., 1989). These mutations are interesting because they uncouple true late gene expression from viral DNA replication and suggest that late gene expression is not purely replication driven or facilitated solely by high copy number.

Little is known of the mechanism about action of ICP27, though evidence indicates it regulates gene expression at the post-transcriptional level, possibly by affecting mRNA processing (Chapman et al., 1992; McLauchlan et al., 1992; Sandri-Goldin and Mendoza, 1992; Smith et al., 1992). In transient expression assays sequences involved in 3'-end processing of pre-mRNA appear to mediate transactivation, while intron sequences are required for repression (Sandri-Goldin and Mendoza, 1992). This is further supported by the observation that infection with HSV-1 dramatically alters the intranuclear distribution of cellular splicing machinery and that ICP27 mutants do not facilitate this alteration (Phelan et al., 1993). ICP27 mutants also fail to mediate a second alteration in intranuclear structure: these viral strains do facilitate the production of viral replication compartments (Curtin and Knipe, 1993). Thus in nuclei of
cells infected with ICP27 deficient strains, the viral replicative machinery
does not organize into distinct punctate regions which colocalize with the
site of viral transcription late in infection. It has been suggested that ICP27
acts post-translationally, either to facilitate alterations in protein structure
or localization of viral replicative proteins (Curtin and Knipe, 1993). Since
transient expression of ICP27 in transfection assays is sufficient to induce
snRNP redistribution, it is unlikely that reorganization of splicing
apparatus is a secondary consequence of the formation of viral replicative
compartments (Phelan et al., 1993). Taken in combination these data
indicate that ICP27 functions to cause fairly dramatic alterations in
intranuclear organization of the machinery required for viral DNA
replication, transcription and RNA processing, and suggests that this
reorganization is necessary for the transition from early to late gene
expression.

Though ICP0 is a potent activator of gene expression (reviewed in
Everett, 1987a), it is not essential for HSV-1 replication in cultured cells
(Stow and Stow, 1986; Sacks and Schaffer, 1987). This 110 kDa polypeptide is
necessary for high level expression of all classes of viral proteins during
low multiplicity infections but mutants lacking ICP0 are indistinguishable
from wild type virus at higher multiplicities (Cai and Schaffer, 1992; Chen
and Silverstein, 1992). The ICP0 gene is the first IE gene expressed during
lytic infection, suggesting that ICP0 expression is required to stimulate
further IE gene expression and efficiently initiate the cascade of viral gene
expression (Elshiekh et al., 1991). Indirect evidence supporting this
suggestion comes from the similarity in phenotype between ICP0 mutants
and Vmw65 mutants which lack the transactivation domain (Cai and Schaffer, 1992; Chen and Silverstein, 1992; Greaves and O'Hare, 1989).

Also, the requirement for ICP0 is most pronounced under conditions when virion proteins (presumably Vmw65) are not provided in trans. Thus DNA isolated from ICP0 mutant viruses is less infectious by transfection and mutant strains lacking ICP0 do not reactivate from latency with normal efficiency (Cai et al., 1993; Cai and Schaffer, 1992). These data suggest that ICP0 plays a pivotal role in helping the viral genome to achieve or maintain a transcriptionally active state.

ICP0 is a non-specific transactivator, since it can stimulate the expression of all classes of HSV-1 genes and heterologous eukaryotic genes in transient assays, alone or in a synergistic manner with ICP4 (Everett, 1984; Everett, 1987a; Everett, 1987b; Gelman and Silverstein, 1985; Gelman and Silverstein, 1987; O'Hare and Hayward, 1984; O'Hare and Hayward, 1985; Perry et al., 1986; Quinlan and Knipe, 1985), it is not essential for viral replication in tissue culture (Sacks and Schaffer, 1987; Stow and Stow, 1986). ICP0 has nonspecific DNA binding activity, and a potential Zn-finger domain is essential for ICP0 function in transfection assays and in virus (Everett, 1987b), suggesting that DNA binding and transactivation activity are tightly coupled. Isolation of dominant-negative ICP0 mutant proteins suggest that ICP0 interacts with a cellular factor and that DNA binding activity and transactivation may be mediated by interactions with host proteins (Chen et al., 1992; Weber and Wigdahl, 1992). Cai and Schaffer (1991) have identified a cellular function which can substitute for ICP0, providing further evidence that ICP0 may connect into a cellular pathway.
The host ICP0-like function is apparently regulated in a cell cycle dependent manner, since it is maximally expressed after cells are released from growth arrest. Evidence of direct physical interactions between ICP0 and TBP, and ICP0 and ICP4, further suggest that ICP0 may function as transcriptional activator by modulating the basal transcriptional machinery, either directly or indirectly through ICP4.

ICP22 is a nuclear phosphoprotein which is required for the production of normal levels of late proteins in primary human cell strains and rodent cell lines and is therefore required for optimal viral replication in these cell types (Poffenberger et al., 1993; Sears et al., 1985). ICP22 deletion mutants show slightly extended expression of some early genes and delayed activation of late genes in permissive and non-permissive cells, implying that this gene product is required for transition from the early to the late stage of infection (Poffenberger et al., 1993; Sears et al., 1985). ICP22 deficient mutants also produce reduced levels of ICP0 in non-permissive cells, suggesting that abnormal ICP0 expression may contribute to the ICP22 mutant phenotype (Purves et al., 1993).

Phosphorylation plays a crucial role in regulation of ICP22 activity, since mutants lacking the virion structural protein encoded by the UL13 protein kinase gene produce aberrantly phosphorylated ICP22 and have the same phenotype as ICP22 deletion mutants (Purves et al., 1993).

1.2.v Inhibition of gene expression by viral products

The majority of the regulatory proteins encoded by HSV-1 are transcriptional activators which interact with the basal cellular
transcription machinery to facilitate synthesis of viral mRNAs by host RNA pol II. In addition, HSV-1 specifies post-transcriptional regulators of gene expression: ICP27 is one example. A second instance of virus encoded post-transcriptional regulatory protein is the virion host shutoff (VHS) protein, which is partially responsible for the global shut off of host gene expression which occurs during lytic infection (Fenwick and McMenamin, 1984; Nishioka and Silverstein, 1977; Nishioka and Silverstein, 1978; Read and Frenkel, 1983). Almost immediately upon infection VHS induces a rapid turnover of cellular mRNA and decrease in levels of host protein synthesis. Viral mRNAs are not spared from VHS activity: they are also destabilized and the down-regulation of VHS function may be required for the accumulation viral RNA after the initial wave of VHS activity (Fenwick and Clark, 1983; Kwong et al., 1988; Read and Frenkel, 1983). Mutants which lack the VHS protein are significantly less efficient for host shut off early in infection but still show suppression of host gene expression at late times, suggesting that a second type of inhibition of host gene expression is mediated by viral products other than VHS (Fenwick, 1984; Fenwick and McMenamin, 1984). Transcription of most cellular genes is repressed by HSV-1 infections which are restricted to the IE phase of the lytic cycle (Pizer and Beard, 1976; Smibert and Smiley, 1990; Stenberg and Pizer, 1982), implying that an IE protein acts as an inhibitor of cellular transcription.

Thus the level of expression of any HSV-1 gene is determined by the differential effects of viral transactivators, repressors and post-transcriptional regulators, each of which is modulated in a temporally, and perhaps spatially, distinct manner to allow correct temporal expression of
the viral genome, at the expense of the host. Ad5 is very similar to HSV-1 in that it also expresses a variety of regulatory polypeptides which facilitate viral gene expression by subsuming the cellular transcriptional machinery. Perhaps the most notable difference between HSV-1 and Ad5 is the mechanisms these viruses use to modulate their own developmental programs. Expression of each HSV-1 gene is directed by individual promoters (Godowskki and Knipe, 1986; Homa et al., 1988; Post et al., 1981; Silver and Roizman, 1985; Weinheimer and McKnight, 1987; reviewed in Everett, 1987a), and the vast majority of messages are unspliced (Phelan et al., 1993). In contrast, Ad5 has a limited number of promoters (Berk et al., 1979; Price and Penman, 1972; Sharp et al., 1975; Soderlund et al., 1976; Stillman et al., 1981; Weinmann et al., 1976; Wilson et al., 1979; reviewed in Horowitz, 1990) and depends on alternative splicing to maximize the coding capacity of its small genome (Akusjarvi and Persson, 1981; Berget et al., 1977; Chow et al., 1977; Gelines and Roberts, 1977; Nevins and Wilson, 1981; Ziff and Evans, 1978). The transition from early to late stages of the lytic cycle of Ad5 and HSV-1 requires viral proteins that regulate RNA processing, but they differ in that HSV-1 appears to repress expression of spliced transcripts (Sandri-Goldin and Mendoza, 1992), while Ad5 modulates alterations in the RNA processing or transport apparatus to stimulate late gene expression (Brady et al., 1992; Bridge et al., 1991; Dix and Leppard, 1993; Leppard, 1993; Leppard and Shenk, 1989; Nordqvist and Akusjarvi, 1990; Ohman et al., 1993; Pilder et al., 1986; Sandler and Ketner, 1989).
1.3. Regulation of Ad5 gene expression

1.3.i Ad5 structure and genome organization

The 36 kb genome of Ad5 is packaged in a nucleoprotein core which is surrounded by an icosahedral protein shell (Horowitz, 1990). Ad5 DNA molecules are characterized by inverted terminal repeats with a 55 kDa protein covalently linked to the 5'-ends (Garon et al., 1972; Rekosh et al., 1977; Robinson et al., 1973; Wolfson and Dressler, 1972). The terminal protein is required for initiation of viral DNA replication, a process which requires a combination of cellular transcription factors and virally encoded DNA binding protein and polymerase (Diffley and Stillman, 1986; Friefeld et al., 1983; Lichy et al., 1982; Lichy et al., 1983; Nagata et al., 1982; Pruijn et al., 1986; Pruijn et al., 1987; Stillman, 1983; Stillman and Tamanoi, 1983). The Ad5 genome is organized into transcription units, defined by promoters that each produce multiple mRNAs due to alternative splicing (reviewed in Horowitz, 1990; Persson and Philipson, 1982).

1.3.ii Temporal regulation of Ad5 gene expression

Ad5 contains six class II promoters and two class III promoters and temporal expression is determined in part by these promoter sequences (Berk et al., 1979; Price and Penman, 1972; Sharp et al., 1975; Sodrelund et al., 1976; Stillman et al., 1981; Weinmann et al., 1976; Wilson et al., 1979). The five early transcription units are maximally expressed prior to viral DNA replication and produce pre-mRNAs that undergo extensive processing (Berget et al., 1977; Chow et al., 1979; Chow et al., 1977). The
onset of viral DNA synthesis marks to passage between early and late stages of the lytic cycle (Lucas and Ginsberg, 1971; Shaw and Ziff, 1980; Thomas and Green, 1969; Thomas and Mathews, 1980). Transcription of the majority of late mRNA is directed by the major late promoter, and these transcripts show even more extensive differential splicing and polyadenylation (Akusjarvi and Persson, 1981; Berget et al., 1977; Chow et al., 1977; Gélinas and Roberts, 1977; Nevins and Wilson, 1981; Ziff and Evans, 1978). In addition, the two class III transcription units, the VA1 and VA2 genes, are also expressed late in infection (Soderlund et al., 1976; Weinmann et al., 1976). Ad5 class II and class III promoters lack virus-specific regulatory sequences and consist solely of binding sites for cellular trans-acting factors (reviewed in Berk, 1986) and the virus relies on host machinery for the biogenesis of its RNAs (Price and Penman, 1972).

1.3.iii E1a gene products and viral gene expression

The proteins of early transcription unit 1 (E1a) of Ad5 are the first to be expressed during lytic infection, and these stimulate transcription of the remaining early and late promoters (Berk et al., 1979; Jones and Shenk, 1979a; Montell et al., 1982; Nevins, 1981). E1a gene products have no sequence-specific DNA binding activity (Chatterjee et al., 1988), and can act on a variety of viral and cellular class II and class III promoters that do not share common sequence elements (Gaynor et al., 1985; Gaynor et al., 1984; Green et al., 1983; Hearing and Shenk, 1983; Kingston et al., 1984; Leff et al., 1985; Svensson and Akusjarvi, 1984; Treisman et al., 1983). Analysis of E1a-inducible promoters has demonstrated that the same host specified
factors that are required for constitutive activity of these promoters are also required for transcriptional stimulation by E1a proteins (Lin and Green, 1988; Pei and Berk, 1989; Simon et al., 1988; Wu et al., 1987; reviewed in Berk, 1986). Thus it has been suggested that E1a products act indirectly, by modification or induction of host cell factors, to assemble viral DNA into transcription complexes (reviewed in Berk, 1986; Flint and Shenk, 1989).

At least five differentially spliced E1a mRNAs are produced and the two predominant forms, 12S and 13S mRNAs, encode 234 and 289 amino acid proteins (Berget et al., 1977; Chow et al., 1977; Esche et al., 1980; Perricaudet et al., 1979). These E1a proteins differ only in the presence of an internal 46 residue element with zinc binding activity, designated as conserved region 3 (CR3) due to its high degree of sequence conservation between different adenovirus serotypes (Culp et al., 1988; Kimelman et al., 1985; Moran and Mathews, 1987). Both 12S and 13S gene products are transactivators of viral gene expression and they differ in their potency and targets (Ferguson et al., 1985; Leff et al., 1984; Winberg and Shenk, 1984). This suggests that conserved regions 1 and 2, which are present in both 12S and 13S mRNAs, may be important for transactivation functions which are shared between these polypeptides, while CR3 may account for the differences. The larger E1a protein has more potent transactivating capability and a broader range of target sequences and is primarily responsible for initiating viral transcription during lytic infection (Carlock and Jones, 1981; Montell et al., 1982; Ricciardi et al., 1981). Analysis of the CR3 region suggests that it carries out two mutationally separable functions: the amino-terminal portion of CR3 provides a transactivation
domain while the carboxy-terminal region is thought to be necessary for protein-protein interactions that facilitate assembly of transcription complexes (Fahnestock and Lewis, 1989; Webster et al., 1991).

E1a gene products activate class II transcription by a variety of mechanisms that result in changes in the amount or activity of cellular transcription factors (reviewed in Flint and Shenk, 1989). These mechanisms include post-translational modifications, such as alterations in protein phosphorylation, disruption of protein-protein interactions or formation of E1a-containing multi-protein transcription complexes. E1a-associated kinases are thought to be responsible for alterations in the phosphorylation states of cellular transcription factors E2F and E4F seen upon Ad5 infection (Cortes et al., 1988; Hardy and Shenk, 1988; Raychaudhuri et al., 1987). E1a gene products bind a number of polypeptides that are not themselves transcription factors, but which regulate the activity of this class of polypeptides. For example, the products of the 12S and 13S mRNAs interact with the retinoblastoma protein (Rb), disrupting an Rb-E2F complex, presumably freeing E2F to stimulate viral transcription (Bandara and LaThangue, 1991; Chellappan et al., 1991). Direct physical interaction with a variety of cellular transcription factors, including AP-1, CREB-P1/ATF2, and Oct-4, has also been reported (deGroot et al., 1991; Lillie and Green, 1989; Liu and Green, 1990; Maguire et al., 1991; Martin et al., 1990; Scholer et al., 1991). E1a proteins have also been shown to bind TBP (Horikoshi et al., 1991; Lee et al., 1991) and a 300 kDa TBP associated factor (Abraham et al., 1993), suggesting that they directly modulate the basal cellular transcriptional machinery. The observation
that holo-TFIID is able to overcome E1a-mediated squelching in *in vitro* transcription systems indicates that the interaction between E1a products and components of TFIID has functional significance (Boyer and Berk, 1993).

The E1a 289 amino acid protein also stimulates class III transcription of VA genes, by increasing the amount or activity of TFIIC, a limiting pol III transcription factor (Fuhrman et al., 1984; Hoeffler et al., 1988; Hoeffler and Roeder, 1985; Yoshinaga et al., 1986). TFIIC functions as an assembly factor, which facilitates the interaction of TFIIB with DNA: TFIIC binds intragenic pol III promoter sequences and positions TFIIB on the template by activating its DNA binding potential (Camier et al., 1990; Geiduschek and Tocchini, 1988; Kassavetis et al., 1991; Kassavetis et al., 1990; Kassavetis et al., 1989; Klekamp and Weil, 1987). TFIIB is a multisubunit factor which contains TBP and stably binds the upstream regions of pol III genes without any apparent sequence specificity (Kassavetis et al., 1992; Lobo et al., 1992; Taggart et al., 1992; White and Jackson, 1992; White et al., 1992). It is unclear whether the E1a/TBP interaction plays a role in regulation of class III transcription, though this seems a likely scenario.

1.3.iv E1 proteins and cellular gene expression

In addition to transactivation functions, E1a gene products are able to transform cells in cooperation with other oncogene products, including proteins encoded in the adenovirus E1b region (Graham et al., 1978; Haley et al., 1984; Jones and Shenk, 1979; Ruley, 1983; Zerler et al., 1986; van de Eb
et al., 1977; reviewed in Branton et al., 1985; Shenk and Flint, 1991). CR1 and CR2 domains are required for immortalization and transformation by E1a polypeptides (Schneider et al., 1987; Whyte et al., 1988). E1a proteins associate with a number of cellular proteins (Harlow et al., 1986; Yee and Branton, 1985) and formation of these complexes is thought to play a role in transformation. Two of the proteins which associate with E1a polypeptides, the 105 kDa product of the retinoblastoma anti-oncogene (Whyte et al., 1988) and cyclin A (Faha et al., 1992), are involved in regulation of cell cycle progression, suggesting that E1a polypeptides interfere with normal cell cycle regulation. E1a gene products are able to stimulate cells to enter S phase (Kaczmarek et al., 1986; Stabel et al., 1985) and progress through mitosis (Zerler et al., 1987) supporting this hypothesis. It seems possible that E1a-induced stimulation of expression of some cellular housekeeping genes, such as proliferating cell nuclear antigen (Morris and Mathews, 1991) and brain creatine kinase (Kaddurah et al., 1990) (both by a mechanism that requires CR1 and CR2), may contribute to cell cycle deregulation and transformation. E1a gene products also inhibit some instances of tissue-specific gene expression: the down regulation of members of the metalloprotease gene family requires CR1 and involves repression of AP-1 activity (Offringa et al., 1990), while inhibition of muscle gene expression is achieved by inactivation of the muscle specific transcription factor, Myf-5, and is also dependent on CR1 (Braun et al., 1992). The domains required for E1a-mediated transcriptional repression are not mutationally separable from those required for transformation (Lillie et al., 1986; Schneider et al., 1987; Velcich and Ziff, 1985), however it
is not clear whether repression is a prerequisite for transformation or an incidental side effect of this process.

Ad5 E1b encodes two major gene products of 19 kDa and 58 kDa (Bos et al., 1981; Perricaudet et al., 1979), and each of these proteins transforms primary cells in cooperation with E1a proteins (Barker and Berk, 1987; Graham et al., 1978; van der Eb et al., 1977). The 19 kDa protein is required for viral DNA synthesis, the protection of newly synthesized viral DNA from degradation and cellular transformation (Barker and Berk, 1987; Chinnadurai, 1983; Pilder et al., 1984; Stillman, 1986; Subramanian et al., 1984; White et al., 1984; White and Stillman, 1987). Transient expression of the 19 kDa protein results in alterations of the organization of nuclear lamina and intermediate filaments, suggesting that the cellular target of the E1b 19 kDa protein is the cytoskeleton (White and Cipriani, 1989; White and Cipriani, 1990). The function of the larger E1b protein may be related to its ability to bind and interfere with the tumour suppressor gene product p53 (Kao et al., 1990; Sarnow et al., 1982; Yew and Berk, 1992). The p53 protein is a transcription factor that appears to act as a chaperone of the cell cycle: it is a non-essential inhibitor of cell cycle progression, whose failure to function results in the propagation of cells displaying genomic instability associated with transformation (Farmer et al., 1992; Oliner et al., 1992; reviewed in Lane, 1992). Analysis of a number of adenovirus type 2 55 kDa E1b mutant proteins (homologous to the Ad5 58 kDa E1b protein) demonstrated that the ability of E1b proteins to inhibit p53 transactivation correlates with their transforming activity (Yew and Berk, 1992). Therefore, it seems probable that the larger E1b gene product modulates
host gene expression in a manner similar to E1a proteins, by interacting
with an key cellular regulatory protein.

1.3.5 Regulation of viral gene expression by E1b and E4 gene products

At least 7 species of mRNA are encoded in E4 (Freyer et al., 1984;
Tigges and Raskas, 1984; Virtanen et al., 1984), and the proteins produced
from of four of these mRNA's play regulatory roles during Ad5 infection
(Huang and Hearing, 1989a; Huang and Hearing, 1989b; Sandler and
Ketner, 1989). The 19 kDa protein encoded by E4 ORF6/7 complexes with
cellular transcription factor E2F, stabilizing its interaction with DNA and
enhancing transcription from the E2a promoter (Huang and Hearing,
1989b; Marton et al., 1990; Raychaudhuri et al., 1990). The mode of action of
this E4 gene product recapitulates the theme that viral regulatory proteins
function by modulating the activity of host transcriptional apparatus.

The product of E4 ORF4 appears to have multiple functions in
regulation of viral gene expression. Mutants lacking E4 ORF4 accumulate
hyperphosphorylated forms of c-FOS and E1a and show an increase in AP-1
dNA-binding activity (Muller et al., 1992). Since E1a and cyclic AMP
cooperate to induce AP-1 activity during lytic infection, one possible role for
E4 ORF4 proteins is to down-regulate the induction of AP-1, perhaps by
altering phosphorylation of E1a and c-FOS (Kleinberger and Shenk, 1993).
The E4 ORF4 polypeptide may also regulate viral DNA replication in
combination other E4 gene products, and thus indirectly modulate the
transition from early to late gene expression (Bridge et al., 1993). Viral
mutants lacking all E4 ORF's except E4 ORF4 show a profound reduction in
levels of viral DNA replication, while mutants lacking the entire E4 region show a much slighter decrease in levels of DNA synthesis. Polypeptides expressed from E4 ORF3 and ORF6 antagonize the effect of E4 ORF4 proteins, suggesting that the products of these three E4 ORF's cooperate in the regulation of viral DNA replication, and therefore, indirectly in viral late gene expression (Bridge et al., 1993).

The products of E4 ORF3 and ORF6 also have a more direct effect on viral late gene expression. Genetic and biochemical evidence indicate the products of E4 ORF3 and 6, in addition to the 58 kDa protein encoded by E1b, are required for the cytoplasmic accumulation of late viral mRNA (Bridge and Ketner, 1989; Huang and Hearing, 1989a; Leppard and Shenk, 1989; Sandler and Ketner, 1989). Viral mutant strains which do not produce these gene products show similar phenotypes: decreased or delayed viral DNA replication, reduction in cytoplasmic levels of late mRNA with an accompanying decrease in viral late protein synthesis, and failure to shut off host protein synthesis (Babiss et al., 1985; Halbert et al., 1985; Pilder et al., 1986; Weinberg and Ketner, 1983; Weinberg and Ketner, 1986; Williams et al., 1986; Yew et al., 1990). E4 and E1b gene products appear to regulate Ad5 gene expression by increasing the intranuclear stability or facilitating the processing or transport of late viral mRNA precursors. Evidence suggests that E4 gene products are required for nuclear transport of viral late mRNA and correct splicing of the major late promoter tripartite leader, while the E1b 58 kDa protein is necessary for processing of late mRNAs (Brady et al., 1992; Bridge et al., 1991; Dix and Leppard, 1993; Leppard, 1993; Leppard and Shenk, 1989; Nordqvist and Akusjarvi, 1990; Ohman et al.,
1993; Pilder et al., 1986; Sandler and Ketner, 1989). Transcription and RNA processing appear to be coupled in vivo (Jimenez-Garcia and Spector, 1993) and it has been suggested that there are a limited number of sites at which active genes are gated to nuclear transport/processing machinery and that viral proteins increase the number of these sites to facilitate expression of late genes from newly replicated viral templates (Leppard and Shenk, 1989). Therefore, it is interesting to note that the E1b 58 kDa protein and the 34 kDa product of E4 ORF6 form a nuclear complex (Cutt et al., 1987) which is thought to be essential for their function and that this complex is located in virus specific inclusions that are thought to be sites of DNA replication and transcription late in infection (Hasson et al., 1992).

Differential splicing contributes to the temporal control of Ad5 gene expression, since splicing patterns of viral transcripts alters during the different stages of the lytic cycle (Gattoni et al., 1991; Larsson et al., 1991; Montell et al., 1984; Wilson et al., 1992; Wilson and Darnell, 1981; Zerivitz et al., 1992). Splicing of non-adenoviral genes inserted into the Ad5 genome is regulated, and these patterns are imposed on foreign genes, which do not normally demonstrate regulated splicing (Adami and Babiss, 1991). Progression into the late phase of infection causes the nuclear redistribution of the splicing machinery, snRNPs concentrate in discrete clusters which are distinct from the sites of DNA replication and late transcription (Bridge et al., 1993). However, staggered infections with two differentially marked genomes revealed that correctly regulated expression could occur simultaneously from early and late genomes in the same nucleus, indicating that early splicing can occur in the presence of global
alterations in trans-acting splicing factors associated with the change to a late splicing pattern (Adami and Babiss, 1991). Therefore the regulation of the transition late stage post-transcriptional RNA processing is most likely due to changes in the state of the DNA template, or perhaps its intranuclear localization. It is unclear whether E1b and E4 gene products play a direct role in formation of viral replicative compartments or reorganization of splicing apparatus, though their function as post-transcriptional regulators of late gene expression supports this possibility. Finally, cellular messages are transcribed at nearly normal levels during Ad5 infection but are not exported from the nucleus (Beltz and Flint, 1979), suggesting that the modulation of host splicing apparatus may inhibit the processing of cellular mRNA precursors.

1.4 **Statement of purpose**

The enigma provided by both HSV-1 and Ad5 regulators is that though they interact with cellular regulatory apparatus, their net affects are the specific activation of viral gene expression. Both viruses cause the relocalization of RNA pol II to nuclear compartments that are sites of viral transcription and DNA replication. Both viruses also induce the alteration in the intranuclear organization of RNA processing apparatus. These changes in intranuclear organization associated with transcription and processing of viral genes may, directly or indirectly, adversely affect host gene expression. Therefore the ultimate activity of viral regulatory proteins is the facilitation viral gene expression by altering the nuclear environment
such that viral genomes out compete cellular templates for the general transcription apparatus. The purpose of the experiments which were compiled to produce this thesis was to study two unusual instances of activation of cellular gene expression during HSV-1 or Ad5 infection, with the eventual aim of learning more about the mechanisms by which viral regulatory proteins function.
Chapter 2: Methods and Materials
2.1. **Manipulation of plasmids and phage**

2.1.i. **Bacterial culture**

HB101, JM109, MV1190 and Dh5α, strains of *E. coli*, were grown in 37 °C continuous agitation suspension cultures of Luria broth or Terrific broth (Sambrook *et al.*, 1989). When required, the antibiotics ampicillin (Sigma; 20 μg/ml) or tetracycline (Sigma; 10 μg/ml) were used to select for bacteria containing plasmids bearing antibiotic resistance genes. Bacteria were stored frozen at -70 °C in 1 ml aliquots in cryovials (Nunc). DMSO (Bethesda Research Laboratories) was added to log phase cultures (OD₆₀₀=0.4 to 0.6) at a final concentration of 7% prior to freezing.

2.1.ii. **Transformation of bacteria**

Two procedures were used to produce competent cells: the CaCl₂ and RbCl₂ techniques. Both these procedures alter the physiology of the bacterial cell wall to allow entry of DNA molecules. The RbCl₂ was used more regularly since it was less time consuming and consistently produced a higher yield of competent bacteria which remained viable at 4 °C for at least one week. The CaCl₂ technique is described in detail in Sambrook *et al.* (1989). The RbCl₂ technique diverges from the CaCl₂ technique once bacteria have been pelleted for the first time. The bacterial pellet is resuspended in 20 ml of 4 °C MOPS I buffer (50 mM MOPS pH 7.0, 70 mM CaCl₂, 10 mM RbCl₂) and immediately centrifuged (1000 g, 4 °C, 10 min). The pellet is then resuspended in MOPS II buffer (100 mM MOPS pH 6.5, 10
mM RbCl₂) and promptly centrifuged once more. The bacterial pellet is resuspended in a final volume of 2 ml of MOPS II.

Competent bacteria were incubated with DNA for 30 to 60 min on ice, then subjected to a 42 °C heat shock for 2 min and plated on 10 cm Luria agar plates containing antibiotics. Plates were incubated at 37 °C for 8 to 16 hours before colonies were visible.

2.1.iii. DNA manipulations and cloning

All DNA manipulations required for cloning, including small scale isolation of plasmid DNA, determination of DNA concentration, restriction enzyme digests, techniques to remove 5' and 3' overhangs, ligations, phenol:chloroform extractions, ethanol precipitations and gel electrophoresis (polyacrylamide or agarose) are described by Maniatis et al. (1982) or Sambrook et al. (1989).

2.1.iv. Purification of DNA fragments

DNA was isolated from polyacrylamide gels for cloning or S1 analysis. Isolation from polyacrylamide required excising the appropriate band, mashing it with a teflon plunger and incubating the acrylamide pieces in elution buffer [300 mM sodium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% SDS] overnight at 37 °C or several hours at 65 °C. Polyacrylamide was removed by spinning the mixture through siliconized glass wool, and DNA was ethanol precipitated from the elution buffer. This procedure was also used to purify synthetic oligonucleotides, which were resolved on 10%, 7 M Urea sequencing gels.
DNA fragments were isolated from standard agarose and low melting point agarose gels using techniques described in Sambrook et al. (1989). Quiaxe (Quiagen) and GeneClean (Bio 101) kits were also employed to purify DNA fragments from agarose gels.

2.1.v. Polymerase Chain Reaction (PCR) procedure

PCR (Mullis et al., 1986) was carried out with Taq polymerase (Pharmacia) using buffers supplied by the manufacturer. Oligonucleotide primers were synthesized at the Central Facility of the Institute for Molecular Biology and Biotechnology, McMaster University. Reaction conditions varied between primer sets and were determined by calculating annealing temperature of the primers (based on length and GC content) and length of the PCR product as described in PCR protocols: a guide to methods and applications (Innes, 1990). The PCR products, which were purified by phenol:chloroform extraction and ethanol precipitation, were analyzed by gel electrophoresis.

2.1.vi. Large scale isolation of plasmid DNA

A number of methods were used to purify large amounts of plasmid DNA and all procedures generally yielded approximately 1 mg of plasmid DNA from 1 l of bacterial culture. Banding in a cesium chloride density gradient (Maniatis et al., 1982) resulted in the purest DNA preparations and these were used for transfections into mammalian cells and in vitro transcription reactions. Two additional techniques were employed to purify plasmid DNA for all other applications: the LiCl method as described in
Sambrook *et al.* (1989) and Qiagen (Quiagen) columns, as per the manufacturer's instructions.

2.1.vii. **Large scale isolation of M13 phage DNA**

Large volumes (1-2 l) of log phase cultures of the *E. coli* strain MV1190 were infected with bacteriophage inoculum (Sambrook *et al.*, 1989) and grown for 8 to 10 hours. The infected bacteria were first removed by two consecutive 20 min centrifugations at 5 000 rpm in a Sorvall GSA rotor. The M13 phage was then precipitated from the culture medium using PEG. The volume of the culture supernatant was determined, and 1/4<sup>th</sup> that volume of a 20% PEG/2.5 M NaCl solution was added. After 15 minutes stirring at room temperature, the phage was pelleted by a 20 min centrifugation at 8 000 rpm in a GSA rotor. The precipitated phage, clearly visible as a creamy white smear on the sides of the centrifuge bottle, was resuspended in 50 to 100 ml of a solution containing 10 mM NaCl, 10 mM Tris-HCl pH 7.5, and 1 mM EDTA. The volume of the resuspended phage was measured, and 1/8<sup>th</sup> that volume of 20% PEG/2.5 M NaCl was added. This mixture was incubated at room temperature for 15 minutes and the phage pelleted by centrifugation for 20 min at 10 000 rpm in an Sorvall SS34 rotor. The PEG-precipitated bacteriophage was resuspended in 5 to 10 ml of suspension media [2.9 g NaCl, 1 g MgSO<sub>4</sub>, 25 ml 1 M Tris-HCl pH 7.5 and 0.5 g gelatin/500 ml] and KCl was added to a final concentration of 1 M. After a 1 hour incubation on ice, the phage proteins were spun down at 10 000 rpm for 10 min in a SS34 rotor. An equal volume of Urea-SDS [7 M Urea, 1% SDS] was added to the supernatant and this mixture was
repeatedly extracted with phenol:chloroform until there was no longer an interphase (usually 4-5 times). NaCl (final concentration of 0.15 M) was added to the aqueous phase from the final extraction and the single stranded phage DNA collected by ethanol precipitation.

2.1.viii. **DNA sequencing**

DNA was sequenced using the chain termination method developed by Sanger *et al.* (1977). The Pharmacia T7 DNA sequencing kit was utilized, following manufacturer’s instructions. The radiolabelled nucleotides employed were \([\alpha^{32}\text{P}]dCTP\) or \([\alpha^{32}\text{P}]dATP\) (New England Nuclear, 3000 Ci/mM). Good results were achieved using single stranded phage DNA and plasmid DNA isolated using all large scale and small scale preparation techniques.

2.2 **Cells and Virus**

2.2.i. **Tissue culture**

All cell lines were maintained in monolayer cultures using standard sterile tissue culture techniques. Medium (Gibco) was supplemented with fetal bovine or newborn calf serum, 2 mM L-glutamine (Gibco), 100 U/ml penicillin G (Gibco), and 100 \(\mu\)g/ml streptomycin sulphate (Gibco).
2.2.ii. **Virus culture**

HSV-1 strains were grown on Vero cells (Flow Laboratories), except when complementing cell lines were required. Stocks were routinely produced by infection of 10 to 20 175 cm² flasks at an MOI of 0.1 PFU per cell. Virus was adsorbed to cells in small volumes (1-5 mls) of serum-free medium for 2 hours before fresh, complete medium was added. After 48 to 72 hours, when there was dramatic CPE, cells were collected by centrifugation (1000 X g, 10 min, 4 °C). The cell pellet, resuspended in 1 ml of medium/plate harvested, was burst by sonication. Supernatant was cleared of cell debris by centrifugation (2000 X g, 10 min, 4 °C) and stored in 0.5 ml aliquots at -70 °C. Viruses were titred by serial dilution, and spread by secondary infection was inhibited by the addition of 0.05% human immune serum (Connaught Laboratories) after the adsorption period. Viral titres varied between strains, but were generally 0.5 to 1.0 \( \times \) 10⁹ PFU per ml.

Ad5 strains were propagated on 293 cells (Graham et al., 1977), except when a complementing cell line was indicated. Stocks were generated by infection of semi-confluent 10 to 20 175 cm² flasks at an MOI of 0.01 PFU per cell. Several days (3-4) were usually required for total CPE, at which point cells were collected by centrifugation (1000 X g, 10 min, 4 °C) and resuspended in 1 ml of culture medium per flask of infected cells harvested. Ad5-infected cells were lysed by sonication and after stocks were clarified of cell debris glycerol was added to a final concentration of 10%. Alternatively, virus was purified by banding on a cesium chloride gradient as described by Graham and Prevec (1991). Banded virus was
dialysed extensively at 4 °C with PBS containing 10% glycerol. Ad5 stocks were stored in 0.2 ml aliquots at -70 °C. Viruses were titred by serial dilution, and overlayed with media solidified with 0.5% agar. Titres were routinely 1 to 5 X 10^{10} PFU per ml.

2.2.iii. Construction of recombinant viruses

Recombinant HSV-1 strains were generated by standard techniques (Smiley, 1980; Smiley et al., 1981). Cellular sequences were introduced into a number of locations in the HSV-1 genome by in vivo homologous recombination between infectious HSV-1 DNA and plasmids bearing cellular genes flanked by viral sequences.

A variety of selection techniques, determined by the site of insertion, were employed to isolate viral recombinants. Insertions into the tk locus were selected by plaque purification in the presence of 20 μm acycloguanosine (Sigma). Insertions into the gI locus were selected on the basis of recovery of the linked, essential gD gene (Ligas and Johnson, 1988; Salloukh et al., 1993). Insertions in the gC locus were identified by plaque hybridization (Homa et al., 1986), a general technique that can be used to identify any virus that has taken up foreign DNA sequences.

The plaque hybridization technique involves transferring infected cell monolayers onto nitrocellulose filters and screening the filters for sequences of interest by hybridization with radiolabelled probes. Confluent 150 cm² dishes of Vero cells were infected with 100 to 500 PFU of a transfection yield. After well spaced plaques began to form, medium was removed and cells were gently covered with an agarose overlay [200 ml 2X
Hanks medium (Gibco) supplemented with 12 ml sodium bicarbonate (Gibco), 4 ml of 20 mM L-glutamine (Gibco), 4 ml of 1000 U/ml penicillin G (Gibco) and 1000 mg/ml streptomycin sulphate (Gibco), 4 ml BME amino acids (Gibco), 4 ml vitamins (Gibco), 40 ml fetal bovine serum, 22 ml 1% protamine sulphate (filter sterilized; Sigma) and mixed with 290 ml of sterilized 1.5% agarose. Plates were left at 37 °C for 24 to 48 hours to allow plaques to enlarge. The plates were then inverted and virus infected cells, which were no longer strongly adherent, were removed with the overlay. The inverted overlay was stored at 4 °C; under these conditions virus remained viable for at least one week. All manipulations were carried out in a laminar flow hood and inverted overlays were tightly packaged in plastic wrap to ensure that the virus remained sterile.

Sterile, notched nitrocellulose filters (sterilized by autoclaving in water, then stored in PBS), were used to lift the cells remaining on the culture dishes. The DNA from the cells and virus which adhered to the filter was released and denatured by two 1 min treatments with a mixture of 0.5 M NaOH and 0.5 M NaCl. The filters were then treated with 1 M Tris-HCl pH 7.0, followed by a treatment with 20X SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0). DNA was noncovalently linked to the filters by baking at 80 °C for 2 hours. Filters were subject to standard hybridization techniques (Maniatis et al., 1982) to identify plaques which had taken up the insert of interest. The autoradiographic exposure was aligned to inverted overlay by matching the positions of notches labelled on the tissue culture plate to the notches on the nitrocellulose filters, and plugs of infected cells and overlay were removed at locations indicated to contain a recombinant virus. This
plug was used to inoculate cells and propagate virus, which was subject to a second screening by plaque hybridization. Stocks which were sufficiently enriched for recombinant viruses were selected for plaque purification.

2.2.iv Screening of recombinant viruses by Southern blot hybridization

Recombinant viruses were considered to be pure following 2 rounds of plaque purification without any wild type contamination detectable by Southern blot analysis (Southern, 1975). DNA for Southern blots was isolated by the method of Hirt (1967) or following treatment of infected cell pellets with proteinase K (Sigma) as outlined by Sambrook et al. (1989). DNA fragments, generated by cleavage with appropriate restriction enzymes and resolved on agarose gels, were transferred to nitrocellulose using standard techniques (Maniatis et al., 1982). Briefly, DNA contained in the agarose matrix was subject to acid-catalysed hydrolysis and depurination, followed with denaturation by base, and equilibration to neutral pH and high salt. Transfer was carried out by capillary action in 20X SSC. DNA was immobilized on filters by baking at 80 °C for 2 hours or UV crosslinking using a Stratalinker (Stratagene). Sequences of interest were detected by specific hybridization to radiolabelled probes. All buffers and washing solutions used for Southern blotting are described in Sambrook et al. (1989).

Radiolabelled probes were generated by either nick translation (Rigby et al., 1977) or random priming. Both these techniques use the Klenow fragment of E. coli DNA polymerase I to incorporate [α-32P]dNTP into newly synthesized DNA. Random priming generally produced probes
of higher specific activity and can therefore be considered the method of choice. For nick translation 500 ng of plasmid DNA is incubated at 16 °C for 1 hour in a 50 ml volume containing 10 ml of [α-32P]dNTP (New England Nuclear, 3000 Ci/mM), 0.02 mM each of the three remaining dNTP's, 1X nick translation buffer (0.05 M Tris-HCl pH 7.2, 0.01 M MgSO4, 0.1 mM dithiothreitol, 50 μg/ml BSA), 0.5 μl DNAase I (Sigma; 1 mg/ml made up in 0.15 M NaCl, 50% glycerol), and 5 μl of the Klenow fragment of *E. coli* DNA polymerase I. For random priming, 50 to 200 ng of plasmid DNA is denatured by boiling in a 40 μl volume containing 2 μl of pd(N)₆ primers (Pharmacia), then quickly cooled to 4 °C on ice. The polymerization reaction is initiated by adding 6 μl 10X nick translation buffer (or Klenow buffer sold with the enzyme by most suppliers), 5 μl of [α-32P]dNTP (New England Nuclear, 3000 Ci/mM), 3 μl of 1 mM stocks of the remaining dNTPs and 2 ml of Klenow fragment. The reaction mixture was incubated at room temperature for 2 to 48 hours. Unincorporated dNTP's were removed from using a G-50 Sephadex (Pharmacia) spun column assembled following the procedure described by Sambrook *et al.* (1989).

### 2.3 Analysis of gene expression

#### 2.3.i. Isolation of RNA

Total cytoplasmic RNA was isolated by the method of Berk and Sharp (1977). RNA was quantitated by UV absorbance, using a Beckman DU-7 spectrophotometer. The OD₂₆₀ was used to calculate concentration of
diluted RNA samples (1 OD$_{260}$ = 42 µg RNA/ml). All glassware used in extraction of RNA was baked for at least 4 hours at 250 °C. Solutions used in this procedure, with the exception of Tris-HCl, were treated with 0.1% diethylpyrocarbamate (Sigma) for at least 12 hours, followed by autoclaving. Tris-HCl was made up in diethylpyrocarbamate treated ddH$_2$O, and autoclaved a second time. All solutions were handled with disposable borosilicate serologic pipettes (Corning).

2.3.ii. **Primer extension analysis**

Primer extension analysis was carried out as described previously (Smiley et al., 1987). Synthetic oligonucleotide primers were purchased from the Central Facility of the Institute for Molecular Biology and Biotechnology, McMaster University. Oligonucleotides were 5'-end labeled with [$\gamma$-$^{32}$P]ATP (New England Nuclear, 3000 Ci/mM) using T4 polynucleotide kinase following the procedure of Maxam and Gilbert (Maxam and Gilbert, 1977) or using reagents and protocols supplied by the manufacturer.

2.3.iii. **S1 nuclease analysis**

DNA probes for S1 mapping were radiolabeled by 5'-end labeling with [$\gamma$-$^{32}$P]ATP (New England Nuclear, 3000 Ci/mM) using T4 polynucleotide kinase following the procedure of Maxam and Gilbert (1977) or 3'-end labeling with the appropriate [$\alpha$-$^{32}$P]dNTP (New England Nuclear, 3000 Ci/mM) using the Klenow fragment of *E. Coli* DNA polymerase 1 as
described in Maniatis et al. (1982). Single stranded probes were prepared by electrophoretic separation as described by Maxam and Gilbert (1977). The procedure for S1 mapping has been detailed previously (Smiley et al., 1987).

2.3.iv. **Nuclear run-on assays**

Nuclei for run-on transcription assays were isolated on sucrose cushions. Cells (2-4 X 10^7) were pelleted by centrifugation at 1000 X g for 10 min at 4 °C. The cell pellet was resuspended in 1 ml of buffer I (0.32 M sucrose, 3.0 mM CaCl₂, 2.0 mM magnesium acetate, 0.1 mM EDTA, 0.1% Triton X-100, 1.0 mM dithiothreitol, 10 mM Tris-HCl pH 8.0, 40 U/ml RNAsin (Promega)) and transferred to a Dounce homogenizer. The cells were broken with 8 to 10 strokes with a tight fitting pestle (B). The homogenate was diluted with 2 ml of buffer II (1.8 to 2.0 M sucrose [the appropriate concentration of sucrose varies and must be determined for each new cell type]), 5.0 mM magnesium acetate, 0.1 mM EDTA, 5.0 mM dithiothreitol, 10 mM Tris-HCl pH 8.0, 40 U/ml RNAsin). Buffers I and II were freshly prepared each time nuclei were isolated, and were always stored on ice. The 3 ml of diluted homogenate were carefully layered over 1.8 ml of buffer II in a 5 ml polyallomer tube (Beckman), and spun at 37 000 rpm (130 000 X g) for 45 min in a precooled Sw50.1 rotor (Beckman). The pelleted nuclei were resuspended in 100 μl of nuclei storage buffer (25% glycerol, 5.0 mM magnesium acetate, 0.1 mM EDTA, 5.0 mM dithiothreitol, 50 mM Tris-HCl pH 8.0, 400 U/ml RNAsin) and immediately frozen and stored in liquid nitrogen.
In preparation for run-on transcription reactions, nuclei were thawed, and a 10 µl diluted in 500 µl of PBS and the concentration of nuclei determined by counting in a hemacytometer. All preparations were diluted to the same concentration, usually 1 X 10^7 nuclei per 100 µl, using nuclei storage buffer. Run-on transcription reactions were carried out in a mixture consisting of 100 µl of nuclei, 30 µl of NTP mix (3.33 mM ATP, 3.33 mM CTP, 3.33 mM GTP), 40 µl of 0.6 M KCl-12.5 mM magnesium acetate, 25 µl [α-32P]UTP (New England Nuclear, 3000 Ci/mM) and 5 µl dd H2O. Parallel reactions were carried out with α-amanitin (Sigma). Nuclei were incubated at 30 °C for 35 min to allow transcription to proceed. E. coli tRNA (final concentration of 100 µg/ml) and 5 µg pancreatic DNAase I (Boehringer Mannheim) were then added to the nuclei, and this mixture was incubated at 37 °C for 10 min. This was followed by protease treatment with proteinase K (Sigma, final concentration of 300 µg/ml) in 0.5 % SDS (final concentration) for 20 min at 37 °C. A 25 µl aliquot was removed and these aliquots were assayed for incorporation of radioactive isotope by determining Cerenkov counts of acid insoluble material. An equal volume of 10% trichloroacetic acid/60 mM sodium pyrophosphate was added to the 25 µl aliquot and incubated on ice for 30 min. Acid insoluble material was precipitated onto nitrocellulose filters and Cerenkov counts on the filters were measured in a scintillation counter (Beckman).

An equal volume of Urea/SDS (7 M Urea, 1% SDS) was added to the remainder of the nuclei, and this mixture was repeatedly extracted with phenol:chloroform until there was no material at the interface. The aqueous layer was collected and an equal volume of 10% trichloroacetic
acid/60 mM sodium pyrophosphate was added, and this mixture was incubated on ice for 30 min. Acid insoluble material was collected by precipitation on nitrocellulose filters and washed extensively with cold 5% trichloroacetic acid/30 mM sodium pyrophosphate, followed by a final wash with 95% ethanol. The precipitate was eluted from nitrocellulose filters by two successive incubations in 2 ml of RNA elution buffer (1% SDS, 10 mM Tris-HCl pH 7.5, 5 mM EDTA) for 10 min at 65 °C. The pooled eluate was extracted with phenol:chloroform twice. The aqueous phase was recovered, NaCl added to a final concentration of 0.15 M, and the RNA was precipitated with ethanol. Cerenkov counts of each sample was determined using a scintillation counter (Beckman) and typically 1 X 10^8 cpm were incorporated per 10^7 nuclei. The cpm per nuclei in the 25 μl aliquot were used to calculate the expected number of counts in the corresponding processed sample. When the observed and expected counts were compared it was clear that elution from the nitrocellulose filters was variable between samples. The percentage of sample loss due to processing was calculated for each sample, and all samples were normalized to the sample which showed the greatest disparity between observed and expected counts: thus if only 50% of one sample was recovered, then the counts equivalent to a 50% recovery of all the other samples were used for further analysis.

Specific run-on transcripts were detected by hybridization to DNA immobilized on nitrocellulose filters. Single stranded DNA (M13 phage or denatured plasmid) was applied to filters soaked with 5X SSC using a slot blot apparatus (Bethesda Research Laboratories), and immobilized by UV crosslinking using a Stratalinker (Stratagene). Generally 10 μg of DNA
was transferred to each slot in a 100 μl volume of 5X SSC, containing 0.01% bromophenol blue (Sigma), to mark the location of the slots. Filters were incubated at 55 °C for 6 hours in small volumes (400 to 600 μl) of hybridization buffer (5X SSC, 0.1% SDS, 1 mM EDTA, 10 mM Tris-HCl pH 7.5, 2X Denhardt's solution, 50% formamide, 250 mg/ml salmon sperm DNA, 250 mg/ml E. Coli tRNA). Radiolabeled RNA, resuspended in small volume (10 μl to 25 μl) of dd H₂O, was added to the prehybridization solution and the blots were hybridized at 55 °C for 36 to 48 hours in a shaker waterbath. Following hybridization filters were washed extensively in 2X SSC/0.1% SDS at 65 °C, and then treated with 10 μg/ml of RNAase (Sigma) for 30 min at 37 °C, with gentle agitation. Filters were exposed to XAR5 film (Kodak) and overnight exposures were generally adequate to visualize signals from actively transcribed genes, such as viral genes.

2.3.iv. **Sequencing of RNA using reverse transcriptase**

5.2), washed with 70%, then 95% ethanol, and dried under vacuum for 1 hour. The sample was redissolved in 8 μl of TE (10 mM Tris-HCl pH 7.9, 1 mM EDTA) containing 0.25 ng of labeled primer, followed by the addition of 2 μl of TKE (10 mM Tris-HCl pH 7.9, 1.25 M KCl, 1 mM EDTA). The mixture was incubated at 62 °C-63 °C for one hour, then cooled to room temperature. An additional 45 μl of TE/TKE (4:1) was added, and the sample divided into five 10 μl aliquots. Each sample was incubated with 10 U of avian myeloblastosis virus reverse transcriptase (Joseph Beard Life Sciences) in 25 μl of reverse transcriptase buffer (20 mM Tris-HCl pH 8.7, 0.33 mM of
each dNTP, 5 mM dithiothreitol, 10 μg/ml actinomycin D (Sigma), 10 mM MgCl₂, or reverse transcriptase-sequencing buffer (20 mM Tris-HCl pH 8.7, 0.33 mM of each three dNTP’s, 0.10 mM of the fourth dNTP, 0.2 mM of the equivalent ddNTP, 5 mM dithiothreitol, 10 μg/ml actinomycin D (Sigma), 10 mM MgCl₂) for 60 min at 37 °C. The volume of the samples was increased with 75 μl of TE, followed by extraction once with phenol:chloroform. The aqueous phase was removed, sodium acetate was added to a final concentration of 0.3 M, after which 10 μg E. Coli tRNA and 400 μl of ethanol were added. The cDNAs were precipitated and resolved on 8% sequencing gels.
Chapter 3: Regulation of cellular genes transduced by herpes simplex virus
Regulation of Cellular Genes Transduced by Herpes Simplex Virus

BARBARA PANNING AND JAMES R. SMILEY*

Molecular Virology and Immunology Program, Pathology Department, McMaster University,
Hamilton, Ontario, Canada L8N 3Z5

Received 21 October 1988/Accepted 27 January 1989

Previous studies demonstrated that the rabbit β-globin gene is transcribed from its own promoter and regulated as a herpes simplex virus (HSV) early gene following insertion into the early HSV thymidine kinase gene in the intact viral genome (J. R. Smiley, C. Smiberti, and R. D. Everett, J. Virol. 61:2368–2377, 1987). We report here that the β-globin promoter remained under early control after insertion into the late HSV gene encoding glycoprotein C. On the basis of these findings, we concluded that the β-globin promoter is functionally equivalent to an HSV early-control region. We found that a transduced human α-globin gene was also regulated as an early HSV gene, while two linked Alu elements mimicked the behavior of HSV late genes. These results demonstrate that certain aspects of HSV temporal regulation can be duplicated by cellular elements and provide strong support for the hypothesis that the regulation of HSV gene expression can occur through mechanisms that do not rely on recognition of virus-specific temporal control signals.

The 70 herpes simplex virus (HSV) genes are transcribed by RNA polymerase II in a regulatory cascade driven by viral products (33). Five immediate-early (IE) genes are expressed first (1, 6, 43, 51, 68), and four of the IE polypeptides play crucial roles in activating transcription of the remaining early (E) and late (L) genes (8, 9, 16, 18, 24, 47, 49, 50, 53–55, 61, 66; reviewed in reference 19). E genes are maximally expressed before the onset of viral DNA replication, while two subclasses of L genes require DNA replication for high-level expression. Promoter transept experiments have shown that the temporal regulation of individual HSV genes during infection is dictated mainly by sequence-dependent elements present in their respective promoter regions (31, 48, 57), and nuclear run-on transcription assays suggest that this control occurs largely at the transcriptional level (25, 69).

The detailed mechanisms of action of the HSV IE proteins remain unknown. Although the IE polypeptide ICP4 binds directly to specific sequences present in HSV and some heterologous DNAs (2, 20, 21, 38, 45, 46), the role of sequence-specific DNA binding in the transcription activation mediated by this polypeptide remains unclear; for example, an ICP4-binding site located in the upstream region of the glycoprotein D gene (2, 20) does not appear to be required for transcription of this gene by ICP4 (14, 15). Extensive studies of several HSV E and L promoters have indicated that many of the cis-acting regulatory sequences required for activation by IE proteins and temporal regulation during infection correspond to the binding sites of cellular transcription factors (7, 12, 13, 15, 31, 32, 35, 56). These results suggest that the temporal control of HSV E and L genes relies at least in part on changes in the activity of cellular transcription factors that recognize distinctive constellations of binding sites in E and L promoters.

An independent line of evidence supporting this view comes from studies of the regulation of cellular promoters by HSV products. The rabbit β-globin gene is activated by HSV IE polypeptides when it is newly introduced into fibroblasts by transfection (15–17) or as part of an infecting HSV genome (60). In the latter case, β-globin is regulated as an HSV E gene following insertion into the E thymidine kinase gene (tk). A straightforward interpretation of these results is that the β-globin control region is functionally equivalent to a bona fide HSV E promoter. An alternative explanation is that cellular genes resident in the viral genome are regulated by HSV-specific temporal control signals present in the flanking viral DNA sequences. According to this hypothesis, β-globin is controlled as an E gene following insertion into the tk locus because it falls under the influence of putative HSV-specific E-regulatory signals that govern tk gene expression.

The hypothesis that the β-globin promoter is equivalent to an HSV E-control region predicts that its regulation does not depend on the temporal class of the viral gene into which it is inserted. We tested this prediction by inserting the β-globin gene into the HSV L gene encoding glycoprotein C (gC) and found that the β-globin promoter remained under E control in this situation. From these results, we concluded that the temporal regulation of β-globin expression by HSV products depends on features of the β-globin promoter rather than on the nature of the flanking viral sequences.

The finding that the β-globin promoter was regulated as an HSV E promoter in a context-independent fashion prompted us to study the control of several additional cellular genes following incorporation into the HSV genome. We found that the human α-globin gene was also regulated as an HSV E gene, while two linked Alu elements closely mimicked the behavior of HSV L genes. These results demonstrate that a variety of cellular promoters are able to function efficiently in the context of the HSV genome and lend further support to the hypothesis that temporal regulation during HSV infection can occur through mechanisms that do not involve recognition of virus-specific signals.

MATERIALS AND METHODS

Viruses and cells. HSV type 1 (HSV-1) strain KOS PAA5 (27) was used throughout this study. Virus stocks were propagated and titers were determined on monolayers of Vero cells. Expression of virally transduced cellular genes was monitored following infection of Vero or BHK21 cells with a multiplicity of 10 PFU per cell. When indicated, cycloheximide (100 μg/ml) or aphidicolin (10 μg/ml) was added 30 min prior to infection and maintained continuously.

Construction of recombinant viral strains. Viral strains
bearing inserts of cellular sequences were derived by in vivo recombination following cotransfection of Vero cells with KOS PAA5 DNA and plasmids bearing the desired insertion mutation, as previously described (58-60).

Strain gC-beta, bearing the rabbit β-globin gene inserted into the HSV gene encoding glycoprotein C, was constructed by converting a previously described 3.7-kilobase (kb) globin-bearing SstI fragment (60) into an XbaI fragment, then inserting this into the XbaI site within gC-coding sequences (22) present on pHindIII L (Fig. 1; pHindIII L was provided by E. K. Wagner). Globin-bearing viral clones were identified by plaque hybridization (32), plaque purified, and then screened for the desired insertion by Southern blot hybridization.

Strain tk-alpha, bearing the human αβ globin gene inserted into the viral gene encoding thymidine kinase (tk), was constructed by inserting a 4.3-kb αβ globin SstI fragment (provided by A. Bernstein) into the SstI site located within the tk coding sequences present on pTK173 (65). tk-deficient viral recombinants were selected by plaque purification on Vero cells in the presence of 20 μM acycloguanosine.

Primer extension and S1 nuclease analysis. Cytoplasmic RNA was extracted by the method of Berk and Sharp (3). Primer extension and S1 nuclease analysis were performed exactly as previously described (60).

The following synthetic primers were purchased from the Central Facility of the Institute for Molecular Biology and Biotechnology, McMaster University: (i) gC, 5'-AAACGACCTCCACACGGCCACCCGG3'; predicted extension product of 79 nucleotides (nt) (22, 32); (ii) USII, 5'-GAGGCCGCTTTGACGGTTCTTTC3', predicted extension product of ca. 80 nt (36); (iii) glycoprotein D (gD), 5'-CCCCATACGGGAAACGACCAACCCAA-3'; predicted extension product of ca. 80 to 90 nt (67); (iv) α2 globin, 5'-AGGCCGCTTTGACGGTTCTTTC3', predicted extension product of 80 nt (41); (v) AhII, 5'-TTAGTATAACGTCCTCCCCAT-A3', predicted extension product of 120 nt (11); and (vi) AhII, 5'-TTAGTATAACGTCCTCCCCAT-A3', predicted extension product of 120 nt.

**RESULTS**

Insertion of the rabbit β-globin gene into the HSV gC locus. Previous studies demonstrated that the intact rabbit β-globin gene is transcribed from its own promoter and regulated as an HSV E gene following insertion into the E tk gene in the viral genome (60). We wished to determine whether the β-globin promoter remained under E control when the globin gene was placed within the body of a true L HSV gene. We constructed a plasmid in which a 3.7-kb XbaI fragment bearing the rabbit β-globin gene and 1.200 nt of 5' globin-flanking sequences was inserted into the XbaI site within the dispensable L gene encoding gC (22, 32), and then we transferred the resulting insertion mutation into the viral genome by in vivo recombination to produce strain gC-beta (Fig. 1). The introduction of β-globin sequences into the gC gene resulted in the replacement of a wild-type 2.2-kb HindIII-EcoRI gC fragment with the expected gC-globin fusion fragments of 3.3 and 1.9 kb (Fig. 2) and the acquisition of a 590-base-pair internalglobin EcoRI fragment (data not shown).

Expression of rabbit β-globin in strain gC-beta. We first tested whether the insertion of β-globin sequences disrupted the regulation of transcripts initiated from the gC promoter in strain gC-beta by studying the effects of inhibiting viral DNA replication with aphidicolin. Cytoplasmic RNA extracted from Vero cells infected with gC-beta and the parental PAA5 strain was analyzed by primer extension by using a probe complementary to residues +54 to +79 relative to the gC mRNA cap site. Accumulation of correctly initiated gC RNAs was strongly inhibited by blocking viral DNA replication in both viral strains (Fig. 3). We therefore concluded that expression from the gC promoter remained highly dependent on DNA replication in gC-beta.

We studied the regulation of the inserted β-globin gene by S1 nucleases analysis of globin transcripts produced during lytic infection of Vero cells. The S1 probe was derived from a previously described gD-globin fusion (14) and allowed differentiation of globin RNAs initiated from the globin
FIG. 2. Southern blot analysis of strain gC-beta DNA. The indicated DNAs were cleaved with a mixture of HindIll and EcoRI, transferred to nitrocellulose, and then probed with pHindIII L (lacking the β-globin insert). Insertion of globin sequences disrupted the 2.2-kb PAA/S HindIll-EcoRI fragment and led to the appearance of two gC-globin fusion fragments (3.3 and 1.9 kb).

promoter from those arising by readthrough from upstream sequences (Fig. 4D). Correctly initiated β-globin transcripts were detectable at 3 h postinfection and did not increase in abundance thereafter. These globin RNAs accumulated with

FIG. 3. Expression of gC RNAs in strain gC-beta. Vero cells were infected with 10 PFU of the indicated viral strain per cell. Cytoplasmic RNA (20 μg), harvested at the indicated times postinfection, was analyzed by primer extension with a 5' labeled synthetic gC primer. Following treatment with reverse transcriptase, products were displayed on an 8% sequencing gel. Where indicated, 10 μg of aphidicolin (Aph) per ml was added 30 min prior to infection and maintained continuously. Markers (M) were 3' labeled HpaII fragments of pBR322 DNA. Marker fragment sizes (in nucleotides) are shown at the right.

FIG. 4. Time course of β-globin RNA expression during infection with gC-beta. Cytoplasmic RNA (20 μg), prepared from Vero cells infected with gC-beta (10 PFU per cell), was hybridized to the 5' labeled probe diagrammed in panel D to detect globin transcripts. Following treatment with S1 nuclease, digestion products were displayed on an 8% sequencing gel. These RNA samples were also analyzed for gD and US11 transcripts by primer extension with 5' labeled synthetic 25-mers. (A) Time course of β-globin RNA accumulation. AT1 and AT2. Aberrant transcripts (described in the text). A portion of the globin probe was also hybridized to 1 ng of purified rabbit globin mRNA (alpha plus beta) to provide a marker for the position of correctly initiated globin RNAs. Markers (M) for panels A to C were 3' labeled HpaII fragments of pBR322 DNA. Marker fragment sizes (in nucleotides) are shown at the right. (B) Time course of gD RNA accumulation. (C) Time course of US11 RNA accumulation. (D) Probe fragment used in panel A. The fragment was derived from a gD-globin fusion (14) and allows transcripts initiated from the globin promoter to be distinguished from those arising by splicing (AT1) and readthrough (AT2) from upstream sequences.
replication with aphidicolin (Fig. 5). These data suggest that globin expression required viral IE polypeptide synthesis and was largely independent of DNA replication. On the basis of these results, we concluded that the rabbit β-globin gene remained under E control when it was embedded within a viral L gene.

Two additional globin-related transcripts, AT1 and AT2 (Fig. 4 and 5), accumulated at later times postinfection. One of these, AT2, gave rise to an S1 signal mapping to the site of sequence divergence between gC-beta DNA and the probe used and must therefore arise by readthrough from the upstream globin sequences. The second (AT1) generated an S1 signal mapping to a previously described cryptic splice-acceptor site located at approximately +46 in the globin-coding sequences (26, 60) and most likely originates through alternative splicing of RNAs initiated at one or more upstream promoters. It seems likely that some of the RNAs that give rise to the AT1 and AT2 S1 signals initiate at the gC promoter. However, we found that the AT1 and AT2 S1 signals were considerably less sensitive to inhibition of viral DNA replication than transcripts driven from the gC promoter (compare Fig. 3 and 5). We therefore suspect that transcripts initiated at one or more additional promoters, perhaps located upstream of the gC promoter or within the 5' flanking globin sequences, also contributed to the AT1 and AT2 S1 signals.

Insertion of the human α2 globin gene and two Alu elements into the tk locus. Tackney et al. (62) reported that the intact Chinese hamster adenine phosphoribosyltransferase (aprt) gene was not detectably transcribed following incorporation into the HSV genome and interpreted these results as indicating that HSV regulators distinguish between cellular and viral promoters. Because of the contrasting behavior of the rabbit β-globin and hamster aprt genes, we wished to learn whether the β-globin gene was unique among cellular elements in its ability to be expressed to high levels when transduced by HSV. We therefore studied the regulation of the human α2 globin gene and two linked Alu elements present on a 4.3-kb SstI fragment following insertion into the viral tk gene (Fig. 6). These particular cellular elements were

FILE. 5. Effects of inhibitors on β-globin expression. Cytoplasmic RNA (20 μg), prepared from Vero cells at the indicated times postinfection with gC-beta (10 PFU per cell), was hybridized to the β-globin probe described in the legend to Fig. 4, and the hybrids were digested with S1 nuclease. Cycloheximide (C): 100 μg/ml or aphidicolin (Aph): 10 μg/ml was added 30 min prior to infection. Markers (M) were 3'T-labeled HpaII fragments of pBR322 DNA. Marker fragment sizes (in nucleotides) are shown at the right.

FIG. 6. Structure of strain tk-alpha. A 4.3-kb SstI fragment of human DNA bearing the α2-globin gene and two Alu repetitive elements was cloned into the SstI site within tk coding sequences, and the resulting tk-deficient insertion mutation was transferred into the tk locus in the viral genome by recombination in vivo. The intron-exon arrangement of the α-globin gene and the transcriptional polarities of the Alu elements are indicated. RI. EcoRI.
chosen for two reasons. First, the α- and β-globin promoter regions have little primary nucleotide sequence homology, and these two genes are regulated very differently in transfection assays (5, 34, 63). Thus, if the α-globin gene was also activated by HSV products, the result would reduce the likelihood that this control results from recognition of "virus-specific" signals accidentally present in globin DNA. Second, Alu elements are transcribed by RNA polymerase III in vitro (10, 11: reviewed in reference 52), and we wished to learn whether certain cellular pol III-transcribed genes can also be expressed at high levels following transduction by HSV.

Insertion of the 4.3-kb α-globin fragment into the tk gene of strain tk-alpha resulted in loss of the wild-type 3.5-kb BamHI and 2.4-kb EcoRI tk fragments and acquisition of the predicted 7.8-kb BamHI and 6.7-kb EcoRI fragments bearing the globin insert (Fig. 7).

E expression of the α-globin gene. We studied the regulation of the inserted α-globin and Alu transcription units by primer extension analysis of cytoplasmic RNAs produced during lytic infection of Syrian hamster BHK21 cells with tk-alpha. BHK21 cells were chosen instead of Vero cells to reduce the risk of cross-hybridization between primers designed to detect transcripts arising from the virally transduced human genes and the closely related endogenous primate sequences present in Vero cells. Control experiments demonstrated that RNA prepared from PAA's-infected BHK21 cells did not react with the α-globin and Alu primers (data not shown).

Correctly initiated α-globin transcripts accumulated with an early time course: transcripts were detected at 3 h postinfection, reached maximal levels by 6 h, and remained relatively constant in abundance thereafter (Fig. 8). In addition, α-globin expression was blocked by inhibiting protein synthesis with cycloheximide but was not greatly affected by suppressing viral DNA replication with aphidicolin (Fig. 9). Similar results were obtained during infection of Vero cells with strain tk-alpha (data not shown).

I. expression of Alu elements. The two Alu elements transduced by strain tk-alpha differ in size: AluI is a standard human monomer, while AluII is a dimer composed of two fused Alu elements. In addition, AluI and AluII differ significantly in their primary sequences, necessitating the use of separate primers to detect their respective transcripts. The primers were complementary to residues 95 to 120 of the Alu elements and were designed to prevent cross-hybridization to the related 7SL RNA (64).

Both Alu elements gave rise to abundant cytoplasmic transcripts initiated at the first residue of the Alu repeat, i.e., the initiation site of RNA polymerase III in vitro (Fig. 8 and 9) (11). Alu transcripts were first detected 6 h postinfection, and the levels of Alu RNA increased at later times. Accumulation of Alu transcripts was completely suppressed by blocking DNA replication with aphidicolin and by inhibiting protein synthesis with cycloheximide (Fig. 9). In these respects, the Alu elements were regulated in a fashion that closely mimics the behavior of HSV true L genes.

RNA polymerase III terminates transcription immediately following a run of four or more T residues in the non-template strand (4, 11, 40). As an indirect test of whether the Alu transcripts arising from the virally transduced elements were transcribed by polymerase III, we mapped the 3' end of the AluI transcript by S1 nuclease protection analysis. Using a 3'-labeled Aval-Ncol probe fragment labeled at an Aval site ca. 100 nt upstream of the 3' end of the element, we detected a protected fragment of ca. 162 nt (Fig. 10). This result maps the 3' end of the AluI transcript within a run of 6 T residues (Fig. 10C), which corresponds to the first run of four or more T residues in the downstream flanking human sequences. Thus, the position of the 3' end of this AluI transcript provides indirect evidence that it is transcribed by RNA polymerase III.
FIG. 9. Effects of inhibitors on expression of α-globin and Alu RNAs. Cytoplasmic RNA (20 µg), prepared at the indicated times (hours) postinfection of BHK21 cells with tk-alpha (10 PFU per cell), was analyzed for α-globin and Alu transcripts by primer extension. Cycloheximide (Cox; 100 µg/ml) or Aphidicolin (Aph; 10 µg/ml) was added 30 min prior to infection. Markers (M) were 3'-labeled HpaII fragments of pBR322 DNA. Marker fragment sizes (in nucleotides) are shown at the left.

DISCUSSION

Previous studies demonstrated that the rabbit beta-globin gene was transcribed from its own promoter and regulated as an HSV E gene following insertion into the E tk gene in the intact viral genome (60). The results presented in this paper show that the globin promoter remained under E control when it was inserted into the body of the true L viral gene encoding gC. Thus, in this instance, the regulation of a cellular gene residing in the HSV genome was not dictated by the temporal class of the viral gene into which it was inserted. These data strongly suggest that the β-globin control region provides the functional equivalent of an HSV E promoter and demonstrate that E regulation can occur through mechanisms that are not restricted to viral promoters. Consistent with this view, we found that the highly diverged human α-globin gene was also expressed under E control in a viral recombinant. This latter finding reduces the likelihood that the regulation of globin genes by HSV products relies on recognition of virus-specific temporal control sequences accidentally present in the upstream regions or transcribed bodies of these genes. Rather, it seems much more likely that this control results at least in part from virus-induced modifications that facilitate the interaction of one or more cellular factors with the globin control regions. Interpreted in this way, our data support the hypothesis that HSV regulates alter the activity of cellular transcription factors (44) in a fashion similar to that of the adenovirus Eia proteins (reviewed in reference 37).

It is intriguing that the expression of an HSV-transduced α-globin gene required viral polypeptide synthesis. In contrast to β-globin genes, the human α-globin gene is efficiently expressed following transfection into a variety of cell types (5, 34, 63). Thus, one might have anticipated that the α-globin gene would be directly expressed upon infection in the absence of viral regulators. Indeed, Hearing and Shenk (28) found that this gene was expressed in the absence of IE polypeptides during infection with an Eia-deficient adenovirus recombinant. One interpretation of our data is that the α-globin gene is somehow prevented from interacting with the cellular transcriptional apparatus when it is placed in the HSV genome and that one or more viral products are required to overcome this negative control. If this idea is correct, then it seems possible that similar mechanisms contribute to the severe restriction of viral E and L gene expression that is observed in the absence of HSV IE polypeptides.

Expression of the Alu elements transduced by strain tk-alpha stringently required viral protein and DNA synthesis—in these respects mimicking the behavior of viral true L genes. It is not yet clear whether the requirement for protein synthesis reflects a direct effect of viral polypeptides on Alu expression: an equally plausible explanation is that Alu expression is driven entirely by viral DNA replication. One hypothesis to explain the requirement for viral DNA replication postulates that Alu promoters are very weak in vivo and that detectable expression therefore requires template amplification. Alternative explanations include (i) the generation of a transcriptionally permissive, altered DNA conformation during replication and (ii) a replication-dependent segregation of DNA molecules into specialized nuclear compartments containing the necessary transcription factors.

Epstein-Barr virus encodes small polIII-transcribed RNAs (42), but HSV is not known to bear polIII genes. While we have no direct evidence that the virally transduced Alu elements are transcribed by RNA polymerase III during infection, our observation that the Alu transcripts start at the predicted polIII initiation site and end at a classical polIII terminator provides indirect evidence that this is the case. Adenovirus and pseudorabies virus IE proteins can activate the transcription of certain polIII genes that have been newly introduced into cells (23, 30) by modifying the activity of the polIII transcription factor IIIC (29, 70). We are testing whether HSV IE products are also able to directly stimulate transcription by polIII.

Our results strongly suggest that some cellular cis-acting elements are able to specify control patterns that closely resemble those of HSV E and L genes. This in turn suggests that certain aspects of the temporal regulation of HSV genes
REGULATION OF CELLULAR GENES BY HSV 1935


47. O'Hare, P., and G. S. Hayward. 1985. Evidence for a direct role for both the 175,000 and 110,000 molecular weight immediately-early proteins of herpes simplex virus in the transcription of delayed-early promoters. J. Virol. 53:751–760.
49. Presten, C. M. 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK. J. Virol. 29:275–284.
56. Shapira, M., F. L. Homa, J. C. Glorioso, and M. Levine. 1987. Regulation of the herpes simplex virus type 1 late (Y2) glycoprotein C gene: sequence between base pairs −34 to +29 controls translational expression and responsiveness to transcription by the products of immediate-early (α) and 0 genes. Nucleic Acids Res. 15:3097–3111.

3.2 **Segue**

In this paper we demonstrated that two cellular genes, human α-globin and rabbit β-globin, were expressed with early kinetics when rescued into the genome of HSV-1. Smiley et al. (1987) had previously demonstrated that the rabbit β-globin gene was also expressed with early kinetics when inserted into the viral tk locus and since tk is an early gene, it was possible that elements in viral flanking sequences may have determined the kinetics of expression of this cellular gene. The results presented in this paper demonstrate that the rabbit β-globin was also expressed with early kinetics when present in the late gc locus. Thus temporal regulation of β-globin is not determined by the kinetic class of viral gene into which it was inserted, indicating that the β-globin promoter is equivalent to a viral early promoter. The human α-globin gene, inserted into the tk locus was also expressed as an early gene, indicating that α-globin cis-regulatory sequences, like those of β-globin, are functionally equivalent to a viral early promoter. The likelihood that both globin genes serendipitously contain virus specific cis-regulatory sequences is small, suggesting that temporal control was mediated by cellular promoter elements. The subsequent demonstration that ICP4 was absolutely required for transcription α- and β-globin genes residing in the HSV-1 genome (Smiley and Duncan, 1992) provided more evidence that these promoters are analogous to a viral early promoter and suggests that temporal regulation of genes present in the HSV-1 genome is not mediated through virus specific regulatory motifs.
The α- and β-globin genes are normally expressed only in cells of the erythroid lineage. In transient assays the human β-globin gene (which is almost identical in sequence to the rabbit gene) is transcribed only in erythroid cells while α-globin expression is not lineage restricted, suggesting that the β-globin gene contains regulatory elements that impose requirements for tissue-specific factors (Humpheries et al., 1982; Treisman et al., 1983). In the HSV-1 genome the β-globin gene was able to overcome or bypass the requirement for erythroid-restricted factors, suggesting that erythroid-specific promoter elements are not required for viral β-globin expression. Since non-erythroid β-globin expression occurs when an SV40 enhancer is linked to the β-globin gene (Treisman et al., 1983), it is possible that there are enhancer-like sequences in viral DNA and that these putative enhancers require ICP4 in trans. This is consistent with models that postulate that ICP4 behaves as a trans-acting enhancer. It would be interesting to infect MEL cells with globin/ICP4- recombinants genes to determine whether tissue-specific transcription factors can bypass the requirement for ICP4 when the α- and β-globin genes are delivered to erythroid cells as part of a viral genome.

The recombinant strain tk-α provided another example of the transcription of virally transduced cellular sequences which normally show limited or restricted expression: two Alu repetitive elements linked to the α-globin gene were expressed at high levels from HSV-1 recombinants. The sites of transcript initiation and termination indicated that transcription was driven by Alu pol III promoters, providing the first indication that class III transcription could occur from the HSV-1 genome.
Though Alu sequences are transcribed by pol III in vitro (Fuhrman et al., 19), the majority of the nearly one million Alu pol III promoters present in HeLa cells are inactive (Paulson and Schmid, 1986; Paulson et al., 1987). It has been suggested that the majority of Alu elements are pseudogenes since intragenic Alu pol III promoters are thought to be insufficient to drive transcription in vivo (Matera et al., 1990). Thus the expression of Alu sequences from HSV-1 recombinants was the first reported instance of robust transcription of Alu sequences in vivo and the first indication that Alu class III promoters are sufficient to drive transcription in cultured cells. Since Alu repeats present in HSV-1 recombinants responded to viral transactivators we wished to determine whether expression of Alu sequences in their cellular loci could also be affected by viral factors. The following paper describes the investigation of the effects of HSV-1 infection on the expression of this unusual class of DNA elements.
Chapter 4: Activation of RNA polymerase III transcription of human Alu elements by herpes simplex virus
Activation of RNA Polymerase III transcription of Human Alu Elements by Herpes Simplex Virus

Barbara Panning and James R. Smiley*

Molecular Virology and Immunology Program,
Pathology Department,
McMaster University,
Hamilton Ontario,
CANADA L8N 3Z5

* corresponding author
(416) 525-9140 ext 2695, FAX: (416) 546-9940,
Electronic mail: Smiley@fhs.mcmaster.ca

running title: HSV-induced activation of Alu transcription

key words: Herpes simplex virus, transcription, Alu elements, RNA polymerase III.
Abstract

We found that HSV infection of HeLa cells strongly induces RNA polymerase III transcription of endogenous human Alu elements, resulting in the accumulation of high levels of cytoplasmic RNAs initiated from Alu pol III promoters. Induction required viral protein synthesis, and occurred during infection with a viral mutant bearing a null mutation in the immediate-early (IE) gene encoding ICP4, suggesting that one or more IE proteins are sufficient for activation. However, mutations in each of the other four IE genes had no effect on activation of Alu expression. We therefore conclude that HSV most likely encodes at least two proteins that are each sufficient to activate Alu transcription, and that at least one of these is an IE protein other than ICP4.
Introduction

The *Alu* family of repetitive elements consists of over 500,000 dispersed copies of a ca. 280 nucleotide sequence that makes up approximately 5% of human DNA (Deininger, 1989; Deininger and Schmid, 1979; Weiner et al, 1986). *Alu* sequences preferentially localize with segments of GC-rich DNA that contain a high density of genes (Bernardi et al, 1985; Soriano et al, 1983), and cluster in reverse bands of metaphase chromosomes (Korenberg and Rykowski, 1988). The elements are mobile (Deininger et al, 1992; Matera et al, 1990; Wallace et al, 1991), and are thought to amplify and disperse by retroposition, the process of transposition through an RNA intermediate (Deininger, 1989; Deininger et al, 1992; Jagadeeswaran et al, 1981; Rogers, 1983).

*Alu* elements have a heterodimeric structure consisting of a tandem arrangement of two related but nonidentical monomers that each appear to have arisen from the 7SL RNA gene (Ullu et al, 1982; Ullu and Tschudi, 1984; Weiner, 1980). Each *Alu* monomer bears a variable 3′ A-rich region that is thought to arise by processing of retroposition intermediates (Deininger and Schmid, 1979; Rogers, 1985; Weiner et al, 1986). The elements are flanked by short direct repeats that vary in length and sequence between elements (Van Ardsell, 1981), and the left *Alu* monomer contains a bipartite RNA polymerase III (pol III) promoter that drives transcription of the entire element (Fuhrman et al, 1981).
In vivo, Alu repeats are transcribed from nearby pol II promoters (Matera et al., 1990; Paulson et al., 1987; Paulson and Schmid, 1986), presumably because of the high density of Alu elements located in the vicinity of pol II-transcribed genes (Bernardi et al., 1985; Flemington et al., 1987; Korenberg and Rykowski, 1988; Lee et al., 1984). However, the majority of Alu elements are not expressed from their pol III promoters (Matera et al., 1990; Paulson et al., 1987; Paulson and Schmid, 1986). The related 7SL RNA gene requires upstream transcriptional control elements in addition to the internal pol III promoter for efficient transcription (Ullu and Weiner, 1985), and it has therefore been suggested that most Alu elements are inactive pseudogenes (Matera et al., 1990; Paulson and Schmid, 1986; Ullu and Weiner, 1984). According to this hypothesis, only those rare Alu elements that are located in the vicinity of an appropriate upstream activation sequence are transcriptionally competent.

Several nuclear DNA viruses encode proteins that regulate pol III transcription (Aufiero and Schneider, 1990; Carey et al., 1986; Gaynor et al (Aufiero and Schneider, 1990; Carey et al., 1986; Gaynor et al., 1985; Singh et al., 1985). For example, the adenovirus E1a protein and the Pseudorabies virus IE protein activate expression of class III genes that have been newly introduced into cells by transfection or infection (although the majority of endogenous pol III-transcribed genes are not induced) (Gaynor et al., 1985); in addition, transformation by SV40 stimulates pol III transcription of B2 sequences, a murine family of repetitive DNA elements (Carey et al., 1986; Singh et al., 1985). Adenovirus E1a and SV40 T antigen appear to activate
pol III transcription by increasing the amount or activity of the limiting transcription factor TFIIIC (Hoeffler et al., 1989; White et al., 1990; Yoshinaga et al., 1986). We have previously demonstrated that adenovirus infection activates pol III transcription of endogenous human Alu sequences in HeLa cells (Panning and Smiley, 1993). In contrast to other examples of adenovirus-induced activation of RNA pol III transcription, induction required the E1b 58 kDa protein and the products of E4 open reading frames 3 and 6 in addition to E1a. E1a appeared to play an indirect role in Alu activation, because E1a-deficient mutants induced Alu expression during infection at high input multiplicities. In addition, 293 cells which constitutively express E1a (Graham et al., 1977; Lassam et al., 1979) did not display increased levels of Alu RNAs. These results strongly suggest that E1a-induced alterations in TFIIIC are insufficient to stimulate expression of endogenous cellular Alu elements.

We have previously shown that two Alu elements derived from the human α-2-globin locus are efficiently transcribed when they are delivered into cells by a herpes simplex virus (HSV) vector (Panning and Smiley, 1989); furthermore, the sites of transcription initiation and termination suggested that these elements were transcribed by RNA pol III. These data raised the possibility that HSV regulators are also able to activate RNA polymerase III transcription of Alu elements. Consistent with this possibility, Jang and Latchman reported that HSV stimulates pol III transcription of endogenous human Alu sequences during infection of HeLa cells (Jang and Latchman, 1989). This group also reported that the
HSV immediate-early (IE) protein ICP27 is necessary and sufficient for *Alu* activation, and that ICP27 acts by increasing the activity of TFIIC in a fashion analogous to Ad E1a (Jang and Latchman, 1992). This latter conclusion is in apparent conflict with our finding that E1a is not sufficient to activate *Alu* transcription during adenovirus infection.

In this paper we confirm that HSV infection stimulates pol III transcription of endogenous human *Alu* elements, and that this activation requires de novo viral gene expression. Induction occurred when viral protein synthesis was largely restricted to IE polypeptides; however, deletions which inactivated each of the five HSV IE genes had no effect on *Alu* activation. In particular, three ICP27 null mutants displayed wild-type activity. We therefore conclude that HSV most likely encodes two or more functionally redundant proteins that each suffice to stimulate *Alu* expression, and that ICP27 is not required for induction.
Methods and Materials

Cells and Virus

With the exceptions listed below, HSV1 strains were propagated in Vero cells. The ICP4 null mutant d120 was grown on complementing E5 cells (DeLuca et al, 1985), the ICP27 mutant s 5dl1.2, d27-1, and n59R were grown on 3-3 cells (McCarthy et al, 1989; Rice et al, 1989), and the gD-deficient mutant FgDβ was grown on VD60 cells (Ligas and Johnson, 1988). Vero, E5, 3-3, and VD60 cells were maintained in α-MEM supplemented with 5% fetal calf serum. 293 cells (Graham et al, 1977) were grown in α-MEM containing 10% newborn calf serum. Human embryonic kidney cells and human adult fibroblasts (obtained from S. Bacchetti) were maintained in α-MEM plus 10% fetal calf serum.

Primer extension analysis

Cytoplasmic RNA was isolated by the method of Berk and Sharp (Berk and Sharp, 1977). Primer extension analysis was carried out using oligonucleotide primers specific for Alu, 7SL, and gD RNAs (Panning and Smiley, 1989; Panning and Smiley, 1993), as previously described (Smiley et al, 1987). All oligonucleotides were synthesized at the Central Facility for the Institute of Molecular Biology and Biotechnology, McMaster University.

Nuclear Run-on transcription assays

HeLa cells were harvested at 6 hours post-infection. Isolation of nuclei and run-on transcription in the presence of α-32P-UTP were as
previously described (Panning and Smiley, 1993). Probes for detecting Alu, 7SL, and rRNA transcripts have been previously described (Panning and Smiley, 1993): Alu 18 and Alu 19 are M13 clones that detect Alu sense and antisense transcripts respectively. HSV-1 tk RNAs were detected using an M13mp19 recombinant containing 1.3 kb of HSV-1 tk DNA. DNA was applied to nitrocellulose filters using a slot-blot apparatus and immobilized by UV crosslinking with a Stratalinker 2000 (Stratagene), following manufacturer's instructions. Slot blots were hybridized and processed as previously described (Panning and Smiley, 1993).
Results

**HSV infection induces expression of endogenous Alu elements in HeLa cells.**

We used a primer extension assay to monitor accumulation of cytoplasmic transcripts of endogenous Alu elements during lytic infection of HeLa cells with HSV1 strain KOS PAA5 (Hall et al, 1984). As previously reported (Panning and Smiley, 1993), two classes of Alu cDNAs are produced in primer extension assays using RNA extracted from uninfected cells: a heterogeneous smear that extends into the high molecular weight range of the gel, and a unique species of approximately 120 bp (figure 1A). The high molecular weight cDNAs most likely arise from Alu elements embedded in pol II transcripts, while the 120 nt product corresponds to Alu RNAs driven from the Alu pol III promoter (Panning and Smiley, 1993). The profile of Alu cDNAs was greatly altered after HSV infection: the high molecular weight cDNAs were virtually eliminated (most likely due to degradation of host mRNAs by the HSV vhs function (Fenwick, 1984; Fenwick and Walker, 1978; Kwong and Frenkel, 1987; Shek and Bacchenheimer, 1985), while the 120 nt signal increased 30-50 fold. Alu elements are homologous to the 7SL RNA gene, the presumed progenitor of Alu sequences (Ullu et al, 1982; Ullu and Tschudi, 1984; Weiner, 1980). However, HSV infection had no effect on the levels of 7SL RNA (figure 1b). This observation demonstrates that the signal observed with the Alu primer does not stem from cross-hybridization with 7SL RNAs. Taken in combination, these data indicate that HSV infection stimulates expression
FIGURE 1. Primer extension analysis of Alu RNAs produced during HSV infection. HeLa cells were infected with 10 PFU/cell with HSV1 strain KOS PAAr5, and cytoplasmic RNA harvested at the indicated times post infection was analyzed by primer extension using 5' 32P-labeled synthetic 25-mers designed to detect Alu and 7SL transcripts. The resulting cDNAs were resolved on an 8% sequencing gel. Molecular weight markers were 3' end-labelled Hpa II fragments of pBR322 DNA. Fragment sizes are indicated in nucleotides. (A) Primer extension analysis of Alu transcripts using 20 µg of cytoplasmic RNA. (B) Primer extension analysis of 7SL RNAs, using 0.5 µg of cytoplasmic RNA.
of Alu RNAs arising from the Alu pol III promoter, and are consistent with an earlier report that HSV activates pol III transcription of Alu elements (Jang and Latchman, 1989).

**Induction requires de novo viral gene expression.**

HSV virions contain several proteins that modulate viral and cellular gene expression in the absence of de novo viral protein synthesis (Batterson and Roizman, 1983; Campbell et al., 1984; Fenwick and Walker, 1978; Nishioka and Silverstein, 1977; Nishioka and Silverstein, 1978; Read and Frenkel, 1983). In order to determine if virion proteins are sufficient to activate Alu expression, we infected HeLa cells under several conditions which prevented viral gene expression (figure 2). We first tested the effect of blocking viral protein synthesis with cycloheximide (figure 2A). As previously reported (Panning and Smiley, 1993), treatment of uninfected cells with cycloheximide leads to a small and variable increase in the levels of Alu RNAs. However, HSV infection gave rise to a much larger increase in Alu RNAs, and this further increase did not occur when cells were infected in the presence cycloheximide. These data indicated that HSV-mediated activation requires de novo viral protein synthesis. Consistent with this hypothesis, induction was eliminated by UV-inactivation of the virus (figure 2A), and did not occur when cells were inoculated with noninfectious virions of the gD-null mutant FgDb (Ligas and Johnson, 1988) (these particles lack gD, and therefore cannot penetrate cells). Finally, induction was blocked when cells were infected with tsB7 at the non-permissive temperature (figure 2B). tsB7 bears a temperature-sensitive
FIG. 2. Requirement for viral gene expression and protein synthesis.

Cytoplasmic RNA extracted from HeLa cells 12 hours postinfection with the indicated viruses was analyzed by primer extension using synthetic 25-mer oligonucleotides specific for Alu, 7SL, and gD transcripts. Analysis of 7SL and Alu RNAs was carried out as described in the legend to figure 1; 20 μg of cytoplasmic RNA was used to detect gD transcripts. Sizes of marker fragments are indicated in nucleotides. (A) mock, mock-infected cells; mock + cyclo, mock-infected cells treated with 100 μg/ml cycloheximide; PAAr5, infected with HSV1 PAAr5; PAAr5 + cyclo, infected with PAAr5 in the presence of 100 μg/ml cycloheximide; PAAr5 + UV, infected with PAAr5 that had been UV-irradiated to produce a 3 log reduction in titre; FgDβ, infected with the burst from an equivalent number Vero cells that had been infected with 1 PFU/cell with FgDβ. (It was not possible to titre this preparation of FgDβ, because it is unable to penetrate cells). (B) HeLa cells were infected with 10 PFU/cell of PAAr5 or TsB7 and maintained at the permissive (34 °C) or nonpermissive (39 °C) temperature.
mutation that prevents release of the viral genome into the nucleus (Batterson et al, 1983; Knipe et al, 1981); as a result, viral gene expression is blocked at the non-permissive temperature. Controls indicated that in each case, expression of the early viral gD gene was eliminated by the treatment, and that levels of 7SL RNA were unaffected. Taken in combination, these data establish that induction of Alu expression requires de novo viral gene expression.

E and L gene expression is not required for activation

To determine which class of HSV gene products is required for activation, we carried out infections under conditions which largely restrict viral gene expression to the IE or E/E-L stages of the lytic cycle (figure 3). Induction was not prevented when viral DNA replication was blocked by aphidicolin, a condition that strongly inhibits L gene expression. However, the levels of Alu RNAs were somewhat reduced, raising the possibility that one or more L proteins contributes to activation. Activation also occurred when cells were infected with the ICP4 null mutant d120 (DeLuca et al, 1985); in this case the levels of Alu RNAs were not detectably reduced relative to PAAR5. ICP4 is required for expression of most E, E-L and L genes (Preston, 1979; Watson and Clements, 1980), and d120 infections are limited to the IE phase of infection (DeLuca et al, 1985). Taken together, these data indicate that wild type levels of E, E-L and L gene expression are not required to activate Alu expression, and raise the possibility that one or more IE gene products other than ICP4 are sufficient. Perhaps the overproduction of IE proteins that occurs during infection with d120
FIG. 3. Requirements for induction of Alu RNAs. HeLa cells were infected with 10 PFU/cell of the indicated virus in the presence or absence of 10 µg/ml aphidocolin, and cytoplasmic RNA harvested 12 hours postinfection was analyzed by primer extension as described in the legend to figure 1. Sizes of marker fragments, produced by 3'-labeling Hpa II fragments of pBR322 DNA, are indicated in nucleotides.
mock
mock+aph
d120
PAAr5
PAAr5+aph
(DeLuca et al, 1985) compensates for the lack of viral DNA replication during infection with this mutant.

**Role of Individual IE gene products.**

The wild-type activity of d120 raised the possibility that one or more IE proteins other than ICP4 are sufficient to induce *Alu* expression. We therefore examined the effects of null mutations in each of the four other IE genes (ICP22, ICP47, ICP0, and ICP27). Strain R325tk+ (a derivative of HSV1 strain F) bears a deletion in the ICP22 gene (Post and Roizman, 1982), N38 lacks US9, US10, US11 and ICP47 coding sequences (Umene, 1986), dlx3.1 bears a deletion in both copies of the ICP0 gene (Sacks and Schaffer, 1987), and 5dl1.2 carries a deletion that removes the promoter and first 1/3 of the coding sequences of ICP27 gene (McCarthy et al, 1989). We found that all of these mutants induced *Alu* RNAs to wild-type levels (figure 4). The E protein ICP6 is also expressed to high levels during infection with d120 (DeLuca et al, 1985). We therefore examined hrR3, bearing a deletion-insertion mutation that replaces ICP6 coding sequences with the *lacZ* gene (Goldstein and Weller, 1988) and found that this strain also induced *Alu* RNAs. These data establish that no single HSV IE protein is essential for induction of *Alu* RNAs. One interpretation is that two or more proteins are each sufficient for activation, and at least one of these is an IE protein other than ICP4. Another possibility is that induction is mediated by an E or E-L protein, and that the low levels produced during infection with d120 are sufficient for full activation. Further studies are required to distinguish between these possibilities.
FIG. 4. Role of IE gene products in activation of Alu expression. HeLa cells were infected with 10 PFU/cell of the indicated HSV strain, and cytoplasmic RNAs analyzed by primer extension as described in the legend to figure 1. Sizes of marker DNA fragments, generated by 3'-end labeling Hpa II cleaved pBR322, are indicated in nucleotides.
ICP27 is not required for activation of Alu transcription.

Jang and Latchman reported that the IE protein ICP27 is necessary and sufficient for HSV to stimulate pol III transcription of endogenous human Alu elements in a nuclear run-on transcription assay (Jang and Latchman, 1989; Jang and Latchman, 1992), a result that is in apparent conflict with our finding that the ICP27-deficient mutant 5dl1.2 induced accumulation of Alu transcripts driven from the Alu pol III promoters (figure 4). Jang and Latchman implicated ICP27 on the basis of the phenotype of an HSV1 strain 17 mutant, 17X2D, which bears a large deletion downstream of ICP27 coding sequences (Maclean and Brown, 1987, figure 5B). Although 17X2D produces reduced levels of ICP27 (Maclean and Brown, 1987), the protein is functional, and this strain can replicate in non-complementing cells. In contrast, viral mutants bearing ICP27 null mutations can only be propagated on cells that provide ICP27 in trans (McCarthy et al, 1989; Rice et al, 1989). In order to more closely examine the role of ICP27 in induction of Alu expression, we examined a panel of ICP27 mutant strains for their ability to induce Alu RNAs (figure 5). Strain 5dl1.2 was described above; d27-1 bears a deletion that removes the ICP27 promoter and most of the coding sequences (Rice et al, 1989), while n59R carries a nonsense codon near the amino terminus and produces a truncated protein of 59 aa (Rice et al, 1989). All of these strains differ from 17X2D in that they must be propagated on complementing cell lines. We found that all of the ICP27 mutants (including 17X2D) activated Alu expression, as judged by the accumulation of Alu RNAs in infected cells.
FIG. 5. Analysis of additional ICP27 mutants. (A) HeLa cells were infected with 10 PFU/cell of the indicated virus, and cytoplasmic RNA harvested 12 hours postinfection was analyzed as described in the legend to figure 1. Molecular weight markers were produced by cleavage of pBR322 with Hpa II, followed by 3'-end labeling. Fragment sizes are indicated in nucleotides. (B) Diagram of mutants employed.
(figure 5A). These results indicate that ICP27 expression is not required to mediate an increase in the levels of cytoplasmic Alu RNAs.

We then used the nuclear run-on transcription assay to determine if HSV infection activates Alu pol III transcription, and if so, whether ICP27 is required. As previously described (Panning and Smiley, 1993), nuclei prepared from uninfected HeLa cells displayed only low levels of Alu run-on transcripts (figure 6A). In contrast, HSV infection induced a large increase in transcripts that hybridized to the Alu non-coding strand probe (Alu 18). This asymmetric increase in Alu transcription was resistant to 2.5 μg/ml α-amanitin, but was eliminated by the addition of 200 μg/ml α-amanitin, suggesting that HSV infection induces pol III transcription of Alu elements (Kedinger et al, 1970; Lindell et al, 1970). Controls confirmed that pol II transcription of the HSV tk gene was strongly inhibited by 2.5 μg/ml α-amanitin, while pol I transcription of rDNA was resistant to 200 μg/ml. Similar assays carried out in the presence of 2.5 μg/ml α-amanitin revealed that 5dl1.2 and 17X2D induced Alu pol III transcription to roughly the same levels observed with PAAR5. These results conflict with those reported by Jang and Latchman (Jang and Latchman, 1989), and establish that ICP27 is not required for transcriptional activation of Alu elements by HSV.
FIG. 6. Nuclear run-on analysis of Alu transcription in infected HeLa cells. HeLa cells were infected with 10 PFU/cell of the indicated virus, and nuclei were isolated 6 hours later. Nuclear run-on transcription was carried out with $^{32}$P-UTP and the resulting RNAs were hybridized to nitrocellulose filters bearing M13, Alu18, Alu19 and tk19 constructs as described in the legend to figure 2. In addition, 10 μg of a plasmid bearing 7SL sequences and 50 μg of a plasmid containing rat rDNA sequences were used to detect 7SL transcripts and rRNAs (Panning and Smiley, 1993). (A) Results of run-on transcription assays carried out in the presence of the indicated concentrations of α-amanitin. (B) Results of run-on transcription assays carried out in the presence of 2.5 mg/ml α-amanitin.
A.

<table>
<thead>
<tr>
<th></th>
<th>0µg/ml</th>
<th>2.5µg/ml</th>
<th>200µg/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>m13</td>
<td>Alu18</td>
<td>Alu19</td>
<td>Alu19</td>
</tr>
<tr>
<td>mock</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rDNA</td>
<td>7SL</td>
<td>tk19</td>
<td></td>
</tr>
<tr>
<td>m13</td>
<td>Alu18</td>
<td>Alu19</td>
<td></td>
</tr>
<tr>
<td>PAArs5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rDNA</td>
<td>7SL</td>
<td>tk19</td>
<td></td>
</tr>
</tbody>
</table>

B.

<table>
<thead>
<tr>
<th></th>
<th>PAArs5</th>
<th>5dl1.2</th>
<th>17X2D</th>
</tr>
</thead>
<tbody>
<tr>
<td>m13</td>
<td>Alu18</td>
<td>Alu19</td>
<td>Alu19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rDNA</td>
<td>7SL</td>
<td>tk19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HSV activates Alu expression in a variety of cell types.

We used the primer extension assay to determine if HSV-induced activation of Alu expression is restricted to HeLa cells (figure 7). Induction was observed during infection of early passage human embryonic kidney cells and human fibroblasts, demonstrating that transformation is not a prerequisite for activation.

293 cells constitutively express the adenovirus type 5 E1a and E1b proteins (Graham et al, 1977; Lassam et al, 1979). E1a activates RNA pol III transcription in vitro by increasing the amount or activity of TFIIC (Hoeffler et al, 1989; Yoshinaga et al, 1986), and extracts of 293 cells show increased levels of pol III transcription (Hoeffler et al, 1989; Yoshinaga et al, 1986). However, transcription of most endogenous class III genes is not activated in 293 cells (Gaynor et al, 1985), and we have previously reported that these cells display only low levels of constitutive Alu expression (Panning and Smiley, 1993). We found that HSV infection strongly induced Alu expression in 293 cells (figure 7). Thus, constitutive activation of TFIIC by adenovirus E1a does not stimulate Alu expression to the levels seen upon HSV infection.

V27 and 3-3 cells are Vero cell derivatives that contain integrated copies of the ICP27 gene (McCarthy et al, 1989; Rice et al, 1989), and express ICP27 inducibly upon infection. Jang and Latchman reported that uninfected V27 cells show increased expression of endogenous Alu sequences in nuclear run-on assays, and that this increase is due to
FIGURE 7. HSV induces Alu expression in multiple cell types. The indicated cells were infected with 10 PFU/cell of HSV1 KOS PAAr5, and cytoplasmic RNA was harvested 12 hours postinfection. Levels of Alu and 7SL transcripts were determined using primer extension analysis as described in the legend to figure 1. Molecular weight markers were produced by cleavage of pBR322 with Hpa II, followed by 3'-end labeling. Sizes are indicated in nucleotides. HF, early passage human fibroblasts; HEK, early passage human embryonic kidney cells; 293, Ad5 transformed HEK cells; 3-3 and V-27, Vero cell derivatives containing integrated copies of the ICP27 gene.
increased levels of TFIIC (Jang and Latchman, 1992). However, we found that 3-3 and V27 cells contain the same low levels of Alu RNAs as Vero cells, and that infection with HSV strongly induces Alu expression in both cell types. These data indicate that increased levels of TF IIIC in V-27 cells are insufficient to induce expression of Alu RNAs.
Discussion

The results described in this paper confirm that HSV infection activates RNA polymerase III transcription of endogenous human Alu elements (Jang and Latchman, 1989). Induction occurred at the transcriptional level, and required de novo viral gene expression. The ICP4 deletion mutant d120 displayed wild-type activity, implicating one or more IE proteins other than ICP4 in induction. However, viral mutants bearing mutations in each of the HSV IE genes showed wild type induction of Alu pol III RNAs. The simplest interpretation of these data is that two or more proteins are each capable of activating Alu expression and that one of these is most likely an IE gene product. Alternatively, it is possible that induction is mediated by an E or E-L protein, and that the low levels produced during infection with d120 are sufficient for full activity.

Jang and Latchman reported that the IE protein ICP27 is necessary for activation of Alu pol III transcription during infection of HeLa cells (Jang and Latchman, 1989). However, we found that three ICP27 null mutants induced the accumulation of high levels of cytoplasmic RNAs initiated from the Alu pol III promoter, and that the ICP27-deficient mutant 5dl1.2 induced Alu transcription as measured in the nuclear run-on transcription assay. In addition, the HSV1 strain 17 mutant 17X2D studied by Jang and Latchman (Jang and Latchman, 1989) displayed close to wild-type activity in our hands. We therefore conclude that ICP27 is not required for induction of Alu transcription. Jang and Latchman concluded that ICP27 induces Alu transcription by increasing the activity of TFIIIC
(Jang and Latchman, 1992). This conclusion was based on the observation that a Vero cell derivative containing integrated ICP27 genes (V27) displays both elevated levels of Alu transcription, and an increase in TFIIIC activity (Jang and Latchman, 1992). However V27 cells do not express detectable quantities of ICP27 unless they are superinfected with HSV (S. Rice, pers. comm.), and they display the same low levels of cytoplasmic Alu RNAs as Vero cells (figure 7). Adenovirus E1a has been shown to activate TF IIIC activity (Hoeffler et al, 1989; Yoshinaga et al, 1986) yet expression of E1a is not sufficient to stimulate Alu transcription (Panning and Smiley, 1993). We therefore consider it quite unlikely that activation of TF IIIC is solely responsible for the induction of Alu expression that occurs during HSV infection. Indeed, we found that HSV greatly stimulated Alu expression in the 293 cell line that displays constitutively activated TF IIIC (Hoeffler et al, 1989; Yoshinaga et al, 1986).

Although our data clearly indicate that ICP27 is not required for induction of Alu transcription, they leave open the possibility that it is one of several HSV proteins that are each sufficient for activation. In this context, it is interesting to note that the functional properties of ICP27 are reminiscent of the E1b 58Kda and E4 ORF 3 and 6 proteins that are required for induction of Alu transcription by adenovirus (Panning and Smiley, 1993). Thus, ICP27 is required for efficient viral L gene expression (McCarthy, 1989; Rice et al, 1989; Sacks et al, 1985), and it appears to act, at least in part, by modulating intranuclear RNA processing events (Sandri-Goldin and Mendoza, 1992; Smith et al, 1992). Along the same lines, the
E1b and E4 proteins are post-transcriptional regulators of adenovirus L gene expression, and are required for the accumulation of normal amount of cytoplasmic L mRNAs (Babiss et al, 1985; Halbert et al, 1985; Sandler and Ketner, 1989; Weinberg and Ketner, 1986; Williams et al, 1986). The mechanism by which E1b and E4 proteins regulate levels of L cytoplasmic messages is unclear; however, evidence suggests that they facilitate the nuclear transport or processing of late RNAs (Leppard and Shenk, 1989; Nordqvist and Akusjarvi, 1990; Pilder et al, 1986). Further studies are required to determine if ICP27 can induce Alu transcription, and if so, whether the mechanism involved is the same as that employed by adenovirus regulators.

The mechanism(s) by which HSV and adenovirus regulators induce Alu transcription remain unknown. Possibilities include relief of global repression, stimulation of the activity of a required activator, or redistribution of components of the pol III transcriptional apparatus. Alternatively, it has been suggested that methylation plays a role in tissue specific Alu expression (Hellmann et al, 1993): perhaps viral proteins induces alterations in cellular Alu sequences similar to those caused by changes in methylation patterns.

At least three viruses induce transcription of endogenous Alu elements (Jang et al, 1992; Jang and Latchman, 1989; Panning and Smiley, 1993), suggesting that induction serves a biological function in viral replicative cycles, or that activation is part of a general cellular response to
virus infection. Transfected \textit{Alu} and 7SL RNA genes have been shown to suppress HeLa cell proliferation, indicating that \textit{Alu} expression may be involved in regulation of the cell cycle (Sakamoto et al, 1991); perhaps virus-induced alterations in cellular signaling pathways indirectly stimulate \textit{Alu} transcription. It has also been suggested that \textit{Alu} RNAs might perform an antiviral function (Jurka and Milosavljevic, 1991). The \textit{Alu} domain of 7SL RNA is responsible for elongation arrest of nascent proteins prior to their cotranslational insertion into microsomal membranes (Siegel and Walter, 1986), and \textit{Alu} sequences display an evolutionary conserved secondary structure similar to that of 7SL RNA (Sinnett et al, 1991). It is therefore possible that \textit{Alu} transcripts assemble into a ribonucleoprotein particle that functions in translational control during infection.

The ability of adenovirus and HSV proteins to induce high level transcription of \textit{Alu} elements provides a novel approach for the study of the biology of this important class of transposable elements, and raises the possibility that virus infection may contribute to \textit{Alu} dispersion by increasing the concentration of the presumed transposition intermediate.
Acknowledgments

We thank Bernard Roizman, Priscila Schaffer, Neal DeLuca, David Knipe, Kenichi Umene, and Moira Brown for providing HSV mutants. This research was funded by the National Cancer Institute of Canada. JRS is a Terry Fox Senior Scientist, and B.P was a research student, of the NCI(C).
References


Preston, C. M. (1979). Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature sensitive mutant tsk. J. Virol. 29, 275-284.


distribution of interspersed repeats is non-uniform and conserved in the
mouse and human genomes. Proc. Nat. Acad. Sciences (USA) 80, 1816-
1820.

Ullu, E., Murphy, S., and Melli, M. (1982). Human 7SL RNA consists of a
140-nucleotide middle repetitive sequence inserted in an Alu sequence. Cell


internal promoter of the human 7SL RNA gene. Nature (London) 318, 371-
374.

Umene, K. (1986). Conversion of a fraction of the unique sequence to part of
the inverted repeats in the S component of herpes simplex virus type 1


Williams, J., Karger, B. D., Ho, Y. S., Castiglia, C. L., Mann, T. and Flint, S. J. (1986). The adenovirus E1B 495 R protein plays a role in regulating the transport and stability of the viral late messages. Cancer Cells 4, 275-284.

4.2. **Segue**

*Alu* elements are thought to be retropositionally active, yet there was little evidence for the presumed pol III transcribed transposition intermediate: the majority of *Alu* elements were thought to be inactive pseudogenes. In this paper we demonstrate that there is low level *Alu* expression in a variety of human cell types, and that in every instance expression was stimulated by HSV-1 infection. In infected HeLa cells *Alu* elements were transcribed by RNA pol III. Though these results do not indicate the number of individual *Alu* sequences which are activated, they do indicate that some of these sequences possess pol III promoters which are functional *in vivo* and suggest that these promoters may be activated by viral regulatory proteins.

We demonstrate that HSV-1 mediated induction of *Alu* expression required viral protein synthesis and occurred upon infection with d120, when viral gene expression was limited to the IE phase, indicating IE proteins are required for this effect. However viruses bearing mutations in each of the IE genes were able to induce *Alu* elements, indicating that HSV-1 encodes two or more proteins which are functionally redundant for *Alu* transactivation. One of these is most likely an IE protein other than ICP4, and the other is most likely an IE or early gene product.

Ad5 recombinants bearing HSV-1 IE genes offered an alternative manner for the delivery and expression of HSV-1 IE gene products. The recombinant virus 0pro-0 was generated by introduction of the ICP0 gene into a virus which lacked the E1 and E3 regions (Zhu *et al.*, 1988). AdICP4
Figure 9. Alu expression in HeLa cells infected with Opro-0 and AdICP4. Cytoplasmic RNA was isolated from HeLa cells infected for 24 hr with the virus indicated (25 PFU per cell) and 10 μg aliquots were analyzed by primer extension using the Alu oligonucleotide primer. The strain dIE1,3 is an Ad5 mutant lacking E1 and E3 regions, while Opro-0 and AdICP4 have copies of the ICP4 and ICP0 genes inserted into vectors lacking E1 and E3 regions. Marker fragments were generated by 3'-end labeling pBR322 cleaved with HpaII and sizes are indicated in nucleotides. (A) Primer extension analysis of Alu RNA produced upon infection with AdICP4 at 34 °C. (B) Primer extension analysis of Alu RNA produced during infection with Opro-0.
contains the temperature sensitive ICP4 gene from the HSV-1 strain tsK in a Ad5 vector lacking E1 and E3 sequences (Spessot et al., 1989). These recombinant viruses both produce the proteins encoded in their respective HSV-1 IE genes, and these proteins function in trans to rescue HSV-1 ICP0 and ICP4 deletion mutants (Zhu et al., 1988; Spessot et al., 1989). However, the ICP0 and ICP4 produced in these Ad5 recombinants do not functionally substitute for E1a gene products: Ad5 expression is limited to the E2a gene. These results suggest that ICP4 and ICP0 can stimulate transcription from the E2a promoter, but do not detectably affect expression of the remaining Ad5 genes. The ICP0-expressing recombinant 0pro-0 stocks contain a low level of wild type virus which could not be eliminated by plaque purification (Zhu et al., 1988), suggesting that E2a expression may be activated by E1a from the wild type contaminant rather than ICP0.

We infected HeLa cells 0pro-0, AdICP4, Ad5, or the virus dIE1,3, which contained the deletions in the E1 and E3 regions similar to those used to allow insertion of HSV-1 IE genes into the Ad5 vectors (Haj-Ahmad, and Graham, 1986) and assayed steady state levels of Alu RNA by primer extension to determine whether ICP4 and ICP0 could stimulate Alu expression in the absence of other HSV-1 proteins (figure 9). Infections were carried out at 37 °C in the series for 0pro-0 and 34 °C in the series for AdICP4, since this is the permissive temperature for tsK ICP4 (Preston, 1979). Mock infected cells and those infected with dIE1,3 did not show a detectable difference in abundance of Alu transcripts, indicating that the Ad5 gene products expressed in the absence of E1 and E3 had little effect on Alu expression. Cells infected with AdICP4 or 0pro0 produced amounts of
Figure 10. IE gene expression in HeLa cells infected with 0pro-0 and AdICP4. HeLa cell RNA was isolated from cells infected with Ad-IE recombinants (25 PFU per cell) for the time indicated (hours). HeLa cells were also infected with HSV-1 PAA'5 (10 PFU per cell), and cytoplasmic RNA isolated at times indicated. Where indicated by + cyclo, infections were carried out in 10 μg/ml of cycloheximide. 10 μg of RNA was used for each primer extension analysis. Sizes of marker fragments (3'-end labeled pBR322 cleaved with HpaII) are indicated in nucleotides. (A) ICP0 RNA levels determined by primer extension using the synthetic oligonucleotide primer (5'-CGTAGGCGCGCTTCTGTGGTGATG), which is expected to yield a 95 nucleotide ICP0 cDNA. (B) ICP4 RNA levels determined by primer extension using the ICP4 oligonucleotide (5'-CGAGCGTCTGACGGTCTGTCTCTCTGG), which hybridizes from +53 to +77 of the ICP4 mRNA.
Alu RNA comparable to those seen with dE1,3. Primer extension analysis was also carried out with ICP4 and ICP0 specific oligonucleotides to determine the levels of expression of HSV gene products (figure 10). Both Ad5 recombinants produced amounts of viral IE transcripts similar to those seen early in HSV-1 lytic infection, indicating that the failure of these two Ad-HSV recombinants to stimulate Alu expression was a result of failure to produce sufficient IE protein, or that ICP0 and ICP4 are in themselves insufficient to induce expression of Alu elements. Since both 0pro-0 and AdICP4 have previously been shown to express functional IE proteins it seems most likely that these HSV-1 IE proteins are insufficient to stimulate Alu expression when expressed from Ad5 vectors.

In these experiments, Ad5 was able to induce high level expression of Alu sequences (figure 9). Infection with the strain dE1,3 infection did not result in the same increase in Alu RNAs, suggesting that E1 or E3 gene products might be involved in Ad5 mediated Alu expression. The 289 residue E1a protein has been demonstrated to be a promiscuous activator of class II and class III transcription in transient transfection assays (reviewed in Flint and Shenk, 1989), and it was known that it activated pol III transcription by altering amounts or activity of TFIIIC (Hoeffler et al., 1988; Yoshinaga et al., 1986). In addition, Jang and Latchman (1992) had unconvincingly demonstrated HSV stimulated Alu transcription by a mechanism that involved modulation of TFIIC activity. The following chapter is a description of the investigation of Ad5-mediated Alu transcriptional activation.
Chapter 5: Activation of RNA polymerase III transcription of human \textit{Alu} repetitive elements by adenovirus type 5: requirement for the E1b 58-kilodalton protein and the products of E4 open reading frames 3 and 6.
Activation of RNA Polymerase III Transcription of Human \textit{Alu} Repetitive Elements by Adenovirus Type 5: Requirement for the E1b 58-Kilodalton Protein and the Products of E4 Open Reading Frames 3 and 6

BARBARA PANNING AND JAMES R. SMILEY*

Molecular Virology and Immunology Program, Pathology Department, McMaster University,
1200 Main Street West, Hamilton, Ontario, Canada L8N 3S5

Received 12 January 1993/Accepted 2 March 1993

We found that transcription of endogenous human \textit{Alu} elements by RNA polymerase III was strongly stimulated following infection of HeLa cells with adenovirus type 5, leading to the accumulation of high levels of \textit{Alu} transcripts initiated from \textit{Alu} polymerase III promoters. In contrast to previously reported cases of adenovirus-induced activation of polymerase III transcription, induction required the E1b 58-kDa protein and the products of E4 open reading frames 3 and 6 in addition to the 289-residue E1a protein. In addition, E1a function was not required at high multiplicities of infection, suggesting that E1a plays an indirect role in \textit{Alu} activation. These results suggest previously unsuspected regulatory properties of the adenovirus E1b and E4 gene products and provide a novel approach to the study of the biology of the most abundant class of dispersed repetitive DNA in the human genome.

\textit{Alu} elements are the single most abundant class of dispersed repeated sequences in the human genome, comprising 5 to 10\% of the mass of human DNA (17, 19, 67, 70, 75, 77). These elements have a dimeric structure that consists of two related but nonidentical \textit{Alu} monomers that are each homologous to an internally deleted 7SL RNA gene (Fig. 1C and D) (13, 86, 87, 97). \textit{Alu} elements are mobile (51, 52, 93), and several lines of evidence suggest that they transpose through an RNA intermediate transcribed by RNA polymerase III (Pol III): they are flanked by direct repeats that vary in length and sequence between elements, end in a 3' A-rich tract, and contain an internal RNA Pol III promoter which directs transcription initiation to the first residue of the \textit{Alu} element (Fig. 1C) (for reviews, see references 17, 18, 37, and 90). \textit{Alu} elements have no known function (34, 103); however, it has been suggested that they may provide cis-acting sequences that serve as modulators of chromatin structure (20), sites of initiation of cellular DNA replication (2, 41), hotspots for recombination (14, 38, 69), inhibitors of gene conversion (30), stabilizers of cytoplasmic RNA (11, 68), negative transcriptional regulators (72, 85, 100), or modulators of intranuclear processing of mRNA precursors (40). Alternatively, it has been proposed that they may encode the RNA component of a cytoplasmic ribonucleoprotein similar to the signal recognition particle (80).

\textit{Alu} elements bear an internal Pol III promoter that resembles those found in rRNA and adenovirus (Ad) VA genes (24, 63). Although Ad promoters are active in vitro (20, 21, 31, 63), \textit{Alu} elements do not appear to be efficiently transcribed by Pol III in vivo. Rather, the \textit{Alu}-related transcripts found in HeLa cells are transcribed primarily by RNA Pol II, presumably reflecting the fact that many \textit{Alu} elements are embedded within Pol II transcription units (11, 61, 62, 77). On the basis of studies of the promoter of the related 7SL RNA gene, it has been suggested that the inactivity of most \textit{Alu} Pol III promoters in vivo stems from the absence of required upstream transcriptional control elements (62, 88, 89); the 7SL promoter is heavily dependent on sequences located upstream of the transcription initiation site, and 7SL RNA pseudogenes lacking these upstream sequences are apparently inactive in vivo. According to this hypothesis, only a limited subset of \textit{Alu} elements that are located in proximity to favorable 5' flanking sequences are transcriptionally competent, and the remainder are inactive pseudogenes. This hypothesis implies that flanking sequences could impose distinct patterns of regulation on individual \textit{Alu} elements, a possibility that is supported by two observations. First, transcripts of a single \textit{Alu} element are abundant in primate brain (53, 94). Second, Matera et al. have provided evidence that a small subset of \textit{Alu} elements are transcriptionally competent and that members of this class are preferentially transcriptional in cultured human cells (51, 52).

Several nuclear DNA viruses encode regulatory proteins that modulate Pol III transcription (3, 23, 33, 79). In the best-characterized case, Ad infection activates transcription of transcribed copies of class III genes in vivo (23) and stimulates in vitro transcription by RNA Pol III in HeLa cell extracts by altering the activity of transcription factor IIIC (TFIIC) (32, 33, 101). The 289-residue protein encoded by Ad early region 1a (E1a) is required for activation (23, 33), and purified E1a is capable of stimulating Pol III transcription in vitro (16). E1a-induced activation of Pol III transcription in vivo appears to be largely restricted to genes that have been newly introduced into cells, inasmuch as transcription of most resident cellular class III genes is not affected (23). Recently, Jang andLatchman reported that herpes simplex virus induces Pol III transcription of endogenous \textit{Alu} elements (39). In this report, we demonstrate that Ad5 infection also strongly stimulates Pol III transcription of human \textit{Alu} elements in HeLa and 293 cells. In contrast to the cases of Ad5-induced Pol III transcriptional activation mentioned above, this process requires the E1b 58-kDa protein and the products of E4 open reading frames (ORFs) 3 and 6.

* Corresponding author.
FIG. 1. Induction of Alu expression upon Ad5 infection. Cytoplasmic RNA, harvested from HeLa cells 24 h postinfection with 25 PFU per cell, was analyzed by primer extension using 5'-labeled synthetic 25-mer designed to detect Alu and 7SL transcripts. Samples were then resolved on an 8% sequencing gel. Markers were 3'-end-labeled HpaII fragments of pBR322 DNA; fragment sizes in nucleotides are indicated. (A) Primer extension analysis of 20 µg of cytoplasmic RNA, using the Alu-specific primer. (B) Primer extension analysis of 1 µg of cytoplasmic RNA, using the 7SL-specific primer. (C) Diagram of an Alu element, indicating the location of the primer relative to site of initiation of Pol III transcription. (D) Comparison of Alu and 7SL sequences. The boxed regions indicate areas of homology between Alu and 7SL RNAs.

in addition to the Ela 289-residue product. These findings suggest novel regulatory properties of the Ad5 Elb and E4 proteins and raise the possibility that analogous cellular trans-acting factors serve to modulate Alu expression in vivo.

MATERIALS AND METHODS

Cells and virus. Table 1 lists the Ad5 mutants that were used in this study; detailed descriptions of the structure of each mutation are provided in the legends to Figs. 3, 6, 7, 8, and 9. Ad5 mutants bearing lesions in the E4 region were propagated and titered on the E4-complementing line W162 (95), and all other Ad5 mutants were grown on 293 cells (26). 293 cells were maintained in α-minimal essential medium supplemented with 10% newborn calf serum; W162 and HeLa cells were grown in α-minimal essential medium containing 5% fetal calf serum. With the exception of ts125, pm975, and dl520, all virus preparations were purified by cesium chloride equilibrium density gradient centrifugation (25). Where indicated, infections were carried out in the presence of 20 µg of cytosine arabinofuranoside hydrochloride (araC; Sigma) per ml or 100 µg of cycloheximide (Sigma) per ml.

The titers of stocks of E4 mutants that are obtained by plaque assays on complementing W162 cells cannot be directly compared with the titers of other Ad5 stocks determined on 293 cells, because W162 cells are of simian origin and support plaque formation by wild-type Ad5 at a substantially reduced efficiency relative to human cells. To estimate the true titer of E4 stocks, we measured the virion concentrations in purified wild-type and E4 mutant stocks by measuring their A260 and then titered the stocks on 293 cells and W162 cells in parallel. Wild-type and E4 mutant strain stocks had titers of approximately 10¹⁰ PFU per optical density unit when titered on W162 cells, indicating that the E4 mutations did not greatly alter virus particle-to-PFU ratios; however wild-type stocks gave a 25-fold-higher titer on 293 cells. Therefore, the titers of E4 mutant stocks were estimated on the basis of their A260 relative to that of wild-type stocks titered on 293 cells.

Primer extension analysis. HeLa or 293 cells were infected at a multiplicity of 25 PFU per cell, and cytoplasmic RNA was isolated 24 h later by the method of Berk and Sharp (7). Primer extension was performed as previously described (82). Synthetic oligonucleotide primers were purchased from the Central Facility of the Institute for Molecular Biology and Biotechnology, McMaster University. The Alu primer, 5'-TTTAGTAGAGA(C/G)GGGTTTCCAAGGTCACATG-3', is predicted to give rise to a ca. 122-nucleotide (nt) extension product from transcripts initiated from the Alu Pol III promoter. The indicated degeneracy was incorporated into the primer at position 11 in order to maximize the number of Alu elements detected. The 7SL RNA primer, 5'-AAGTACGTTCGGTCAAGCAAGG-3', is predicted to give rise to a 122-nt product from 7SL RNA. Both primers were chosen to minimize cross-hybridization due to the homology between Alu RNAs and 7SL RNA (see Results).

Nuclear run-on transcription assays. Nuclear run-on assays were carried out as described previously (81), except that nuclei were isolated as follows. Infected or uninfected HeLa cells (3 x 10⁶ to 5 x 10⁶) were washed extensively in cold phosphate-buffered saline, pelleted, suspended in 1 ml of buffer 1 (0.32 M sucrose, 3.0 mM calcium chloride, 2.0 mM magnesium acetate, 0.1 mM EDTA, 0.1% Triton X-100, 1 mM diithiothreitol, 10 mM Tris-HCl [pH 8.0], 40 U of RNAse [Promega Biotech Corp.] per ml), and disrupted in a Dounce homogenizer, using a wide-bore pestle. Two volumes of buffer 2 (1.85 M sucrose, 5.0 mM magnesium acetate, 0.1 mM EDTA, 0.1 mM dithiothreitol, 10 mM Tris-HCl [pH 8.0], 40 U of RNAse per ml) was added to the cell lysate, and the mixture was layered over 1.8 ml of buffer 2 and centrifuged at 130,000 x g for 45 min in an SW50.1 rotor. The nuclear pellet was resuspended in 100 µl of nuclei storage buffer (5.0 mM magnesium acetate, 0.1 mM EDTA, 5.0 mM dithiothreitol, 50 mM Tris-HCl [pH 8.0], 200 U of RNAse per ml, 25% glycerol) per 1.5 x 10⁶ cells in the original cell pellet and stored in liquid nitrogen. Run-on transcription in isolated nuclei was carried out in the presence of α-amanitin (Sigma) at the concentrations indicated.

Alu run-on transcripts were detected by using M13mp18 and M13mp19 constructs bearing a 1,300-nt Neo1-Pol fragment containing the 600-bp duplicated Alu element present in the human α-2 globin gene cluster (31). The Neo1 end was filled in with Klenow DNA polymerase, and the blunt Neo1-Pol fragment was cloned into M13mp18 and M13mp19 digested with HindIII and PstI. Constructs Alu18 and Alu19 digest cDNA and noncoding-strand Alu DNA and is therefore predicted to hybridize Alu Pol III transcripts, while Alu19 produces the

<table>
<thead>
<tr>
<th>TABLE 1. Mutant strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viral strain</td>
</tr>
<tr>
<td>dl309</td>
</tr>
<tr>
<td>dlEL13</td>
</tr>
<tr>
<td>dl313</td>
</tr>
<tr>
<td>dl520</td>
</tr>
<tr>
<td>pm975</td>
</tr>
<tr>
<td>dl50</td>
</tr>
<tr>
<td>dl55</td>
</tr>
<tr>
<td>1969/7</td>
</tr>
<tr>
<td>1989/7</td>
</tr>
<tr>
<td>ts125</td>
</tr>
<tr>
<td>pm 201S/2250</td>
</tr>
<tr>
<td>dl1007</td>
</tr>
<tr>
<td>dl355+</td>
</tr>
<tr>
<td>dl356+</td>
</tr>
<tr>
<td>E4ΔORF1-4</td>
</tr>
<tr>
<td>E4ΔORF6-8</td>
</tr>
<tr>
<td>E4ΔORF3+</td>
</tr>
<tr>
<td>dl355+ΔE4ΔORF3</td>
</tr>
<tr>
<td>ts125</td>
</tr>
</tbody>
</table>
opposite strand of Alu DNA. To detect 7SL-specific transcripts, a 125-bp Sau3A fragment of 7SL DNA (+101 to +226) was cloned into the BarnHI site of pBR322. This fragment was generated by polymerase chain reaction-mediated amplification of 7SL cDNA, using the synthetic oligonucleotides 5’-GGTCACTGTATCCACTGGAGG (–245 to +26) and 5’-CCCTGCGCATGGTGGTCCGG (+93 to +112) followed by cleavage with Sau3A, and contains a region of the 7SL gene that is lacking from Alu elements (86). Transcription of the E1A gene was detected by using pKH101 (6), a plasmid bearing the Ad5 KpnI H fragment (0 to 5.7 map units). VA RNAs were detected by using a pUC19 plasmid bearing the 100-bp BamHI- XbaI fragment spanning the 5’ end of the Ad2 VA1 gene from –30 to +70. Human rRNAs were detected by using p2.0, a plasmid bearing the 5’ end of the rat rDNA cluster (71). Ten micrograms of denatured plasmid DNA or single-stranded M13 DNA was bound to nitrocellulose filters in a slot blot apparatus and immobilized by UV cross-linking with a Stratalinker 2400 (Stratagene) according to the manufacturer’s instructions. Slot blots were hybridized as previously described (81) for 36 h at 58°C. Following hybridization, filters were washed extensively at 68°C and then treated with 10 μg of RNase A per ml for 30 min at 37°C.

Metabolic labeling of proteins. HeLa cells were infected with the appropriate virus in the presence or absence of 20 μg of araC per ml. Cells were then labeled with [35S]methionine from 24 to 25 h postinfection, harvested, and analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis as previously described (1).

RESULTS

Activation of expression of human Alu elements by Ad5. We used a primer extension assay to monitor changes in the abundance of Alu-related transcripts in HeLa cells after infection with Ad5. To eliminate interference from cross-hybridization with the related 7SL RNA, we designed an Alu-specific primer complementary to nt 98 to 122 of the Alu consensus sequence (Fig. 1C) (44). This region extends into the A-rich sequences connecting the two Alu monomers (Fig. 1D) and displays only limited homology to the 7SL gene. As we also used an oligonucleotide primer designed to specifically detect 7SL RNA. The 7SL RNA primer was complementary to residues 98 to 122 of 7SL RNA (Fig. 1D), the region immediately downstream of the breakpoint in homology between Alu elements and the 7SL RNA gene (86). Cytoplasmic RNA from uninfected HeLa cells gave rise to two types of primer extension products with the Alu primer: a heterogeneous smear that extended into the high-molecular-weight range of the gel, and a faint band of ca. 122 nt (Fig. 1A; note: the autoradiogram was overexposed to allow visualization of the 122-nt product). The heterogeneous primer extension products are most likely derived from Alu elements that are embedded within transcripts of Pol II-transcribed genes, as reported by other workers (61, 62), while the 122-nt product corresponds to that predicted for transcripts initiated at the Alu Pol III promoter (20, 21, 60) (Fig. 1A). We found that the 122-nt primer extension signal was greatly augmented following infection with Ad5, while the heterogeneous products remained relatively constant in abundance. In contrast, 7SL RNA levels did not change after Ad5 infection (Fig. 1B). These data indicate that Ad5 infection strongly stimulates the accumulation of cytoplasmic RNAs initiated from Alu Pol III promoters.

FIG. 2. Alu transcription in Ad5-infected cells. HeLa nuclei were harvested 24 h postinfection, and in vitro run-on transcription was carried out in the presence of α-amanitin at the concentrations indicated. Radiolabeled run-on transcripts were hybridized to 10 μg of M13 replicative-form DNA, single-stranded DNA from M13 constructs bearing Alu DNA (Alu18 and Alu19), and double-stranded plasmids encoding E1a, 7SL, and VA sequences and 50 μg of a plasmid bearing rDNA sequences fixed to nitrocellulose filters. Alu18 and Alu19 bear the Alu noncoding and coding strands, respectively.

Matara et al. (52) have reported that a distinct subset of Alu elements that have served as donors in recent retroposition events are transcriptionally active in HeLa cells, while the elements belonging to the major consensus class are silent. The Alu primer used in the present work is an exact match to the Alu consensus and does not cross-react with transcripts derived from the recently transposed class (60a). Our data therefore suggest that uninfected HeLa cells contain detectable levels of transcripts derived from Alu elements belonging to the major consensus class and that Ad5 infection activates expression of at least some of these elements. However, in other experiments, we have found that transcripts derived from the recently transposed class are also strongly induced during Ad5 infection (60a).

Increased Pol III transcription of Alu elements following infection with Ad5. We used the nuclear run-on transcription assay to determine whether the increase in cytoplasmic Alu RNAs seen after Ad5 infection reflected an increase in Pol III transcription of Alu sequences (Fig. 2). In the first experiment (Fig. 2A), run-on transcription assays were performed in the presence and absence of 2.5 μg of α-amanitin per ml to preferentially inhibit Pol II transcription (45, 48), and the labeled RNA products were hybridized to M13 clones bearing the Alu noncoding and coding strands (Alu18 and Alu19, respectively). As controls, the RNA samples were also hybridized to plasmids bearing the Ad5 E1a region (which is transcribed by Pol II) and a portion of the human 7SL gene and the Ad5 VA1 gene (both transcribed by Pol
Mock-infected nuclei displayed low levels of transcription from both Alu DNA strands, while infected nuclei displayed a large increase in Alu run-on transcripts that hybridized to the noncoding (Alu18) strand probe. In contrast, no significant change in hybridization to the coding (Alu19) strand probe was observed after infection. This asymmetric increase in Alu transcription was resistant to 2.5 \( \mu \text{g} \) of \( \alpha \)-amanitin per ml, consistent with the notion that Ad5 infection stimulates Pol III transcription of Alu elements. Controls indicated that, as expected, transcription of the Ad5 E1a region was sensitive to 2.5 \( \mu \text{g} \) of \( \alpha \)-amanitin per ml, while transcription of the Ad5 VA gene was largely unaffected. To confirm that the Alu transcripts were produced by RNA Pol III (as opposed to RNA Pol I), we tested the sensitivity of Ad5-induced Alu transcription to higher concentrations of \( \alpha \)-amanitin (Fig. 2B). As expected, transcription of Alu elements and the Ad5 VA gene was resistant to 5 \( \mu \text{g} \) of \( \alpha \)-amanitin per ml, while transcription of the E1a gene was strongly inhibited. However, the Alu and VA signals were eliminated by 200 \( \mu \text{g} \) of \( \alpha \)-amanitin per ml, while RNA Pol I transcription of rDNA was not affected. These data therefore indicate that Ad5 infection stimulates transcription of Alu elements by RNA polymerase III.

Because of the extensive homology between Alu elements and the 7SL RNA gene, the Alu noncoding-strand probe (Alu16) should also detect 7SL transcripts. To determine whether the increased Alu18 hybridization signal was due primarily to Alu transcripts, a probe corresponding to the region of the 7SL RNA gene that is deleted from Alu elements was used to monitor the rate of transcription of this gene (and its pseudogenes) following infection. We found that the 7SL signal did not differ between infected and uninfected nuclei (Fig. 2A). These results suggest that the sequences detected by the Alu noncoding-strand probe in this assay are predominantly Alu transcripts.

Requirement for the Ad5 E1A 289-aa gene product. We surveyed the ability of a variety of Ad5 mutants to induce Alu expression to determine which Ad5 functions are required for this process. The products of the E1a gene seemed likely candidates, because they have been previously shown to be required for stimulation of transcription of Pol III-transcribed genes in vivo and in vitro (16, 25, 32, 33, 101). The E1a gene produces two major proteins of 289 and 243 amino acids (aa) by differential splicing (5, 64). To determine whether either or both of these products mediate activation, HeLa cells and 293 cells were infected with viruses bearing lesions in the E1A gene and then assayed for levels of Alu RNA by primer extension (Fig. 3A). Deletion mutants that lack either the E1a gene (d312) or the entire E1 region (d1E1,3) did not induce Alu expression in HeLa cells, suggesting that one or more E1a products are required. Further analysis indicated that pm5975, which cannot produce the 243-aa E1A gene product, induced wild-type levels of Alu RNAs, while d3520, which does not produce the 289-aa protein, was defective. Taken in combination, these data establish that the E1a 289-aa protein is required for activation of Alu elements by Ad5 in HeLa cells and that the 243-aa protein is dispensable.

293 cells constitutively express E1a gene products (46), but uninfected 293 cells displayed the same low levels of Alu transcripts as do uninfected HeLa cells (Fig. 3A). However, Alu transcripts were induced to much higher levels when 293 cells were infected with either wild-type or E1-deficient Ad5. The simplest interpretation of these data is that the E1a (and E1b) proteins produced in 293 cells are not sufficient for high-level expression of Alu elements and that the superinducing Ad5 provides one or more additional required proteins. Direct evidence supporting this hypothesis is presented below.

The 289-aa E1a protein is required for the efficient expression of the other Ad early genes (42). It was therefore possible that E1a plays only an indirect role in Alu activation, by stimulating the expression of additional required Ad function(s). To address this question, we took advantage of the observation that Ad early gene expression can be achieved in the absence of E1a during infection at high multiplicity (56). We found that the E1a deletion mutant d312 was able to induce Alu expression following infection of HeLa cells with 2,500 PFU per cell (Fig. 4). Although not definitive, these data raise the possibility that the E1a 289-aa protein is required primarily to activate expression of other viral gene products, which in turn serve to stimulate Alu transcription.

FIG. 3. Role of E1a gene products in activation of Alu expression. Cytoplasmic RNA was isolated from cells infected with the indicated Ad5 mutants and then analyzed by primer extension as described in the legend to Fig. 1. Sizes of marker fragments (3'-end-labeled pBR322 cleaved with HpaII) are indicated in nucleotides. (A) Primer extension analysis of Alu and 7SL RNAs from 293 or HeLa cells infected with various E1a mutants. (B) Diagram of mutants used, showing the E1 region and indicating the locations of the lesions. Mutant d312,3 (27) bears deletions extending from 1.0 to 10.6 map units and from 78.5 to 84.7 map units, eliminating E1 and E3 coding regions. d312 (43) lacks nt 448 to 1349 (1.2 to 3.7 map units), eliminating the E1a region. d520 (29) has a deletion in the splice donor of the 135 E1a message and produces only the 243-aa 12S product. pm5975 (55) encodes a point mutation in the splice donor of the E1a 12S species and as a result specifies only the 135 289-aa polypeptide.
Role of Ad5 DNA replication. Because the data outlined above suggested that Ela products were not sufficient for activation of Alu expression, we examined whether viral DNA replication, and therefore wild-type levels of late gene expression, was required. We found that ts125, bearing a temperature-sensitive mutation in the E2a 72-kDa DNA-binding protein that is essential for viral DNA replication, was able to stimulate Alu expression at both the permissive and nonpermissive temperatures (Fig. 5A). In addition, induction was not prevented by inhibiting viral DNA replication with araC (Fig. 5B) or hydroxyurea (data not shown). Control experiments confirmed that the treatment with araC effectively blocked Ad5 DNA replication (Fig. 5C). Taken together, these results indicate that viral DNA replication, and therefore wild-type levels of late gene expression, is not required for activation of Alu transcription.

Requirement for Elb and E4 gene products. The simplest interpretation of the data presented above was that one or more Ad5 early proteins in addition to Ela were required for activation of Alu expression. To test this hypothesis, we infected HeLa cells with viral mutants bearing deletions in early regions Elb, E3, and E4 (Fig. 6). dElb, bearing a large deletion in the E3 region, showed wild-type levels of Alu RNA, indicating that E3 products are not required. This observation is consistent with the wild-type activity of dElb,3 in 293 cells (Fig. 3). dElb,3 bears the same E3 deletion as does dElb3). In contrast, mutants bearing large deletions in Elb (dIs0) or E4 (dI1007) were strongly impaired, indicating that one or more products of the Elb and E4 genes are necessary for efficient induction of Alu expression. Additional control experiments confirmed that dIs0 and dI1007 complemented each other for Alu activation during coinfection of HeLa cells and that dIs0 was able to efficiently induce Alu expression during infection of 293 cells, while dI1007 was inactive in this cell type (data not shown).

The Elb gene products act as posttranscriptional regulators of viral gene expression (4, 65, 99), while the E4 region encodes at least one protein that acts as a transcriptional regulator and at least two others that serve as posttranscriptional modulators of gene expression (10, 28, 35, 36, 50, 66, 73, 96). Analysis of run-on transcription in nuclei isolated from HeLa cells infected with dI1007 or dIs0 showed that there was no detectable increase in Pol III transcription of Alu elements upon infection with either mutant strain (Fig. 7). These data therefore indicate that products of the Elb and E4 genes are required for transcriptional activation of Alu elements. The levels of expression of the viral Ela and VA genes seen upon infection with dI1007 and dIs0 were comparable to those seen in nuclei from wild-type-infected cells, demonstrating that comparable multiplicities of infection were used.

Requirement for the Elb 496-aa protein. The Elb region encodes two major products, a 19-kDa, 176-aa protein and a 58-kDa, 496-aa protein (8, 64). To determine which of these proteins was required to activate Alu expression, we analyzed RNA extracted from cells infected with pm2015/2250, which does not produce the 58-kDa protein, and the viruses 1569(-), 1893(-), and 1772(-), which fail to specify the 19-kDa protein (Fig. 8). Strains 1569(-), 1893(-), and 1772(-) induced levels of Alu RNA similar to those observed during infection with wild-type Ad5, while pm2015/2250 was strongly impaired. As expected, dIs0 and dI55 (defective for
both E1b products) also failed to induce Alu RNAs. Controls indicated that dl50, dl55, and pm2015/2250 could be efficiently complemented by coinfection with the E1a deletion mutant dl312, demonstrating the presence of biologically active virus in the mutant stocks. These data indicate that the 58-kDa 496-aa E1b protein is required for the activation of Alu transcription. Although both dl50 and pm2015/2250 displayed greatly reduced activity relative to that of wild-type Ad5, they did show slight induction above the levels seen in uninfected or mock-infected cells in some experiments (data not shown), suggesting that other viral proteins are able to induce low levels of Alu expression in the absence of the 58-kDa protein.

Requirement for the products of E4 ORF3 and ORF6. The E4 region of Ad5 contains at least six ORFs and appears to encode at least seven distinct polypeptides through differential RNA splicing (22, 84, 92). Three of the E4 polypeptides have been shown to serve as regulators of gene expression. The ORF6/7 polypeptide forms a complex with the cellular E2F transcription factor, thereby stimulating the activity of the E2 promoter (36, 50, 66), while the ORF3 and ORF6 proteins serve as functionally redundant posttranscriptional regulators of viral late gene expression (10, 35). The 34-kDa ORF6 protein forms a complex with the E1b 58-kDa protein (74), which is apparently required for the posttranscriptional regulatory activity of the 58-kDa protein (15). To determine which E4 gene products are required for stimulation of Alu expression, we analyzed cytoplasmic RNA isolated from HeLa cells infected with mutant viruses bearing lesions in various E4 ORFs (Fig. 9A). The E4 mutants fell into three categories with respect to the ability to induce Alu expression. One mutant, dl355*, showed wild-type activity. This result demonstrates that the ORF6/7 protein that binds the E2F transcription factor is not required. A second group of mutants were partially impaired but displayed significantly higher levels of Alu RNAs than did the E4 null mutant dl1007. This group included two mutants that inactivate ORF6 (dl355*/E4inORF6,67), one mutant that deletes ORF1 through ORF4 (E4delORF1-4), and one that inactivates ORF3 (E4inORF3). These data therefore implied the existence of at least two partially redundant E4 functions that are required for Alu activation, one encoded by ORF6 and the other specified by ORF3. To test this hypothesis, we examined a double mutant, dl355*/E4inORF3, that bears inactivating mutations in both ORF3 and ORF6. This mutant showed the same extreme phenotype as did the E4 null mutant dl1007, suggesting that the ORF3 and ORF6 products together account for most or all of the activity of the E4 region. Control experiments demonstrated that all of the E4 mutants were capable of complementing the E1a deletion mutant dl312 for activation of Alu elements in coinfected cells, indicating that infectious virus was present in each E4 mutant stock. In summary, the data obtained in these experiments indicated that the products encoded by E4 ORF3 and ORF6 are both required for wild-type levels of Alu activation and suggest that these proteins display partial functional redundancy in this process.

Variable induction of Alu RNAs by cycloheximide. As described in Discussion, the E1b 58-kDa protein and the products of E4 ORF3 and ORF6 contribute to the shutoff of
Indeed, analysis of proteins labeled from 24 to 25 h postinfection confirmed that araC treatment prevented the onset of late viral protein synthesis and the inhibition of host translation under our experimental conditions (Fig. 11). Second, we found that the E4 deletion mutant d5007 induces considerable (albeit reduced) shutoff of host translation at late times postinfection (data not shown), yet this mutant fails to activate Alu expression. Taken in combination, these observations raise the possibility that Ad5-induced activation of Alu expression stems from selective depletion of a labile inhibitor of Alu transcription rather than from global inhibition of cellular protein synthesis (see Discussion).

E1B- and E4-independent expression of the Ad5 VA gene. Previous studies have not documented a requirement for E1B and E4 function in Ad-induced activation of Pol III transcription; indeed, earlier reports concluded that the 289-aa E1a protein is sufficient to activate Pol III transcription in extracts prepared from infected cells (16, 33). Thus, our data raised the possibility that only a subset of class III genes rely on E1B and E4 for transcriptional activation during infection. Consistent with this conclusion, the nuclear runon transcription assays presented in Fig. 7 suggested that the Ad5 VA1 gene was efficiently transcribed in the absence of E1B and E4 products. As an independent test of the requirements for VA gene expression, we monitored the accumulation of VA RNAs in cells infected with various Ad5 mutants. Infected HeLa cells were incubated in the presence of araC to block viral DNA replication and labeled with [32P] from 2 to 18 h postinfection. Cytoplasmic RNA was then examined for the presence of labeled VA RNA by electrophoresis through an 8% sequencing gel. As previously reported, the E1a-deficient mutant d312 failed to accumulate VA RNA (42). However, deletion of the E1B (d50) or E4 (d1007) regions did not prevent VA RNA synthesis (Fig. 12). These data confirm that E1B and E4 gene products are not required for all cases of activation of Pol III-transcribed genes in Ad5-infected cells.

**DISCUSSION**

The data presented in this report demonstrate that Ad5 infection activates Pol III transcription of endogenous human Alu elements in HeLa and 293 cells, resulting in the accumulation of high levels of cytoplasmic RNAs initiated from Alu promoters. Induction occurred in the absence of viral DNA replication, suggesting that activation is mediated by one or more early gene products. Further analysis revealed that at least four Ad5 early polypeptides are required for this process: the E1a 289-aa protein, the E1b 58-kDa protein, and the products of E4 ORF3 and ORF6. Although it is not yet clear what proportion of the ca. 500,000 Alu elements residing in the human genome are induced upon Ad5 infection, data to be presented elsewhere indicate that at least two distinct subclasses of Alu elements are stimulated (60a), suggesting that Ad5 gene products provoke a relatively global activation of Alu transcription.

Previous reports have established that the 289-aa E1a protein activates in vitro transcription of class III genes in the absence of other Ad5 proteins, by modifying the activity of TFIIC (16, 32, 33, 101). However, although the 289-aa E1a protein was required for activation of Alu transcription under our standard conditions of infection, additional proteins encoded by the E1b and E4 regions were also required. These results indicate that E1a-induced modification of TFIIC is not sufficient for activation of Alu elements in vivo. Consistent with this conclusion, 293 cells did not show...
FIG. 9. Effects of deletion of various E4 gene products on activation of Alu expression. Cytoplasmic RNA isolated from HeLa cells 24 h after infection with mutants bearing lesions in the E4 coding region was analyzed by primer extension as described in the legend to Fig. 1. Single infections were done at a multiplicity of infection of 25, while coinfections used 25 PFU of each virus per cell. Sizes of marker HpaII fragments of pBR322 DNA are indicated in nucleotides. (A) Levels of Alu transcripts after infection with E4 mutants. (B) Locations of the mutations used in these experiments. d11007 contains a deletion eliminating expression of all E4 gene products. d355* (35) has a 14-bp deletion at 93.5 map units in the sequences unique to the ORF6 coding region and fails to produce the ORF6 protein product. d356* (36) contains a 12-bp deletion at 91.8 map units and fails to produce the ORF6/7 protein product. E4d/ORTF1-4 (35) lacks sequences between 94.9 and 98.4 map units and does not produce the products encoded by ORF1 through ORF4. Mutant strain E4in/ORTF6,6/7 (35) contains a 2-bp insertion at 94.6 map units, a region common to both ORF6 and ORF6/7, and thus fails to produce both these proteins. E4in ORF3 (35) introduces 8 bp at 96.4 map units, in sequences encoding ORF3, and as a result fails to produce the ORF3 gene product. d355*/E4in ORF3 (35) combines the lesions in d355* and E4in ORF3 and therefore produces neither the ORF3 nor the ORF6 protein product.
high levels of constitutive expression of Alu RNAs. Furthermore, we found that the requirement for E1a could be at least partially bypassed during infection at high multiplicities, a condition that permits E1a-independent expression of other Ad5 early genes. Thus, E1a-mediated modification of TFIIC is not stringently required for Alu activation. In our view, the simplest interpretation of the available data is that the 289-aa E1a protein plays an indirect role in Alu activation, by facilitating expression of additional required Ad5 proteins, including the E1b and E4 gene products.

It is not yet known to what extent the requirement for E1b 58-kDa and E4 ORF3 and ORF6 proteins extends to other cases of Ad-induced activation of Pol III transcription in vivo. Gaynor et al. (23) reported that transfected copies of tRNA and Ad VA genes are expressed to high levels in 293 cells in the absence of E4 functions, but the lack of a matched control cell line makes it difficult to determine whether this response is directly mediated by the resident E1a and/or E1b proteins. A more direct indication that the regulatory requirements of various class III genes can differ is the observation that Ad5 VA gene expression occurred in the absence of E1b or E4 products (Fig. 7 and 12). The additional requirement for E1b and E4 in Alu activation is intriguing in view of the similar organization of the Alu and VA RNA gene promoters. Perhaps Alu elements bear E1b- and E4-responsive regulatory sequences. Alternatively, the requirement may reflect some feature of the higher-order organization of Alu elements within the nucleus. In this regard, it is worth noting that previously documented cases of in vivo activation of Pol III transcription by Ad products have been restricted to templates that had been newly introduced into cells by transfection or virus infection.

How do the E1b 58-kDa and E4 ORF3 and ORF6 proteins contribute to transcriptional activation of Alu elements? The E1b 58-kDa protein binds the cellular p53 tumor suppressor protein (75, 102) and can cooperate with E1a to transform primary cells in culture (reviewed in reference 8). In addition, it plays a critical role in the mediating the onset of the late phase of lytic infection, by facilitating the accumulation of cytoplasmic late viral mRNAs (4, 65, 99). This latter process also requires the products of E4 ORF3 and ORF6 (28, 35, 73, 95, 96). The E1b 58-kDa and the E4 ORF6

---

**TABLE 2. Levels of induction of steady-state amounts of Alu RNA**

<table>
<thead>
<tr>
<th>Expt</th>
<th>Relative amount of Alu RNA*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cycloheximide/mock</td>
</tr>
<tr>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>2</td>
<td>1.8</td>
</tr>
<tr>
<td>3</td>
<td>2.6</td>
</tr>
<tr>
<td>4</td>
<td>2.9</td>
</tr>
<tr>
<td>5</td>
<td>6.6</td>
</tr>
<tr>
<td>6</td>
<td>10.9</td>
</tr>
</tbody>
</table>

* Determined by primer extension followed by quantitation with a PhosphorImager (Molecular Dynamics).
products associate to form a complex that localizes to the nucleus (74, 83), and mutations that eliminate the 58-kDa protein or E4 region result in similar defects in the accumulation of cytoplasmic late mRNAs, viral DNA replication, and shut off of host protein synthesis (4, 28, 65, 95, 96, 99). The E4 ORF3 and ORF6 proteins appear to be functionally redundant, in that mutations that inactivate either protein have relatively minor effects, while mutations that simultaneously eliminate both display an exaggerated phenotype with profoundly reduced levels of virus replication (10, 35). These data have been interpreted to indicate that the E1b 58-kDa protein functions as a complex with ORF6 and that E4 ORF3 performs a similar or partially overlapping function. The precise mechanism by which these proteins stimulate the accumulation of cytoplasmic late viral mRNA remains unknown; however, the available evidence indicates that they function within the nucleus at the posttranscriptional level to facilitate the transport or processing of mRNA precursors. Thus, mutants lacking the 58-kDa E1b protein display defects in the transit of late mRNA precursors through nuclear subcompartments, leading to reduced levels of cytoplasmic mRNAs (47), while mutants lacking all of the E4 region exhibit reduced stability of intranuclear late mRNA precursors (73). In addition, evidence has been presented that E4 products may act at the level of mRNA processing (57). The striking correlation between the Ad5 gene products required for the accumulation of late viral mRNAs, suppression of host protein synthesis, and activation of Alu elements suggests that these processes are mechanistically related.

One possibility is that activation is a part of cellular response to virus infection that is induced by events well downstream of the immediate targets of E1b and E4 activity: for example, the accumulation of late viral mRNA or the attendant global shut off of host protein synthesis. Although this possibility cannot be discounted, several observations appear to exclude some of the more obvious potential inducing signals. First, activation was not affected by blocking viral DNA replication (Fig. 5), a condition that prevents late viral protein synthesis and the suppression of host translation (Fig. 11), and the accumulation of double-stranded RNA capable of activating the PKR protein kinase (49, 58, 59). Second, mutant strain E4d/ORF1-4 shows wild-type levels of host translational shutoff (35) but is partially impaired for induction of Alu expression (Fig. 9). A second possibility is that activation is the indirect consequence of the same intranuclear events that coordinate the transport and/or processing of late viral mRNA precursors. For example, these alterations might block the export of an unstable cellular mRNA encoding a labile repressor of Alu transcription (note that this hypothesis can account for the stimulation of Alu expression by cycloheximide and the observation that Alu activation can be uncoupled from global translational shutoff). Alternatively, activation might result from Ad5-induced changes in intranuclear architecture. It has been proposed that the E1b 58-kDa protein acts by increasing the number of intranuclear sites from which newly transcribed mRNA precursors can engage the RNA transport machinery (47). Perhaps this alteration involves changes in intranuclear organization or chromatin structure that allow Alu elements to gain access to required transcription factors or Pol III. In this context, it is interesting to note that the Ad12 E1b 55-kDa protein plays a role in decondensing a limited number of specific sites in human metaphase chromosomes, termed Ad12 modification sites (78). Although this phenomenon is not observed during infection with Ad5, these observations provide a precedent for E1b-dependent alterations in chromatin organization. Finally, it is possible that the E1b and E4 proteins exert their effects on the transport and/or processing of viral mRNA precursors by inducing transcription of one or more endogenous cellular genes. Perhaps the relevant targets are transcriptional activators such as Pol III, and Alu elements are coincidentally induced; alternatively, it is conceivable that Alu elements play some role in mediating E1b and E4 function. Further work is required to distinguish between these possibilities.

The ability of Ad5 gene products to induce abundant expression of Alu transcripts offers a novel approach to the study of the biology of this important class of retrotransposons. For example, we have found that Alu elements falling into several distinct subclasses are induced, implying that many Alu elements are transcriptionally competent, given the presence of appropriate trans-acting factors (60a). RNA polymerase chain reaction-based approaches should allow the identification of a large number individual Alu elements that respond to Ad5 infection, and it will be interesting to learn whether these elements share common features that might account for their activation. In addition, Ad5-infected cells can be used to examine whether Alu transcripts are assembled into ribonucleoproteins similar to the signal recognition particle. Finally, it is possible that Ad5 infection increases the frequency of Alu transposition, by increasing the concentration of the presumed RNA transposition intermediate.

ACKNOWLEDGMENTS

We thank P. E. Branton, P. Hearing, G. Ketner, and F. L. Graham for generous gifts of Ad5 mutants and F. L. Graham for valuable advice and criticism. This research was supported by the National Cancer Institute of Canada. B.P. is a Research Student, and J.R.S. is a Terry Fox Senior Scientist, of the National Cancer Institute of Canada.

REFERENCES


The adenovirus E4 17-kilodalton protein complexes with the cellular transcription factor E2F, altering its DNA-binding properties and stimulating E1A-independent accumulation of E2 mRNA. J. Virol. 64:2345–2359.


5.1.ii. **Appended data**

Figure 13  Effect of inhibition of DNA replication with hydroxyurea
Figure 13. *Alu* expression in HeLa cells treated with DNA synthesis inhibitors during Ad5 infection. Infections were carried out at an MOI of 25 PFU per cell and RNA was harvested 24 hours post infection. Where indicated cells were treated with 20 μg per ml of ara C or 100 mM hydroxyurea. Sizes of marker fragments, generated by 3'-end labeling pBR322 cleaved with *Hpa*II, are indicated in nucleotides.
5.2. **Segue**

In this paper we have demonstrated that Ad5 infection stimulates \textit{Alu} pol III transcription in HeLa cells. \textit{Alu} transcriptional activation was dependent on the activity of the 289 amino acid E1a protein, a potent transactivator of viral class II and class III genes. Though the stimulation of TFIIIC activity induced by the E1a 289 residue protein may contribute to the increase in levels of \textit{Alu} pol III transcription during Ad5 infection, it seems unlikely that it is the sole cause: E1a gene products are not required during high MOI infections, and 293 cells, which have constitutively activate TFIIIC (Hoeffler and Roeder, 1985; Yoshinaga \textit{et al.}, 1986; Hoeffler \textit{et al.}, 1988), do not show a significant increase in steady state levels of \textit{Alu} RNA. The E1b-58 kDa protein and the products of E4 ORF 3 and 6 are also required for the stimulation of \textit{Alu} transcription, suggesting that the primary contribution of the E1a 289 residue protein to \textit{Alu} induction may be the activation E1b and E4 expression.

The E1b and E4 proteins have pleiotropic effects on viral and cellular gene expression. The E1b 58-kDa protein transforms primary cells in cooperation with E1a gene products, a process which may involve complex formation between the E1b protein and the cellular p53 tumor suppressor protein. Viruses which do not produce the E1b 58-kDa protein or the E4 ORF 3 or 6 gene products have similar phenotypes: they are impaired in DNA replication, late gene expression and the shutoff of host protein synthesis (Babiss \textit{et al.}, 1985; Halbert \textit{et al.}, 1985; Huang and Hearing, 1989; Pilder \textit{et al.}, 1986; Sandler and Ketner, 1989; Weinberg and Ketner, 1983,
1986; Williams et al., 1986). The correlation between viral gene products required for shutoff of host protein synthesis and accumulation of viral late mRNA, and the activation of Alu transcription suggests that these are related processes.

The observation that E1b and E4 proteins can activate expression of Alu sequences from their cellular loci raises some interesting questions. First, the expression of the VA genes, which have split internal class III promoters very similar to those present in Alu elements (Geidushek and Tocchini-Valentini, 1988), requires E1a proteins but not E1b or E4 products (Hoeffler and Roeder, 1985), suggesting that differences between the viral and cellular genome determine which factors are required in trans for pol III transcriptional activation. Ad5 recombinants carrying Alu sequences could be used to determine if the requirement for viral gene products alters when Alu elements are present on the viral genome. These results may also suggest that transient assays would not be the best system to investigate Alu transcriptional activation, as higher order structure or intranuclear localization of these sequences in their cellular loci may be a significant factor in determining the mechanism of activation.

Second, the precise mechanism by which E1b and E4 proteins activate class III transcription remains unclear. All evidence suggests that these proteins regulate gene expression post-transcriptionally, by facilitating the transport and processing of viral hnRNA while inhibiting cytoplasmic accumulation of cellular mRNA (Leppard and Shenk, 1989; Nordqvist and Akusjarvi, 1990; Sandler and Ketner, 1989; Williams et al., 1986). Though shutoff of host gene expression may play some role in Alu
transcriptional activation, since treatment with the protein synthesis inhibitor cycloheximide causes a slight increase in steady state levels of *Alu* transcripts, it is unlikely that it is the sole cause, since E4 mutants showed no correlation between levels of host shutoff and abundance of *Alu* RNA. Therefore it seems likely that E1b and E4 gene products stimulate *Alu* transcription by a mechanism which is related to their role in RNA maturation. Analysis of viruses bearing mutations in, rather than deletion of, E1b and E4 sequences may useful in identifying which regions of these proteins are necessary for induction of *Alu* sequences, and may begin to address the mechanisms by which they act.

Finally, there are nearly one million *Alu* repeats in the human genome and it may be possible to determine whether a small, reproducible set of these are activated by viral infection, suggesting that there are *cis*-acting sequences which are important for regulation of *Alu* expression. The division of *Alu* sequences into subfamilies which were thought to be generated by transposition of a limited number of sequences which were retropositionally active during different periods of human evolution (Britten *et al.*, 1988, 1989; Deininger and Slagel, 1988; Jurka and Smith, 1988; Labuda and Striker, 1989; Matera *et al.*, 1990; Quentin, 1988) suggests that there may be transcriptionally active subsets of *Alu* elements. As described in the next chapter, we were able to exploit the conserved nucleotide changes between different *Alu* subfamilies to design primers to detect transcripts from these subfamilies and analyze amount of RNA produced by members of each subgroup. In addition we were able to directly sequence *Alu* RNAs by incorporating dideoxynucleotides into reverse
transcriptase reactions. These results demonstrated that a large number of Alu sequences are induced upon infection, suggesting that transcriptional activation is not restricted to a very limited subset of these elements.
Chapter 6: Expression of multiple subfamilies of Alu elements is activated by herpes simplex virus and adenovirus infection.
Abstract

The nearly one million Alu repetitive elements in the human genome can be grouped into a number of subfamilies. Comparisons between subfamily consensus sequences suggest that Alu evolution is characterized by the sequential amplification and dispersal of a limited number of Alu founder sequences. Mutation events are thought to result in the transpositional inactivation of parental founder elements and the generation of new founders among their progeny. The S, Sb and Sb1 subfamilies provide an example of such a related series of Alu subfamilies. We have previously demonstrated that adenovirus type 5 and herpes simplex virus type 1 activate RNA polymerase III transcription of endogenous Alu elements in HeLa cells. Here we report that expression of Alu sequences belonging to the S, Sb and Sb1 subfamilies was activated following infection with these viruses. The data indicate that transpositionally inactive Alu elements can give rise to high levels of pol III transcripts in the presence of appropriate trans-acting factors, and demonstrate that the class III promoters of a significant number and variety of Alu sequences are functional in vivo. Multiple subfamilies of Alu sequences were induced in transformed and non-transformed cell types, suggesting that induction of Alu expression may be part of the normal cellular response to viral infection.
**Introduction**

*Alu* repeats are the most abundant mobile short interspersed elements (SINEs) in the human genome (8, 47, 66). They are present in nearly one million copies, comprising over 5% of human DNA (19, 46). The majority of *Alu* elements are approximately 300 nucleotides in length (10, 46), and consist of a tandem arrangement of two related but non-identical monomers (47, 66) that each appear to have been derived from an internally deleted 7SL RNA gene (59, 60, 65). The existence of free right and left *Alu* monomers suggests that the more common dimeric elements arose from fusion of monomers (25,45). The left monomer lacks 31 nucleotides present in the right (47, 66), and contains an intragenic RNA polymerase III (pol III) promoter that drives transcription of the entire element *in vitro* (14).

*Alu* elements are thought to transpose through an RNA intermediate: their structure is suggestive of a cDNA copy of an *Alu* pol III transcript inserted into a staggered break in cellular DNA (47, 66). These SINEs contain 3' A-rich regions which are thought to be produced by RNA processing prior to cDNA synthesis (43, 66) and are flanked by direct repeats created by duplication of sequences at the site of integration (63). Although there is no obvious sequence homology between *Alu* integration sites, the elements display a markedly nonrandom distribution in the human genome. These SINEs preferentially localize to R bands of metaphase chromosomes (27) and are disproportionately abundant in GC rich isochores isolated using density gradient centrifugation (56). These
data are consistent with the tendency of Alu elements to cluster with class II genes (13, 29, 31, 37), which are also preferentially found in R bands and GC rich isochores (27, 56). The targeting of Alu elements into transcriptionally active regions of the genome suggests that some feature of chromatin which is active in somatic cells increases that likelihood of integration of Alu sequences. Such features may include timing of DNA replication, chromatin conformation, CpG content or local increases in sequence homology due to the abundance of Alu sequences.

Alu repeats display substantial heterogeneity in nucleotide sequence, such that pairs of elements diverge at ca. 14% of their residues on average (8, 47, 66). Despite this heterogeneity, Alu elements can be divided into subclasses which are characterized by the clustering of diagnostic alterations relative to the Alu consensus sequence. Several classification schemes have been proposed which group Alu elements into subfamilies that encompass specific Alu sequence variants (6, 11, 24, 44, 67). For example, in the system of Jurka and coworkers (23, 24), Alu elements are first divided into the J and S subfamilies (referred to as the diverged and major subfamilies by other groups). The J subfamily (comprising 20-30% of the elements in the human genome) is a relatively heterogeneous group of Alu sequences that show homology to the 7SL RNA gene at a number of diagnostic positions in the subfamily consensus sequence. The S subfamily (comprising 70-80% of Alu elements) differs from the J subfamily at these diagnostic residues, and shows less sequence divergence between its members. The S subfamily can be subdivided into a number of smaller
groupings, which are themselves further divisible (figure 1): each progressive subdivision is characterized by a new cluster of diagnostic sequence alterations, a higher degree of sequence conservation among its members and a decrease in number of individual members (23).

The differences in the degree of evolutionary divergence within Alu subfamilies led to the suggestion that each subfamily was produced by the radiation of a limited number of master sequences, and that different parental sequences were transpositionally active at different times during evolution (2, 6, 7, 24, 28, 35). This model predicts that recently transposed Alu elements should belong to relatively small and homogeneous subfamilies and analysis of the recently transpositionally active Sb1 (23) (also referred to as HS or PV (3,35)) and Sb2 subfamilies (18, 22) confirms this prediction. Both subfamilies are subgroups of the relatively small Sb family which contains approximately 10% of Alu elements and there is minimal sequence variation among the members of these subfamilies which have been identified to date (3, 35, 67). The Sb1 subfamily is currently transpositionally active (64), and it is among the smallest subgroupings of Alu sequences in the human genome, consisting of approximately 2000 members (3, 35).

The retroposition model predicts that mobile Alu sequences should be transcribed by pol III for two main reasons: class III transcription of Alu elements initiates at the first nucleotide of these sequences in vitro (14), and intragenic pol III promoters are thought to facilitate sequential
transposition events by driving Alu transcription regardless of the site of integration (47, 66). Yet the majority of Alu-containing RNAs detected in HeLa cells are transcribed from non-Alu promoters (41, 42), implying that most Alu pol III promoters in are non-functional tissue culture cells. Transcription of the closely related 7SL RNA gene requires class II-like cis-regulatory sequences in addition to the intragenic pol III promoter (61, 62), leading to the suggestion that the majority of Alu elements may be inactive pseudogenes due to the absence of appropriate promoter elements at the site of integration (23, 36). According to this scenario, only a limited number of Alu sequences are transcribed, and transpositional efficiency is directly correlated with transcriptional activity (9, 36, 52). Consistent with this view, RNAs derived from of the transpositionally active Sb1 subfamily have been detected in a variety of human tissues and cells (36). However, a wide range of Alu elements are expressed in pluripotent NTera2D teratocarcinoma cells (albeit with disproportionate representation of the youngest subfamilies), indicating that Alu transcription does not occur from a single master sequence and that transpositional selection may occur post-transcriptionally (54). The relevance of these efforts to correlate Alu transcriptional activity with transpositional competence is questionable, because only germ line transposition events have been detected to date, and most studies have focused on Alu transcription in tissue culture. However, spermatazoa also contain Alu specific RNAs which initiate at the pol III promoter (26). It is therefore interesting to note that the recent transposition of an Alu element into the NF1 gene occurred in the paternal germline (64).
Adenovirus type 5 (Ad5) and herpes simplex virus type 1 (HSV-1) activate pol III transcription of endogenous Alu repetitive elements during infection of HeLa cells (23, 39), indicating that these viruses encode trans-acting factors that perturb one or more cellular systems that regulate Alu expression. In this report we demonstrate that Ad5 and HSV-1 activate expression of members of the S, Sb and Sb1 subfamilies. These data indicate that both transpositionally competent and incompetent elements can give rise to high levels of pol III transcripts in the presence of appropriate trans-acting regulators in tissue culture, and demonstrate that the class III promoters of a significant number and variety of Alu sequences are functional in vivo. Induction of expression of multiple Alu subfamilies occurred in non-transformed cell strains, suggesting that this phenomenon may be part of a general cellular response to viral infection.
**Materials and Methods**

**Cells and Virus**

Human embryonic kidney cells and human fibroblasts were maintained in α-MEM (GIBCO laboratories) with 10% fetal bovine serum. 293 (14) cells were grown in α-MEM supplemented with 10% newborn calf serum. HeLa cells, BHK21 cells and Vero cells were carried in α-MEM with 5% fetal calf serum. Ad 5 was propagated and titred on 293 cells and HSV-1 strains PAAr5 (16), F and F-Alu2 were grown on Vero cells.

**Primer extension analysis**

Cells were infected with Ad5 at a multiplicity of 25 PFU per cell and cytoplasmic RNA harvested 24 hours later by the method of Berk and Sharp (4). HSV-1 infections were carried out at 10 PFU per cell for 12 hours prior to isolation of cytoplasmic RNA. Synthetic oligonucleotide primers were purchased from the Central Facility of the Institute for Molecular Biology and Biotechnology, McMaster University. Table 1 lists the primers and their sequences. The primers Alu1, Alu2, deg and 7SL have been described previously (38, 39). The primers S, Sb and Sb1 are predicted to produce 115 nucleotide cDNAs from transcripts initiated from Alu pol III promoters of members of the S, Sb and Sb1 subfamilies. Oligonucleotides were designed to include sequences which diverge between Alu subfamilies, particularly at the 3' ends, to maximize their specificity (see figure 1a).
Primer extension analysis was performed as previously described (55), with the exception of the annealing step. Hybridization of primers (1 X $10^6$ Cerenkov cpm in 0.10 to 0.50 ng of oligonucleotide per primer extension reaction) to cytoplasmic RNA was carried out at 63 °C, the measured $T_m$ of hyrbids between the primers and Alu transcripts. In competition assays RNA was preannealed with a 100-fold excess of cold competitor oligonucleotide at 63 °C, then labeled primer was added to the hybridization mixture for a second annealing reaction, also at 63 °C. Reverse transcriptase reactions were carried out for 1 hour at 38 °C.

**Construction of strain F-Alu2**

F-Alu2 was generated by inserting the human Alu element proximal to the 5' end of the human α-2-globin gene (53) into the glycoprotein I (gI) gene of HSV-1 strain F, using a previously described strategy for constructing HSV recombinants (50). A 1.2 kb PstI-NcoI fragment bearing the Alu sequence was subcloned from pStst1b (30) into pUC19 (between the PstI and HincII sites), then excised as Pst I-Hind III fragment. This Alu fragment was then inserted into a polylinker that had been introduced into the BamI site in gI coding sequences in pSS17 (32). The resulting plasmid was contransfected into Vero cells with DNA from HSV-1 strain FgDβ (32), which bears a lethal deletion substitution mutation of the essential gene encoding glycoprotein D (gD). Viable gD+ plaques were then screened for acquisition of gD and Alu sequences by Southern hybridization (57).
Direct sequencing of Alu RNAs

RNA was isolated as described for primer extension analysis. Primers (5'-end labeled, 1 X 10^6 Cerenkov cpm in 0.10 to 0.50 ng of oligonucleotide per sequencing reaction) were hybridized to RNA at 63 °C. Sequencing reactions were carried out using the same procedure as primer extension analysis, except that the dNTP concentrations were altered in the reverse transcriptase buffer. Reverse transcriptase sequencing buffer contained 0.33 mM of three dNTPs, 0.1 mM of the fourth, and 0.2mM of the equivalent dideoxyNTP. Primer Alu2 was used to sequence RNA produced by the HSV-1 strain F-Alu2 and primers Alu2 and Sb were used to sequence HeLa cell RNAs.
Results

Primer extension assay for transcripts arising from various Alu subfamilies.

We wished to determine whether Ad5 and HSV-1 activate expression of Alu elements belonging to multiple subfamilies. To this end, we generated oligonucleotide primers designed to hybridize to Alu transcripts in a region containing sequence differences diagnostic for the S, Sb, and Sb1 Alu subclasses (figure 2a), then tested these for specificity using competition primer extension assays (figure 2b). Primers S, Sb, and Sb1 are complementary to residues 91 to 115 of the S, Sb, and Sb1 subfamily consensus sequences. The Sb primer differs from the S primer at two positions, and the Sb1 primer bears 2 additional changes relative to Sb (i.e., it differs from S at 4 positions). The deg primer is complementary to residues 98 to 122 of the S consensus sequence, and is identical to the S primer in the 15 nt region of overlap (except for the indicated degeneracy at position 115, included to maximize the number of elements detected). Primer Alu1 is complementary to residues 98 to 122 of an Alu sequence linked to the human α-2-globin gene (53), and diverges considerably from all of the other primers used.

We tested the specificity of the deg, S, Sb and Sb1 primers by determining whether pre-annealing the template RNA with a 100-fold excess of cold competitor oligonucleotide interfered with the ability of
Figure 1. Relationship between Alu subfamilies identified by Jurka and coworkers.
labeled primers to generate \textit{Alu} cDNAs in primer extension assays (figure 2b). RNA isolated from HeLa cells infected with Ad5 was used in these assays, because it is an abundant source of \textit{Alu} pol III RNAs which can be detected using the deg oligonucleotide primer (39). Controls demonstrated that each primer gave rise to a prominent extension product of the size predicted for \textit{Alu} pol III transcripts, that preincubation with excess unlabelled oligonucleotide complementary to 7SL RNA caused only a modest decrease in signal intensity, and that homologous competitor oligonucleotide virtually eliminated the signal. The various \textit{Alu} oligonucleotides displayed a highly specific pattern of competition under the conditions used. Thus the deg and S primers were competed only by the deg and S oligonucleotides (both designed to detect S subfamily transcripts), and the Sb and Sb1 primers were competed only by the corresponding homologous oligonucleotide. The failure of the S and Sb1 competitors to block the Sb primer illustrates that pairs of oligonucleotides differing by only two residues did not compete under the stringent hybridization conditions used. Taken in combination, these results indicate that the deg, Sb and Sb1 primers hybridize to non-overlapping sets of \textit{Alu} transcripts, and suggest that these primers provide adequate specificity to distinguish between RNAs of different subfamilies.

\textbf{Multiple \textit{Alu} subfamilies are activated upon infection with Ad5 and HSV-1.}

The results presented in the preceding section established that the
Figure 2. Competition between various Alu oligonucleotide primers. (A) Comparison of sequences of primers designed to detect Alu RNAs from various subfamilies. (B) Primer extension analysis was carried out using 10 μg of cytoplasmic RNA from Ad5 infected HeLa cells (25 PFU per cell, 24 hour duration of infection). RNA was first hybridized to a 100-fold excess of cold competitor oligonucleotide, then hybridized to labeled primer and treated with reverse transcriptase. The resulting cDNAs were displayed on an 8% acrylamide sequencing gel. The labeled primer used is indicated on the right of each panel. Sizes of molecular weight markers, in nucleotides, are indicated on the left.
A

<table>
<thead>
<tr>
<th></th>
<th>+122</th>
<th>+98</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alu1</td>
<td>5'-TTAGTATAAAC</td>
<td>TGGGGTTTCT</td>
</tr>
<tr>
<td>deg</td>
<td>5'-TTAGTAGAGA</td>
<td>SGGGGTTTCA</td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>CCATA CCATG</td>
</tr>
<tr>
<td>Sb</td>
<td></td>
<td>CCGGTGGGCA CA</td>
</tr>
<tr>
<td>Sb1</td>
<td></td>
<td>CCGGTGGGCA CA</td>
</tr>
</tbody>
</table>

S = C or G

B

<table>
<thead>
<tr>
<th>cold competitor</th>
</tr>
</thead>
<tbody>
<tr>
<td>deg</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>124</td>
</tr>
<tr>
<td>112</td>
</tr>
</tbody>
</table>

labeled primer

deg

S

Sb

Sb1
deg, Sb and Sb1 primers detect largely non-overlapping sets of *Alu* transcripts, and that the S primer hybridizes to largely the same set of transcripts as the deg primer. We therefore used these primers to determine if *Alu* transcripts belonging to the S, Sb and Sb1 subfamilies were induced upon infection with Ad5 and HSV-1. The S, deg, Sb, and Sb1 primers were labeled to the same specific activity and used for primer extension of cytoplasmic RNA isolated from HeLa cells before and after infection with Ad5 and HSV-1 (figure 3a and b). As previously described (39, 40), two classes of extension products were detected when RNA from uninfected cells was analyzed with the deg primer: a smear of high molecular weight cDNA, indicative of *Alu* sequences embedded in transcripts of class II genes, and unique cDNAs approximately 122 nucleotides in length, indicating pol III transcripts (figure 3a). (A similar pattern was observed in the experiment displayed in figure 3b, but the high molecular weight region of the gel is not shown). The S, Sb and Sb1 primers also detected both these species of RNA, but differed from the deg primer in that the cDNAs generated from pol III transcripts were approximately 115 (as was predicted). In addition, the quantity of high molecular weight extension products varied greatly depending on the primer used, and roughly corresponded to the relative abundance of each *Alu* subfamily in the human genome. Thus the deg and S primers produced largest quantities of high molecular weight cDNAs, the Sb primer an intermediate amount and the Sb1 primer produced only barely detectable signal (figure 3a is over exposed to show Sb1 high molecular weight cDNAs). All four primers detected *Alu* pol III transcripts in mock infected cells (figure 3a
Figure 3. Analysis of Alu RNA using subfamily specific primers. 10 µg of cytoplasmic RNA from HeLa cells was analyzed by primer extension using the deg, S, Sb and Sb1 oligonucleotides. Molecular weight marker fragment sizes are indicated in nucleotides. (A) RNA was isolated from mock infected cells or cells infected with Ad5 at an MOI of 25 for 24 hours. (B) Cells were mock infected or infected with 10 PFU per cell of HSV-1 for 12 hours prior to isolation of RNA.
### A

<table>
<thead>
<tr>
<th>deg</th>
<th>S</th>
<th>Sb1</th>
<th>Sb</th>
</tr>
</thead>
<tbody>
<tr>
<td>mock</td>
<td>Ad5</td>
<td>mock</td>
<td>Ad5</td>
</tr>
</tbody>
</table>

![Image](image1.png)

- 629
- 629
- 406
- 311
- 244/242
- 219
- 203
- 192
- 182
- 162
- 149
- 124
- 112
- 79

### B

<table>
<thead>
<tr>
<th>deg</th>
<th>Sb1</th>
<th>Sb</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>mock</td>
<td>HSV-1</td>
<td>mock</td>
<td>HSV-1</td>
</tr>
</tbody>
</table>

![Image](image2.png)

- 124
- 112
and 3b), indicating that members of the S, Sb and Sb1 subfamilies are constitutively expressed in HeLa cells. The intensity of the pol III signal generated with the various primers did not vary as greatly as that arising from pol II-embedded Alu RNAs (table 2), implying that Alu elements belonging to the smaller Sb and Sb1 subfamilies are preferentially expressed relative to the larger S family.

Infection with Ad5 and HSV-1 induced a large increase in the levels of Alu pol III transcripts detected with each of the primers. In contrast, the high molecular weight signal was not altered by Ad5 infection (figure 3a), and was reduced by HSV-1 infection, though the reduction was variable between experiments (see figure 1 chapter 4 for one example). This latter effect presumably arises as a result of HSV-1-mediated degradation of cellular mRNAs. Quantitation of the 115 and 122 nucleotide cDNAs indicated that RNAs detected with each of these primers were induced 30-50 fold (table 3), indicating that all subfamilies were activated to a similar degree.

**Direct Sequence Analysis of S and Sb subfamily Alu RNAs**

The preceding analysis indicated that Alu elements belonging to the S, Sb, and Sb1 subfamilies are activated by Ad5 and HSV-1, but gave no information about the number of elements within each subfamily that are induced. This question is particularly relevant in the case of the large and heterogenous S subfamily. In order to determine if the induced S subfamily transcripts are representative of the full range of sequence variants present
Table 2: Comparison of quantities of different classes of Alu RNA in mock infected HeLa cells

<table>
<thead>
<tr>
<th>Primer</th>
<th>counts in pol II cDNA</th>
<th>counts in pol III cDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>deg</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Sb</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>Sb1</td>
<td>0.9</td>
<td></td>
</tr>
</tbody>
</table>

Counts were determined by PhosphorImager (Molecular Dynamics) analysis of cDNAs.

Table 3: Levels of induction of steady state RNA from various Alu subfamilies.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Infection</th>
<th>counts in infected cell cDNA</th>
<th>counts in mock infected cell cDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>deg</td>
<td>HSV-1</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>deg</td>
<td>Ad5</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>HSV-1</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Ad5</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Sb</td>
<td>HSV-1</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Sb</td>
<td>Ad5</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Sb1</td>
<td>HSV-1</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Sb1</td>
<td>Ad5</td>
<td>41</td>
<td></td>
</tr>
</tbody>
</table>

Counts were determined by PhosphorImager (Molecular Dynamics) analysis of 115 or 122 nucleotide cDNAs.
**Figure 4.** Structure of F-Alu2. (A) Diagram of the structure of F-Alu2 containing a human Alu element the HSV-1 gI locus. F-Alu2 was generated by rescue of the lethal deletion/substitution of the gD gene in the viral strain FgDβ by recombination with pgDAlu2, a plasmid containing the gD gene and the neighbouring gI gene disrupted by an Alu element. (B) Southern blot analysis of F-Alu2 DNA. Insertion of the Alu element created a new SphI site, resulting in the loss of the 5.2 kb gD SphI fragment and the production of two new fragments of 3.9 and 1.2 kb.
A

B

pSpgD  pgDAlu2  F-strain  fgDβ  F-Alu2

5.2
4.0
3.9
1.2
in this class of elements, we directly sequenced the \textit{Alu} transcripts present before and after infection with Ad5 and HSV-1 by adding dideoxynucleotides to primer extension reactions. A test of the feasibility of this technique was the primer extension sequencing of transcripts from a single \textit{Alu} element. Thus we first generated HSV-1 strain F-Alu2 bearing the \textit{Alu} element proximal to the 5' end of the human \(\alpha\)-2 globin gene (53) inserted into the \(gI\) locus (figure 4a), using a previously described system for generating viral recombinants (32, 50). Insertion of 1.2 kb of \textit{Alu} DNA into the \(gI\) gene introduced a new Sph I site into the 3.9 kb HSV F-strain \(gD/gI\) Sph I fragment, producing two novel Sph I fragments of 4.0 and 1.2 kb (figure 4b). Previous work has shown that the \textit{Alu} element present in this vector is expressed to high levels from its pol III promoter during lytic infection (38). RNA was first analyzed by primer extension using homologous (\textit{Alu}2) and heterologous (\textit{Sb}) oligonucleotides to verify that F-Alu2 infected BHK21 cells provide an abundant source of transcripts that were specifically detected by the \textit{Alu}2 primer (figure 5a). RNA isolated from mock infected or HSV-1 strain F infected BHK21 cells did not provide templates for cDNA synthesis using either primer. In contrast, the \textit{Alu}2 primer, but not the \textit{Sb} primer, detected RNA from F-Alu2 infected cells, indicating that these oligonucleotides display the appropriate template specificity. Using chain terminating nucleotide analogues to sequence F-Alu2 RNA, it was possible to read sequences from +4 to +94 (figure 5b). A parallel reaction without dideoxynucleotides was run to identify regions of reverse transcriptase pausing in the absence of chain terminators. The \textit{Alu}2 RNA sequence matched the DNA sequence of this element at the 90 nucleotides that could
Figure 5. Testing the feasibility of reverse transcriptase sequencing using F-Alu2. (A) Cytoplasmic RNA (10 μg) was isolated from BHK21 cells infected with F-strain or F-Alu2, or mock infected and subject to primer extension analysis using Alu2 and Sb oligos. Molecular weight marker sizes are indicated in nucleotides. (B) Cytoplasmic RNA was analyzed by reverse transcriptase dideoxynucleotide sequencing, using 5'-end labeled Alu2 primer. Each sequencing reaction was carried out using 10 μg of RNA and lanes are labeled according to the dideoxynucleotide which was used. RNA was isolated from BHK21 cells infected for 12 hours with 10 PFU per cell of the HSV-1 strain F-Alu2.
be read and the sequence was unambiguous at all positions other than sites of spontaneous pausing.

The *Alu2* primer is identical to the deg primer at 24 of 25 positions, and is predicted to detect S subfamily sequences. As expected, primer extension analysis of HeLa cell RNA isolated before and after infection with Ad5 and HSV-1 gave rise to patterns that were very similar to those observed with other S subfamily primers. Thus both viruses strongly induced the 122 nucleotide signal corresponding to *Alu* pol III transcripts (see differences of intensity of signals in figure 7). We used the *Alu2* primer to sequence S subfamily transcripts for a number of reasons. First, different primers varied in the amount of spontaneous pausing, background, and purity (relative contribution of 24-mers, which were virtually impossible to purify away from 25-mers and generated two staggered, overlapping sequencing ladders) and *Alu2* had already been used successfully for reverse transcriptase sequencing indicating that it was free of most of these problems. In addition *Alu2*, unlike deg, was not degenerate at any position, increasing the possibility of detecting homologous rather than heterologous transcripts. Finally, *Alu2* produced cDNA ladders 7 nucleotides longer than S and therefore had to potential to yield more sequence information.

In contrast to the results obtained using homogenous *Alu* RNA, the sequence produced using uninfected HeLa cell RNA was degenerate at a number of positions, suggesting that a diverse group of transcripts were
Figure 6. Comparison of Sb and S subfamily consensus sequences. Italics indicate diagnostic sequence differences between subfamilies. Dashes at positions +65 and +66 indicate sequences deleted in the conserved subfamily sequence. The locations of primers Alu2 and Sb are indicated.
detected using the Alu2 primer (figure 7a). Despite this degeneracy, it was possible to read sequences from +55 to +90 and in this region the sequence at the most abundant nucleotide at the majority of positions matched that of the S subfamily consensus. At three diagnostic positions which differ between the S and Sb subfamily consensus sequences (figure 6) the cDNA sequence matched the S subfamily consensus: position +88 of the cDNA produced using the Alu2 oligo was consistently a C, position +78 was consistently an A, and the GA present at positions +65 and +66 are indicative of S subfamily species as they are absent in the Sb subfamily consensus sequence. The amount of sequence heterogeneity increased after position +60, perhaps due to the significant proportion of S subfamily members which show single nucleotide deletions between positions +62 and +67 (23). Therefore the Alu2 primer detects a diverse group of transcripts which basically align with the S subfamily consensus sequence, suggesting that this primer detects RNA species of the S subfamily and that the sequence heterogeneity seen upon alignment of individual members of the S subfamily is directly reflected in the degeneracy of S subfamily RNA sequences.

The general profile of sequences produced using the Alu2 primer did not vary between mock, HSV-1 and Ad5 infected cell RNA preparations (figure 7a-e), indicating that there was no detectable difference in the composition of Alu transcripts in infected and uninfected cells. Uninfected HeLa cell RNA consists largely of pol II-embedded Alu transcripts, and therefore presumably represent a relatively random sample of Alu
Figure 7. Sequencing of Alu RNA using the primer Alu2. Cytoplasmic RNA was used for reverse transcriptase sequencing of Alu transcripts which hybridized to the Alu2 primer. Lanes are labelled with the dideoxynucleotide incorporated. Arrows and brackets indicate the position of residues which are diagnostic of the S subfamily sequence. The brackets flank two residues which are deleted among conserved subfamily members. (A) 50 μg of RNA isolated from mock infected HeLa cells was used for each reaction. (B) RNA from HeLa cells infected with HSV-1 at an MOI of 10 was harvested 12 hours post infection, and 10 μg of RNA was analyzed using each dideoxynucleotide. (C) RNA from HeLa cells infected with 25 PFU per cell of Ad5 was isolated 24 hours post infection, and 10 μg of RNA was used in each reverse transcriptase reaction.
sequences. In contrast, RNA from Ad5 and HSV-1 infected cells contains a mixture of pol II and pol III transcripts of S subfamily elements. Therefore the similarity in sequence profile of Alu RNAs present in mock and virus infected HeLa cells suggests that virally-induced Alu RNAs represent an essentially random sampling of those present in the human genome.

In order to test the validity of the conclusion that the sequence degeneracy of the S subfamily of Alu elements is reflected in the heterogeneous mixture of RNA detected using the Alu2 primer, we sequenced the transcripts detected using the Sb primer. The Sb subfamily is smaller and shows greater sequence conservation among its members (23). The cDNA sequence ladder produced using the Sb primer demonstrates this homogeneity: sequencing of RNA from mock infected, Ad5 infected or HSV-1 infected HeLa cells with the Sb oligonucleotide revealed a family of transcripts which showed little sequence degeneracy (figure 8). The sequences produced from RNA isolated from mock infected or virus infected cells were the same, indicating that the Sb subfamily Alu sequences activated by Ad5 and HSV-1 infection do not differ dramatically from those transcribed in mock infected cells. The Sb primer is homologous to sequences 7 nucleotides closer to the initiation of Alu pol III transcription than Alu2, and cDNA sequences consistently matched the Sb subfamily consensus sequence from positions +15 to +82. Two diagnostic Sb subfamily sequence characteristics were detected: the T residue present at position +78 of the Sb cDNA and 2 nucleotide deletion present in Sb subfamily RNAs (deletion of +65 and +66 of the S subfamily consensus).
Figure 8. HeLa cell cytoplasmic RNA was analyzed by reverse
transcriptase sequencing using the Sb primer. Lanes are labeled according
to the dideoxynucleotide used. Arrows and open triangle indicate sites of Sb
subfamily diagnostic residues. The open triangle indicates the site of a 2
nucleotide deletion relative to the major subfamily consensus sequence. (A)
Each track represents the cDNAs produced by analysis of 50 μg of mock
infected cell RNA. (B) RNA was isolated from Ad5 infected cells (25 PFU
per cell) 24 hours post infection, and 10 μg was used for each reaction. (C)
HSV-1 infections were carried out at an MOI of 10 for 12 hours, and 10 μg of
RNA was used for sequencing reactions.
Thus the \textit{Alu}2 and Sb primers detect mixtures of transcripts which differ in primary nucleotide sequence and degree of degeneracy and directly reflect the predicted nucleotide sequence and amount of heterogeneity of each subfamily. This sequencing data reinforces the conclusion that a wide variety of \textit{Alu} sequences are transcribed in uninfected cells, and that infection with HSV-1 and Ad5 stimulates expression of an equally broad distribution of \textit{Alu} elements.

\textbf{Multiple subfamilies are induced in multiple cell types}

We used primer extension analysis to demonstrate that the S, Sb and Sb1 subfamilies are induced in a variety of cell types. Vero, 293, human fibroblasts and human embryonic kidney cells infected with HSV-1 or Ad5 showed an increase in steady state levels of \textit{Alu} RNAs detected using deg, Sb and Sb1 oligos (figure 9). Vero cells are a transformed african green monkey kidney cell line, indicating that multi-subfamily \textit{Alu} expression can be activated in at least one non-human primate line. 293 cells, like HeLa cells, are transformed human cell lines which express DNA tumour virus gene products. Human embryonic kidney cells and human fibroblasts are strains derived from tissue explants, indicating that multiple \textit{Alu} subfamilies can be induced to high level expression following viral infection of non-transformed cells. Longer exposures show that RNA from all three subfamilies can be detected in mock infected cells (data not shown). Together, these results indicate that activation of \textit{Alu} expression
Figure 9. Levels of Alu RNA from various subfamilies in different cell types. 20 µg of cytoplasmic RNA from the cell type indicated below each panel was analyzed by primer extension using the oligonucleotides indicated. cDNA produced by analysis of RNA from human fibroblasts is labeled HF and HEK indicates that human embryonic kidney cell RNA was used. RNA from HSV-1 infected cells was isolated 12 hours post infection with 10 PFU per cell of HSV-1, while Ad5 infections were carried out at an MOI of 25 for 24 hours prior to RNA harvest. Sizes of molecular weight markers are indicated in nucleotides.
Discussion

In this paper we demonstrate that Ad5 and HSV-1 infection of a variety of human cultured cells activates expression of members of the S, Sb and Sb1 subfamilies of Alu repetitive elements. These subfamilies represent successive steps in Alu family evolution: the S subfamily is old and shows considerable sequence heterogeneity among its members, the Sb subfamily is younger, smaller and shows greater sequence identity among its members, while Sb1 is among the youngest and smallest subfamilies in the genome and includes currently transpositionally active elements (3, 23, 36). Using cytoplasmic RNA from mock infected cells, pol III transcripts from the younger Sb and Sb1 Alu subfamilies were found to be more abundant relative to genomic copy number than those from the older S subfamily, indicating that in uninfected cells members of the Sb and Sb1 subfamilies are expressed at higher levels than those of the S subfamily. The increase in quantity of Alu pol III RNAs induced by infection did not vary dramatically between subfamilies, indicating that Alu subpopulations do not differ in the magnitude of their response to viral transactivators.

Sequencing of Alu RNAs demonstrated that an oligonucleotide primer specific for the S subfamily detected a heterogeneous mixture of RNAs, while the Sb subfamily primer detected a relatively homogenous group of transcripts, indicating that these primers detected species of RNA with roughly the same degree of sequence heterogeneity as the subfamilies they were designed to detect. Using each primer to sequence mock infected or virus infected RNA showed that the profile of cDNAs remained the same
regardless of whether the RNA was isolated from mock, HSV-1 or Ad5 infected cells. Thus pol III transcribed Alu RNAs induced upon Ad5 or HSV-1 infection show the same composition of sequences as the presumably random mixture of pol II-embedded transcripts, indicating that infection activates expression of a wide variety of Alu sequences rather than a unique subset.

Though the cDNA sequences produced from Alu RNA showed heterogeneity, particularly using the Alu2 oligo, it was possible to determine which residue was most abundant at many positions. Using the Sb oligo, it was possible to read sequences from +82 to +15, and in this region residues matched the predicted sequence of the Sb subfamily consensus. The cDNA sequence produced using the Alu2 oligo could be read from positions +90 to +55, and though there was greater variability at a number of positions, where sequences could be unambiguously read they matched the S subfamily consensus. Diagnostic differences between the S and Sb subfamily consensus sequences were detected using the Alu2 and Sb primers, indicating these oligos detect RNAs of the subfamily for which they were designed. These data support the conclusion that viral activation of Alu expression is a relatively global effect, acting on a significant number of elements in the genome.

The mechanisms which maintain the majority of Alu elements in a transcriptionally silent state are not well understood. In addition to the internal pol III promoter, it is thought that Alu elements require cis-regulatory motifs provided by flanking cellular DNA and that only a limited of Alu repeats are expressed due to fortuitous insertion adjacent to these
motifs (23, 36). Our data indicate that many Alu elements are capable of being transcribed from their pol III promoters, suggesting that viral induction of Alu expression is not mediated through regulatory sequences present in a limited set of Alu elements. It seems more probable that regulatory motifs required for Alu induction are present in a large number of Alu sequences, and that these motifs are not associated with particular Alu subfamilies. Though the intragenic class III promoters present in the majority of Alu elements fit these criteria, Ad5-induced transcriptional activation of Alu sequences is not mediated by increases in amount or activity of the rate limiting pol III transcription factor TFIIC, suggesting that other factors are involved in regulation of Alu expression (39). Alu elements also contain motifs which allow interaction with a number of proteins, including the transcription factor RSRF (43), methyl-C binding proteins (5), the SV40 T-antigen (21), and a number of as yet unidentified proteins some of which have been suggested to be high mobility group proteins and subunits of human chromatin assembly factor (48, 58). None of these motifs or their binding factors have been demonstrated to regulate Alu or other class III transcription in vivo, or interact with the Ad5 regulatory proteins which are required to mediate Ad5-induced Alu transcriptional activation. Only the identification of specific, individual Alu sequences which are reproducibly induced after viral infection will begin to provide answers as to which regulatory motifs are necessary for transcriptional activation.

Epigenetic effects, such as DNA methylation and nucleosome positioning are also thought to regulate Alu expression (12, 17, 26). Alu
elements have a high concentration of CpG dinucleotides and methylation inhibits transcription of these sequences \textit{in vitro} (17, 26). \textit{In vivo}, \textit{Alu} repeats show disproportionately high levels of methylation in most cell types, suggesting that their transcription may be modulated by DNA methylation (11, 25, 46). \textit{Alu} templates reconstituted with nucleosomes are transcriptionally repressed \textit{in vitro}, indicating that structure of chromatin in which \textit{Alu} elements are embedded may also determine whether they are expressed (12). Allan and Paul report that transcripts from a \textit{Alu} element distal to the human \textepsilon-globin gene accumulate \textit{in vivo} (1). These transcripts, which initiate at the \textit{Alu} pol III promoter, are only detectable in association with the expression of the \textepsilon-globin gene, implying that the transcriptional status of neighbouring genes may affect \textit{Alu} expression. Thus it is possible that HSV-1 or Ad5 may induce alterations in cellular chromatin structure, which result directly or indirectly in demethylation or alteration of nucleosomal patterns and incidental \textit{Alu} transcriptional activation. Certainly, the heterogeneity of \textit{Alu} elements which are induced upon infection suggests that a more global phenomenon may be responsible for their transcriptional activation.

Expression of the S, Sb and Sb1 subfamilies was detected in multiple cell types, including non-transformed cell strains, indicating that the transformed phenotype was not a prerequisite for \textit{Alu} expression. The cell strains examined, embryonic kidney cells and adult fibroblasts, are not pluripotent or of germline derivation, supporting earlier observations that \textit{Alu} expression is not limited to germ cells (1, 26, 34, 36, 39, 54). HSV-1 and Ad5 were able to stimulate expression of members of all three subfamilies
in all cell types tested, indicating that activation of \textit{Alu} transcription may be a general cellular response to viral infection. Sakamoto and Howard report that pol III transcription of \textit{Alu} sequences transfected into HeLa cells inhibits cellular proliferation (19), suggesting that \textit{Alu} transcription may be involved in regulation of cell cycle progression. \textit{Alu} transcriptional activation may be an indirect result of viral perturbations of cellular growth signaling pathways. Alternatively, it could be part of a host response to minimize cellular growth and limit viral replication.

The observation that a wide variety of \textit{Alu} elements appear to be transcriptionally competent, even in the absence of viral \textit{trans}-acting factors provides a novel approach to the study this unusual class of DNA sequences. It should be possible to use representational difference analysis (33) on cDNAs produced from mock and virally infected cells to identify individual elements which are activated during infection. Sequence comparisons of these \textit{Alu} elements and the flanking cellular DNA may reveal \textit{cis}-regulatory motifs required for induction by viral infection. Identification of the elements which direct expression of \textit{Alu} sequences may contribute to our understanding of the mechanisms which determine the mobility of \textit{Alu} elements.
References


6.2. Segue

The results presented in this paper demonstrate that HSV-1 and Ad5 infection increase levels of cytoplasmic RNA from three Alu subfamilies, S, Sb and Sb1. The Sb1 subfamily is among the smallest in the human genome and members of this subfamily are mobile (Jurka and Milosavljevic, 1991; Wallace et al., 1991). The founder element(s) of the Sb1 subfamily were probably derived from the Sb subfamily, which is larger and is thought to be transpositionally inactive (Jurka and Smith, 1988). The S subfamily contains the largest proportion of Alu sequences present in human DNA and is thought to have produced the Sb founder element(s) but has probably long been retropositionally inactive (Jurka and Smith, 1988). Expression of members of all three subfamilies was detected in uninfected cells and was induced by viral infection, indicating that there is no correlation between mobility and levels of induced or uninduced transcription. Infection increased levels of S, Sb and Sb1 RNA in all cell types examined, including primary cell strains. Thus, Alu transcriptional activation in response to infection with HSV-1 or Ad5 is not a phenomenon unique to transformed cell types, such as HeLa, and may be part of a general cellular response to viral infection.

Sequencing of Alu RNA demonstrated that the species detected by S and Sb subfamily specific oligonucleotide primers showed the characteristics which would have been predicted based on the size and sequence heterogeneity of these subfamilies. S subfamily RNAs were a heterogeneous mixture, the predominant proportion of which matched the S subfamily consensus sequence, while the Sb subfamily transcripts were
much more homogeneous and the large majority matched the subfamily consensus sequence. Comparison of sequences of infected and uninfected \textit{Alu} S and Sb subfamily sequences revealed that infection did not detectably alter the appearance of RNA sequencing ladders, suggesting that infection induced expression of a relatively broad spectrum of \textit{Alu} sequences. Thus, HSV-1 and Ad5 must activate \textit{Alu} transcription by a mechanism that can on a large number of elements, distributed throughout the human genome and which show 14\% sequence variation. These result argue in favour of the role of abundant, general transcription factors rather than more specialized sequence specific factors.

\textit{Alu} elements are not unique among cellular sequences in their ability to be activated by HSV-1 or Ad5 infection. Ad5 and HSV-1 proteins can stimulate expression of a number of cellular genes when they are introduced into cells in transient assays or as part of the viral genome (reviewed in Everett, 1987; Smiley \textit{et al}., 1990; Flint and Shenk, 1989; Graham and Prevec, 1992). In addition, Ad5 infection activates transcription of a number of genes from their cellular loci, including \textit{HSP70} and \textit{FOS} (Nevins, 1982; Wu \textit{et al}., 1985; Sassone-Corsi and Borelli, 1987; Rosahl and Doerfler, 1992), while HSV-1 stimulates stress protein and p40 expression and transcription of \textit{AP1} and the U4 SnRNA gene (Notarianni and Preston, 1982; Eastridge \textit{et al}., 1989; Latchman, 1991; Eastridge \textit{et al}., 1989). The following chapter describes another addition to this list. We have observed an unusual instance of HSV-1 and Ad5 induced activation of tissue specific gene expression: \textalpha-globin sequences are transcribed in non-erythroid cells following viral infection.
Chapter 7: Expression of human α-globin, but not β-globin, is induced by herpes simplex virus type 1 and adenovirus type 5 infection.
7.1 Forward

This study of the expression of the endogenous human α- and β-globin genes in adenovirus type 5 (Ad5) and herpes simplex type 1 (HSV-1) infected cells began as a series of control experiments. I had been using Vero cells to determine levels of expression of the human α-globin gene and two linked Alu elements which were present on an HSV-1 recombinant virus. As a control I infected Vero cells with wild type virus and assayed for α-globin RNA, and found that endogenous α-globin sequences were expressed. The identification of HSV-1 gene products which were required for activation of α-globin gene expression became a secondary focus of my Ph. D. thesis, carried out in conjunction with the Alu project. A new graduate student, Peter Cheung, has reproduced and continued much of this work.
7.2 Introduction

The globin genes are among the best characterized tissue specific genes and the developmentally regulated expression of the α-like and β-like globin gene clusters provides a good model for the study of coordinate regulation of unlinked loci. The human α-globin gene is part of a cluster of three functional α-like globin genes, located near the tip of chromosome 16p, while the β-globin gene is one of five related genes located within the terminal band of chromosome 11p. Both genes are expressed only in cells of the erythroid lineage: α-globin expression begins very early in development, at 5 to 6 weeks post conception, while β-globin gene is fully activated only in the final stages of prenatal development (reviewed in Bunn and Forget, 1986). Structural and functional comparisons of α- and β-globin genes has revealed a number of important differences and similarities that may explain their temporally distinct, erythrocyte-specific developmental regulation.

The globin promoters are characterized by binding sites for tissue specific and ubiquitous trans-acting factors. The TATA, CCAAT and CACC box motifs interact with widely expressed transcription factors that presumably facilitate efficient assembly of the basal transcriptional machinery (Evans et al., 1990; Raich and Romeo, 1993). The largely erythroid-restricted zinc finger DNA binding protein GATA-1 binds the sequence WGATAR, present in the promoters of a number of erythroid specific genes, and this interaction is crucial for normal erythroid
differentiation (Orkin, 1992). In stable transformants of mouse erythroleukemia cells human α-globin genes are expressed constitutively while β-globin gene products are synthesized only after terminal differentiation and in both instances levels of expression vary with site of insertion and are generally much lower than those of the endogenous globin genes (Charnay et al., 1984). In human X MEL cell hybrids both globin genes can be induced to levels approximating that of the endogenous mouse genes (Forrester et al., 1987). Taken in combination these observations imply that distant sequences are required for correct developmental regulation and high level expression of human globin genes. HS-40, a tissue specific DNAse I hypersensitive site located 40 kb upstream of the α-globin cluster has been identified as such a cis-regulatory region (Higgs et al., 1990; Jarman et al., 1991). In the β-globin cluster four elements (β-HS1-4), collectively referred to as the β-locus control region (LCR), are located 6 to 18 kb 5' of the globin coding sequences (Tuan et al., 1985; Forrester et al., 1986). These elements confer tissue specific, position independent expression on linked genes in transgenic mice, and naturally occurring deletions of these regions inactivate globin expression in vivo (Grosveld et al., 1987; reviewed in Grosveld et al., 1993). The position independent transcription conferred by the α- or β-LCR is dependent on sequences containing GATA, AP-1-like and CACC motifs. The AP-1-like site recruits a transcription factor, NF-E2, present only in erythroid, megakaryocytic and mast cells (Andrews et al., 1993). Naturally occurring deletions of the β-LCR affect on long range chromatin structure of the β-
cluster, suggesting that it may have a chromatin opening function (Tuan et al., 1985).

Though there is no dramatic difference in the cis-regulatory motifs present in the promoters and far distal LCR of the globin genes, the α- and β-globin clusters are different their higher order structure. The β-globin gene cluster, like most other developmentally regulated genes, shows tissue specific alterations in chromatin structure and timing of DNA replication: in cells of the erythroid lineage, the β-like cluster is not extensively methylated and DNA replication occurs early in S phase, in other cell types the β-globin genes are methylated and late replicating (Forrester et al., 1986; Epner et al., 1988; Dhar et al., 1989; reviewed in Townes and Behringer, 1990). The β-globin regulatory domain is further characterized by 8 nuclear matrix attachment regions (Jarman and Higgs, 1988). In contrast, the α-globin genes are associated with unmethylated CpG-rich islands and replicate early in all cell types (reviewed in Higgs et al., 1989) and no nuclear matrix attachment sites have been identified in the α-globin region (Jarman and Higgs, 1988). The α-globin cluster lies in a region of constitutively open chromatin and HS-40 is located within an intron of a widely transcribed gene (Vyas et al., 1992). Thus, the developmentally regulated expression of the globin genes appears to involve two types of events: the synthesis of tissue-specific transcription factors and the alteration of chromatin structure. The α-globin and β-globin genes appear to differ in manner in which an open, transcriptionally active chromatin conformation is established and maintained.
Though there has been considerable alteration of globin sequences since the avian and mammalian lineages have diverged, the chicken \( \beta \)-globin gene has provided a good model system to study the role of chromatin structure in the modulation of gene expression. Tissue and developmental stage-specific DNAase I hypersensitive sites can be generated on the promoter and 3' enhancer of the cloned chicken \( \beta \)-globin gene using \textit{in vitro} chromatin assembly systems and reconstituted chromatin is transcribed \textit{in vitro} in a tissue and stage specific manner, suggesting that chromatin reconstitution techniques can be used to generate nucleosomal structures resembling those observed \textit{in vivo} (Emerson \textit{et al.}, 1989). Recombinant GATA-1 and NF-E4 were sufficient to assemble \( \beta \)-globin genes into erythroid stage specific transcriptionally active chromatin using non-erythroid reconstitution extracts, suggesting that tissue specific expression of reconstituted \( \beta \)-globin chromatin templates is determined by erythroid restricted transcription factors (Barton \textit{et al.}, 1993). Since recombinant TBP was able to derepress reconstituted chromatin if the GATA-1 motif (situated at -30) was mutated to a canonical TATA box, it is possible that the components of the basal transcription machinery are also involved in altering chromatin into, or maintaining it in, the transcriptionally active state. In addition these observations suggest that tissue specific transcription factors, like GATA-1, may play a similar role in determining the transcriptional state of chromatin (Barton \textit{et al.}, 1993).
Despite the requirements for tissue specific transcription factors for developmentally regulated expression, the globin genes can be transcribed in non-erythroid cells. The human α-globin gene can be transcribed in COS cells in transient assays, while the β-globin gene requires an enhancer for high level non-erythroid expression (Humpheries et al., 1982; Treisman et al., 1983). The globin genes are also differentially expressed from Ad5 vectors: α-globin is transcribed in the absence of the viral transactivator E1a, while the β-globin gene requires E1a proteins, or a viral enhancer provided in cis, for expression (Hearing and Shenk, 1985; Karlsson et al., 1986). Thus the α-globin gene is transcriptionally competent when introduced into cells as part of an Ad5 genome or as naked DNA in transient transfections, while the β-globin requires additional regulatory sequences or transacting factors. Since the α-globin gene can be transcribed in non-erythroid cells all the factors required for α-globin expression must be present in these cells. Therefore, it seems probable that more distal regulatory sequences, such as HS-40, control the developmental timing of α-globin expression in erythroid cells and that transcription of the cellular copies of the α-globin gene is actively repressed in non-erythroid cells (Chernay et al., 1984). The human α-globin present in HSV-1 recombinants required the viral trans-acting factor ICP4 for detectable expression (Smiley and Duncan, 1992) suggesting that in the context of the HSV-1 genome the α-globin gene is also transcriptionally repressed and that ICP4 relieves this repression. Though the original purpose of assaying α-globin expression in Vero cells infected with the HSV-1 strain PAAr5 was not to test the hypothesis that infection might derepress expression of the
cellular copies of the α-globin gene, in retrospect, this seems as good a reason as any for doing the experiment.

In the experiments described in this chapter I demonstrate that HSV-1 infection activated expression of the endogenous Vero and HeLa α-globin genes and that the HSV-1 regulatory protein ICP0 is necessary and sufficient for this process in HeLa cells. In addition infection with Ad5 also stimulated expression of the α-globin gene in HeLa cells and that E1a gene products were required to induce production of α-globin RNA. The endogenous copies of the β-globin gene were not transactivated in HeLa cells, providing another example of the disparate regulation of these two genes.

7.3 Methods and Materials

Cells and Virus

HeLa, Vero, BHK21, E5 (DeLuca et al., 1985), 3-3 (McCarthy et al., 1989) and 293 (Graham et al., 1977) cells were maintained in α-minimal essential medium (Gibco) supplemented with 5% fetal bovine serum. All Ad5 strains were propagated and titred on 293 cells. HSV-1 strains PAA5 (Hall et al., 1984), dlx3.1 (Sacks and Schaffer, 1987), F, R325tk+ (Post and Roizman, 1982) and N38 (Umene, 1986) were grown on Vero cells. The HSV-1 strain d120 (DeLuca et al., 1985) was carried on the complementing
cell line E5 and the strain 5dl1.2 (McCarthy et al., 1989) was grown and
titred on 3-3 cells.

Primer extension analysis

Cytoplasmic RNA was isolated by the method of Berk and Sharp
(1977) and primer extension analysis was carried out as described
previously (Smiley et al., 1987). Oligonucleotides were synthesized by the
Central Facility of the Institute for Molecular Biology and Biotechnology at
McMaster University. The α-globin primer, 5'-
AGGGCGGCCCTTGACGTGGTCTTGTC-3', was designed to hybridize
from +55 to +80 of the human α-globin transcripts. The β-globin primer, 5'-
ACAGGGCAGTAACGCCCAGACTTCTC-3', is homologous to sequences
from +61 to +86 of human β-globin RNA. The is predicted to produce a
nucleotide cDNA product from ICP0 RNA. The E2a primer, 5'-
TCACAGAGATCATGTAATAA-3', is homologous to nucleotides 5 bp to 24
bp 3' of the E2a splice acceptor and produces a 190 nucleotide cDNA. The
Alu primer, 5'-TAAAGTAGAGA(C/G)GGGTTCACCATG-3', is predicted
to give rise to a ca. 122 nucleotide extension product.

7.4. Results

HSV activates α-globin, but not β-globin, expression in HeLa cells

The human α- and β-globin genes both show developmentally
regulated erythroid specific expression and the mechanism by which this
Figure 1. Analysis of globin expression in HeLa cells. Cytoplasmic RNA (20 μg) isolated from mock infected cells or cells infected with PAA5 (10 PFU per cell) was analyzed by primer extension, using 5'-labeled synthetic oligonucleotide primers. The resulting cDNAs were displayed on an 8% sequencing gel. Marker fragment sizes (in nucleotides) are indicated on the right.
stage and tissue specific timing is regulated seems to differ between the two genes. Since I had continued the investigation of the induction of Alu expression in HeLa cells, I had an abundant source of infected HeLa cell RNA which could also be used for the analysis of the expression of human globin genes. To determine whether the human α-globin and β-globin genes in non-erythroid cells were activated by HSV-1 infection, cytoplasmic RNA isolated from HeLa cells infected with PAAr5 was assayed by primer extension using α-globin and β-globin oligonucleotides (figure 1). Both the α- and β-globin primers were able to generate cDNA products of the predicted sizes using RNA isolated from adult human blood (P. Cheung, personal communication). RNA from mock infected cells did not give rise to extension products using either globin oligonucleotide, indicating that there was no detectable expression of either globin gene in HeLa cells. RNA harvested from cells infected for 12 hours with PAAr5 yielded an 80 nucleotide cDNA using the α-globin primer, indicating that HSV-1 infection induces the accumulation of cytoplasmic transcripts initiated at the α-globin promoter. No extension products were detected using the β-globin oligonucleotide, indicating that β-globin expression is not detectably affected by HSV-1 infection. Thus the human α- and β-globin genes, which show similar patterns of tissue specific expression, are not coordinately regulated by HSV-1 infection.

Expression of the α-globin gene peaks early in infection

The time course of α-globin expression was determined by primer extension analysis using samples of HeLa cell RNA isolated at three hour
Figure 2. Primer extension analysis of α-globin RNA at various times post-infection. Cytoplasmic RNA (20 μg) isolated from HeLa cells infected with PAA'5 at an MOI of 10 for the time indicated (hours) was analyzed using the α-globin primer. Infections were carried out in 100 μg/ml of cycloheximide where indicated by + cyclo. The lane labelled tkα displays the cDNA produced by primer extension analysis of RNA isolated from BHK21 cells infected with the recombinant tkα. The sizes of marker fragments, generated by labeling of HpaII cleaved pBR322, are indicated in nucleotides.
intervals for 12 hours after infection with PAA^o5. α-globin expression peaked between 6 and 9 hours post-infection, and dropped off dramatically by 12 hours post-infection (figure 2). The reason for this decrease in α-globin RNA levels late in infection is not clear: it may be a result of increased RNA turnover mediated by the delayed host shut off function (Fenwick and Clark, 1982; Read and Frenkl, 1983) or due to a decrease in rate of globin transcription late in infection. To determine whether a viral infection which was not allowed to proceed into the late phase showed this drop in α-globin expression, levels of α-globin RNA levels were assayed when infections were restricted largely to the early phase (figure 3).

Infections which were carried out using 10 μg/ml aphidicolin, which inhibits DNA synthesis and late gene expression, and RNA was harvested 12 hours post infection. Cells infected with PAA^o5 under an aphidicolin block showed a dramatic increase in abundance of α-globin transcripts relative to those infected without the drug, suggesting that progression into the late phase of infection is necessary for the decrease in α-globin expression seen 12 hours post infection.

**IE gene expression is required to induce α-globin**

The HSV-1 immediate early (IE) protein ICP4 was necessary to activate transcription of α-globin in HSV-1/globin recombinants, since the α-globin gene was not expressed when in was rescued into an ICP4-null background (Smiley and Duncan, 1992). To determine whether the cellular copies of the α-globin gene required ICP4 for transactivation HeLa cells were infected with the strain d120, a virus which contains deletions
Figure 3. Requirement for viral IE and early gene expression. RNA prepared from cells infected under conditions of limited viral gene expression was subject to primer extension analysis using the α-globin specific oligonucleotide. Where indicated, cells were infected with UV inactivated virus (UV-irradiated to produce a 3 log reduction in titre) to eliminate de novo viral gene expression. Viral gene expression is limited to the IE phase during infection with d120, since this virus does not produce ICP4. Infection in the presence of 10 μg/ml aphidicolin, an inhibitor of viral and cellular DNA synthesis, is indicated by + aph. The lane labeled tkα displays the primer extension analysis of α-globin RNA produced by infection of BHK21 cells with the HSV-globin recombinant tkα. Extension products were resolved on an 8% sequencing gel. Size markers were generated by 3'-end labeling PBR322 cleaved with HpaII. Fragment sizes are indicated in nucleotides.
removing both copies of the ICP4 gene (DeLuca et al., 1989), and RNA harvested 12 hours post infection was analyzed for α-globin expression (figure 3). HeLa cells infected with d120 showed slightly reduced levels of α-globin induction relative to PAAr5, indicating that ICP4 expression is not absolutely necessary for the activation of endogenous cellular α-globin expression. The time course (figure 2) clearly indicated that the 12 hour time point was not the ideal time to assay for the highest levels of α-globin expression, however these experiments were carried out primarily to study Alu induction and this was the time of maximal Alu expression. It would be very informative to do a time course with d120 to determine whether levels of α-globin RNA peak early in infection with this virus, and whether the peak is comparable to that seen with PAAr5. No α-globin RNA was produced during infection with UV inactivated virus (figure 3), which had been UV-irradiated to produce a 3-log reduction in viral titre or during infection under a cycloheximide blockade (figure 2), indicating de novo viral protein synthesis was required to activate α-globin expression. Taken in combination these results suggest that the HSV-1 gene products expressed in the absence of ICP4 are sufficient to activate α-globin expression, though they do not eliminate the possibility that gene products which are not expressed during a d120 infection may further stimulate α-globin transcription.

ICP0 is necessary for α-globin expression
d120 infections are limited largely to the IE stage of lytic cycle suggesting that another IE protein may be sufficient for the induction of α-
Figure 4. Role of IE gene products in activation of α-globin expression. Primer extension analysis was carried out on cytoplasmic RNA (20 μg) isolated from HeLa cells infected with 10 PFU per cell of the virus indicated. The strain dlx3.1 contains a deletion removing both copies of the ICP0 gene. The mutant 5dl1.2 fails to produce ICP27, due to a deletion removing the coding sequences for the COOH-terminal two-thirds of the protein. R325tk+ bears a deletion removing the ICP22 coding region and N38 lacks all ICP47 coding sequences. After reverse transcription reactions were complete, α-globin cDNAs were displayed on an 8% sequencing gel. Marker sizes (3'-end labeled, HpaII-cleaved pBR322) are indicated in nucleotides. (A) cDNA produced using the α-globin primer. (B) Extension products using an Alu primer.
globin expression. HSV strains which carried deletions removing coding sequences for the remaining 4 IE gene products were used to infect HeLa cells for analysis of α-globin RNA levels (figure 4a). The interpretation of these results is complicated by the fact that in all instances RNA was isolated at 12 hours post infection: it is not clear when α-globin expression induced by each mutant strain peaks, and differences in levels of globin RNA induced by these viral strains may indicate different time courses of globin induction rather than altered ability to activate expression. The determination of the time course of α-globin expression following infection with each of these viruses is necessary to provide a more convincing demonstration of the role of individual IE gene products. Nevertheless the results of these experiments are worth sharing, since some conclusions could be drawn. Infection with the strains N38 and 5dl1.2, which bear deletions in ICP47 and ICP27 genes respectively (Umene, 1986; McCarthy et al., 1989), resulted in significant activation of globin expression, though slightly less α-globin RNA was produced than seen upon PAAr5 infection. This slight reduction in levels of globin RNA produced during infection with ICP27 and ICP47 mutant strains was not consistent: in other experiments these viral strains activated α-globin expression to the same levels as PAAr5 (P. Cheung, personal communication). The observation that 5dl1.2, the ICP27 deletion mutant, activated α-globin expression to wild type levels was somewhat surprising, since this mutant is deficient for DNA replication and some late gene expression and could be expected to show a hyperinduction similar to that seen when PAAr5 infections were carried out under a DNA replication blockade. The virus R325tk+ (Post and
Figure 5. Analysis α-globin expression induced by Ad/HSV recombinants. HeLa cells were infected with 25 PFU per cell of the virus indicated, and 20 µg of cytoplasmic RNA from each infection was analyzed by primer extension using the α-globin specific oligonucleotide. The virus dIE1,3 contains a large deletion removing E1 and E3 regions. dl312 lacks E1A coding sequences and fails to produce E1A gene products. 0pro-0 contains the HSV ICP0 gene inserted into an Ad5 vector lacking E1 and E3 sequences. AdICP4 bears the HSV ICP4 gene inserted into the E1 region of dIE1,3. Globin cDNAs were displayed on 8% sequencing gels, and marker fragment sizes are indicated in nucleotides.
Roizman, 1982) was reduced in its ability to stimulate \( \alpha \)-globin expression when compared to PAA\(^5\), but showed levels of induction similar to its parental F-strain virus, indicating that strain differences may account for the reduced ability of this ICP22 deletion mutant to activate globin expression. The ICP0 deletion mutant dlx3.1 (Sacks and Schaffer, 1987) was unable to activate \( \alpha \)-globin expression, since no \( \alpha \)-globin RNA was produced in cells infected with this virus.

The samples which were used to assay \( \alpha \)-globin expression were also used to analyse levels of Alu RNA (figure 4b) and the majority of mutant strains showed activation of Alu expression comparable to PAA\(^5\), indicating that there was viral gene expression during infection with each of these mutant viruses. The strain 5dl1.2 did not show the same levels of Alu induction as the other viruses, though a significant increase in levels of Alu RNA was observed, suggesting that the multiplicity of infection with this ICP27 mutant virus may have been different from the other viruses. In the first series of Alu experiments carried out ICP27 mutants showed greater variability in degree of activation of Alu expression, as is clear in figure 5b. It was eventually determined PAA\(^5\) titres are generally one log lower on the ICP27 complementing cell line: normalization of ICP27 mutant titres by this factor resulted in consistently reproducible wild type levels of Alu induction. Most importantly, the strain dlx3.1, another notoriously difficult virus to propagate and titre, was able to stimulate Alu expression to the same levels as PAA\(^5\) indicating that this virus was able to enter cells and initiate IE transcription. Thus, it seems that activation of endogenous cellular \( \alpha \)-globin expression required ICP0.
Figure 6. Comparison of AdICP4 and O-pro0. Cytoplasmic RNA was analyzed by primer extension analysis to determine levels of viral and cellular RNA: 20 μg of RNA was used in each reaction. All Ad5 infections were carried out for 24 hours at 25 PFU per cell. RNA from HSV infections (10 PFU per cell) was isolated after 12 hours. Size of molecular weight markers are indicated in nucleotides. (A) ICP0 RNA levels. (B) Levels of ICP4 RNA. (C) Levels of E2a transcripts.
Ad5 and 0pro-0 activate α-globin expression in HeLa cells

The Ad5 vectors AdICP4 and 0pro-0, which express ICP4 and ICP0 respectively, provided a system could be used to determine whether these HSV-1 IE gene products were sufficient to stimulate α-globin expression in HeLa cells. RNA harvested from HeLa cells infected with Ad5, 0pro-0, AdICP4, dIE1,3 and dl312 was analyzed using primer extension directed by the α-globin oligonucleotide (figure 5, panels a and b). Mock infected cells did not yield detectable amounts of globin RNA, while infection with Ad5 resulted in the accumulation of cytoplasmic α-globin transcripts. HeLa cells infected with the Ad5 mutant strains dl312 and dIE1,3 failed to stimulate α-globin expression. The strain dIE1,3 contains two deletions, removing the E1 and E3 regions, while dl312 lacks the E1a coding sequences (Jones and Shenk, 1979; Haj-Ahmad and Graham, 1986), indicating that E1a gene products are required for Ad5 mediated induction of α-globin expression. 0pro-0, generated by replacing the E1 sequences of an Ad5 vector lacking E3 region with the HSV-1 ICP0 gene (Zhu et al., 1988), was able to induce α-globin RNA in HeLa cells, while AdICP4, an Ad5 vector generated by rescue of the HSV-1 ICP4 gene into dIE1,3 (Spessot et al., 1989), was not, suggesting that ICP0 or Ad5 gene products expressed during infection with 0pro-0 are sufficient to activate α-globin expression.

0pro-0 stocks are reported to be contaminated with wild type virus, caused by the rescue of E1 genes from the complementing cell line used to propagate the virus (Zhu et al., 1988) and that the 0pro-0 stocks used in these experiments did contain some wild type virus (J. R. Smiley, personal communication). Since Ad5 also induced α-globin expression, it remains
Figure 7. Effect of Ad5 infection on β-globin expression. RNA was isolated from HeLa cells infected with 25 PFU per cell of Ad5 and subject to primer extension analysis. Size markers were generated by cleavage of pBR322 with HpaII, followed by 3'-end labeling. Fragment sizes are indicated in nucleotides. (A) Analysis of 50 μg of RNA using the β-globin oligonucleotide. (B) Analysis of 20 μg of RNA using the Alu primer.
possible that the increase in α-globin RNA levels seen upon infection with 0pro-0 was caused by low levels of wild type virus. 0pro-0 is reported to express low levels of E2a gene products and it is not clear whether this E2a expression occurs as a result of transactivation of the E2a promoter by ICP0 or whether it reflects the levels of expression induced by the E1a gene products from the wild type background (Zhu et al., 1988). Levels of E2a RNA in 0pro-0, AdICP4, dIE1,3 and Ad5 infected cells were assayed by primer extension (figure 6c). 0pro-0 failed to produce detectable quantities of E2a RNA indicating that the amounts of either the E1a gene products or ICP0 which were manufactured during infection with 0pro-0 were insufficient to activate E2a expression. Primer extension analysis of ICP0 RNA demonstrated that ICP0 transcripts were produced at levels comparable to those seen during HSV-1 infection (figure 6a), suggesting that the absence of E2a transcripts was not due to the lack of ICP0 expression. Taken in combination these results suggest that the quantities of E1a gene products produced during 0pro-0 infection were insufficient to activate E2a transcription but provide no indication as to whether they were sufficient to activate α-globin expression. AdICP4 is also reported to express low levels of E2a gene products (Spessot et al., 1989). Infection with this recombinant resulted in the accumulation of ICP4 RNA and E2a RNA (figure 6b and c), but failed to activate α-globin expression (figure 5), suggesting that the E2a expression is insufficient to induce the α-globin gene. Since E2a is the only Ad5 gene product reported to be expressed by 0pro-0 (Zhu et al., 1988), it is most likely that 0pro-0 induces α-globin
expression due to the activity of ICP0, rather than indirectly through other Ad5 gene products.

Ad5 infection fails to activate β-globin expression

Since the α-globin and β-globin promoters are reported to behave differently in Ad5 vectors (Hearing and Shenk, 1985; Karlsson et al., 1986), I wished to determine whether the cellular copies of the globin genes were differentially regulated by Ad5. RNA harvested from cells infected with Ad5 was subjected to primer extension analysis using the β-globin and Alu primers (figure 7, panels a and b). Analysis of Alu RNA levels demonstrated that Ad5 E gene expression occurred in these cells, since E1b and E4 gene products are required for induction of Alu transcription. No β-globin expression was detected, demonstrating that, as was observed with HSV-1, the α- and β-globin loci show different responses to viral regulatory factors.

7.4 Discussion

Infection of HeLa cells with HSV-1 or Ad5 activated the synthesis of α-globin RNAs while β-globin transcripts remained undetectable. There are a number of mechanisms by which viral infection could stimulate the expression of tissue specific genes. Those that explain the differential response of these two coordinately regulated tissue specific genes require that virus induced alterations in the cellular transcriptional apparatus
affect only those factors which regulate α-globin expression. Though it is unlikely that Ad5 and HSV-1 induce the expression erythroid restricted transcription factors such as GATA-1 or NF-E2, since lytic infection results in shut off of host protein synthesis (Fenwick, 1978; Fenwick and Clark, 1982; Shenk, 1990), it is possible that infection stimulates the activity of α-globin-specific transcription factors. The α-globin promoter contains a greater number of Sp1 sites than the β-globin cis-regulatory region and it has been postulated that differential expression of globin genes in transient assays may be mediated through these Sp1 sites (Brickner et al., 1991; Proudfoot et al., 1992). In addition it has been postulated that a tissue specific Sp1-like activity is required for the correctly timed expression of globin genes during hematopoietic cell development (Evans et al., 1990), suggesting that Ad5 and HSV-1 may stimulate α-globin expression by modulating Sp1 or an Sp1-like factor. Alternatively there are additional, non-erythroid-restricted factors required for α-globin and not β-globin production, as is demonstrated by the X-linked α-thalassemia/mental retardation syndrome (Wilkie et al., 1990). Another possibility is that the differences in higher order structure of the α- and β-globin loci account for their different responses to viral infection: the α-globin locus may be more responsive to virus induced alterations in cellular transcriptional apparatus than the tightly packaged β-globin genes (reviewed in Grosveld et al., 1993). A final possibility is that Ad5 and HSV-1 infections release the non-erythroid transcriptional repression of the α-globin gene which is postulated to be involved in tissue specific regulation of α-globin expression.
(Charnay et al., 1984; Brickner et al., 1991): perhaps synthesis of a labile repressor protein is inhibited by viral host shutoff.

Ad5 and HSV-1 induced activation of α-globin expression was unexpected for reasons other than the release from tissue specificity. HSV-1 infection results in the turnover of most cellular RNA due to the shutoff of host gene expression mediated by a virion protein, VHS (Read and Frenkl, 1983). Transcripts produced from the Vero and HeLa α-globin genes must be spared the full effects of VHS, contrasting mouse β-globin mRNA in MEL cells, which is rapidly and completely degraded following infection with HSV-1 (Nishioka and Silverstein, 1978). In addition, ongoing mouse β-globin transcription halts completely following HSV-1 infection, and this transcriptional inhibition is mediated by an IE gene product (Smibert and Smiley, 1990). Finally, HeLa and other cultured non-erythroid human cells are reported to show extensive methylation of α-globin and other non-essential genes, which is thought to contribute to transcriptional inactivity (Antequera et al., 1989). These data suggest that the human α-globin gene shows an unusual characteristic: the ability to be expressed under conditions which are generally considered to inhibit transcription. It may be interesting to determine if this property is unique to the human α-globin gene. The mouse α-globin genes are more similar to the β-globin genes in that they show tissue-specific alterations in timing of DNA replication and are not located in a constitutively open region of chromatin (D. Higgs, personal communication). The availability of mouse MEL transformants containing the human α-globin gene, with and without HS-40, makes it
possible to test whether infection affects human and mouse genes differently and whether the initial transcriptional state of the gene plays any affects its response to viral infection.

The HSV-1 IE protein ICP0 was necessary and sufficient for the activation of α-globin expression, since the ICP0 deletion mutant dlx3.1 was the only IE deletion mutant completely unable to induce globin RNA and ICP0 expressed from an Ad5 vector induced the α-globin gene. ICP0 is a chromatin binding protein that contains a zinc-finger DNA-binding motif (Pereira et al., 1977; Hay and Hay, 1980; Berg, 1986; Perry et al., 1986). Though ICP0 is not essential for viral replication in tissue culture, it is necessary to facilitate productive infection at low input multiplicities and for reactivation from latency (Cai and Schaffer, 1989; Clements and Stow, 1989; Leib et al., 1989; Cai and Schaffer, 1991). Latent HSV genomes are assembled with nucleosomes while lytic genomes are nucleosome-free, suggesting that ICP0 is involved in the maintenance of or the transition into a transcriptionally active, nucleosome-free form of viral chromatin (Deshmane and Fraser, 1989). In addition ICP0 is a promiscuous transcriptional activator that stimulates expression of viral and cellular genes in transient assays (reviewed in Everett, 1987). It is tempting to suggest that reactivation from latency and promiscuous transactivation by ICP0 in transient assays are functionally equivalent, implying that ICP0 stimulates expression of transfected genes by nonspecifically allowing them to assume a transcriptionally active conformation. Perhaps ICP0 activates α-globin expression by a similar mechanism. It should be possible to use
expression vectors [more Ad5 vectors or rescue of mutated ICP0 sequences into dlx3.1 - some mutants are already available from R. Everett and P. Schaffer] to introduce mutant ICP0 molecules into HeLa cells and determine which regions of the protein are essential for activation of α-globin expression. In addition, it would be interesting to determine whether a virally transduced α-globin gene is expressed when it is rescued into an ICP0 null background. If the requirement for ICP0 is determined by properties of the α-globin gene in its chromosomal locus rather than sequences present in the gene itself, then virally transduced copies should be expressed in the absence of ICP0.

The HSV-1 IE protein ICP4 was not necessary for expression of the endogenous cellular α-globin gene during HSV-1 infection, which was a surprising observation for two reasons. First, both ICP4 and ICP0 were required for the stimulation stably integrated non-viral promoters, including the rabbit β-globin promoter, in biochemically transformed cell lines (Everett, 1985), suggesting that ICP4 and ICP0 can function in combination to activate expression of genes at some cellular loci. These proteins also stimulate transcription of heterologous eukaryotic genes in transient assays in a synergistic manner (Everett, 1984; O'Hare and Hayward, 1985; Quinlan and Knipe, 1985; reviewed in Everett, 1987) and the observation that ICP0 can form a complex with ICP4 and TBP suggests that a functional physical interaction between these molecules may be involved in their ability to synergistically activate transcription (Feng and Schaffer, 1993). Second, ICP4 was required for α-globin transcription in HSV-
1/globin recombinants, since rescue of the α-globin gene into a d120 background eliminated all globin expression and expression could be restored when ICP4 was provided by a complementing cell line (Smiley and Duncan, 1992). In short, viral copies of the α-globin gene require ICP4 for expression while the cellular copies do not, suggesting that some property of the viral chromosome bestows ICP4 dependence on the α-globin gene in the viral genome. Like ICP0, ICP4 is required to initiate an alteration in organization of viral DNA: it is necessary for viral DNA replication and the production of replicative compartments, sites of viral transcription and DNA replication which colocalize with sites of cellular DNA replication (Knipe et al., 1987). It is possible that some properties (intracellular localization, high concentration of transcription factors or active polymerases, transcriptionally active chromatin structure mediated by number, location, acetylation state or other modifications of histones or other chromatin proteins) of sequences present in the viral replicative compartments are essential for expression of early and late genes present on the viral chromosome and that a limited number of cellular loci, including α-globin, have similar properties.

The expression of the cellular copies of the α-globin gene peaked between 6 and 9 hours post-infection with HSV-1 and then it dropped quite dramatically. HSV-1 infection under a DNA replication blockade did not show this decrease in α-globin. The increased level of α-globin RNA seen during infections limited to the early phase relative to those which were allowed to progress to the late phase suggests that increased expression of
an early protein may further stimulate globin expression or, alternatively, that a late gene product may inhibit α-globin expression. The ICP27 null mutant strain 5dl1.2 shows reduced levels of viral DNA replication and abnormal expression of some early and late genes which are regulated in concert with DNA replication (McCarthay et al., 1989): it could be naively predicted that it should behave as PAAR5 does under an aphidicolin blockade. However 5dl1.2-induced globin expression at 12 hours post infection did not exceed that seen at late times during PAAR5 infection, suggesting that although 5dl1.2 is deficient in viral DNA replication, sufficient replication occurred to allow the late reduction in α-globin RNA levels. Since the failure of DNA replication due to the absence of ICP27 (and also ICP4) is not equivalent to the failure of DNA replication due to addition of aphidicolin, it is also possible that aphidicolin itself may directly affect the expression of the α-globin gene, perhaps by interfering with cellular DNA replication.

Ad5 infection activated α-globin expression and E1a gene products were required for this activation. Since E1a also stimulates transcription of all other Ad5 genes (Jones and Shenk, 1979), it is possible that E1a affects globin expression indirectly by facilitating the production of the viral protein which is directly involved. It is interesting to note that, as with HSV-1, expression of virally transduced and cellular copies of the α-globin gene showed different requirements viral gene products. Hearing and Shenk (1985) demonstrated that α-globin is transcribed when HeLa cells are infected with the strain in340α-globinEnh-, in which E1a sequences were
replaced by the α-globin gene, suggesting that the α-globin gene present in the Ad5 genome is introduced into cells in a form that is transcriptionally competent. Since in340α-globinEnh- lacks E1a coding sequences, it is unlikely that any Ad5 gene products are expressed during infection with this virus, indicating that E1a and other viral gene products are not necessary for α-globin transcription from the Ad5 genome. This indicates that the transcriptional machinery which is present in HeLa cells is sufficient to express α-globin sequences and suggests that it is not available to the cellular copies of the α-globin gene. It is unclear why expression of E1a or other Ad5 gene products is required for activation of expression of the host copy of the α-globin gene, but not the viral copy. It is possible that differences in the higher order structure or intranuclear localization of viral and cellular α-globin genes account for their different E1a requirements.

The strain in340α-globinEnh- produces two species of α-globin transcripts, which show differential accumulation in the presence E1a gene products (Hearing and Shenk, 1985). A smaller transcript, indicating a species which used the globin polyadenylation signal, accounted for the majority of α-globin RNA in in340α-globinEnh- infected cells, while a second species, which represented readthrough to the E1a polyadenylation signal, was far less abundant. Expression of the readthrough product was stimulated dramatically when E1a was provided in trans by coinfection with dl309, while levels of the bona fide α-globin RNA did not alter significantly, indicating that E1a induced an overall increase in α-globin
expression which was largely reflected by an increase aberrantly processed α-globin RNA. The use of alternate polyadenylation signals suggests that the production of the longer RNA may be due to regulation at the level of transcription termination or 3'-end processing. There is little evidence that E1a regulates gene expression at the level of transcription termination, while E1b and E4 gene products are post-transcriptional regulators of viral late gene expression. E1b and E4 act by a mechanism that affects processing or transport of newly transcribed viral mRNA precursors (Leppard and Shenk, 1989; Sandler and Ketner, 1989; Nordqvist and Akusjarvi, 1990), and it is possible that these gene products also are required for the accumulation of α-globin RNAs which are polyadenylated at the E1a site. Therefore further investigation of Ad5-induced host α-globin expression requires a more detailed examination of the structure of the α-globin RNAs which are produced and the precise identification of which viral gene products are absolutely required.

The non-erythroid expression of the human α-globin gene induced by HSV-1 or Ad5 provides a novel opportunity to study the mechanisms by which viral regulatory proteins act, particularly since there has been extensive analysis of the cis-regulatory motifs which regulate α-globin transcription. To begin to examine the mechanisms by which these viruses activate globin expression nuclear run-on assays should be used to determine whether class II transcription of α-globin sequences is affected by infection. Since tissue specific expression of the α-globin gene is corelated with the presence of the erythroid-specific DNAase I
hypersensitive sites it would also be interesting to determine whether these sites are induced in infected cells. It would also be worthwhile to attempt in vivo footprinting of α-globin promoters to determine precisely which sequences interact with proteins and if there is an alteration in factor binding during infection. Viral vectors bearing globin genes would provide valuable tools for these experiments: in vivo footprinting of the α-globin promoter in HSV-1 and Ad recombinants could be used to practise footprinting techniques on high copy number templates and should indicate which cis-regulatory sequences are important for expression in viral genomes. Finally it may be worthwhile to use in situ hybridization to determine the intranuclear localization of the α-globin gene relative to Ad5 and HSV-1 viral replicative compartments.
7.5 References


Forrester, W. C., S. Takegawa, T. Papayannopoulou, G.


Chapter 8: Discussion
8.1 *Alu* repetitive elements

8.1.1 Virus-induced expression of *Alu* repetitive elements

The results presented demonstrate that infection with HSV-1 and Ad5 activates RNA pol III transcription of endogenous cellular *Alu* repetitive elements. Transactivation is a relatively global phenomenon, occurring at multiple loci and in all cell types examined, suggesting that it is part of a general response to viral infection. It is unclear whether Ad5 and HSV-1 stimulate *Alu* expression by similar mechanisms. A number of nuclear DNA viruses encode factors that activate transcription of class III genes. These include Ad5 E1a 289 residue protein (Gaynor *et al.*, 1985), the pseudorabies IE protein (Gaynor *et al.*, 1985), the hepatitis B virus X-gene product (Aufiero and Schneider, 1990), and SV40 T-antigen (White *et al.*, 1990). There are several striking similarities between the activities of these proteins: they stimulate transcription of class II and class III genes, they activate viral or transfected genes, but have little effect on endogenous class III genes, and they function by modulating the activity of TFIIC (Aufiero and Schneider, 1990; Gaynor *et al.*, 1985; White *et al.*, 1990). Therefore, *trans*-acting proteins from a number of different viruses are able to affect relatively similar changes in cellular transcription. This suggests that pathways that regulate gene expression intersect at one or a few points, or that there are a number of regulatory events that can be manipulated to produce a similar results.
The observation that Ad5 infection could stimulate robust pol III transcription of Alu sequences was not entirely surprising since viral class III genes, the VA RNA genes, are highly transcribed during lytic infection (Bogenhagen et al., 1980; Soderlund et al., 1976). VA RNAs are abundant in the cytoplasm of infected cells, where they facilitate the production of L proteins by preventing the general inhibition of translation induced by interferon and double stranded RNA (O'Malley et al., 1986). VA RNA gene expression is E1a-dependent and the only essential cis-regulatory sequences are the A-box and B-box motifs also present in Alu promoters (Bogenhagen et al., 1980; Hoeffler and Roeder, 1985). In contrast E1a gene products were not absolutely required for Alu transactivation, suggesting that expression of these sequences is not regulated by factors that directly interact with the pol III promoter. TFIIC is the only factor (identified to date) which binds the regulatory elements common to Alu and VA RNA gene promoters (Geiduschek and Tocchini, 1988). It is a rate limiting assembly factor for class III transcription complexes in vitro and E1a-induced activation of pol III transcription correlates with an increase amounts or activity of TFIIC (Fuhrman et al., 1984; Hoeffler et al., 1988; Hoeffler and Roeder, 1985; Yoshinaga et al., 1986). Alu expression is not activated in 293 cells providing further evidence that it is unlikely that E1a-induced alterations in TFIIC are entirely responsible for the increase in expression of Alu sequences seen during Ad5 infection.

Though Jang and Latchman report that the HSV-1 IE protein ICP27 activates Alu transcription by increasing amounts or activity of TFIIC (Jang and Latchman, 1989; Jang and Latchman, 1992), we have been
unable to reproduce their observations. In addition, the ICP27 mutant strain they employed is not a null mutant: it produces reduced quantities of the protein, and is therefore an inappropriate substrate on which to base conclusions about an ICP27 requirement (Maclean and Brown, 1987). Finally, the cell line used to demonstrate ICP27 is sufficient for increased TFIIC activity does not constitutively express ICP27 and does not show increased levels of Alu RNA in our experiments. These data suggest that ICP27 is not necessary for the Alu transcriptional activation induced by HSV-1 and strongly suggest that ICP27 is not responsible for the modulation of TFIIC activity reported by Jang and Latchman (Jang and Latchman, 1989; Jang and Latchman, 1992).

It was surprising that Alu elements could be transcribed from HSV-1 recombinants and that HSV-1 infection could activate expression of endogenous sequences, since HSV-1 does not contain any class III genes and there was no indication the HSV-1 regulatory proteins could affect pol III transcription. The pseudorabies virus IE protein may provide a precedent: it can activate transcription of class III genes in transient assays (Gaynor et al., 1985), even though this porcine herpesvirus does not contain any class III transcription units. Other members of the herpes family encode small pol III transcribed RNAs, suggesting that they may also produce proteins that regulate class III transcription. The best characterized, the EBERs, were first detected in B lymphocytes infected with Epstein-Barr virus (Jat and Arrand, 1982; Lerner et al., 1981). An interesting feature of the EBER genes is that they are regulated by class II and class III promoter elements. In addition to the split internal pol III
promoter, the EBER upstream region contains sequences with strong TATA homology, and Sp1 and ATF sites (Howe and Shu, 1989; Howe and Shu, 1993). The 7SL gene is also regulated by upstream cis-regulatory sequences in combination with a bipartite intragenic class III promoter and the 7SL promoter contains an ATF site and a TATA-like sequence which is similar in location and sequence to the EBER TATA box (Howe and Shu, 1989). The similarity in regulatory sequences of the 7SL RNA gene and the EBER genes suggests that this combination of class II and class III motifs, in combination with the similarity between 7SL RNA genes and Alu sequences suggests that EBV may also encode proteins which stimulate Alu expression. It may be worth examining whether EBV infection of human lymphocytes results in activation of Alu transcription.

EBERs form a stable complex with the La autoantigen (Howe and Shu, 1988), a class III transcription termination factor that normally transiently associates with nascent pol III RNAs by their 3'-U rich sequences (Gottlieb and Steitz, 1989; Gottlieb and Steitz, 1989). It has been suggested that the La phosphoprotein is part of a feedback loop which coordinates class II and III gene expression by coupling class III transcription rates with levels of a protein that binds the newly synthesized RNA (Bachmann et al., 1989). EBERs may sequester La autoantigen during EBV infection, thus deregulating pol III transcription. Infection with Ad5 and HSV-1 causes a redistribution of cellular pools of La autoantigen (Bachmann et al., 1992; Bachmann et al., 1989; Bachmann et al., 1990). Given that this phosphoprotein is postulated to play a role in regulation of class III transcription, and since both these viruses show
some deregulation of Alu expression it is tempting to suggest that redistribution of La autoantigen during viral infection may be one of the cellular events that contributes to Alu transcriptional activation.

Ad5-induced Alu transcriptional activation provides the system with the greatest potential to investigate the molecular mechanisms which underpin Alu transactivation, since several viral gene products proved to be essential in this process. The E1b 58 kDa protein and the products of E4 ORF 3 and 6 were required for the Ad5-induced activation of Alu pol III transcription. These gene products act as post-transcriptional regulators of viral gene expression by facilitating transport or processing of late mRNA precursors (Leppard and Shenk, 1989; Nordqvist and Akusjarvi, 1990; Sandler and Ketner, 1989). The mechanisms by which E1b and E4 gene products may stimulate Alu transcription can be broadly divided into two classes: those which predict that transactivation is mediated by specific interactions with E1b and E4 responsive cis-regulatory sequences or those which predict that transactivation is an indirect consequence of E1b or E4 activity. Ad5 (and HSV-1) were able to effect expression of a large number of Alu sequences, so it seems unlikely they are stimulating transcription by acting through cis-regulatory sequences which are present on a very limited subset of Alu repeats. A large number of Alu elements contain motifs that correlate with binding of a number of fairly ubiquitous cellular proteins and Alu transactivation may be mediated through one or more of these motifs (Tomilin et al., 1992). Alternatively, transcription of Alu sequences may be regulated by factors with less specificity, perhaps by general changes affecting transcriptional status of large segments of
cellular chromatin, such as alterations in DNA methylation patterns or modulation of histones. Transcription of Alu sequences in vitro has been shown to be inhibited by nucleosome binding and DNA methylation (England et al., 1993), suggesting that these mechanisms may regulate Alu expression in vivo.

8.1.ii Future experiments

a/ Are E1b and E4 sufficient for Alu transcriptional activation?

E1b and E4 gene products are necessary for the Ad5-induced stimulation of Alu element transcription. Expression systems, which allow high level production of E1b and E4 proteins in the absence of other Ad5 gene products, should be explored to determine whether they are sufficient. These include transient transfection assays and production of stable cell lines with E1b and E4 coding sequences under the control of inducible promoters. It is probably essential that the Alu reporter in any of these instances is endogenous cellular expression levels, since transiently transfected Alu elements might not respond to viral factors in the same manner as their genomic counterparts. Since individual cells vary in efficiency of expression of transfected sequences, it may be possible to use in situ hybridization techniques to demonstrate colocalization of Alu and E1b or E4 transcripts. If an efficient transient expression assay can be established, further work could include the mutational analysis of E1b and E4 coding sequences. Finally, it would be interesting to determine whether co-expression of E1b and E4 proteins induces alterations in distribution of
splicing apparatus, and to determine whether these alterations correlate with \textit{Alu} transcriptional activation.

\textbf{b/Are specific} \textit{Alu} \textbf{sequences reproducibly transactivated?}

In order to determine which, if any, \textit{cis}-regulatory sequences are required for \textit{Alu} element transcriptional activation, it will be necessary to identify individual elements which are consistently activated during Ad5 infection. One approach to identify elements which are transcriptionally active during viral infection would be to sequence the heterogeneous 3' ends of \textit{Alu} transcripts (due to pol III transcription termination in flanking DNA sequences). Briefly, a primer would be ligated to the 3' end of total cytoplasmic RNA (or size fractionated to enrich for \textit{Alu} RNA) and reverse transcription reactions carried out with the complimentary primer (with 3-4 extra T residues on the 3'-end of this primer, to select for RNA species that terminate in a run of U residues). PCR reactions would then be carried out with this primer and an \textit{Alu}-specific primer (specific for sequences relatively close to the 3'-end of the \textit{Alu} consensus sequence). The resulting PCR products could be cloned and sequenced, and presumably mapped back to their genomic loci. Alternatively, it may be worthwhile to examine expression patterns of a number of individual \textit{Alu} elements, by 3' S1 nuclease analysis (or RNAase protection) to determine if already identified sequences are activated by infection. Good targets may include sequences belonging to the most recently transpositionally active \textit{Alu} subfamily, since there are a limited number of these in the human genome, and those present in the human \textit{\alpha}-globin gene, since the probes are already
available in the laboratory. Identification of a number of Alu repeats which are induced upon Ad5 infection would allow a few important basic questions to be addressed: are these elements consistently induced, to what levels are they activated and are there sequence similarities between Ad5 responsive and non-responsive elements?

Can Alu transcriptional activation be reproduced in vitro?

RNA pol III transcription is relatively well defined in vitro. It may be possible to identify factors in either an in vitro transcription extract generated from HSV-1 infected cells, or in chromatin reconstituted from infected cell extracts that stimulate class III transcription of Alu sequences.

8.2 α-globin

8.2.i Expression of human α-globin

α-globin, but not β-globin, transcription can be induced in HeLa cells by infection with HSV-1 or Ad5. This provides further evidence for the differences in regulation between these two loci. It further demonstrates that expression of globin genes is not entirely coordinately regulated: that α-globin is regulated in manner that can be perturbed by viral infection, while β-globin expression is unaffected. A number of tissue specific genes are expressed at very low levels (1 RNA molecule/500-1000 cells) in inappropriate tissues (Chelly et al., 1989), suggesting that any gene may be transcribed in any cell type and that ubiquitous transcription factors may be
sufficient for the minimal activation of a number of tissue specific promoters. This ectopic expression of genes which are normally regulated in a developmentally specific manner has been termed illegitimate transcription (Chelly et al., 1989) and virus induced activation of α-globin expression clearly appears to be an example of illegitimate transcription. HSV-1 and Ad5 induce expression of α-globin at a much higher level than has been reported in other instances of illegitimate transcription, suggesting that these viruses may activate gene expression by increasing the cells ability for illegitimate transcription. Further characterization of the viral proteins and host cis-regulatory sequences involved in the induction of α-globin expression in non-erythroid cells is necessary before productive speculation on the mechanisms by which HSV-1 or Ad5 stimulate α-globin expression can begin.

8.2.ii Future work with α-globin

Many of the experiments which could be done to address the mechanism of activation of α-globin expression were discussed in detail in the previous chapter. Briefly, it is necessary to carry out nuclear run-on assays to determine whether the α-globin gene is transcriptionally activated, or whether this may be a post-transcriptional effect. Time course analysis of various HSV-1 IE mutants should be carried out to determine whether each IE mutant shows different kinetics of α-globin expression, and to more clearly determine whether ICP0 is absolutely necessary. If ICP0 proves to be both necessary and sufficient, then mutational analysis of this protein may prove informative. Analysis of which Ad5 gene products
are required for stimulation of α-globin expression could be carried out using the selection of mutant viruses available in the laboratory. Finally, in vivo footprinting techniques and use of somatic cell hybrids containing chromosomes with aberrant human globin loci may aid in identification of regulatory sequences important for viral induction of α-globin expression.

8.3 Concluding remarks

The results presented in this thesis demonstrate that HSV-1 and Ad5 infection can activate the expression of Alu repetitive elements and the α-globin gene from their endogenous cellular loci. These unusual instances of virus-induced stimulation of cellular gene expression have provided novel systems to study the role of HSV-1 and Ad5 proteins in transcriptional regulation. In addition we have added to the list of unusual characteristics of the human α-globin gene. Finally, we have identified one manner in which transcription of Alu repetitive sequences can be regulated and provided a new avenue for investigation of these poorly understood DNA elements.
Chapter 9: References


Klessig, D. F. 1977. Two adenovirus mRNAs have a common 5' terminal leader sequence encoded at least 10 kb upstream from their main coding regions. Cell 12:9-21.


O'Hare, P. and G. S. Hayward. 1985a. Evidence for a direct role for both the 175,000 and 110,000 molecular weight immediate-early proteins of herpes simplex virus in transactivation of delayed-early promoters. J. Virol. 53:751-760.


Preston, C. M. 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature sensitive mutant tsk. J. Virol. 29:275-284.


