Structure-Function Studies of the Herpes Simplex Virus Type 1 DNA Polymerase

MARY L. HAFFEY,1,* JIRI NOVOTNY,2 ROBERT E. BRUCCOLERI,2 ROBERT D. CARROLL,1 JOHN T. STEVENS,1 AND JAMES T. MATTHEWS1

Departments of Virology1 and Macromolecular Modeling,2 The Squibb Institute for Medical Research, Princeton, New Jersey 08543-4000

Received 4 April 1990/Accepted 10 July 1990

The analysis of the deduced amino acid sequence of the herpes simplex virus type 1 (HSV-1) DNA polymerase reported here suggests that the polymerase structure consists of domains carrying separate biological functions. The HSV-1 enzyme is known to possess 5'-3'-exonuclease (RNase H), 3'-5'-exonuclease, and DNA polymerase catalytic activities. Sequence analysis suggests an arrangement of these activities into distinct domains resembling the organization of Escherichia coli polymerase I. In order to more precisely define the structure and C-terminal limits of a putative catalytic domain responsible for the DNA polymerization activity of the HSV-1 enzyme, we have undertaken in vitro mutagenesis and computer modeling studies of the HSV-1 DNA polymerase gene. Sequence analysis predicts that the major DNA polymerization domain of the HSV-1 enzyme will be contained between residues 690 and 1100, and we present a three-dimensional model of this region, on the basis of the X-ray crystallographic structure of the E. coli polymerase I. Consistent with these structural and modeling studies, deletion analysis by in vitro mutagenesis of the HSV-1 DNA polymerase gene expressed in Saccharomyces cerevisiae has confirmed that certain amino acids from the C terminus (residues 1073 to 1144 and 1177 to 1235) can be deleted without destroying HSV-1 DNA polymerase catalytic activity and that the extreme N-terminal 227 residues are also not required for this activity.

Several complementary lines of research have been focused on defining the structural-functional organization of the herpes simplex virus type 1 (HSV-1) DNA polymerase (pol). These studies include comparative sequence analysis (1, 2, 4, 5, 18, 19, 21, 26-29, 32, 34, 41, 42, 45, 47, 48, 50), production of antibodies to defined regions of the pol protein (28, 36, 46, 49), and in vitro mutagenesis of the HSV-1 pol gene (35). We have been particularly interested in defining and elucidating features of the HSV-1 pol catalytic domain because it is the major target for a variety of potent antiviral drugs (9, 10, 13, 16, 22, 31, 40).

It has been proposed for several years that the C terminus of HSV-1 pol encodes the substrate-binding domain of the enzyme (19). This proposal was based on comparative sequence analysis of HSV-1 pol drug-resistant point mutations and regions of protein sequence homology among various DNA pol enzymes. Consistent with this proposal, studies utilizing antibodies to defined regions of HSV-1 pol have revealed that antibodies raised against C-terminal HSV-1 pol sequences can neutralize enzyme activity, whereas antibodies directed at N-terminal sequences cannot (28, 36, 46, 49). In addition, a study evaluating truncations or deletions of the HSV-1 pol gene transcribed and translated in vitro revealed that C-terminal truncations could destroy enzyme activity, while N-terminal truncations removing up to 67 residues could retain enzyme activity (14).

Structure-function studies of HSV-1 pol, such as those described above, could benefit from amino acid sequence analysis and a three-dimensional (3D) model of the enzyme or a part thereof. The HSV-1 pol enzyme belongs to a group of eucaryotic and procaryotic polymerases having amino acid sequence homology to the Escherichia coli DNA pol I (1, 2, 4, 5, 26, 34, 35). Only minor amino acid sequence homologies have been reported between the two pol families, primarily in the region of the putative 3'-5'-exonuclease domain (4, 15). Although nothing is known about the 3D structure of HSV-1 pol or the related family of polymerases, the crystal structure of the Klenow fragment of the E. coli pol II enzyme has been resolved (39). The lack of readily distinguishable sequence conservation between pol enzymes from the two families has made extrapolation from the bacterial pol structure difficult; however, computer modeling studies undertaken in our laboratories found a surprising degree of potential 3D structural similarity between the bacterial and the viral enzymes.

In this report, we present an alignment of the amino acid sequence of HSV-1 pol into three distinct catalytic domains (RNase H, 3'-5'-exonuclease, and DNA polymerase, respectively), evaluate the C-terminal limits of the DNA polymerization domain using deletion analysis by in vitro mutagenesis of the HSV-1 pol gene in a Saccharomyces cerevisiae expression system, and present a 3D model of the DNA polymerase catalytic domain. The deletion studies described in this report reveal that previously unidentified portions of both the N and C termini are not required for HSV-1 pol catalytic activity, consistent with predictions of the sequence analysis and model.

MATERIALS AND METHODS

Protein sequence analysis and computer modeling studies. The computer program of Novotny and Auffray (38) was used to calculate and display the amino acid sequence profiles of hydrophobicity, formal side chain electric charges, and secondary structure propensities (a-helix, $\beta$-sheet, and turn). The amino acid sequences of the E. coli pol I Klenow fragment and the HSV-1 (strain KOS) DNA pol gene were obtained from the Protein Identification Resource

* Corresponding author.
Database, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C., and were aligned on the basis of sequence homology. Special attention was paid to amino acid side chain identities, side chain chemical similarities, and the correspondence of secondary structural elements (α-helices and β-strands) between the two sequences.

The atomic coordinates of the backbone α-carbons of the Klenow fragment were obtained from the Brookhaven Protein Data Bank, Brookhaven, N.Y. Building of the HSV-1 catalytic domain structure, structural manipulations, chain insertions, and deletion were done by using the computer programs CONGEN (6) and PEER (R. E. Brucoleri, unpublished work). Color-coded screen representations of backbone models were displayed with the use of the PEER interactive graphics program and the Silicon Graphics Iris 4D/80GT workstation.

Materials. Molecular cloning procedures utilized T4 DNA ligase, T4 DNA polymerase, and restriction endonucleases purchased from New England BioLabs or Bethesda Research Laboratories. [3H]HTP (15 Ci/mmol) was purchased from ICN Biomedicals, Inc. Chemicals were obtained from Sigma Chemical Co. unless otherwise noted. The plasmid pNN3 was obtained from M. Challberg, National Institutes of Health, Bethesda, Md., and has been described previously (8).

Strains and media. The E. coli strains used for this study were XL-1 Blue (Stratagene Cloning Systems) and CJ236 (Bio-Rad Laboratories). The yeast strain was Y294, which has been previously described (7, 20). Bacteria were grown in Luria broth (37) containing 50 µg of ampicillin per ml or 30 µg of chloramphenicol per ml where appropriate; yeast cells were grown in minimal medium (0.7% yeast nitrogen base [Difco Laboratories]) containing required amino acids (20 to 30 µg/ml) and 2% raffinose or 2% galactose. Transformation of yeast cells with the recombinant plasmids was accomplished as previously described (3, 20, 35) or by an alternative technique (23) using lithium acetate.

Oligonucleotides. Oligonucleotides used for this study were synthesized on an Applied Biosystems model 380B DNA synthesizer using β-cyanoethyl phosphoramidite chemistry. Oligonucleotides were purified by chromatography on oligonucleotide purification cartridges (Applied Biosystems Product Bulletin, January 1988). The sequences of the oligonucleotides used were as follows: pRC204, 5′-ccgg ggcctcagctcgg-3′; pRC210, 5′-ctttgccttatccctcatcag ggctctagttt-3′; pRC211, 5′-cctcacaacggctggatatcatgactactataa gtcgcc-3′; pRC212, 5′-ggcgctgcttcgcttgctcgggctcgggggctc gg-3′; pRC213, 5′-cgctccccggcctgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg
with bacteriophage T4 DNA polymerase and religiating the blunt ends to maintain the correct reading frame while deleting residues 1073 to 1144. The prC215 vector contained the same deletion of residues 1073 to 1144 and was prepared in a similar manner as the prC214 vector, except that the starting vector was prC210, a C-terminal deletion mutant described in the preceding paragraph which contains a termination codon after residue 1177.

The prC216 vector was prepared by restriction endonuclease digestion of prC205 with PstI and BamHI and religation, using a synthetic oligonucleotide adaptor designed to maintain the correct HSV-1 pol reading frame, while introducing an internal deletion of residues 1009 to 1144. Finally, the prC218 vector was created from prC205 by oligonucleotide-directed deletion of the sequence between residues 884 and 892, while maintaining the correct reading frame.

Protein analysis and immunoblotting. Protein determinations were performed by using a Coomassie blue protein assay kit (Bio-Rad). Polyacrylamide gel electrophoresis and immunoblots were performed as previously described (20, 35, 36). The filters were processed, using the Immuno-Blot assay kit (Bio-Rad), essentially as instructed by the manufacturer except that bovine serum albumin was substituted for gelatin. The dilution of primary antibody was 1:200, and the secondary antibody was diluted 1:1,000 (alkaline phosphatase-conjugated swine anti-rabbit immunoglobulin G [IgG] [Dako Corporation]). The antigen used for preparation of the anti-P5 rabbit antiserum was a synthetic oligopeptide, Tyr-Gly-Asp-Thr-Asp-Ser-Ile-Phen-Val-Tyr, passively absorbed to polyvinylpyrrolidone; otherwise, preparation of the P5 antiserum was as described for preparation of the P3 antiserum (35, 36).

Preparation of enzyme extracts. Yeast extracts were prepared as described previously (20, 35) from 100-ml overnight cultures grown at 30°C. Yeast cells were pelleted at 4°C, frozen on dry ice, and suspended in 50 mM Tris hydrochloride (pH 8.0)–50 mM NaCl–1% dimethyl sulfoxide–10% (vol/vol) glycerol–5 mM β-mercaptoethanol–10 mM EDTA–2 mM benzamidine–1 mM phenylmethylsulfonyl fluoride at 4°C. Lysis was achieved by repeated vortexing of the cell suspensions, in the presence of acid-washed glass beads, for 30-s intervals followed by cooling on ice until >70% lysis was achieved as judged by light microscopy. The lysed extracts were transferred to 1.5-ml microcentrifuge tubes and centrifuged at 12,000 × g for 10 min. Extracts were stored at −80°C in 15% glycerol until use. The final protein concentrations were approximately 5 mg/ml.

DNA pol assays. The assays for endogenous yeast DNA pol and HSV-1 DNA pol were performed as described previously (20, 35). Endogenous yeast DNA pol reaction mixtures contained 5 µl of cell lysate (25 to 30 µg of protein) in 10 mM Tris hydrochloride (pH 7.5)–10 mM MgCl2–10 mM ammonium sulfate–0.1 mM diethiothreitol–100 µM (each) dATP, dCTP, dGTP, and dTTP (27 cpm/µmol)–40 µg of nicked calf thymus DNA per ml–2 mM spermidine in a final volume of 50 µl. HSV-1 DNA polymerase assays contained 5 µl of enzyme extract (about 25 µg of protein) in 50 mM Tris hydrochloride (pH 8.0)–5 mM MgCl2–100 mM ammonium sulfate–1 mM diethiothreitol–5 µM (each) dATP, dCTP, dGTP, and dTTP (540 cpm/µmol)–30 µg of nicked calf thymus DNA per ml–100 µg of bovine serum albumin per ml in a final volume of 50 µl. All DNA polymerase assays were conducted in duplicate and incubated at 37°C for 15 min.

Neutralization studies. Enzyme extracts (5 µl) were incubated with the appropriate IgG (4 µg per ml) for 1 h on ice before the addition of the other reaction components of the

FIG. 1. Amino acid sequence profiles of the HSV-1 DNA polymerase computed and smoothed as described previously (38). In each panel, the upper profile shows reverse-turn propensity (T); the middle profiles are those of α-helix (thin line, A) and β-sheet (heavy line, B) propensities; and the lower profile shows hydrophobicity (HB). The solid bars under the panels denote the putative location (within 10 residues at either end) of the RNase H, 3′–5′-exonuclease (3′–5′ exo), and polymerase catalytic (DNA pol) domains.

HSV-1 DNA polymerase assay, as previously described (36).

RESULTS

Structural analysis and computer modeling studies. Structural analysis of the HSV-1 pol enzyme was performed on the assumption that structure-function studies under way in our laboratory using in vitro mutagenesis of the HSV-1 pol gene in a yeast expression system could benefit from such information. Figure 1 shows the resultant profiles of the HSV-1 pol amino acid sequence. The large size of the protein (1,235 residues) (19) suggested that its structure would consist of several independent domains, and sequence profiles were used to identify polypeptide chain segments that separate the putative domains. Generally, interdomain segments of proteins are expected to consist of mostly polar side chains and are conspicuous by a lack of hydrophobicity. In contrast, the hydrophobic side chains are believed to be primarily responsible for protein folding and domain organization (43). Prominent breaks in the hydrophobicity profiles of the HSV-1 pol were found around residues 60, 660 to 690, and 1100 to 1140. Of these, the segment spanning residues 660 to 690 is the most remarkable, containing clusters of negatively charged residues (aspartates and glutamates) interspersed with positively charged lysines and arginines. This segment is likely to exist outside compact globular domains because of its extreme charge.

The domain organization of the HSV-1 pol which emerged from this analysis can be summarized as follows. (i) The N-terminal 65 residues are probably not integrated into the rest of the structure. (ii) The C-terminal 130 to 135 residues probably represent a separate domain. (iii) Two segments, 65

Downloaded from http://jvi.asm.org/ on April 29, 2021 by guest
to 660 and 690 to 1100, are the best candidates for the principal structural domains. (iv) It is possible to subdivide the segment from residues 65 to 660 into two smaller domains and hypothetically assign the 5'-3' - (RNase H [11]) and 3'-5'-exonucleolytic activities (28, 40, 49) of HSV-1 pol to residues 65 to 390 and 390 to 660, respectively, on the basis of size and structural similarities of the corresponding domains of the E. coli pol I Klenow fragment (25). (v) The 690 to 1100 domain corresponds in size and structural profile to the catalytic domain of the Klenow fragment, containing short runs of positively charged side chains (lysine and arginine) which are prerequisites for binding the strongly negatively charged polyphosphate backbone of the DNA. This domain (HSV-1 residues 690 to 1100) is the most likely candidate for the catalytic domain of HSV-1 pol. Consistent with this assignment, many point mutations conferring resistance to antiviral drugs (nucleotide analogs) map to this region of the HSV-1 sequence (18, 28, 32).

Operating on the assumption that conservation of DNA pol function might reflect conservation of structure, the HSV-1 pol sequence was compared with that of the E. coli pol I enzyme. After secondary structural analysis, an amino acid alignment of the E. coli and HSV-1 putative catalytic domains was performed. Although the alignment showed only a low level of sequence conservation (about 10% identical residues and about 21% homologous residues) as previously reported (1, 2, 4, 5, 34), some of its features are supportive of a distant structural homology between the two domains. For example, (i) the majority of the conserved residues occur in α-helices and β-strands, as opposed to unstructured loops; (ii) conservative amino acid replacements tend to cluster around invariant residues; (iii) virtually all insertions or deletions occur in surface loops; and (iv) the patterns of amino acid conservation and homologous replacements conform to the character of the secondary structure found at that place in the X-ray crystallographic structure of the Klenow fragment (i.e., n+1 pattern of residue conservation in the β-strands, conservation of prolines and glycines in loops) (Fig. 2).

With use of the amino acid sequence alignment, a 3D tracing of the polypeptide chain was constructed in the computer, on the basis of the known α-carbon coordinates of the Klenow fragment (Brookhaven Database). In the polypeptide chain of the Klenow fragment of the E. coli pol I sequence, two segments of the backbone (residues 574 to 622 and 780 to 787) are not visible, probably because of local crystalline disorder. The X-ray structure of the Klenow fragment thus technically consists of three separate polypeptides, as does our HSV-1 pol catalytic domain model (Fig. 2 and 3). Specifically, the Klenow fragment polypeptides containing residues 521 to 574, 620 to 774, and 790 to 928 approximately align with HSV-1 residues 698 to 749, 794 to 961, and 969 to 1106, respectively. Using the programs CONGEN and PEER, we first generated coordinates for the conserved parts of the catalytic domain backbone by copying the Klenow fragment coordinates into the HSV-1 pol catalytic domain sequence. After completion of this step, the interactive graphics program PEER was used to model the rest of the backbone (insertions and deletions). The next steps involved (i) changes of chain topology involving the Klenow fragment region around α-helices J to K. In HSV-1 pol, the longer J helix connects to a β-strand homologous to strand 13 of the Klenow fragment and proceeds, via a turn, into strand 9. (ii) N-terminal extension of α-helix L by about two helical turns. (iii) Deletion of the irregular loop at residues 742 to 753 in the Klenow fragment. (iv) Introduction of a few minor deletions and insertions in α-helix Q and the surrounding loops. The final approximate model of the HSV-1 DNA pol major catalytic domain is displayed in Fig. 3. The prominent upward groove is the potential DNA-binding site.

Expression of mutant pol proteins. Predictions of the sequence analysis described above were evaluated against results obtained in our laboratories by using in vitro mutagenesis of the HSV-1 pol gene. The S. cerevisiae expression system used for the mutagenesis studies has been described in detail elsewhere and has been shown to produce active HSV-1 pol (20). In vitro mutagenesis of the HSV-1 pol produced in yeast cells has demonstrated that a point mutation introduced into the gene can affect the drug resistance or sensitivity of HSV pol enzyme activity and that the HSV-1 pol expression system in yeast cells can be used for structure-function analysis of the enzyme (35).

A summary of the mutant expression vectors used for this study is shown in Fig. 4. Each of the expression vectors was evaluated for production of HSV-1 pol protein and for HSV-1-specific DNA polymerase activity. The preparation and analysis of yeast lysates was performed as previously described (20, 35). Expression of the HSV-1 pol gene in yeast vectors is under the tight control of the yeast GAL1 promoter (20, 24, 35). The GAL1 promoter is active only when cultures are grown in the presence of the inducing sugar, galactose. When yeast cultures are grown in the presence of the noninducing sugar, raffinose, the GAL1 promoter is not active and no HSV-1 pol protein is produced.

Figure 4 summarizes the results of immunoblot assays run on lysates from galactose-induced cultures of the deletion mutants in order to detect the expression of HSV-1 pol. The immunoblot assays were performed as previously described (35, 36), using an antipeptide antiserum to the HSV-1 pol epitope, P3 (residues 548 to 557). An alternate antipeptide antiserum (P5) directed against residues 884 to 892 (unpublished data) was used to evaluate expression of the HSV-1 pol gene in the pRC213 vector which lacked the P3 sequence and to confirm the deletion of the P5 epitope in pRC218 as well as expression of HSV-1 pol in all of the other lysates studied. Examples of the immunoblot assays analyzing lysates from various recombinant yeast cultures are shown in Fig. 5. Panels A and B demonstrate that (i) the P5 antiserum appeared to recognize the same protein as the P3 antiserum in lysates from wild-type pRC205 cells grown under inducing conditions (in galactose), (ii) as expected, the P5 antiserum (but not the P3 antiserum) failed to detect the HSV-1 pol protein expressed in the pRC218-derived lysate because of the specific deletion of the P5 sequence, and (iii) neither the P3 nor the P5 antiserum could detect HSV-1 pol expression in control lysates of yeast cells grown under noninducing conditions (raffinose) or in the lysate of the pRC213-transformed cells grown under inducing conditions. Panel C reveals that lysates from pJS1, pRC303, pRC204, pRC210, pRC214, pRC215, pRC216, and pRC218 grown in the presence of the inducing sugar, galactose, produced immunoblot-reactive HSV-1 pol polypeptides.

Somewhat surprisingly, four of the C-terminal deletions (pJS2, pJS3, pRC211, and pRC212) failed to produce detectable HSV-1 pol protein as measured by immunoblot assay using either P3 or P5 antiseras (data not shown). Of particular interest was the finding that termination of the HSV-1 pol at residue 1177 (pRC210) did not grossly affect protein expression, whereas removal of 15 or 47 additional residues (termination after residue 1162 [pRC211] or 1130 [pRC212],...
FIG. 3. (A) Stereoscopic representation of the computer-generated model of the putative HSV-1 DNA polymerase major catalytic domain (refer to DNA pol domain in Fig. 1). Amino acid residues are schematically represented as spheres with diameters comparable to the length of a peptide unit (0.38 nm). Sequence regions (I, II, III, and V) known to be conserved in HSV-1 and other DNA polymerases described by Gibbs et al. (18) and Wang et al. (48) are color coded as follows: green, region II (residues 694 to 756); yellow, region III (residues 805 to 845); red, region I (residues 882 to 899); and pink, region V (residues 949 to 963). The prominent upward groove in the model is the putative DNA-binding site. Conserved region VI (residues 772 to 791) (18, 48) is not included in the model because it aligns with a region not resolved in the Klenow fragment X-ray structure. (B) Alpha-carbon tracing of the model of the putative HSV-1 DNA pol major catalytic domain presented in panel A. The three separate polypeptide chain segments are color coded red, green, and black. The consecutive numbering of every tenth residue is indicated and is consistent with the numbering for HSV-1 pol residues in Fig. 2 and reference 19.

respectively) resulted in lack of detectable HSV-1 pol protein. Taken together, the results suggest that expression and/or detection of HSV-1 pol protein was affected by deletion of sequences between residues 1162 and 1177, at least in the yeast expression system. Therefore, vectors pRC214 and pRC215 were created, leaving intact residues 1162 to 1177 and adjacent upstream residues (1145 to 1162), which are conserved among the herpesvirus group (29), while deleting less-conserved sequences further upstream, residues 1073 to 1144. Both the pRC214 and pRC215 vectors produced HSV-1 pol proteins recognized by immunoblot assay.

The pJS1, 227-residue N-terminal deletion produced immunoreactive protein but, disappointingly, an additional N-terminal deletion, pRC213, failed to produce detectable HSV pol protein as measured in our assay (Fig. 5).

One unresolved result which emerged from the immunoblot analysis was a lack of apparent shift in electrophoretic mobility of HSV-1 pol immunoreactive proteins detected in the lysates derived from the HSV-1 pol deletion mutants (Fig. 5C). We are attempting to alter the gel conditions and partially purify the expressed pol proteins in order to investigate this anomaly. Restriction endonuclease and/or DNA sequence analysis of the expression plasmids.
hybridization studies on RNA produced from the mutated vectors, and in vitro transcription and translation of the mutated pol genes does, however, reveal the expected size variations between wild-type and deleted constructs (data not shown). In addition, immunologic data using antisera to P5 in an immunoblot assay (Fig. 5) or antisera to P6 and P7 in the neutralization studies discussed below (Fig. 6) revealed that deletion of the cognate peptide sequence in the gene (pRC218, pRC214, or pRC204, respectively) results in expression of HSV-1 pol proteins which no longer are recognized by their respective antipeptide sera.

**Detection of HSV-1-specific DNA pol activity.** The HSV-1-specific DNA pol activity in yeast lysates was measured in the presence of 100 mM ammonium sulfate, as previously described; under these high salt conditions the endogenous yeast DNA polymerases are not detected (20, 35). Table 1 and Fig. 4 summarize the results of the DNA pol assays run on the uninduced (raffinose) and induced (galactose) yeast cultures, as well as control lysates from uninfected and HSV-1-infected HeLa cells. All of the lysates analyzed contained comparable amounts of total protein and produced comparable amounts of DNA pol activity when assayed under conditions in which the endogenous yeast enzymes were active (α-pol assay). As expected, none of the four yeast lysates (pJS2, pJS3, pRC211, and pRC212) which failed to produce detectable HSV-1 pol protein by immunoblot assay contained HSV-1-specific DNA pol enzyme activity. Two yeast lysates which produced HSV-1 pol protein by immunoblot assay failed to produce detectable HSV-1-specific DNA pol enzyme activity (pRC218 and pRC216). pRC218 contains an internal in-frame deletion of residues 884 to 892, a region contained in the highly conserved region I of DNA polymerases described by Wang et al. (48) refer to

Fig. 3), and the P5 antisera epitope. pRC216 contained an internal in-frame deletion of residues 1009 to 1144.

All of the remaining constructs which produced detectable HSV-1 pol protein also produced readily detectable levels of HSV-1-specific DNA pol activity. Taken together, these results demonstrate that large regions of the N terminus and certain regions of the C terminus are not required for HSV-1 DNA pol enzyme activity. In contrast, sequences between residues 1009 and 1073 (comparing pRC215 with pRC216) and between 884 and 892 (pRC218) are apparently required for HSV-1 pol-specific activity, as measured in our assay.

**Neutralization studies.** Previously, we had described the production of two antipeptide antisera capable of neutraliz-
ing the enzyme activity of purified HSV-1 DNA pol from infected HeLa cells (36). Peptide sequences P6 and P7 contain residues 1100 to 1108 and 1216 to 1224, respectively. These sequences were chosen in an effort to develop HSV-specific reagents and were not reported in other DNA polymerases. The structural analysis and modeling studies predicted, and the deletion studies clearly showed, that the P6 and P7 sequences could be deleted without destroying HSV-1-specific pol catalytic activity. Thus, these sequences were not required for substrate binding. We tested the prediction that loss of these sequences would prevent the ability of the pRC214- and pRC204-encoded polymerases to be neutralized by the P6 and P7 IgGs, respectively. As previously reported for both crude and purified HSV-1 pol produced from HSV-1-infected HeLa cells (36), Fig. 6 reveals that purified IgGs from both the P6 and P7 antisera were able to neutralize the HSV-1-specific pol enzyme activity in lysates from yeast cells producing the pRC205-encoded, wild-type HSV-1 pol. However, the HSV-1 pol activity in lysates from the yeast strain containing pRC214, which lacks P6 by virtue of a deletion of residues 1073 to 1144, and the strain containing pRC204, which lacks the P7 sequence because of premature termination of the protein at residue 1196 (refer to Fig. 4), was not neutralized by antibodies directed at the deleted sequences. pRC214-encoded pol can be neutralized by P7 but not by P6 IgG; conversely, pRC204-encoded pol can be neutralized by P6 IgG but not by P7 IgG. These studies confirm both the presence of the predicted deletions and the sequence specificity of neutralization by these IgGs. In addition, these results demonstrate that although the P6 and P7 sequences are specific targets for neutralizing antibodies, they are not strictly required for catalytic activity as measured in our assay.

**DISCUSSION**

Structural analysis, computer modeling, and in vitro mutagenesis studies described in this report begin to define the C-terminal limit and structure of the HSV-1 DNA pol catalytic domain. The model of this catalytic domain presented in Fig. 3 is based on the crystal structure of the Klenow fragment of *E. coli* pol I and shares many features in common with the Klenow enzyme. Previous studies have proposed the involvement of this region of HSV-1 pol in substrate binding and catalysis. Gibbs et al. (18, 19) proposed that the region from residues 597 to 961 might fold to form part of a domain containing substrate and drug binding sites, on the basis of mapping of HSV-1 pol drug resistance mutations. A cluster of acyclovir-resistant mutations is found in the lower right hand corner of the model (HSV residues 719, 724, 813, 815, 821, 841, and 842 [28] Fig. 3). Unfortunately, the model is not of sufficient resolution to evaluate the effects of point mutations on the overall structure of the enzyme. On the basis of amino acid sequence homology, Wang et al. (48) identified six segments that show a high degree of sequence conservation among viral and eucaryotic DNA polymerases. Five of these segments map to the putative HSV-1 pol catalytic domain described in this report, and four of them (regions I, II, III, and V) participate in formation of the DNA-binding cleft (refer to Fig. 3). Conserved region VI (HSV residues 772 to 791) was not included because it aligns with a discontinuity in the Klenow structure around residue 600. While our manuscript was in preparation, Bernad et al. (4) predicted a similar location for the HSV-1 pol catalytic domain, as well as an assignment of the 3'-5' and 5'-3' exonuclease domains comparable to that described in further detail here. Conserved regions IV (48) and A (18), which are not included in our model of the HSV-1 pol catalytic domain, are predicted to be part of the 3'-5'-exonuclease site (this report and Bernad et al. [4]).

The modeling studies are supported by the view that conservation of function reflects 3D structure in lieu of extensive sequence conservation. The validity of this viewpoint with respect to pol enzymes is confirmed by the recent report of the 3D structure of the *E. coli* RNA polymerase holoenzyme determined by electron crystallography (12),

**TABLE 1. Results of pol assays**

<table>
<thead>
<tr>
<th>Extract*</th>
<th>DNA pol activity as determined by*</th>
<th>HSV pol assay</th>
<th>α-pol assay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raffinose</td>
<td>Galactose</td>
<td>Raffinose</td>
</tr>
<tr>
<td>Yeast cells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRC205</td>
<td>0.28</td>
<td>9.88</td>
<td>117.7</td>
</tr>
<tr>
<td>pJS1</td>
<td>0.32</td>
<td>12.51</td>
<td>145.7</td>
</tr>
<tr>
<td>pJS2</td>
<td>0.36</td>
<td>0.31</td>
<td>138.7</td>
</tr>
<tr>
<td>pJS3</td>
<td>0.52</td>
<td>0.36</td>
<td>148.5</td>
</tr>
<tr>
<td>pRC204</td>
<td>0.24</td>
<td>10.00</td>
<td>137.5</td>
</tr>
<tr>
<td>pRC210</td>
<td>0.67</td>
<td>10.21</td>
<td>141.3</td>
</tr>
<tr>
<td>pRC211</td>
<td>0.29</td>
<td>0.37</td>
<td>156.1</td>
</tr>
<tr>
<td>pRC212</td>
<td>0.41</td>
<td>0.56</td>
<td>124.7</td>
</tr>
<tr>
<td>pRC213</td>
<td>0.37</td>
<td>0.41</td>
<td>143.0</td>
</tr>
<tr>
<td>pRC214</td>
<td>0.37</td>
<td>12.00</td>
<td>131.5</td>
</tr>
<tr>
<td>pRC215</td>
<td>0.33</td>
<td>12.14</td>
<td>135.9</td>
</tr>
<tr>
<td>pRC216</td>
<td>0.30</td>
<td>0.48</td>
<td>139.5</td>
</tr>
<tr>
<td>pRC218</td>
<td>0.44</td>
<td>0.44</td>
<td>151.0</td>
</tr>
</tbody>
</table>

HeLa cells

| 0.33      | 8.95*     | 109.5*    |            |

* Extracts were prepared from yeast cultures grown in either 2% raffinose or 2% galactose medium or from uninfected or HSV-1-infected HeLa cultures, as described in Materials and Methods.

* DNA pol assays were performed as described in Materials and Methods (20). All values are the averages of duplicate assays and are expressed as picomoles of 3H[TMP] incorporated per reaction.

* Uninfected.

* Infected.
which exhibits remarkable similarity to the 3D structure of the *E. coli* Klenow fragment despite only a small degree of amino acid sequence conservation. Consistent with the structural analysis and the model for HSV-1 pol proposed in this report, we have demonstrated that regions N terminal to residue 228 and certain regions C terminal to residue 1073 are not essential for HSV-1 pol catalytic activity as measured in our assay. Because of difficulties in detection of a larger N-terminal deletion, we have not been able to further define the N-terminal limit of the catalytic domain. In fact, it may not be possible to delete some of these sequences and maintain protein conformation and stability, as was experienced in deletions of the C-terminal residues between 1162 and 1177 described in this report. It is noteworthy that separation of the 3'-5'-exonuclease and polymerase domains of the Klenow fragment adversely affected the stability of the polymerase domain (15) and that the T4 DNA polymerase cannot yet be physically separated into exonuclease and polymerase domains, and an overlap of the T4 domains has been suggested (45). One could imagine that the polymerization and exonuclease catalytic sites are separate but that the sites required for binding the DNA substrate or folding the pol protein may overlap. Consistent with this concept, Becker has reported that HSV-1 pol residues between 431 and 638 may be involved in binding to the DNA template (2). An N-terminal sequence proposed from the structural analysis to be a separate folding domain (residues 1 to 65) is not essential for catalytic activity. This is evidenced by the predicted 227-residue, N-terminal (pJ51) deletion in this study and the predicted 27- and 67-residue deletions in an earlier in vitro transcription-translation study (14).

Sequence structural analysis revealed that the C-terminal 130 to 135 residues probably represent a separate domain. In contrast to results from a previous study (14) and consistent with this prediction, certain C-terminal sequences downstream of residue 1100 (residues 1073 to 1144 and 1177 to 1235) can be deleted without grossly affecting HSV-1 pol catalytic activity despite the fact that they are recognized by antibodies capable of neutralizing pol activity. Antibodies to these regions (P6 and P7 sequences) may accomplish enzyme inactivation by sterically blocking access of DNA substrate to the active site or by perturbing the 3D structure of the enzyme. If the former is true, it is possible that antibodies which prevent access of DNA substrate could affect all three catalytic activities of the enzyme. This hypothesis could be tested with the P6 and P7 antibodies and might explain results of a previous study (28, 49) in which antibodies directed at C-terminal sequences of the HSV-1 pol neutralized both pol and 3'-5'-exonuclease activities. Additional work will be required to determine whether P6 or P7 or both neutralizing antibodies sterically block binding of DNA substrate or whether deletion of these sequences affects exonuclease activity.

Conservation of certain sequences between HSV-1 and HSV-2 and not other viruses (including the P6 and P7 epitopes) suggests a requirement for these regions in the specific context of HSV replication in vivo. The possible involvement of these sequences in interaction with the other viral or cellular components involved in origin-specific replication (8) will be investigated using the deletion mutants. We have also shown that sequences between the structural and 1177 may be required for stable expression of the pol protein, at least in the yeast expression system. This sequence is not a classic nuclear localization or protein stability sequence, is generally hydrophobic, and could, alternatively, be involved in intranuclear localization or intramolecular protein folding. A requirement for this region for correct intramolecular protein folding could explain the difference between our result and that of a previous study by Dorsky and Crumpacker (14). In that study, using in vitro transcription-translation, it was reported that termination of the HSV pol gene at a BamHI restriction endonuclease site coincident with residue 1073 of the HSV pol protein produced a polypeptide lacking HSV pol enzyme activity. It is possible that the HSV pol protein expressed in their system was not properly folded because of the loss of residues between 1162 and 1177 which were found, in the study reported here, to be required for detection of stable HSV pol expression and activity. The region between residues 1162 and 1177 has previously been reported to be conserved among DNA polymerases from such members of the herpesvirus group as Epstein-Barr virus, human cytomegalovirus, and varicella-zoster virus (2, 29; M.L.H., unpublished observations), suggesting a conservation of function, albeit not necessarily for catalysis.

In general, regions which are poorly conserved among the herpesvirus group (29) at both the N and C termini can be deleted without destroying HSV-1 pol catalytic activity or the ability to detect protein expression, consistent with previous proposals that sequence conservation is related to conservation of function. While deletion of the sequences upstream of residue 228 which are conserved among the herpesvirus group (residues 176 to 221) does not affect pol catalytic activity, it is possible that this deletion may affect the 5'-3'-exonuclease activity of the HSV-1 pol which is predicted to map to this region (4; this report). A modest sequence homology of this region of the HSV-1 pol protein (residues 171 to 214) to the 5'-3'-exonuclease domain of *E. coli* pol I has also been reported previously (2).

The loss of enzyme activity, as measured in our system, by deletion of residues 1009 to 1144 (pRC216) but not residues 1073 to 1144 (pRC215) suggests that the region between 1009 and 1073 is required for catalysis, and the structural analysis and the model would also predict this. This region of the HSV-1 sequence aligns roughly with helix Q and strands 12 and 13 of the Klenow fragment. This region of the *E. coli* enzyme has been proposed as part of the polymerase active site (25). Conservation of this region in the sequences of the herpesvirus polymerases also supports a functional role for this region (29). In order to definitively address the requirement for this region, it will be necessary to purify the expressed pRC216 protein and evaluate its pol activity and ability to bind substrate under a variety of assay conditions. Finally, deletion of the highly conserved region I sequence (48) (HSV-1 residues 884 to 892) destroyed catalytic activity, as measured in our assay. This result is consistent with predictions of an essential role for this sequence (1, 48, 50) and results of a previous study using in vitro translation (14) in which deletion of residues 881 to 958 also destroyed HSV-1 pol activity.

While this report focused on defining the HSV-1 pol catalytic domain, the structural analysis also predicted an assignment for the 5'-3' and 3'-5'-exonuclease domains by analogy with the *E. coli* pol I holoenzyme. A number of the constructs created for this study will be analyzed for these additional activities. The structural analysis reported here might predict that the pJ51 N-terminal deletion would result in loss of 5'-3'-exonuclease activity or that the pRC216 and pRC218 C-terminal deletions which lack pol catalytic activity might still retain both exonuclease activities. Mutagenesis of the putative 3'-5'-exonuclease domain is in progress to directly address predictions of this functional assignment.
Several of the mutations created for this study should also be tested for the ability to be stimulated by another HSV-1 gene required for origin-specific replication, UL42 (17).

Because the model of the HSV-1 pol enzyme presented in Fig. 3 is 3D and closely resembles the Klenow fragment 3D structure, additional predictions and comparisons of the two enzymes can be made. Specific residues in the Klenow enzyme have been identified (25, 39) which bind nucleotide or DNA substrates (i.e., Arg-690, Lys-758, and Tyr-766); although these residues do not directly align with identical residues in the HSV-1 sequence, mutagenesis of the analogous regions of HSV-1 pol could be performed to determine whether the HSV-1 sequences have similar properties. Other HSV-1 regions may be predicted from the model to be accessible to DNA or exposed on the exterior of the molecule and available for interaction with other components of the replication system. Both sorts of prediction are amenable to experimental testing. The 3D model of HSV-1 pol reported here is an approximate one and will certainly need refining; for example, we have shown that HSV-1 sequences between residues 1073 and 1100 are not essential for catalysis and thus need not be included. It is hoped that the model may serve as a starting point for further studies designed to elucidate the structure and function of this vital HSV-1 replicative enzyme. In addition, through a similar process, approximate 3D structures of other eucaryotic pol enzymes could be derived. Comparisons of the predicted 3D structure of the HSV-1 pol with other viral or cellular polymerases could also prove to be enlightening.

ACKNOWLEDGMENTS

We gratefully acknowledge A. K. Field for support, encouragement, and critical reading of the manuscript. We thank M. Hagen and K. Mazina for preparation of the oligonucleotides and J. Alcantara and J. Danielsen for creating some of the illustrations. We thank M. D. Challberg for the generous gift of the pN3J vector and B. J. Terry for supplying HSV-1-infected and uninfected HeLa cell extracts.

LITERATURE CITED