JOURNAL OF VIROLOGY, May 1983, p. 434–445
0022-538X/83/050434-12$02.00/0

Vol. 46, No. 2

Processing of Herpes Simplex Virus Proteins and Evidence that Translation of Thymidine Kinase mRNA Is Initiated at Three Separate AUG Codons

H. S. MARSDEN,[1]* L. HAARR,[2] AND C. M. PRESTON[1]

Medical Research Council Virology Unit, Glasgow, G11 5JR, Scotland,[1] and Department of Biochemistry, University of Bergen, Bergen, Norway[2]

Received 8 September 1982/Accepted 14 February 1983

The role which post-translational modification plays in the genesis of herpes simplex virus-induced polypeptides was investigated. Two-dimensional gel electrophoresis was used to identify those polypeptides (i) synthesized in vitro, (ii) labeled in vivo during a pulse, and (iii) labeled after a chase. Excluding glycoproteins, we detected 36 precursor or short-lived polypeptides, 8 polypeptides which were generated by post-translational modification, 46 polypeptides which were apparently not modified after synthesis, and 19 polypeptides which were either transient intermediates or not modified. Comparison of polypeptides synthesized in vitro and during an in vivo pulse showed that translation in vitro resembles quite closely translation in vivo and that amounts of protein synthesized in vivo are determined largely by the levels of mRNA. This analysis provided the basis for an investigation of the suggestion (C. M. Preston and D. J. McGeoch, J. Virol. 38:593–605, 1981) that the two polypeptides of apparent molecular weights of 43,000 (VI 43) and 39,000 (VI 39) encoded by the herpes simplex virus type 1 thymidine kinase gene are translated from a single mRNA by two in-phase initiation codons. Hybrid arrest was used to identify in vitro translation products encoded by the thymidine kinase gene. Two-dimensional gel electrophoresis showed that VI 39 was more acidic than VI 43, consistent with the predicted amino acid composition of a polypeptide whose synthesis was initiated at the second AUG codon, located 135 bases downstream from the first. Furthermore, two-dimensional gels revealed a third polypeptide whose synthesis was arrested by the same fragment. Its pI and apparent molecular weight (38,000) were compatible with initiation of translation at a third AUG codon an additional 42 bases downstream. Our findings provide strong evidence that downstream initiation codons within the thymidine kinase mRNA are used.

It is now well established that many herpes simplex virus (HSV)-induced polypeptides are subject to post-translational modification (12). Such modifications are likely to be important to the normal functioning of the polypeptide. Indeed, Preston (19) showed that the defect in mutant *tsK* which results in over-production of immediate early proteins at nonpermissive temperatures was associated with a failure to fully process the 175,000 (175K)-molecular-weight polypeptide in which the mutation lies. Post-translational modifications are usually detected by shifts in mobility on sodium dodecyl sulfate-polyacrylamide gels. Shifts are often very small and not reproducibly detected. Since many of the modifications which proteins undergo result in charge differences, it seemed appropriate to examine processing by two-dimensional gel electrophoresis in which separation in the first dimension is based on the inherent charge of the protein. To do this, we compared the mobilities of polypeptides (i) synthesized in vitro, (ii) labeled in vivo during a pulse, and (iii) labeled after a chase.

A second objective of this study, made possible by the two-dimensional analysis of in vitro-synthesized polypeptides, was to investigate synthesis of the HSV-induced pyrimidine deoxyribonucleoside kinase (thymidine kinase; TK) (3). This enzyme is of particular interest since DNA sequences containing the TK gene are capable of biochemically transforming cells (11, 15, 21, 25). The gene has been completely sequenced (9, 24; D. J. McGeoch, unpublished results) and is capable of coding for a protein of 376 amino acids with a molecular weight of 40,918. Early after infection, transcription is from one strand only (9, 23, 26).

Recently, Preston and McGeoch (20) identified both in vivo and in vitro two polypeptides of apparent molecular weights of 43K (VI 43) and 39K (VI 39) encoded by the HSV type 1 (HSV-1) TK gene. Since the message does not appear to be spliced (9, 20), they concluded that the coding sequences for the two polypeptides must overlap and that the polypeptides were probably translated from a single message. They suggested three ways in which VI 43 and VI 39 could be related: (i) proteolytic cleavage, (ii) premature termination of protein synthesis, and (iii) secondary initiation of translation downstream from the first AUG codon. The first two alternatives were considered unlikely, the first because no change in the relative intensity of the two polypeptides was detected during a 20-h chase and the second because translation of both fragments was arrested equally well (20) by a fragment (pTK1 *Sma*I B) which overlapped the 3′ coding sequences by only 18 base pairs (9, 24). The third alternative derives support from inspection of the TK DNA sequence which shows a second AUG codon 135 bases downstream from that at which translation of VI 43 is thought to be initiated. If used, this second codon would generate a polypeptide of a size compatible with that of the observed VI 39 (20).

Evidence that synthesis of VI 39 is initiated at the second codon would be provided if VI 43 and VI 39 shared common tryptic peptides. However, we have been unable to test the possibility by this approach because VI 39 has not yet been isolated in sufficient amounts. We have therefore used a different and new approach. In vitro translation products encoded by the TK gene were identified by hybrid arrest, using a DNA fragment mapping entirely within the TK coding sequences. The mobility of these polypeptides on two-dimensional gels was then determined. The results presented here provide strong evidence that initiation of protein synthesis within the TK mRNA occurs not only at the second AUG, but also at a third AUG codon an additional 42 bases downstream from that generating VI 39, to give rise to a previously undetected polypeptide of an apparent molecular weight of 38K.

MATERIALS AND METHODS

Cells. BHK 21 clone 13 (10) or BSC-1 cells were grown in Eagle medium supplemented with 10% calf serum.

Virus. HSV-1 strain 17 syn$^+$ (2) was used throughout the study.

Chemicals. Chemicals were obtained from the sources listed by Haarr and Marsden (4), except that urea was from Bethesda Research Laboratories (Cambridge, England), acrylamide for slab gels was from E. Merck AG (Darmstadt, Germany), *N,N′*-diallyltartardiamide (DATD) was from EGA-Chemie (Steinheim, Albuch, West Germany) and 2,5-diphenyloxazole was from Koch-Light Labortories (Slough, England).

Radioisotope labeling of infected cells. Infected cell monolayers in 30-mm dishes were labeled with [^{35}S]methionine (Amersham International, Amersham, England; SJ 204; 250 μCi in 0.7 ml of phosphate-buffered saline) at 5 h postabsorption (6). After 30 min, the cells were either harvested immediately (pulse) or chased by removing the label, washing the monolayer, and incubating additionally in Eagle medium supplemented with 10% calf serum.

Hybrid arrest. The *Bgl*II-*Sst*I fragment was cleaved from pTK1 DNA, isolated, denatured, and hybridized to cytoplasmic mRNA from either uninfected or HSV-1-infected cells (20). RNA containing polyadenylic acid was selected on an oligodeoxythymidylic acid-cellulose column, precipitated with ethanol, dissolved in water, and used for in vitro translation either directly or after hybridization with, and denaturation from, DNA (20). The samples were frozen at −70°C.

In vitro protein synthesis. RNA samples were translated in a micrococcal nuclease-treated fractionated reticulocyte system (16–18).

Preparation of protein and two-dimensional electrophoretic analysis. Two-dimensional isoelectric focusing (IF) or non-equilibrium pH gradient gel electrophoresis (NEPHGE) was performed essentially as described previously by O'Farrell et al. (13, 14), with slight modifications (4). IF was performed for 23 h at 400 V and NEPHGE for 3 h at 550 V. The second dimension slab gels were either gradients of 5 to 12.5% polyacrylamide cross-linked with 1 in 20 (wt/wt) *N,N′*-methylenebisacrylamide (MBA) (8), or 9% polyacrylamide crosslinked with 1 in 40 (wt/wt) *N,N′*-diallyltartardiamide (DATD) (5).

Measurement of pH gradient in IF gels. We ran IF gels on which no protein was loaded. After electrophoresis, the gels were cut into 4-mm slices which were put into tubes containing 1.5 ml of degassed water. The air was replaced by nitrogen, and the tubes were sealed and left overnight at room temperature with occasional shaking. The pH was measured immediately after removing the seal.

Migration of pH markers in NEPHGE gels. A set of nine protein markers (Pharmacia, Uppsala, Sweden) with pI's ranging from 3.50 (amyloglucosidase) to 9.30 (trypsinogen) were separated by NEPHGE under standard conditions. Proteins were stained with Coomassie brilliant blue to determine the positions to which they migrated.

Fluorography. The slab gels were infused with 2,5-diphenyloxazole (1) and then dried and exposed to Kodirex XR-1 films at −80°C.

Polypeptide nomenclature. Polypeptides synthesized in vitro and in vivo are classified according to their apparent molecular weights ($\times 10^{-3}$) and are prefixed VI or V_{mw}, respectively (8, 17). Polypeptides separated by two-dimensional gel electrophoresis have been identified with a number according to a scheme described later (see Table 2). This number has been related to our two-coordinate system described previously (4). In that study, several very basic polypeptides which migrated off the end of our standard 8-cm NEPHGE gel and which were resolved by using a more basic gel were assigned relative mobilities greater than 1.0. In this study, we have used longer gels (10

TABLE 1. Correlation of polypeptide spot number with coordinates of migration in two-dimensional gels

Polypeptide	Coordinates: 5–12.5% gel cross-linked with MBA	Coordinates: 9% gel cross-linked with DATD	Polypeptide	Coordinates: 5–12.5% gel cross-linked with MBA	Coordinates: 9% gel cross-linked with DATD	Polypeptide	Coordinates: 5–12.5% gel cross-linked with MBA	Coordinates: 9% gel cross-linked with DATD
1	155,0.23–0.47	155,0.23–0.47	61	43,0.655	44,0.655	201	ND[a]	135–139,0.35
2	ND	145,0.26–0.36	62	43,0.71	44,0.71	202[b]	129,0.35	ND
3	130,0.35	130,0.35	63	38.5,0.22	39.5,0.22	203[b]	119,0.31	ND
4	130,0.38	130,0.38	64	40,0.33	47,0.33	204[b]	116,0.30	ND
5	125,0.30	122,0.30	65	ND	43.5,0.33	205	110,0.30	113,0.30
6	ND	124,0.40	66	40,0.35	45.5,0.35	206	108,0.33	108,0.33
7	128,0.25–0.33	119,0.25–0.33	67	39,0.35	43,0.35	207	103,0.31	106,0.31
8	117,0.64–0.83	130,0.64–0.83	68	ND	43,0.375	208	ND	49,0.33
9	100,0.30	100,0.30	69[c]	39,0.38	40,0.38	209	38.5,1.00	39,1.00
10	95,0.30	95,0.30	70[c]	38,0.38	39.5,0.38	210	21,0.87	21,0.87
11	92,0.45	ND	71	38.5,0.71	40,0.71	211	21,0.94	21,0.94
12	ND	91,0.32–0.38	72	38.5,0.79	40,0.79			
13	87,0.22–0.35	87,0.22–0.35	73	38.5,0.825	40,0.825			
14	99,0.5–0.84	88,0.5–0.84	74	38.5,0.86	40,0.86			
15	86,0.41–0.47	91,0.41–0.47	75	38.5,0.915	40,0.915			
16	78,0.365	82.5,0.365	76	ND	39.5,0.85			
17	78,0.385	82.5,0.385	77	38,0.95	38.5,0.95			
18	77,0.365	80,0.365	78	37.5,0.95	ND			
19	77,0.385	80,0.385	79	38.5,0.395	ND			
20	75,0.375	75,0.375	80	38.5,0.44	39,0.44			
21	76,0.385	76.5,0.385	81	38.5,0.48	39,0.48	301[b]	95,0.535	ND
22	ND	83,0.425	82	38.5,0.575	39,0.575	302[b]	95,0.56	ND
23	ND	81.5,0.425	83	38,0.665	39,0.665	303[b]	95,0.58	ND
24	81,0.775	80,0.775	84	38,0.345	39.5,0.345	304	95,0.61	95.5,0.61
25	81,0.83	80,0.83	85	38,0.395	ND	305	97,0.66	97.5,0.66
26	67,0.31	67,0.31	86	ND	38.5,0.385	306	100,0.73	100,0.73
27	ND	72,0.435	87	37.5,0.32	37.5,0.32	307	87.5,0.43	93.5,0.43
28	ND	72,0.45	88	37.5,0.34	37.5,0.34	308	85,0.42	87,0.42
29	67,0.51	70,0.51	89	37.5,0.36	37.5,0.36	309	85,0.45	87,0.45
30	ND	72,0.425	90	37.5,0.835	38,0.835	310	78,0.345	82.5,0.345
31	ND	72,0.45	91	37,0.88	37.5,0.88	311	77,0.35	80,0.35
32	67,0.58	70,0.48	92	30,0.325	30,0.325	312[b]	75.5,0.40	ND
33	68,0.81	ND	93	29,0.09	30,0.09	313[b]	75,0.42	ND
34	65,0.13	65,0.13	94	28.5,0.20	29,0.2	314	62.5,0.195	65,0.195
35	65,0.16–0.23	65,0.16–0.23	95	28.5,0.22	ND	315	62.5,0.22	65,0.22

TABLE 1—*Continued*

Poly-peptide	Coordinates: 5–12.5% gel cross-linked with MBA	Coordinates: 9% gel cross-linked with DATD	Poly-peptide	Coordinates: 5–12.5% gel cross-linked with MBA	Coordinates: 9% gel cross-linked with DATD	Poly-peptide	Coordinates: 5–12.5% gel cross-linked with MBA	Coordinates: 9% gel cross-linked with DATD
36	65,0.435	63,0.435	96	29,0.425	ND	316[b]	38,0.33	ND
37	65,0.45	63,0.45	97	ND	28,0.245	317	38,0.355	38,0.355
38	ND	63,0.15	98	28,0.295	28,0.295	318	37.5,0.36	37,0.36
39	62.5,0.24	65,0.24	99	28,0.33	28,0.33	319[b]	18,0.54	ND
40	62.5,0.27	65,0.27	100	28,0.80	28,0.80			
41	62.5,0.295	65,0.295	101	28,0.84–0.89	28,0.84–0.89	401	86,0.32–0.38	91,0.32–0.38
42	62.5,0.31–0.35	65,0.31–0.35	102	26.5,0.10	ND	402	ND	82,0.74–0.78
43	ND	59.5,0.34	103	26.5,0.43	ND	403	65,0.16	67,0.16–0.23
44	57.5,0.47	ND	104	27,1.00	ND	404[b]	46.5,0.29	ND
45	ND	60,0.48	105	26.5,0.825	27.5,0.825	405	ND	41,0.33
46	ND	59,0.48	106	22.5,0.235	ND	406	38.5,0.35	40,0.35
47	ND	57,0.55	107	22.5,0.425	ND	407	ND	36.5,0.30
48	52,0.51	55.5,0.51	108	21.5,0.27	23,0.27	408	27,0.315	27,0.315
49	52,0.565	ND	109	21.5,0.815	21.5,0.815			
50	51.5,0.60	55,0.60	110	ND	21,0.565			
51	51.5,0.66	55,0.66	111	21,0.815	21,0.815			
52	51.5,0.695	55,0.695	112	19,0.40	ND			
53	51,0.80–0.86	53,0.80–0.86	113	19.5,0.44	ND			
54	51,0.575	53,0.575	114	19,0.50	ND			
55	50,0.625	52,0.625	115	18,0.75	ND			
56	48,0.55	52.5,0.55	116	18,0.80	ND			
57	ND	53,0.25	117	11,0.95	ND			
58	ND	53,0.385	118	11,0.58	ND			
59	43.5,0.48–0.52	47.5,0.48–0.52						
60[d]	43,0.575	44,0.575						

[a] ND, Not detected.
[b] Not seen on the figures shown.
[c] The molecular weight coordinates previously assigned to spots 69 and 70 (39.5K and 39K, respectively) have been changed to fit with the values determined by Preston and McGeoch (20).
[d] Not previously detected.

cm) to retain these polypeptides. Accordingly, spots which previously had relative mobilities between 0.8 and 1.4 (4) have now been given new values which lie between 0.8 and 1.0 (Table 1).

RESULTS

Identification of polypeptides synthesized in vitro. mRNA was extracted from BHK or BSC-1 cells 5 h after infection or mock infection and translated in vitro. Polypeptides were separated by two-dimensional NEPHGE. Virus-induced polypeptides were identified by aligning and comparing fluorographs of the in vitro translation products from infected cells (Fig. 1 and 2) with fluorographs from similar mock-infected cell controls (data not shown) as described previously (4). The data shown were obtained with mRNA from infected BHK cells, but only polypeptides translated with mRNA from both infected BHK cells and infected BSC-1 cells have been included. Excluding the regions of multiple spots which are enclosed in boxes and which appear to be generated by in vitro processing of glycoproteins (14a), we detected 118 virus-induced polypeptides (numbers 1 through 118). Both 5 to 12.5% gradient gels cross-linked with MBA (Fig. 1) and 9% gels cross-linked with DATD (Fig. 2) were used, but only 64 of the 118 polypeptides were detected on both systems.

From non-equilibrium pH gradient gels it is not possible to determine accurately the pIs of polypeptides. However, an estimate can be made by coelectrophoresis with a set of marker proteins. The positions to which they migrated are shown in Fig. 1.

Identification of polypeptides synthesized in vivo. For comparison of in vivo-synthesized polypeptides with those synthesized in vitro, infected cells were pulse-labeled for 30 min at 5 h postinfection and either harvested immediately (Fig. 3A) or chased for 6 h and then harvested (Fig. 3B). Polypeptides induced after infection were again identified by comparison of these fluorographs with fluorographs of similarly

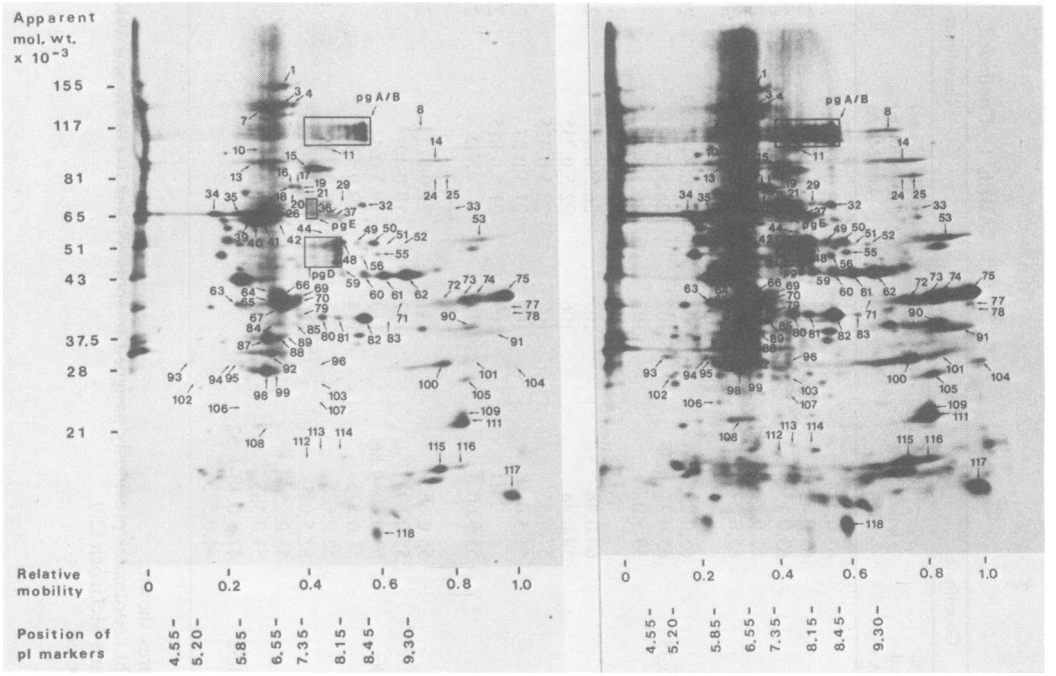

FIG. 1. Polypeptides translated in vitro from mRNA extractd from HSV-1-infected cells. mRNA was extractd from BHK cells 5 h after infection and translated in vitro using [^{35}S]methionine as the label. Separation was by two-dimensional NEPHGE. Arrows show virus-induced proteins, identified as described in the text. In this and the two-dimensional gels shown in Fig. 2, 3, 4, and 6, ampholines pH 3.5 to 10 were used for electrophoresis in the first dimension. A 5 to 12.5% gradient gel cross-linked with MBA was used for the second dimension. [^{35}S]methionine-labeled polypeptides extracted from in vivo-infected cells served as molecular weight markers (4). The position to which proteins of known pI migrated is indicated along the bottom of the figure. The fluorograph on the right is an overexposure of that on the left to show more clearly the minor polypeptides. In this figure and in Fig. 2 and 3, pgA/B, pgC, pgD, and pgE are precursor forms of glycoproteins A/B, C, D, and E, respectively.

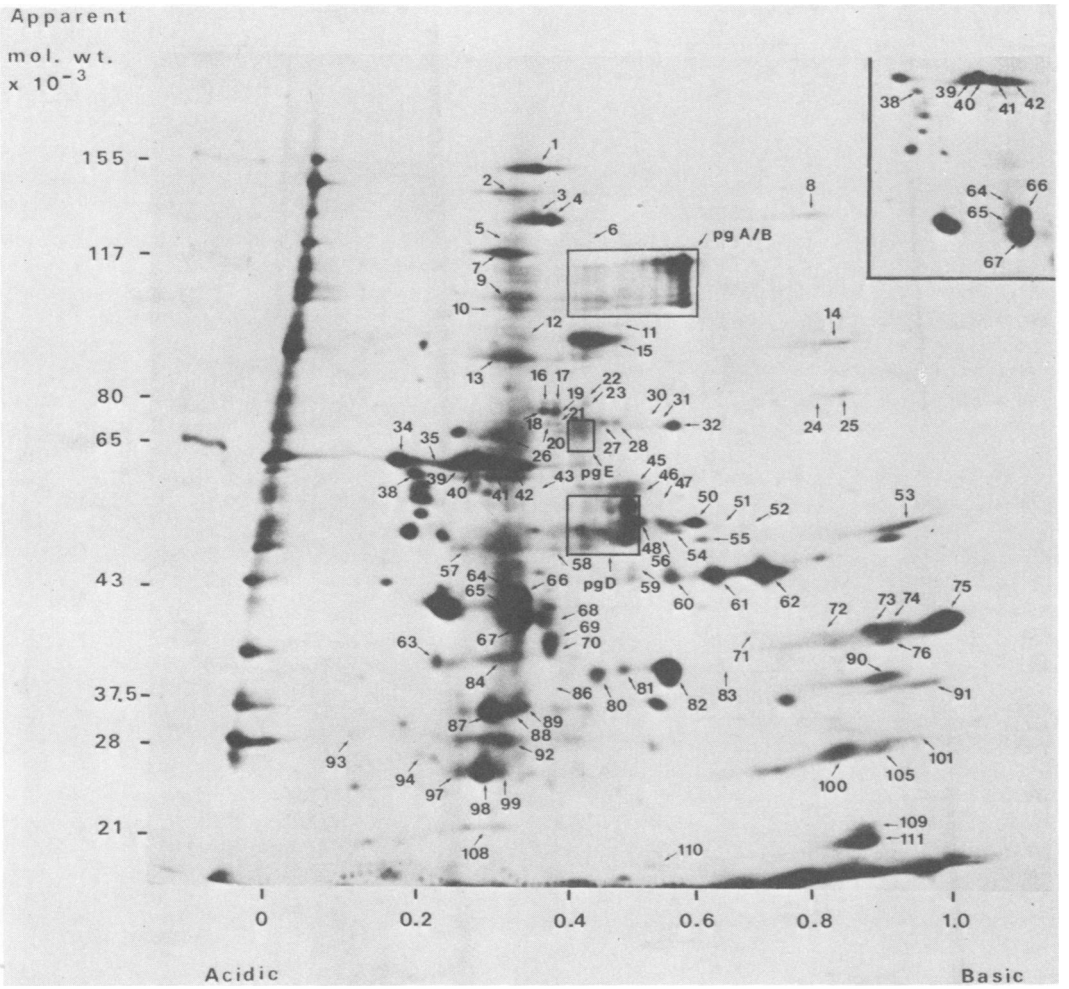

FIG. 2. Polypeptides translated in vitro as described in the legend in Fig. 1, except that the second dimension was a 9% gel cross-linked with DATD. The insert in the upper right hand corner is a less-exposed fluorograph of a portion of the whole gel and shows polypeptides having molecular weights between about 36K and 70K and relative mobilities between about 0.1 and 0.4.

treated mock-infected cell polypeptides (data not shown). Polypeptides were resolved on both 5 to 12.5% gradient gels cross-linked with MBA and 9% gels cross-linked with DATD, although only the latter gels are shown. Not all virus-induced polypeptides detected in the fluorograph of Fig. 3B have been indicated. For clarity only the glycoprotein regions (14a), a few readily identifiable polypeptides and those polypeptides which by criteria discussed later are considered to arise by post-translational modification have been numbered.

Of the 118 polypeptides synthesized in vitro, 69 have been identified among the 101 polypeptides synthesized during a pulse in vivo (Table 2). This suggests that translation in vitro resembles quite closely translation in vivo. Even more striking is the similarity in spot appearance and intensity of polypeptides translated in vitro and during an in vivo pulse (e.g., compare polypeptides 1, 60–62, 72–75, and 80–82 in Fig. 2 and 3A). This indicates that the amounts of protein synthesized in vivo are determined largely by the levels of mRNA and that our mRNA extraction procedure is not selective.

Genesis of polypeptides. Five classes of polypeptides can be recognized (Table 2). All class 0 polypeptides are translated in vitro. This class contains 36 members and is represented by polypeptide 14 (Fig. 1). The genesis of polypeptides in this class could not be determined because their presence or absence on some of the

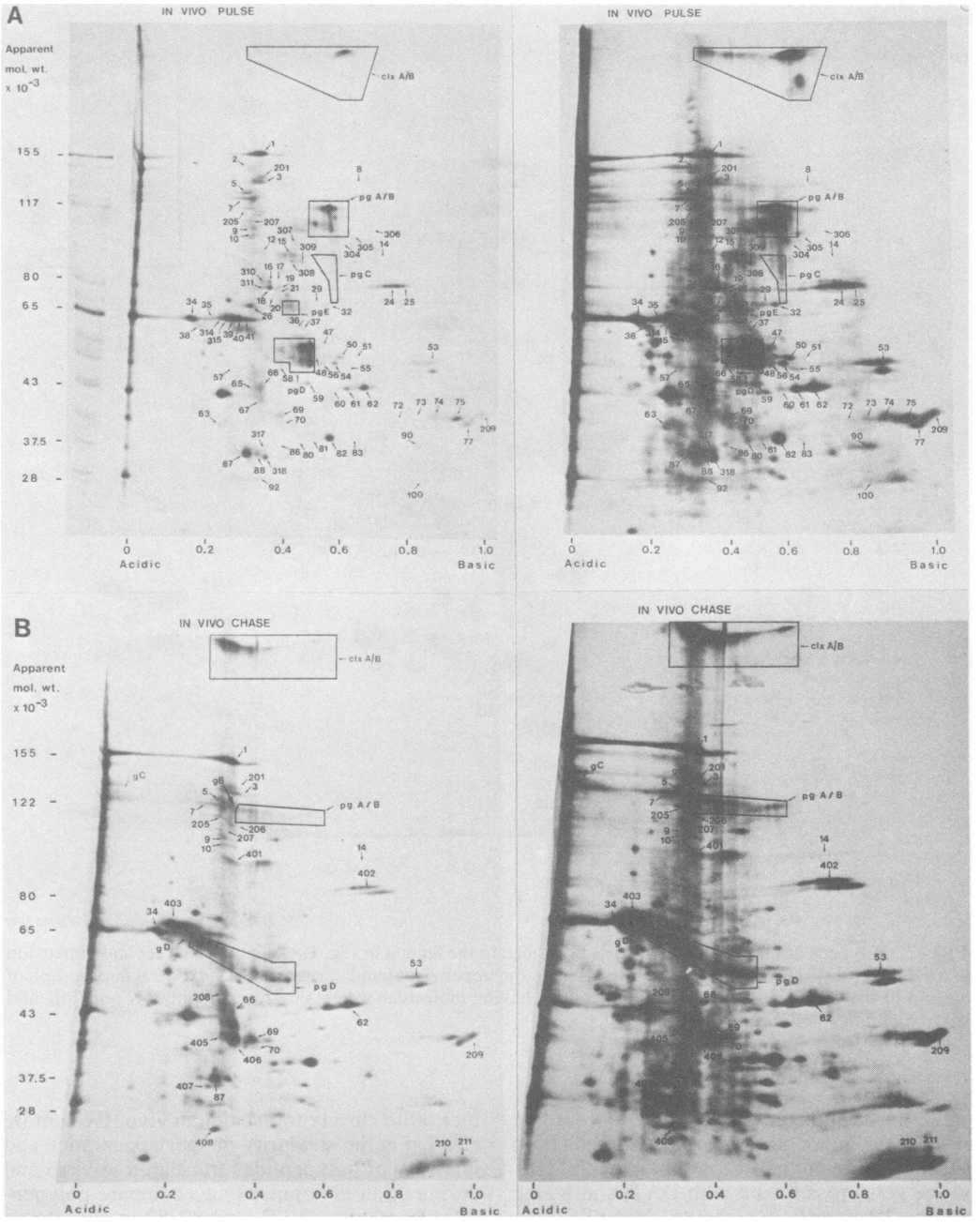

FIG. 3. Synthesis and processing of polypeptides in BHK cells infected with HSV-1 strain 17. Infected cells were labeled with [^{35}S]methionine for 30 min at 5 h after absorption and either harvested immediately (A, in vivo pulse) or chased for 6 h (B, in vivo chase). A 5 to 12.5% gradient gel cross-linked with MBA was used for the second dimension. The fluorographs on the right are overexposures of those on the left to show more clearly the minor polypeptides. The very slowly migrating spots enclosed in a box labeled clx(A/B) are thought to be multimers of glycoprotein A/B (clx, complex).

gels could not be unambiguously established, for a variety of reasons (Table 2). Class 1 contains all other polypeptides synthesized in vitro. It can be subdivided into three subclasses. In subclass 1a are those polypeptides which are detected in vivo during both a pulse and a chase.

TABLE 2. Classification and genesis of polypeptides

Class	Genesis	Detected in vitro	Detected in vivo		Polypeptides
			Pulse	Chase	
0	Unassigned[a]	+	*a*	*a*	8, 12, 14, 24, 25, 29, 30, 31, 35, 45, 46, 51, 56, 64, 65, 66, 71, 77, 78, 79, 86, 90, 91, 93, 97, 98, 100, 104, 107, 110, 111, 112, 113, 114, 117, 118
1a	Unmodified	+	+	+	1, 2, 3, 4, 5, 7, 9, 10, 11, 15, 20, 26, 32, 33, 34, 36, 37, 38, 44, 53, 54, 57, 59, 60, 61, 62, 63, 67, 69, 70, 72, 73, 74, 75, 80, 81, 82, 83, 85, 87, 88, 92, 101, 108, 115, 116
1b	Precursor–short-lived	+	+	−	16, 17, 18, 19, 21, 39, 40, 41, 47, 48, 49, 50, 55, 58, 99
1c	Precursor–short-lived	+	−	−	6, 13, 22, 23, 27, 28, 42, 43, 52, 68, 76, 84, 89, 94, 95, 96, 102, 103, 105, 106, 109
2	Rapidly processed (or unmodified)[b]	−	+	+	201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211
3	Transient intermediate (or precursor–short-lived)[b]	−	+	−	301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319
4	Processed product[b]	−	−	+	401, 402, 403, 404, 405, 406, 407, 408

[a] The genesis of polypeptides in class 0 could not be determined because their presence or absence on some of the gels could not be unambiguously established for the following reasons: too faint, 31, 51, 56, 86, 106, 107, 112, 113, 114; very basic and may have migrated partly off the end of some gels, 77, 78, 104; comigrating cellular protein in vivo, 91, 97; differently shaped spot in vivo and in vitro, 8, 14, 24, 25, 71, 100, 111; insufficiently resolved from other virus-induced spots, 12, 29, 30, 35, 45, 46, 64, 65, 79, 93; resolved in insufficient experiments, 110, 118; intensity reduced but not zero after chase, 66, 90, 98, 117.

[b] It is possible that the mRNA for some or all of the polypeptides in these classes is not translated in vitro, in which case such polypeptides in classes 2 and 3 would be more appropriately assigned to classes 1a and 1b respectively.

There is no evidence that this class is modified after synthesis. The class contains 46 members and is represented by polypeptide 1 (Fig. 1 and 2). Subclass 1b is detected in vivo during a pulse but not during a chase, whereas subclass 1c is not detected at all in vivo. We consider subclasses 1b and 1c to consist of either precursor polypeptides which undergo post-translational modification or polypeptides which are short-lived (precursor–short-lived). There are 15 members in subclass 1b and 21 in subclass 1c, represented by polypeptides 18 (Fig. 1, 2, 3A, and 4A) and 52 (Fig. 1 and 2), respectively. Class 2 is comprised of those polypeptides detected in vivo during both a pulse and a chase but not detected in vitro. The class contains 11 members and is represented by polypeptide 201. Such polypeptides are probably either rapidly processed species or, alternatively, are unmodified species which for some reason are not translated in vitro. With the available data we cannot distinguish between these two possibilities. Polypeptide class 3 (19 members) is comprised of polypeptides which are detected in an in vivo pulse but are not detected during the chase or in vitro. Such polypeptides (e.g., polypeptide 310) probably represent either transient intermediates in the synthesis of other polypeptides or precursor–short-lived polypeptides which, like those in class 2, are not translated in vitro. Again, with the available data we cannot distinguish between these two possibilities. Class 4 contains those polypeptides which are detected only after the chase and which in vitro or in an in vivo pulse are either not detected or only present in substantially reduced amounts. This class contains eight members in addition to the glycoproteins and is represented by polypeptide 401.

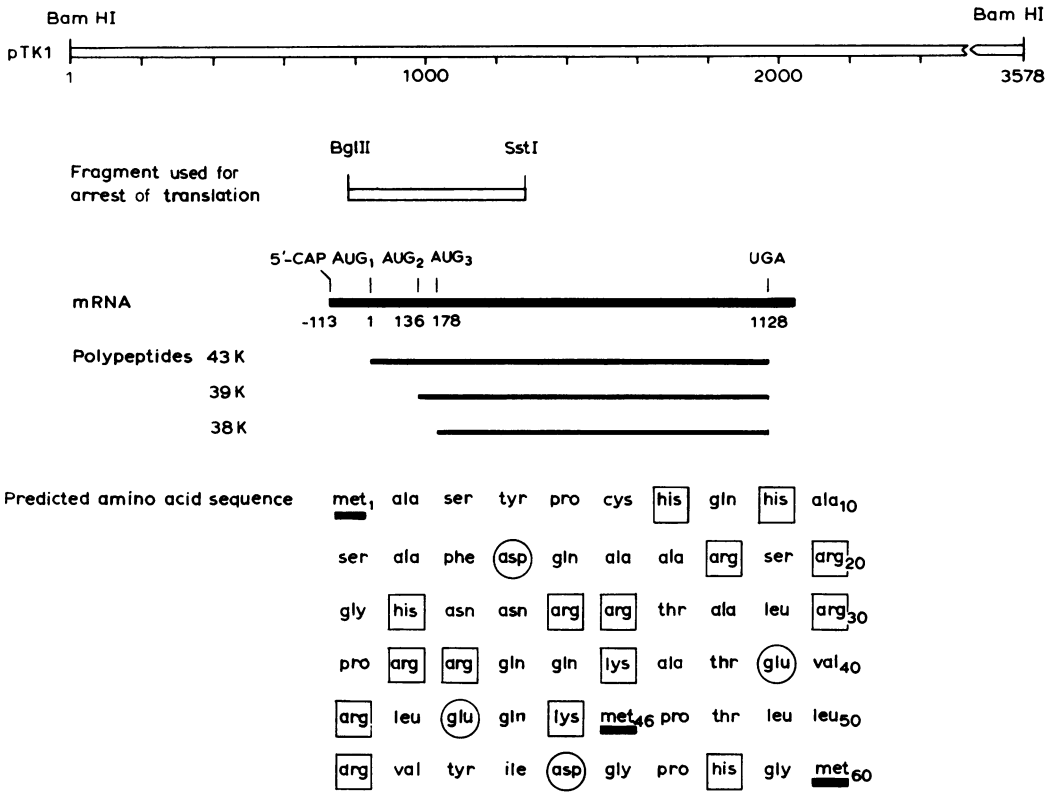

FIG. 4. Location of the first three initiation codons within the HSV TK mRNA and the first 60 amino acids predicted from the DNA sequence. The *BglII-SstI* fragment was used for the hybrid arrest of translation shown in Fig. 5. The AUG codons at nucleotides 1, 136, and 178 give rise to the methionine residues at amino acids 1, 46, and 60, respectively. Basic amino acids are identified by enclosing squares and acidic ones by circles.

We consider polypeptides of this class to be generated by some post-translational modification event.

Translation of TK mRNA. We next investigated the genesis of the 43K and 39K polypeptides encoded by the HSV-1 TK gene (20). In vitro translation products of total cytoplasmic mRNA from infected BHK cells were separated by two-dimensional gel electrophoresis. Those encoded by the TK gene were identified by hybridization of the mRNA with the 497-base-pair *BglII/SstI* fragment (Fig. 4). This fragment maps entirely within the TK gene (Fig. 4) (9, 20, 24; D. J. McGeoch, unpublished data) and has been shown to arrest synthesis of both the 43K and 39K polypeptides (20). Comparison of Fig. 2 and 5A shows that synthesis of five polypeptides was arrested. These were spots 60 with coordinates (43,0.575), 61 (43,0.655), 62 (43,0.71), 69 (39,0.38), which corresponds to VI 39, and 70 (38,0.38). After denaturation of the hybrid, cytoplasmic mRNA could again be translated to give the five polypeptide spots (Fig. 5B), which indicates that they all contain 5′ end sequences coded by the TK gene. Similar results were obtained with mRNA from infected BSC-1 cells (data not shown). Two additional control experiments provided further evidence that the five polypeptide spots were related to the TK gene. First, in vitro translation of mRNA from uninfected cells did not give any of the five spots (data not shown). Second, their translation was not arrested if DNA was omitted from the hybrid arrest experiment (data not shown).

To determine the pIs of the five polypeptides, we separated in vitro translation products using isoelectric focusing for the first dimension rather than NEPHGE (Fig. 6); the pH gradient in the gel was determined as described above. The pIs of the 38K and 39K polypeptides, which were 6.43 ± 0.03 and 6.54 ± 0.02, respectively, were obtained from three experiments and did not depend on whether proteins were loaded at the acidic or basic end of the gel.

DISCUSSION

The reported experiments were designed in part to investigate the genesis of the many HSV-

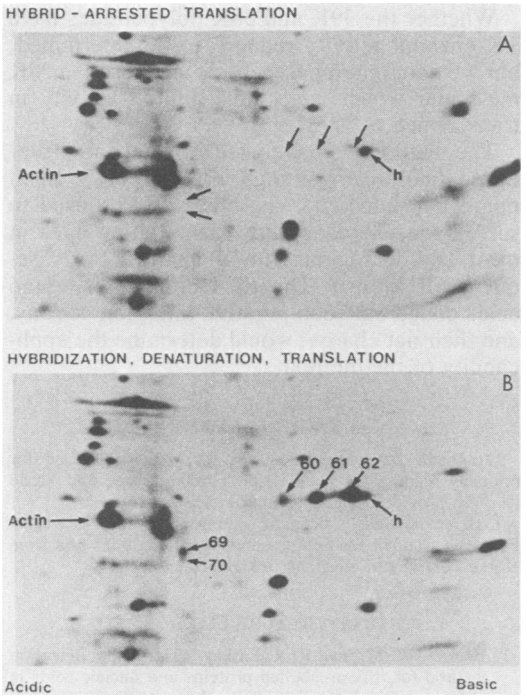

FIG. 5. Hybrid-arrested translation. mRNA extracted from HSV-1-infected cells was hybridized with the *Bgl*II-*Sst*I fragment and translated either directly (A) or after denaturation (B). Polypeptides were separated by two-dimensional NEPHGE, using a 9% gel cross-linked with DATD for the second dimension. Arrows identify the position of polypeptides whose translation is arrested by hybridization and which after denaturation of the hybrid are again translated. Actin and one other host polypeptide (labeled h) have been indicated to facilitate alignment of the fluorographs.

induced polypeptides resolved by two-dimensional gel electrophoresis. We identified 46 polypeptides which did not appear to be modified after synthesis, 36 precursor or short-lived polypeptides, 11 polypeptides which could be either very rapidly processed or unmodified, 19 polypeptides which could be either transient intermediate or precursor–short-lived polypeptides, and 8 polypeptides which were generated by post-translational modification. Thirty-six polypeptides could not be assigned to any class. The analysis did not include glycoproteins, which have been dealt with separately (14a). Our findings show considerably greater complexity in processing of HSV-induced polypeptides than has previously been detected, but the nature of the processing events has, for most proteins, yet to be understood.

We showed previously (4) that the 43K polypeptide band was resolved on two-dimensional gels into three polypeptide spots: (43,0.665), (43,0.71), and (43,0.74). It was not possible from the data presented in that study to distinguish whether these were (i) polypeptides with different primary sequences, (ii) modified forms of the same primary sequence, or (iii) the same polypeptide existing in different states in the infected cell (for example, in a complex with other small molecules). In the present study, we observed the two major spots, (43,0.665) and (43,0.71), and also a more acidic one, (43,0.475), which was previously not resolved from the host spots. The reason why we did not detect the most basic one, (43,0.74), in this study is not yet clear. The results presented here (Fig. 5) show that hybridization of cytoplasmic mRNA with the 497-base-pair *Bgl*II-*Sst*I fragment arrested synthesis of all three 43K polypeptide spots. As the TK gene contains only one long open reading frame (9, 24), and the mRNA is not spliced (9, 20), the three 43K polypeptides must be synthesized with the same primary sequences and possibility (i) can therefore be eliminated. It remains to be determined whether the 43K polypeptides are generated by mechanisms (ii) or (iii) suggested above and also whether all forms have enzyme activity.

Recently, Preston and McGeoch (20) showed that synthesis of not only the 43K protein but also that of a 39K protein was arrested by hybridization with fragments mapping entirely within the coding region of TK RNA. Our data confirm their result and furthermore show that synthesis of a 38K polypeptide was also arrested (Fig. 5). The DNA sequence of the TK gene shows a second and third initiation codon 135 bases and 177 bases downstream from, and in phase with, the first initiation codon (Fig. 4). By initiation from these codons proteins of molecular weights of 35,684 and 34,152 would be generated. These sizes are compatible with the 39K and 38K polypeptide molecular weights estimated by sodium dodecyl sulfate-polyacrylamide

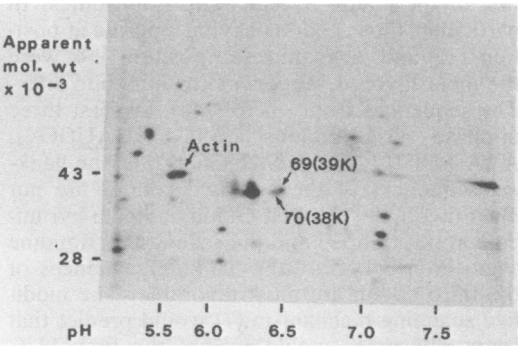

FIG. 6. Measurement of the pI of the 39K and 38K polypeptides. Polypeptides translated in vitro were isoelectrically focused and the pH along the gel determined as described in the text.

gel electrophoresis. A further 75 bases downstream is a fourth in-phase codon which if used would give a polypeptide of a molecular weight of 31,409. Arrest of synthesis of a polypeptide having this approximate molecular weight was not seen.

Both VI 39 (spot 69) and VI 38 (spot 70) are considerably more acidic than the three 43K polypeptides (spots 60, 61, and 62) (Fig. 1 and 5), and VI 38 is slightly more acidic than VI 39 (Fig. 6). These differences in pI are compatible with the predicted amino acid composition of the polypeptides. Figure 4 shows the first 60 amino acids of the TK gene of strain 17 (used in these studies) based on the data of Preston and McGeoch (20) and McGeoch (unpublished data). A protein initiated at the second AUG codon would lose 13 basic amino acids (3 histidine, 9 arginine, and 1 lysine) and 2 acidic amino acids (1 aspartic acid and 1 glutamic acid). It would therefore be considerably more acidic and behave analogously to VI 39. A protein initiated at the third AUG codon would lose an additional two basic amino acids (one arginine and one histidine) and one acidic amino acid (aspartic acid). It would therefore be slightly more acidic than VI 39, as was observed for VI 38.

A trivial explanation for the synthesis of the 39K and 38K polypeptides (spots 69 and 70) would be that they arise by translation of TK mRNA with an abnormal 5'-terminus, perhaps caused by mRNA degradation. Our previously published work (20) rules this out, since DNA fragments containing sequences upstream from the *Bgl*II site of pTK1 will select and hybrid arrest mRNAs encoding both VI 43 and VI 39. This result shows that mRNAs for VI 43 and VI 39 must contain sequences upstream from the *Bgl*II site and therefore that mRNA degradation is not the reason for synthesis of VI 39.

It has been suggested (7) that sequences flanking the AUG initiation codon influence recognition of the codon by eucaryotic ribosomes. In particular, those codons having a purine at position −3 and a guanine at position +4 were the most favored sequences for initiation sites. The sequences from −3 to +4 of the first three in-phase AUG codons are $_{-3}$CGUAUGG$_{+4}$, AAA AUGC, and GGGAUGG. Thus, the flanking sequences of the first two codons are not most favorable, the first codon having a pyrimidine at position −3 and the second a pyrimidine at position +4. Only the flanking sequences of the third codon are most favorable. The modified scanning mechanism (7) would predict that some ribosomes would bypass the first AUG codon to initiate at the second, and some may bypass both the first and the second to initiate at the third, a prediction compatible with our observations.

Whether the 39K and 38K polypeptides have TK enzyme activity remains to be determined, but we note that a polypeptide smaller than 38K molecular weight may have limited activity in transformed cells (22).

The method we have used is faster than tryptic peptide fingerprinting, requires only a small number of infected cells, and could be used to screen for secondary protein initiation sites in most coding regions for which the DNA sequence is known. The total number of amino acids coded between putative initiation codons, and their net charge, would determine the applicability of the method.

ACKNOWLEDGMENTS

We thank J. H. Subak-Sharpe for suggestions and for critically reading the manuscript. Graham Hope and Sigrid Øvernes provided skillful technical assistance.

L.H. acknowledges financial support from the Norwegian Research Council for Science and the Humanities and from Norsk Forening til Kreftens Bekjempelse.

LITERATURE CITED

1. **Bonner, W. M., and R. A. Laskey.** 1974. A film detection method for tritium-labelled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
2. **Brown, S. M., D. A. Ritchie, and J. H. Subak-Sharpe.** 1973. Genetic studies with herpes simplex virus type 1. The isolation of temperature-sensitive mutants, their arrangement into complementation groups and recombination analysis leading to a linkage map. J. Gen. Virol. **18**:329–346.
3. **Dubbs, D. R., and S. Kitt.** 1974. Mutant strains of herpes simplex deficient in thymidine-kinase-inducing activity. Virology **22**:493–502.
4. **Haarr, L., and H. S. Marsden.** 1981. Two-dimensional gel analysis of HSV type 1-induced polypeptides and glycoprotein processing. J. Gen. Virol. **52**:77–92.
5. **Heine, J. W., R. W. Honess, E. Cassai, and B. Roizman.** 1974. Proteins specified by herpes simplex virus. XII. The virion polypeptides of type 1 strains. J. Virol. **14**:640–651.
6. **Hope, R. G., J. Palfreyman, M. Suh, and H. S. Marsden.** 1982. Sulphaged glycoproteins induced by herpes simplex virus. J. Gen. Virol. **58**:399–415.
7. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. **9**:5233–5252.
8. **Marsden, H. S., I. K. Crombie, and J. H. Subak-Sharpe.** 1976. Control of protein synthesis in herpes virus infected cells: analysis of the polypeptides induced by wild type and sixteen temperature-sensitive mutants of HSV-strain 17. J. Gen. Virol. **31**:347–372.
9. **McKnight, S. L.** 1980. Th nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucleic Acids Res. **8**:5949–5964.
10. **McPherson, I., and M. Stoker.** 1962. Polyoma transformation of hamster cell clones. An investigation of genetic factors affecting cell competence. Virology **16**:147–151.
11. **Minson, A. C., P. Wildy, A. Buchan, and G. Darby.** 1978. Introduction of the herpes simplex virus thymidine kinase gene into mouse cells using virus DNA or transformed cell DNA. Cell **13**:581–587.
12. **Morse, L. S., L. Pereira, B. Roizman, and P. A. Schaffer.** 1978. Anatomy of herpes simplex virus (HSV) DNA. X. Mapping of the viral genes by analysis of polypeptides and functions specified by HSV-1 × HSV-2 recombinants. J. Virol. **26**:389–410.
13. **O'Farrell, P. H.** 1975. High resolution two-dimensional

electrophoresis of proteins. J. Biol. Chem. **250**:4007–4021.
14. **O'Farrell, P. Z., H. M. Goodman, and P. H. O'Farrell.** 1977. High resolution two-dimensional electrophoresis of basic as well as acidic proteins. Cell **12**:1133–1141.
14a. **Palfreyman, J. W., L. Haarr, A. Cross, R. G. Hope, and H. S. Marsden.** 1983. Processing of HSV-1 glycoproteins: 2-dimensional gel analysis using monoclonal antibodies. J. Gen. Virol. **64**:873–886.
15. **Pellicer, A., M. Wigler, R. Axel, and S. Silverstein.** 1978. The transfer and stable integration of the HSV thymidine kinase gene into mouse cells. Cell **14**:133–141.
16. **Pelman, H. R. B., and R. J. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
17. **Preston, C. M.** 1977. The cell-free synthesis of herpesvirus-induced polypeptides. Virology **78**:349–353.
18. **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.
19. **Preston, C. M.** 1979. Abnormal properties of an immediate early polypeptide in cells infected with herpes simplex virus type 1 mutant *tsK*. J. Virol. **32**:357–369.
20. **Preston, C. M., and D. J. McGeoch.** 1981. Identification and mapping of two polypeptides encoded within the herpes simplex virus type 1 thymidine kinase gene sequences. J. Virol. **38**:593–605.
21. **Rapp, F., N. Turner, and P. A. Schaffer.** 1980. Biochemical transformation by temperature-sensitive mutants of herpes simplex virus type 1. J. Virol. **34**:704–710.
22. **Roberts, J. M., and R. Axel.** 1982. Gene amplification and gene correction in somatic cells. Cell **29**:109–119.
23. **Smiley, J. R., M. J. Wagner, W. P. Summers, and W. C. Summers.** 1980. Genetic and physical evidence for the polarity of transcription of the thymidine kinase gene of herpes simplex virus. Virology **102**:83–93.
24. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Nat. Acad. Sci. U.S.A. **78**:1441–1445.
25. **Wigler, M., A. Pellicer, S. Silverstein, and R. Axel.** 1978. Biochemical transfer of single copy eucaryote genes using total cellular DNA as donor. Cell **14**:725–731.
26. **Wilkie, N. M., R. P. Eglin, P. G. Sanders, and J. B. Clements.** 1980. The association of herpes simplex virus with squamous carcinoma of the cervix and studies of the virus thymidine kinase gene. Proc. R. Soc. London Ser. B **210**:411–421.