

# Structural Features of Ribonucleotide Reductase

Ioannis Nikas,<sup>1</sup> John McLauchlan,<sup>1</sup> Andrew J. Davison,<sup>1</sup> William R. Taylor,<sup>2</sup> and J. Barklie Clements<sup>1</sup>

<sup>1</sup>MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, Scotland; <sup>2</sup>Department of Crystallography, Birkbeck College, London WC1E 7HX, England

**ABSTRACT** Herpes simplex virus type 1 (HSV-1) encodes a ribonucleotide reductase which comprises two polypeptides with sizes of 136,000 (RR1) and 38,000 mol. wt. (RR2). We have determined the entire DNA sequence specifying HSV-1 RR1 and have identified two adjacent open reading frames in varicella-zoster virus (VZV) which have homology to HSV RR1 and RR2; the predicted sizes for the VZV RR1 and RR2 polypeptides are 87,000 and 35,000 mol. wt. respectively. Amino acid comparisons with RR1 and RR2 polypeptides from other organisms indicate that HSV-1 RR1 contains a unique N-terminal domain which is absent from other RR1 polypeptides apart from HSV-2 RR1. These N-terminal amino acid sequences are poorly conserved between HSV-1 and HSV-2 in contrast to the remainder of the protein which shows greater than 90% homology. Polypeptide structural predictions suggest that the HSV-1 N-terminal domain may be separated into two regions, namely, a  $\beta$ -sheet structure followed by a nonstructured area. Across the remainder of RR1 and RR2, comparisons also reveal blocks of amino acids conserved between the different ribonucleotide reductases, and these may be important for enzyme activity. From predictions on the structure of these conserved blocks, we have proposed that the location of a substrate binding site within RR1 is centered on three conserved glycine residues in a region which is predicted to adopt a  $\beta$ -sheet/turn/ $\alpha$ -helical structure; this approximates to the structure for ADP nucleotide binding folds. Finally, we propose that the promoters for the HSV and Epstein-Barr virus (EBV) RR2 transcripts have evolved by separate evolutionary routes.

**Key words:** ribonucleotide reductase, unique N-terminal domain, nucleotide binding site

## INTRODUCTION

Ribonucleotide reductase is an essential enzyme for the conversion of ribonucleotides to the corresponding deoxyribonucleotides in prokaryotic and eukaryotic cells, and its activity may represent a rate-limiting step in DNA replication.<sup>1</sup> Several herpesviruses, namely herpes simplex virus type 1 (HSV-1)<sup>2</sup> and HSV-2<sup>3</sup> Epstein-Barr virus (EBV),<sup>4</sup> pseudorabies virus<sup>5</sup> and equine herpesviruses types 1 and 3<sup>6</sup>, induce a novel ribonucleotide reductase activity with properties different from those of the endogenous cell enzyme. To date, the virus ribonucleotide reductases studied appear to lack the allosteric control which

regulates activity of the cellular enzymes.<sup>5,7,8</sup> The virus-specified ribonucleotide reductase is essential for HSV replication,<sup>9,10</sup> thus this activity provides an important target for antiviral chemotherapy.

In common with other ribonucleotide reductases which use ribonucleoside diphosphates (rNDPs) as substrates, the HSV enzyme is composed of two non-identical polypeptides.<sup>11</sup> Analysis by marker rescue of the HSV-1 mutant 17tsVP1207<sup>10</sup> has shown that a virus-coded polypeptide of molecular weight 136,000 (Vmw136, RR1) is a component of ribonucleotide reductase. A second HSV-1 polypeptide with a molecular weight of 38,000 (Vmw38, RR2) associates with RR1, and this complex is probably essential for enzyme activity.<sup>11,12</sup> Synthetic peptides corresponding to the C-terminal region of RR2 inhibit enzyme activity and these may act by preventing complex formation.<sup>13,14</sup>

The HSV-1 RR1 and RR2 polypeptides are encoded by two early mRNAs of 5.0 kilobases (kb) and 1.2 kb, respectively, which are unspliced and share a common 3' terminus<sup>15</sup> (Fig. 1). Moreover, the 5' terminus of the 1.2-kb mRNA is located within the C-terminal region of RR1.<sup>16</sup> Two late transcripts of 7.0 and 1.9 kb are also specified by this region of the HSV-1 genome<sup>17</sup>; the 7.0-kb mRNA overlaps the RR transcripts and shares the same 3' terminus as these mRNAs. On the HSV-2 genome, the ribonucleotide reductase locus is similarly arranged,<sup>18,19</sup> and HSV-2 RR1 and RR2 have sizes of 138,000 and 38,000 respectively.<sup>20</sup>

DNA sequences encoding HSV-2 RR1 and RR2<sup>18,21</sup> and HSV-1 RR2<sup>22,16</sup> have been previously obtained as have those encoding the N- and C-terminal portions of HSV-1 RR1. Sequence comparisons have indicated that blocks of amino acid homology within RR1 and RR2 are present in equivalent HSV, EBV, mouse, *Escherichia coli* and surf clam polypeptides;<sup>21,23,24</sup> these conserved blocks are presumed to have functional importance for enzyme activity.

Here, we present the complete sequence of HSV-1 RR1 and show that there is amino acid homology extending throughout RR1 and RR2 between HSV

Received November 12, 1986; accepted January 23, 1987

A. Davison's present address is Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892.

Address reprint requests to J. Barklie Clements, MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, U.K.

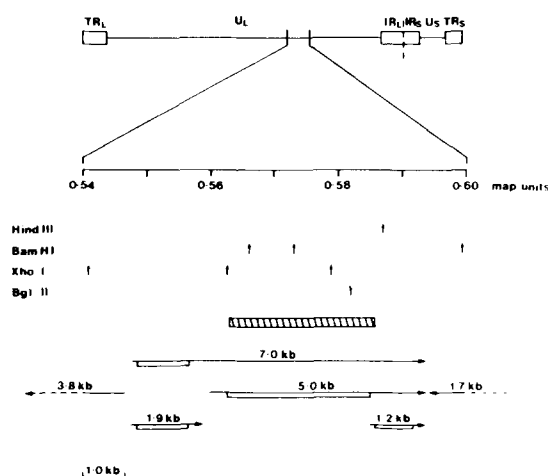


Fig. 1. Locations of restriction endonuclease cleavage sites and arrangement of mRNAs between 0.54 and 0.60 map units on the HSV-1 strain 17 genome. The 5.0-kb and 1.2-kb mRNAs correspond to the transcripts which encode the RR1 and RR2 polypeptides respectively. Open boxes denote the translated portion of each transcript, and the hatched box represents the region for which the nucleotide sequence presented here was determined.

and two adjacent varicella-zoster virus (VZV) open reading frames. From further comparisons with equivalent RR1 polypeptides, there is an additional N-terminal region of some 320 amino acids present in HSV-1 RR1, which, except for HSV-2 RR1, is absent from all other RR1 proteins so far examined; this region shows poor amino acid conservation with its HSV-2 equivalent. From protein secondary structure predictions, this N-terminal region may be divided into two distinct areas, one of which comprises a  $\beta$ -sheet structure, while the second area is nonstructural. Structural predictions for the remainder of RR1 and RR2 are also presented from which we propose a location for a nucleotide binding site on RR1.

Similar to the genomic arrangement for HSV-1 RR mRNAs, the EBV RR transcripts overlap, and the 5' terminus of the EBV RR2 transcript is located within the C-terminal portion of EBV RR1.<sup>25</sup> However, DNA sequences forming the promoter for the EBV RR2 mRNA are absent from HSV, and we suggest that these viral promoters have evolved by separate evolutionary events.

## MATERIALS AND METHODS

### Recombinant Plasmids

The derivation of plasmid pYN1 is described in the Results and Discussion. pYN1 was propagated in *E. coli* strain DH1,<sup>26</sup> and plasmid DNA was prepared as described by Holmes and Quigley.<sup>27</sup>

### Nucleotide Sequence Analysis

Sequence analysis was performed by the M13 chain terminator method. Sets of M13 clones were generated by sonication of the restriction enzyme DNA

fragments of interest<sup>28</sup> and ligation into the Sma I site of M13mp8.<sup>29</sup> The resultant M13 clones were sequenced as described by Sanger et al.<sup>30</sup> Sequence products were fractionated by electrophoresis in 6% denaturing polyacrylamide gradient gels<sup>31</sup> which were covalently bonded to one glass plate. Gels were fixed and dried prior to autoradiography. Sequencing artifacts were resolved as described by McGeoch et al.<sup>32</sup>

### Computer Analysis of Sequences

Computing was performed with a DEC PDP 11/44 computer running under the RSX 11M operating system. Sequencing data were stored and manipulated using the programs devised by Staden.<sup>33</sup> Graphical presentation of similarities between two amino acid sequences were generated with the CINTHOM program,<sup>34</sup> which was modified by P. Taylor, such that nonidentical but biochemically similar amino acids are treated as evolutionarily conserved changes, and these matches are scored higher than completely dissimilar residues. The similarity between sequences was plotted as an uppercase or lowercase letter of the alphabet with A representing the highest homology (100%) and z denoting the lowest homology; areas with conserved sequences form a diagonal on the graph. Optimal alignments and comparisons of amino acid sequences were performed with the HOMOL program.<sup>35</sup>

### Structural Analysis of RR1 and RR2 Amino Acid Sequences

Amino acid sequences of RR1 and RR2 proteins were aligned by a computer program which classifies amino acid type based on a synthesis of physico-chemical and mutation data.<sup>36</sup> The classification is organized in the form of a Venn diagram from which subsets are derived that include groups of amino acids likely to be conserved for similar structural reasons. These sets describe conservation in aligned sequences by allocating to each position the smallest set that contains all the residue types brought together by the alignment. Secondary structure predictions of RR1 and RR2 proteins were performed with the GARNIER program,<sup>37</sup> with no bias towards any structural type.

## RESULTS AND DISCUSSION

### Nucleotide Sequence Encoding HSV-1 RR1

Previously, we have suggested that the ATG codon for the N-terminal methionine residue of RR1 (Vmw136) is located 34 nucleotides downstream from the Xho I site at 0.561 map units<sup>16</sup> (Fig. 1); the C-terminus for this polypeptide was positioned within the untranslated leader of the 1.2-kb mRNA, approximately 290 nucleotides upstream from the Hind III site at 0.587 map units. Therefore, to complete the nucleotide sequence for the RR1 coding region, a 4 kbp Xho I/Hind III fragment (0.561–0.587 map units) from

Hind III *k* was inserted into the Sal I and Hind III sites of vector pUC8. Using the resultant plasmid, pYN1, the complete nucleotide sequence and predicted amino acid sequence of RR1 was determined (see Materials and Methods, Fig. 2). The DNA sequence contains an open reading frame encoding 1,137 amino acids, giving the polypeptide a predicted molecular weight of 124,017. This is less than the molecular weight of RR1 as determined by polyacrylamide gel electrophoresis (136,000), however, RR1 is phosphorylated,<sup>38,39</sup> which probably accounts for this size discrepancy. The G+C content across the HSV-1 RR1 polypeptide coding region is 65.7%, however, in the third base position the G+C content (83.2%) is markedly higher. This is consistent with the codon usage of HSV proteins, which shows that, in most cases, the third base position of amino acid codons has a higher G+C content as compared with first and second base positions.<sup>40</sup>

### VZV Coding Regions With Homology to HSV RR1 and RR2 Proteins

The entire nucleotide sequence of the VZV genome has been determined.<sup>41</sup> Our predicted amino acid sequences for HSV-1 RR1 and HSV-2 RR2 have homology with two adjacent open reading frames on the VZV genome; these open reading frames would specify polypeptides of 87,000 and 35,000 mol. wt.

A feature of the HSV-1/VZV RR1 comparison (Fig. 3) is that significant homology begins around HSV-1 amino acid 384 and VZV amino acid 16; thereafter the two sequences share 43% homology across the remainder of their coding regions, with the stretch from HSV-1 amino acids 650–820 showing 53% homology. In contrast to RR1, HSV-2 RR2 does not possess extensive additional sequences at the N-terminus compared to the VZV 35,000 mol. wt. polypeptide (Fig. 3); homology between HSV-2 and VZV is 53% and extends to the C-terminus of the RR2 polypeptides. Homology between the HSV and VZV sequences is predominantly retained in blocks of conserved amino acids, a number of which are also conserved in other RR1 and RR2 polypeptides (Table I). To date, a novel ribonucleotide reductase activity has not been detected in VZV-infected cells, however, these homologies strongly suggest that a viral ribonucleotide reductase is encoded by VZV.

Amino acid conservation between HSV and VZV sequences is greater than that between HSV and EBV ribonucleotide reductase sequences. For example, the HSV-1/EBV amino acid homology for RR1 begins around HSV-1 residue 480 and EBV residue 90 (Fig. 4B), which is beyond the start of homology between HSV-1 and VZV; homology from this point between HSV-1 and EBV is 38%, which is less than that between HSV-1 and VZV. Moreover, the HSV-1 and EBV sequences are not completely colinear, as there is an insertion of 34 amino acids at position 421 in the EBV sequence and the EBV polypeptide con-

tains a C-terminal region of 75 amino acids which is absent from both HSV and VZV (Fig. 4B). This pattern of amino acid conservation is similarly reflected in nucleotide sequence comparisons (data not shown) despite the wide variation in G+C content of HSV (64%) and VZV DNA (39%) in this region; the corresponding area on the EBV genome has a G+C content of 55%.

### HSV RR1 Proteins Contain a Unique N-Terminal Region Which is Poorly Conserved Between HSV-1 and HSV-2

From the HSV/VZV amino acid comparisons, two main features emerge, namely (1) the presence of an N-terminal region in HSV-1 RR1 which is absent from VZV, and (2) discrete blocks of amino acid homology, certain of which are shared with other RR1 and RR2 polypeptides. The additional domain present in HSV-1 RR1 is illustrated by the CINTHOM comparisons of HSV-1 RR1 with equivalent VZV and EBV polypeptides (Fig. 4A,B); the N-terminal region of HSV-1 RR1 also is absent from mouse and *E. coli* RR1 proteins (data not shown). Further comparisons with other predicted polypeptide sequences of VZV and EBV<sup>42</sup> and from the National Biomedical Research Foundation data base also reveal no homology with the HSV-1 N-terminal region.

A striking feature of amino acid comparisons between RR1 of HSV-1 and its HSV-2 equivalent is the relatively poor conservation of amino acids at the N-terminus (Fig. 5A). At the N-terminal region extending to HSV-1 amino acid 311, the overall HSV-1/HSV-2 amino acid homology is 38%. Moreover, there are a number of insertions and deletions between both sequences and the HSV-2 N-terminal region comprises 318 amino acids. In contrast, from HSV-1 amino acid 312 to the C-terminus of RR1 the HSV-1/HSV-2 homology is 93% with 57 amino acid changes out of 826 residues and both sequences are colinear. A secondary structure for the HSV-1 N-terminal domain was predicted by the method of Garnier et al.<sup>37</sup> (Fig. 5B). The predicted structures emphasized two distinct regions divided by the short stretch of prolines which occurs approximately midway along the domain and is conserved between HSV-1 and HSV-2 (Fig. 5A, residues 150–154).

The first portion of the sequence is punctuated by short stretches of hydrophobic residues, predominantly V and A, which were predicted to be  $\beta$ -sheet structure. The intervening sequences, predicted as either turn or unstructured regions, contain in addition to many hydrophilic residues, a substantial number of A, G, and P residues. A possible structure that could incorporate these features is a globular domain formed of  $\beta$ -strands arranged in one or more  $\beta$ -sheets, with a well-buried core of hydrophobic residues (Fig. 5B). Beyond the stretch of prolines, the sequence is more hydrophilic with few hydrophobic residues; the sequence prediction becomes increasingly  $\alpha$ -helical

TGACGGCCGCGGAAACCCGCGCGCTCTTTGAA ATG GGC AGC GGC CCA GCA TCC TCT CCC GTG GAA GCG GCG 14  
A P V G G G E A G G P S A A T Q E A A G A P A R 75  
GCC CCG GTT GGA GAG CAG GGC GGC CCG AGC GCA GGC ACC CAG GGG GAG GGC GGC GGC GGC CCT 144  
L A H G H Y C Q R V N G V M L S D K T P 60  
CTC GGC CAC GGC CAC GTG TAC TGC CAG CGA ATC AAT GGC GTG ATG GTG CTT TCC AGG AGC GGC 213  
G S A Y R I S N E V D C G S N C T M I I 83  
GGG TCC GCG TCC TAC GAT AGC ATC AAT TTT GTG CAA TGT GGT TCC AAC TGC ATG ATC ATC 282  
D G V V R G R P Q D P G A A S P A P F V A 106  
GAC GAG GAG GTG GGC GGC CCG CCG GAG CAG CCG GGC GGC CCA TCC CCC GCT CTT GTT GGC 351  
V T N I G A G S D G G T A V A V A F G G T P R R 129  
GTG ACA AAT GGA GGC GGC GGC GGC ACC GGC GTG GCA TTC GGG GGA ACC CCA GGT GGC 420  
S A G T S T G T Q T A D V P T E A L G G P P 152  
TGG GGG GGC AGC TCT ACC GGT ACC CAG AGC GGC GAT GTG CCG ACC CAG GGC CTT GGG GGC CCG CCT 489  
P R F T L G G G C C S C R D T R B S A V F 175  
CCT CCC GGC TCC ACC CTG GGT GGC GGC TGT TGT TGT GGC GAC ACA CCG GGC CCG TCT GGC CTA TTC 558  
G G E G D P V G P A E F V S D D R S S D S D 198  
GGG GGG GAG GAG GAT CCA GTG GGC CCG GGC GAG TTC GTG GAC GAG CCG TCG TCC GAT TCC GAC TCG 627  
D S E D T D S E T L S H A S D V S G A T 221  
GAT GAC TCG GAG CAG CAG TCG GAG CCG CTG TCA CAC GGC TCC TCG GAC GTG TCC GGC GGC ACC AGC 696  
Y D D A L D S D S D D S L Q I D G P V C R 244  
TAC GAC GAC GGC CTT GAT TCC TCA TCG GAT GAC TCC CTG CAG ATA GAT GGC CCG GTG TGT GGC 765  
P M S N D T A P L D V C P G T P G P G A D A G 867  
CCG TGG AGC AAT GAC ACC GGC CCG CTG GAT GTT TGC CCC GGG ACC CCG GGC CCG GGC GAC GGC GGT 894  
G P S A V D P H A P T P E A G A L A D P A 290  
GGT CCC TCA GGG GTA CAC CAG CCG AGC CCA GAG GGC GGT CTT GGT CTT GGC GGC GAT CCC GGC 903  
V A R D A E G L S D P R P R T A Y P R 313  
GTG GGC CCG GAC GGC GAG GGC CTT TCG GAC CCG CCG CCA COT CTG GGA AGC GGC AGC GGC TAC CCG 372  
V P L E L T P E N A E A V A R F L G D A V N R 336  
GTG CCC CTG GAA CTC AGC CCG GAG AAC GGC GGC GTG GGC CTT CTC GGA ATC GGT CCG AAC CCG 1041  
E P A L M L E Y F C R C A G A G A C C K V P R 359  
GAA CCG CCG CTC ATG CTG GAG TAT TTT TGC CCG TGC GGC CCG GAG GAA ACC AAG GGT GTC CCC AGC 1110  
T F G S P P L T E D D F G L L N Y A L V E M 382  
ACA TTC GAC AGC CCC CTT GGC CTC AGC GAG GAC GAT TTT GGC CTT CTC AAC TAC GGC CTG GTG GAG ATG 1179  
Q R L C L D V P V P N A Y M P Y Y L R E Y 405  
CAG CCG CTG TGT GTC GAT CTT CCG GTC CCG AAC GCA TAC ATG CCC TAT TAT CTC AGG GAG TAT 1248  
V T R L V N G F K P L V S R A R L Y R I L G 428  
GTG AGC CCG CTG GTC AAC GGG TTC AAG CCG CTG GTG AGC CCG TCC GCT CCG CTT TAC GGC ATC CTG GGG 1317  
V L V H L B I R E A S F E M L R S K E V 451  
GTT CTG GTG CAC CTG ACC CCG GAG CCG TCC TTT GAG TGG CTG GCA TCC AAG GAA GTG 1386  
A L D F G L T E R L R E H E A Q L V I L A Q A 474  
GCC CTG GAT TTT GGC CTG AGC GAA AGG CTT GGC GAG CAC GAA GCC CAG CTG GTG ATC CTG GGC GGT 1455  
L D H Y D C L I H S T P H T L V E R G L Q S A 497  
CTG GAC CAT TAC GAT TGT CTG ATC CAC AGC ACA CCG CAG CCG CTG GTC GAG CCG GGC CTG CAA TCG GGC 1524  
L K B E F Y L K R F G G H Y M E S Y F Q M Y 520  
CTG AAG TAT GAG GAT TTT TAC AAG GGT TTT GGC GGC CAG TAC ATG GAG TCC CTC CAG ATG TAC 1593  
T R I A G F T T G C C A R A T R G M C H A L G G R E 543  
ACC CCG ATC GGC CTT TTG GGC CCG ACC CCG AGC AGC ATG CCG CAC ATC CCG CCG GGC CCA GAG 1662  
G S M W E M F K F F H R L Y D H Q I V P S T 566  
GGG TCG TGG TGG GAA ATG TTT TTT TTT CAC GGC CTC TAC CAC CAG ATC GTC GTC AGC 1731  
P A M L N L G T R N Y Y T S S C Y L V N P Q A 589  
CCC GGC ATG AAC CTG GCG ACC CCG AAC TAC TAC ACC TGC TGC TAC CTG GTA AAC CCG CAG GGC 1800

T T N K A T L R A I T S N V S A I L A R N G G 612  
ACC ACA AAC AAG GCG ACC CTG CCG GGC ATC ACC AAC AGC GTG AGT GGC CTC GGC AAC GGG GGC 1869  
I G L C V Q A F N D S G P G T A S V M P A L K 635  
ATC GGG CTA TGC GTG GGC TTT AAC GAC TCC GGC CCG GGC ACC GGC GTG ATG GGC CCG CTT CAC 1938  
V T L D S L V A A H N K E S A R P T G A C V Y L 658  
GTG CTT GAC TCG CTG GCG GCG CAC AAA GAG AGC GCG CCG GTG GGC GGC GTG TAC GTG 2007  
E P W H T D V R A V L R M K G V L A G E A Q 681  
GAG CCG TGG CAG AAC GGC GGC GTG CCG ATG AAG GGC GTG CTC GGC GGC GAA GAG GGC CAG 2076  
R C D N I F S A L M W P D L F K R L I R H L 704  
CSC TGC GAC AAT ATC TTC AGC GGC TCG ATG CCA GAT TTT TTT AAG GGC CTG ATT CCG CAC CTG 2145  
D G E K N V T W T L F D R D T S M S L A D F H 727  
GAC GGC GAG AAG AAC GTG ACA TGG ACC CTG TTC GAC CCG GAG ACC ACC ATG TCG CTC GGC GAC TTT CAC 2214  
G E E F E K L Y G H L E V M G F G E Q I P I Q 750  
GGG GAG GAT TTT GAG AAG CTC TAC CAG CAC CTC GAG GTG ATG GGG TTC GGC GAG CAG ATA CCG ATC CAG 2283  
E L A Y G I V R S A A T T G S P F V M F K D A 773  
GAG CTG GGC TAT GGC ATT GTG CCG AGT GCG GCG ACC GGC AGC CCG TTC GTG ATG TTC AAA GAG GGC 2352  
V N R H Y I Y D T Q G A A I A G S N L C T E I 796  
GTG AAC CCG CAC TAC ATC TAC GAC ACC CAG GGG GGC GGC ATC GGC GGC TCC AAC CTC TGC ACC GAG ATC 2421  
V H P A S K R S S G V C N L G S V N L A R C V 819  
GTG CAT CCG GGC TCC AAG CGA TCC AGT GGG GTG TGC AAC CTG GGA AGC GTG AAT CTG GGC CGA TGC 2490  
S R Q T F D F G R L R D A V Q A C V L M V N I 842  
TCC AAG CAG AGT TTT GAC TTT GGG GGG CTC CCG GAC GGC GTG CAG GGC TGC GTG ATG GTG AAC ATC 2559  
M I D S T L Q P T P Q C T R G N D N L R S M G 865  
ATG ATC GAC AGC CTA CAA CCC CCG CCG TGC ACC CCG GGC AAC GAC CAG CCG TCC ATG AAA CAG GGC 2628  
I G M Q G L H T A C L K L G L D L E S A E F Q 888  
ATC GGC ATG GGC CTG CAC AGC GGC TCC ATC GAG CTG GGG CTG GAT CTG GAG TCT GGC GAA TTT CAC 2697  
D L N K H I A E V M L L S L S N A L C V 911  
GAC CTG AAC AAA CAC ATC GCG GAG GTG ATG CTG TCG GCG ATG AAG ACC AAC GCG CTG TGC GTT 2766  
R G A R P F N H F K R S M Y R A G R F H M E R 934  
CSC GGC ACC GGT CTT AAC CAG TTT AAC CAG AGC ATG TAT CCG GGC GGC CCG TTT CAC TGG GAG CCG 2835  
F P D A R P R Y E G E M L R Q S M K H G 957  
TTT CCG GAC CCG CCG TAC GAG GGC GAG TGG GAG ATG CTA CCG CAG AGC ATG ATG AAA CAG GGC 2904  
L R N S Q F V A L M P T A A S A Q I S D V S E 980  
CTG GGC AAC AGC CAG TTT GTG GCG CTG ATG CCC ACC GGC CCG TCG GCG CAG ATC TCG GAC GTG 2973  
G F A P L P T N L F S K V T R D G E T L R P N 1003  
GGC TTT GGC CCG TCG ACC AAC CTG TTC AGC AAG GTG ACC CCG GAG GGC GAG AGC CTG CCG CCG 3042  
ACG CTG CTG CTA AAG GAA CTG GAA CCG AGC TTT AGC GGG AAG CCG CTC CTG GAG GTG ATG GAG AGT 3111  
T L L L K E L E R T F S G K R L L E V M D S L 1026  
D A K Q M S V A Q A L P C L E P T H R L R F 1049  
GAC GGC AAG CAG TGG TCC GTG CCG CAG GCG CTC CCG TCG GAG CCG ACC CCG CTC CCG CCA TTC 3180  
K T A F D Y D Q K L L I D L C A D R A P Y V D 1072  
AAG ACC GCG TTT GAC TAC CAG AAG TTT GAT GAC CTG TGT GGG GAG CCG GGC CCG CCG TAC GTG 3249  
W S Q S M T L Y V T E K A D G T L P A S T L V 1095  
CAT AGC CAA TCC ATG ACC TAT GTG AGG GAG AAG GGC AGC CCG CCG CCA GGC TCC ACC CTG 3318  
R L L V H A Y K R G L K T G M Y Y C K V R K A 1118  
CSC CTT CTG GTC CAC GCA TAT AAG CCG GGA CTA AAA ACA GGG ATG TAC TAC TCC AAG GTT CCG AAG GCG 3387  
T N S G V F G G D N I V C M S C A L 1137  
ACC AAC AGC GGG GTT TTT GGC GGC GAC AAC ATT GTG TGC ATG AGC TGC CCG CTG TGA CCGAACCC 3458  
CCTCGCGCCAGGCCCCCGCCACTGTGTGGCGTCCCAAGCTCTCTCCCTGTGCG ATG M 3518

Fig. 2. Nucleotide sequence and deduced amino acid sequence for the HSV-1 RR1 polypeptide. Translational initiation and stop codons for RR1 are underlined, and the ATG triplet for RR2 is indicated by two solid lines. Location of the 5' terminus of the 1.2-kb mRNA also is indicated.

RR LARGE SUBUNIT COMPARISON	
VZV	MEKRIPTVHD IIRKLOHGYKEYIIPRESTTPVELMEYISTIVSKLKAIVTODERVYVCCGEL 65
HSV-1	TGSPFPLATDQFGLNLYLVNQRUC LOVPVPPHAYNPVYLREYVTRLVNGFKLVSRRLYLQVL 430
VZV	INCRINLRSVNETHLPS(LCLCTPPVRAIGRDRDIRRAIEFPLKQYPAL ATL GLOEATKYE 131
HSV-1	VHLRITREASFEENLRSREVAL DFLTERLREHAGLVLAGAL DNYDCLHSTPHTLVERGLQALRYE 501
VZV	DYLTALREGLNLESCQFFULAAVTTEIVNLKAT LEPGNDG YTMTEVCRVFTTALACOKIVPATPV 202
HSV-1	EYLLKRRGCHYNSVFQHYTRIA GFLACRATGRMRHIALGREGSWENKFFPHRLYDGIIVSTPA 568
VZV	MPUGKETGATACCTYNDPESITVGRVRAITGGVGTQLSHGGVQISLQSLNLIPTENTKGLAVLKLDCM 276
HSV-1	MLNUGTRNTYTTSCYLVNFOATTNKATLRATNSVSAILARNSIGLCVQAFN DSGPCTASVMPALKVDSL 640
VZV	VMAINSDEKPTVCVYIPFPIVDELQTVLATAGNLVRDEIFRCNIFCCLTNPILPFRVYLYLKGSNVPQWTL 350
HSV-1	VAAHRESASPTGACVYIPFPIVRAVLRHMGVLAGEAQDNTESALNMPLEFPAALIRHLDGKERRVWTL 714
VZV	PDWRADILRLTHGEAPTSTYLRLREGLGVSSVPIDQIAFTIIRSAAVTGSPTLHFKDCAENRHHYHNTQNAI 423
HSV-1	FQDTSMSLADFHLEEFELVYHGLVGVFG EQIPIDELAYGIVSAAITGSPFVHFKDAVNRHYIYDQAAAI 787
VZV	TGNECTEYVGAADAHQGVNINLSTLTLKSGPVSNLNLQETARTVIFLQVLAAGNPCKKCKGVK 497
HSV-1	AGNELCTEYHFAKRSQSSNLNLSVNLARCVS ROTPDFORLRDAVQACVLNINIDSTLOFTQCTRGND 859
VZV	NNRSLEGTIGTIGTILRLGLDLSOPARLNVDAELMLYETKMTSMCKIKGLAPPKGTSEYAKGLHQ 571
HSV-1	NLRSELEIGDGLTACKLGLDLESAEFODLNKHAIEVNLKANKTSNALC VQGRPFNHPKRSYRAGRPHN 932
VZV	DGFSTISYLDLPKCTTBDICATGLVSDPLALPFISSAQVTECSGQSPFINNMFSTSGELLRPMLDL 644
HSV-1	EAPFDARPRYEGEENLROSHMGLNLSQNALPFISSAQISDVSEGAFLPTNLFSTVNDGETLKPNTLL 1006
VZV	MDLSDMYSCEKRLVINILEKQNSVIRSFQCLNSHPLKLTAFYEGEDLVNCAERAPPIDGQST 717
HSV-1	LKELERTFS GKRLLEVNDSLDKQMSVQALPCLEPTNPLRRKTAFTAFYDQKLLIDLCADRAPYVDG 1078
VZV	LFIEERPDGTPASKINMLIRAYAGELTQRTYKIRKATNSGLFAG GELTCTSCAL
HSV-1	LYVTEKADTLPASTLLVLLVAYKRELKTHYKRYKATNSGVGGDDNIVCMSCAL
RR SMALL SUBUNIT COMPARISON	
VZV	MD OKCSHFYRPECFDINHLRALSISINLSESDFIIEDDYGLCLTE 49
HSV-2	MDPAVSPASTDPLDTHSAGAAPIPVCTPPEFYTTSCFDINHLRSLISINLETELVLVPGDEQSKLSE 74
VZV	DELIFRYFIPTLSAADLVNVLGSLTGLFSCDINHYYTIEDECIEVMARVYSQIQMLFRGDSLRVQYVN 123
HSV-2	GELGTFYRFLFATLSAADLVNVLGSLTGLFSCDINHYYTIEDECIEVMARVYSQIQMLFRGDSLRVQYVN 148
VZV	VTIRKPSIQQYQMLEEKVRDNPSSAKEYILMLTIGDTFVSSFAAIALYLRNGLPVVTCQPNLISRDQVHT 197
HSV-2	RTINHPAIVKVVNMLEARVRECDSVPEFELMLTIGGVFFASFAAIALYLRNGLLVTCQSNLISRDQVHT 222
VZV	SASCCINNY VPERPAITRINQLTSEAVEIECAPKSHAP KTRLVNVOAITQYKFAADRLLSAINVP 265
HSV-2	TASCYIYNKYLGNHAKPEARVYKLFREAVDIEGIFRSQAPTSSILSPGALAAIENYVRFADRLGLLHM 296
VZV	KLFNTFPFGDGFPLAFMIADKNLFFRSTSVAGTVINOL
HSV-2	PLYSAPAFDSFPLSLNLSKDNLFFRSTSVAGTVINOL

Fig. 3. Amino acid comparisons of RR1 and RR2 polypeptides of HSV and VZV. Asterisks denote conserved amino acids, and underlined sequences represent regions which are highly conserved between HSV and VZV. The numbered blocks represent regions which are conserved in other viral and cellular ribonucleotide reductases. Proposed locations for the iron binding site and tyrosine free radical in RR2 are indicated above conserved histidine (●) and tyrosine (◆) residues.

towards the end of the N-terminal region which contains an unusual sequence of approximately alternating D and S residues (shaded in Fig. 5B), conserved between HSV-1 and HSV-2.

The conservation between the HSV-1 and HSV-2 sequences suggests that this area of RR1 is not a vestigial protein remnant. Differences between the HSV-1 and HSV-2 amino acids appear to be conservative, suggesting evolutionary pressure at the level of protein structure and function and not random drift.

### Structures of RR1 and RR2

The sequences of RR1 and RR2 from HSV-1, HSV-2, VZV, EBV, mouse, surf clam, and *E. coli* were aligned and secondary structures predicted<sup>43</sup>; these data are summarized in Figure 6.

TABLE I. Conserved Amino Acid Blocks in Ribonucleotide Reductases\*

Block	Herpesviruses	Eukaryotic	Prokaryotic
RR1			
1	+	—	—
2	+	+	+
3	+	+	+
4	+	+	+
5	+	—	—
6	+	+	+
7	+	+	+
8	+	+	—
9	+	+	+
10	+	—	—
11	+	+	—
12	+	—	—
13	+	+	—
14	+	+	—
15	+	+	+
RR2			
16	+	—	—
17	+	—	—
18	+	+	—
19	+	+	+
20	+	+	—
21	+	+	—

\*Conservation of the amino acid blocks (see Fig. 3) between ribonucleotide reductases from different sources. The presence (+) or absence (—) of a block is indicated.

RR1 can be divided into various regions, the first of which is the N-terminal region (Fig. 6, position 1), as discussed in the previous section. Beyond this domain, the structural predictions are predominantly  $\alpha$ -helical. The remainder of RR1 can be envisaged as two regions of  $\beta$ -sheet/ $\alpha$ -helix which are separated by an insertion of 34 amino acids present only in EBV (Fig. 6, position 3). The position of the *E. coli* allosteric effector site<sup>44</sup> is shown in Figure 6, position 2; this is located between amino acids 321 and 328 in the revised *E. coli* sequence (J.A. Fuchs, personal communication). Interestingly, upstream from this site, there is an insertion of seven amino acids which is conserved only in the allosterically regulated *E. coli* and mouse polypeptides and is absent in all herpesvirus RR1 proteins.

Figure 6, position 4 indicates the location of three conserved glycine residues with the sequence G-x-G-x-x-G (see Fig. 3, block 9); this pattern of conserved glycines which occurs only at this location in RR1 and RR2 is indicative of nucleotide binding sites.<sup>45</sup> In support of nucleotide binding at this site, a secondary structure of  $\beta$ -turn- $\alpha$  is predicted for this region, and this resembles the units which bind ADP. For RR1, this pattern does not extend to the hydrophobic residues expected in the  $\beta$ -strand, which follows the initial  $\beta$ -turn- $\alpha$  in the binding folds.<sup>45</sup> However, studies

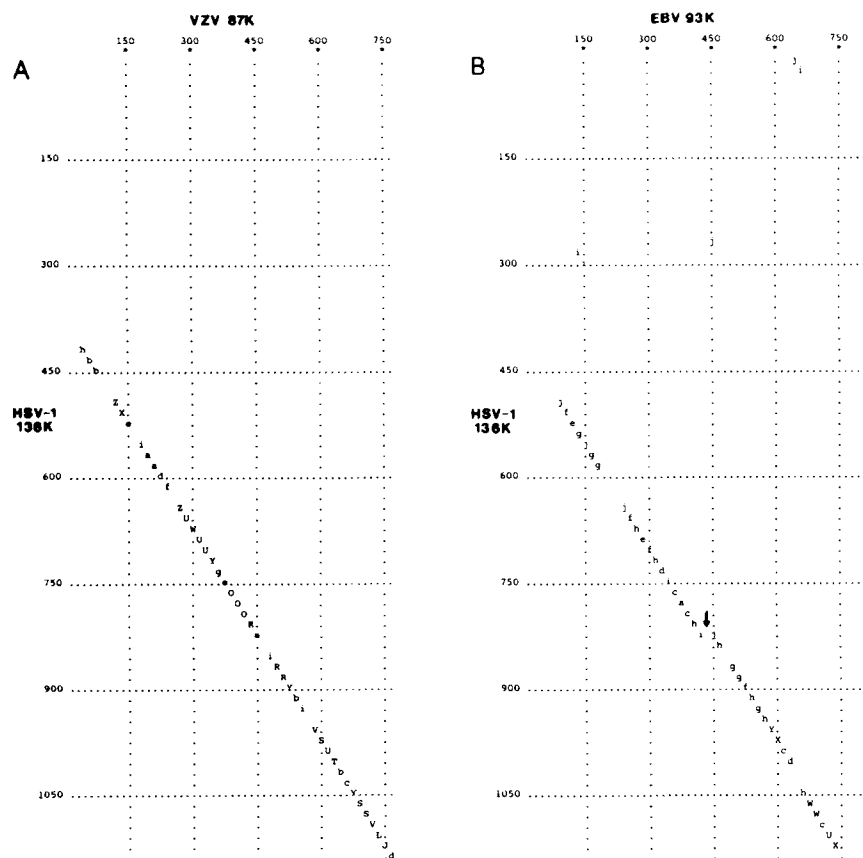


Fig. 4. CINTHOM comparisons of the HSV-1 RR1 amino acid sequence with those of VZV (A) and EBV (B) RR1 polypeptides. In panel B, the arrow indicates a shift in the diagonal due to the insertion of 34 amino acids into the EBV sequence.

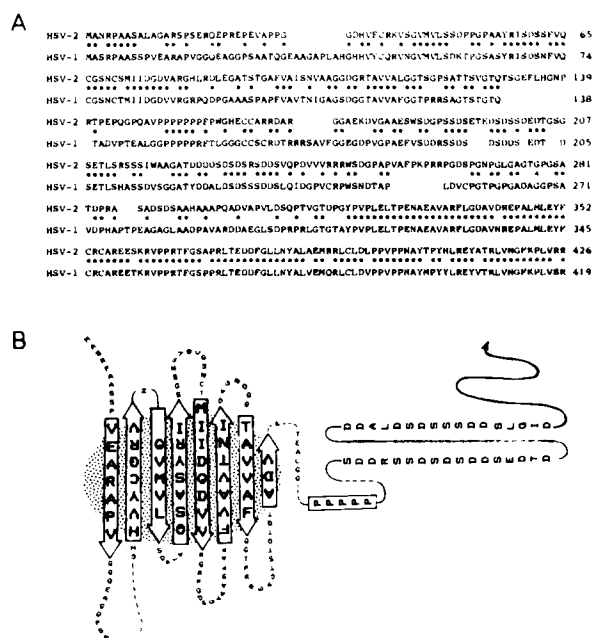


Fig. 5. A: Alignment of the amino acid sequences at the N-termini of the HSV-1 and HSV-2 RR1 polypeptides. Asterisks represent conserved amino acids. B: Schematic representation of a predicted secondary structure for the N-terminal domain of HSV-1 RR1. The oval shaded area represents the hydrophobic core. Amino acids in boxed arrows represent stretches of hydrophobic residues with a predicted  $\beta$ -sheet structure and sequences looping out of the core represent unstructured regions. The run of prolines shown separating the N-terminal domain is boxed, and beyond this the approximate sequence repeat is emphasized by shading.

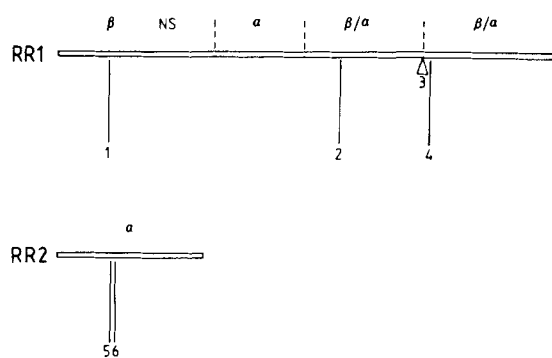


Fig. 6. A summary of the predicted secondary structures for RR1 and RR2. Dashed lines denote boundaries between larger areas of different secondary structure.  $\alpha$  and  $\beta$  represent areas of  $\alpha$ -helix and  $\beta$ -sheet, respectively; NS denotes nonstructured regions. Numbers shown represent positions within RR1 and RR2 and these are referred to in the text.

on the *E. coli* enzyme<sup>46</sup> indicate one substrate binding site for all four rNDPs; hence, this may influence the nucleotide binding fold structure.

For RR2, predictions show predominantly an  $\alpha$ -helical structure. Positions 5 and 6 in Figure 6 show the locations of the iron binding site and tyrosine free radical,<sup>24</sup> these amino acid residues are conserved between all RR2 polypeptides (Fig. 3).

An interesting feature of the RR2 comparisons is that the amino acid sequence, KDILHYYVE, which lies in the N-terminal portion of the HSV-2 polypeptide between residues 108 and 116 (Fig. 3, block 17), is almost perfectly conserved (KDILCQYVE) in the C-terminal region of the equivalent *E. coli* protein.<sup>47</sup> The presence of homologous sequences at different locations in the HSV-2 and *E. coli* polypeptides is not likely to have arisen by chance. It is possible that the N- and C-terminal positions in the HSV-2 and *E. coli* polypeptides may be structurally related, and this sequence similarity may be a result of convergent evolution. Alternatively, the different locations for these conserved residues may be due to movement of

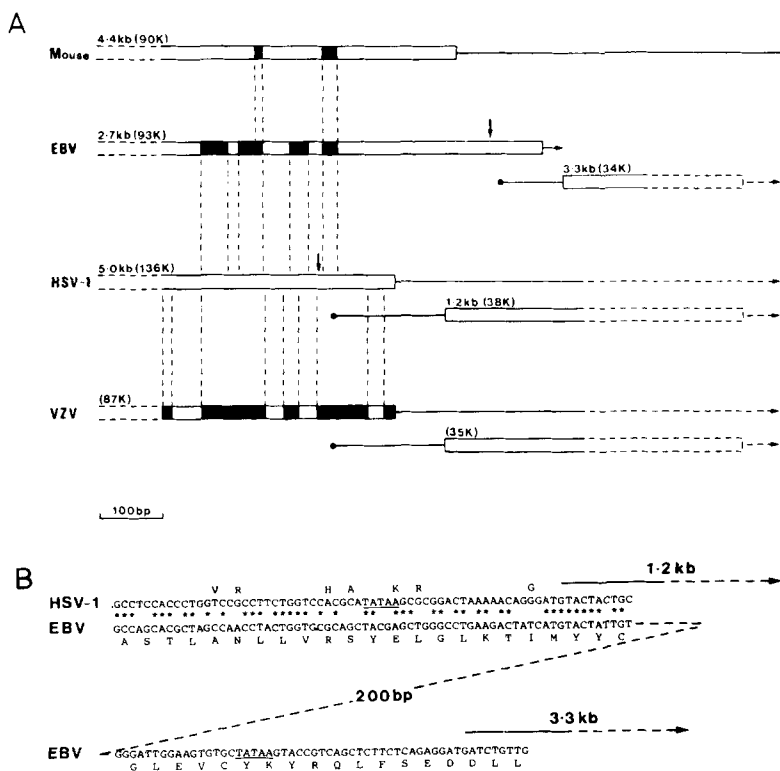


Fig. 7. **A:** Alignment of the C-terminal portions of RR1 polypeptides specified by mouse, EBV, HSV-1, and VZV. The sizes of proteins specified by the various mRNAs are given in brackets. Shaded blocks represent regions of conserved amino acids, and vertical dashed lines indicate their relative positions within equivalent polypeptides. Vertical arrows denote the position of TATA box homologues for the HSV-1 1.2-kb and EBV 3.3-kb mRNAs. **B:** Alignment of the nucleotide sequences specifying the C-terminal portions of HSV-1 and EBV RR1 polypeptides. Predicted RR1 amino acids are shown below the EBV nucleotide sequence, and amino acids not conserved in HSV-1 are shown above the HSV-1 nucleotide sequence. TATA box homologues are underlined, and the locations of the HSV-1 1.2-kb and EBV 3.3 kb mRNA 5' termini are indicated.

protein domains. Recombination of exons into introns has been proposed as a mechanism for shuffling domains and thereby increasing genetic complexity.<sup>48,49</sup> By implication, precursor polypeptides to the HSV-2 and *E. coli* RR2 proteins would have contained introns which have subsequently been deleted; recent evidence indicates that the T4 RR2 gene is spliced.<sup>50</sup>

### Evolution of the Promoters for the HSV and EBV RR2 Polypeptides

Figure 7A shows a representation of DNA sequence alignments at the RR1 C-terminal region of HSV-1, VZV, and EBV. For HSV-1 and VZV, the amino acid homology at this region is high, and the putative transcription regulatory elements upstream for the RR2 mRNA are conserved. In contrast, as can be seen in the CINTHOM comparison between HSV-1 and EBV (Fig. 4B), the C-terminal region of EBV RR1 contains an additional 75 amino acids which are not found in HSV or VZV. Within the sequences encoding these additional 75 amino acids are the upstream transcription regulatory elements for the EBV RR2 mRNA (Fig. 7A); these EBV sequences are also absent from the mouse M1 polypeptide. The nucleotide sequences specifying the C-terminal portions of HSV-1 RR1 and EBV RR1 (Fig. 7B) indicate that the HSV-1 TATA box for the 1.2-kb mRNA is changed to a TACGA sequence in EBV and the cap site of the 3.3-kb mRNA is not positioned within 30 nucleotides of this sequence. Thus the promoters for the HSV-1 1.2-kb and EBV 3.3-kb transcripts appear to have arisen by different evolutionary events.

Novel sequences containing transcription control signals may have been inserted fortuitously into the polypeptide coding region of a precursor EBV RR1. Indeed, insertion of a complete promoter downstream from a primitive promoter has been suggested for the major human chorionic somatomammotropin gene.<sup>51</sup> By contrast, the promoter for the HSV-1 1.2-kb mRNA does not appear to have arisen from insertion of extensive amino acid sequences into the protein coding region. Sequences upstream from the cap site of the 1.2-kb mRNA are present in EBV DNA, although this region does not contain a promoter in EBV; hence, the promoter signals for the 1.2-kb mRNA may have evolved by mutation of preexisting polypeptide coding sequences.

### ACKNOWLEDGMENTS

We thank Professor J.H. Subak-Sharpe and Dr. H.S. Marsden for comments on the manuscript. I.N. was the recipient of a NATO postgraduate studentship. We thank Ms. Linda Shaw for efficient preparation of the manuscript.

### REFERENCES

1. Thelander, L., Reichard, P. Reduction of ribonucleotides. *Annu. Rev. Biochem.* 48:133-158, 1979.
2. Cohen, G.H. Ribonucleotide reductase activity of synchronized KB cells infected with herpes simplex virus. *J. Virol.* 9:408-418, 1972.
3. Cohen, G.H., Factor, M.N., Ponce de Leon, M. Inhibition of herpes simplex virus type 2 replication by thymidine. *J. Virol.* 14:20-25, 1974.
4. Henry, B.E., Glaser, R., Hewetson, J., O'Callaghan, D.J. Expression of altered ribonucleotide reductase activity associated with the replication of Epstein-Barr virus. *Virology* 89:262-271, 1978.
5. Lankinen, H., Graslund, A., Thelander, L. Induction of a new ribonucleotide reductase after infection of mouse L cells with pseudorabies virus. *J. Virol.* 41:893-900, 1982.
6. Cohen, J.C., Henry, B.E., Randall, C.C., O'Callaghan, D.J. Ribonucleotide reductase activity in hydroxyurea-resistant herpesvirus replication. *Proc. Soc. Exp. Biol. Med.* 155:395-399, 1977.
7. Averett, D.R., Lubbers, C., Elion, G.B., Spector, T. Ribonucleotide reductase induced by herpes simplex virus type 1. Characterisation of a distinct enzyme. *J. Biol. Chem.* 258:9831-9838, 1983.
8. Averett, D.R., Furman, P.A., Spector, T. Ribonucleotide reductase of HSV-2 resembles that of HSV-1. *J. Virol.* 52:981-983, 1984.
9. Dutia, B.M. Ribonucleotide reductase induced by herpes simplex virus has a virus-specified constituent. *J. Gen. Virol.* 64:513-521, 1983.
10. Preston, V.G., Palfreyman, J.W., Dutia, B.M. Identification of a herpes simplex virus type 1 polypeptide which is a component of the virus-induced ribonucleotide reductase. *J. Gen. Virol.* 65:1457-1466, 1984.
11. Frame, M.C., Marsden, H.S., Dutia, B.M. The ribonucleotide reductase induced by herpes simplex virus type 1 involves minimally a complex of two polypeptides (136K and 38K). *J. Gen. Virol.* 66:1581-1587, 1985.
12. Bacchetti, S., Eveleigh, M.J., Muirhead, B. Identification and separation of the two subunits of the herpes simplex virus ribonucleotide reductase. *J. Virol.* 57:1177-1181, 1986.
13. Dutia, B.M., Frame, M.C., Subak-Sharpe, J.H., Clark, W.N., Marsden, H.S. Specific inhibition of herpesvirus ribonucleotide reductase by synthetic peptides. *Nature* 321:439-441, 1986.
14. Cohen, E.A., Gaudreau, P., Brazeau, P., Langelier, Y. Specific inhibition of herpesvirus ribonucleotide reductase by a nonapeptide derived from the carboxy terminus of subunit 2. *Nature* 321:441-443, 1986.
15. McLauchlan, J., Clements, J.B. A 3' co-terminus of two early herpes simplex virus type 1 mRNAs. *Nucleic Acids Res.* 10:501-512, 1982.
16. McLauchlan, J., Clements, J.B. Organisation of the herpes simplex virus type 1 transcription unit encoding two early proteins with molecular weights of 140000 and 40000. *J. Gen. Virol.* 64:997-1006, 1983.
17. Anderson, K.P., Frink, R.J., Devi, G.B., Gaylord, B.H., Costa, R.H., Wagner, E.K. Detailed characterisation of the mRNA mapping in the Hind III fragment *k* region of the herpes simplex virus type 1 genome. *J. Virol.* 37:1011-1027, 1981.
18. McLauchlan, J., Clements, J.B. DNA sequence homology between two colinear loci on the HSV genome which have different transforming abilities. *EMBO J.* 2:1953-1961, 1983.
19. Galloway, D.A., Swain, M.A. Organisation of the left-hand end of the herpes simplex virus type 2 Bgl II *n* fragment. *J. Virol.* 49:724-730, 1984.
20. Galloway, D.A., Goldstein, L.C., Lewis, J.B. Identification of proteins encoded by a fragment of herpes simplex virus type 2 DNA that has transforming ability. *J. Virol.* 42:530-537, 1982.
21. Swain, M.A., Galloway, D.A. Herpes simplex virus specifies two subunits of ribonucleotide reductase encoded by 3'-coterminal transcripts. *J. Virol.* 57:802-808, 1986.
22. Draper, K.G., Frink, R.J., Wagner, E.K. Detailed characterisation of an apparently unspliced  $\beta$  herpes simplex virus type 1 gene mapping in the interior of another. *J. Virol.* 43:1123-1128, 1982.
23. Caras, I.W., Levinson, B.B., Fabry, M., Williams, S.R., Martin, D.W. Jr. Cloned mouse ribonucleotide reductase subunit M1 cDNA reveals amino acid sequence homology with *Escherichia coli* and herpesvirus ribonucleotide reductases. *J. Biol. Chem.* 260:7015-7022, 1985.
24. Standart, N.M., Bray, S.J., George, E.L., Hunt, T., Ruderman, J.V. The small subunit of ribonucleotide reductase is encoded by one of the most abundant translationally regu-



- lated maternal RNAs in clam and sea urchin eggs. *J. Cell Biol.* 100:1968-1976, 1985.
25. Gibson, T., Stockwell, P., Ginsburg, M., Barrell, B. Homology between two EBV early genes and HSV ribonucleotide reductase and 38K genes. *Nucleic Acids Res.* 12:5087-5099, 1984.
  26. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580, 1983.
  27. Holmes, D.S., Quigley, M. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* 114:193-197, 1981.
  28. Deininger, P.L. Random subcloning of sonicated DNA: Application to shotgun DNA sequence analysis. *Anal. Biochem.* 129:216-223, 1983.
  29. Messing, J., Vieira, J. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* 19:269-276, 1982.
  30. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., Roe, B.A. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143:161-178, 1980.
  31. Biggin, M.D., Gibson, T.J., Hong, C.F. Buffer gradient gels and <sup>35</sup>S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* 80:3963-3965, 1983.
  32. McGeoch, D.J., Dolan, A., Donald, S., Brauer, D.H.K. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. *Nucleic Acids Res.* 14:1727-1745, 1986.
  33. Staden, R. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Res.* 10:4731-4751, 1982.
  34. Pustell, J., Kafatos, F.C. A high speed, high capacity homology matrix: zooming through SV40 and polyoma. *Nucleic Acids Res.* 10:4765-4782, 1982.
  35. Taylor, P. A fast homology program for aligning biological sequences. *Nucleic Acids Res.* 12:447-455, 1984.
  36. Taylor, W.R. The classification of amino acid conservation. *J. Theor. Biol.* 119:205-218, 1986.
  37. Garnier, J., Osguthorpe, D.J., Robson, B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97-120, 1978.
  38. Marsden, H.S., Stow, N.D., Preston, V.G., Timbury, M.C., Wilkie, N.M. Physical mapping of herpes simplex virus-induced polypeptides. *J. Virol.* 28:624-642, 1978.
  39. Wilcox, K.W., Kohn, A., Sklyanskaya, E., Roizman, B. Herpes simplex virus phosphoproteins. I. Phosphate cycles on and off some viral polypeptides and can alter their affinity for DNA. *J. Virol.* 33:167-182, 1980.
  40. McGeoch, D.J. The nature of animal virus genetic material. In: "The Microbe 1984." Part 1. Viruses. Mahy, B.W.J., Pattison, J.R., eds. Cambridge: Cambridge University Press, 1984:75-107.
  41. Davison, A.J., Scott, J.E. The complete DNA sequence of Varicella-Zoster virus. *J. Gen. Virol.* 67:1759-1816, 1986.
  42. Baer, R., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Sequin, C., Tuffnell, P.S., Barrell, B.G. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature* 310:207-211, 1984.
  43. Taylor, W.R. Identification of protein sequence homology by consensus template alignment. *J. Mol. Biol.* 188:233-258, 1986.
  44. Eriksson, S., Sjöberg, B.M., Jorvall, H., Carlquist, M. A photoaffinity-labeled allosteric site in *Escherichia coli* ribonucleotide reductase. *J. Biol. Chem.* 261:1878-1882, 1986.
  45. Wierenga, R.K., De Maeyer, M.C.H., Hol, W.G.J. Interaction of pyrophosphate moieties with  $\alpha$ -helices in dinucleotide binding proteins. *Biochemistry* 24:1346-1357, 1985.
  46. von Döbeln, U., Reichard, P. Binding of substrates to *Escherichia coli* ribonucleotide reductase. *J. Biol. Chem.* 251:3616-3622, 1976.
  47. Carlson, J., Fuchs, J.A., Messing, J. Primary structure of the *Escherichia coli* ribonucleoside diphosphate reductase operon. *Proc. Natl. Acad. Sci. USA* 81:4294-4297, 1984.
  48. Blake, C.C.F. Do genes-in-pieces imply proteins-in-pieces? *Nature* 273:267-268, 1978.
  49. Gilbert, W. Why genes in pieces? *Nature* 271:501, 1978.
  50. Sjöberg, B.-M., Hahne, S., Mathews, C.Z., Mathews, C.K., Rand, K.N., Gait, M.J. The bacteriophage T4 gene for the small subunit of ribonucleotide reductase contains an intron. *EMBO J.* 5:2031-2036, 1986.
  51. Selby, M.J., Barta, A., Baxter, J.D., Bell, G.I., Eberhardt, N.L. Analysis of a major human chorionic somatomammotropin gene. Evidence for two functional promoter elements. *J. Biol. Chem.* 259:13131-13138, 1984.