dNTP Binding Site in Rat DNA Polymerase β Revealed by Controlled Proteolysis and Azido Photoprobe Cross-Linking†

Deepak K. Srivastava,‡ Robert K. Evans,§ Amalendra Kumar,‡,|| William A. Beard,‡ and Samuel H. Wilson*,‡

Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, Texas 77555-1068, and Department of Chemistry, Williams College, Williamstown, Massachusetts 01267-2692

Received November 6, 1995; Revised Manuscript Received January 16, 1996⊗

ABSTRACT: Mild proteolysis of rat DNA polymerase β (β-pol) generates an N-terminal 8 kDa domain and a C-terminal 31 kDa domain; the 31 kDa domain is degraded to 6 and 27 kDa fragments by further proteolysis [Kumar, A., Widen, S. G., Williams, K. R., Kedar, P., Karpel, R. L., & Wilson, S. H. (1990) *J. Biol. Chem. 265*, 2124−2131]. In the present study, we found that more vigorous trypsin digestion of the 27 kDa fragment of β-pol produces 10 and 12 kDa subdomains. Thus, rat β-pol has four distinct proteolytic fragments of 8, 6, 10, and 12 kDa, extending from the N-terminus to the C-terminus, respectively. To map the location of the dNTP binding site(s), intact β-pol was photoaffinity labeled with 8-azido-ATP or 5-azido-dUTP in presence or absence of competitor dNTP (dATP). The labeled enzyme was subjected to controlled proteolysis, and the resulting labeled peptides were separated and sequenced. Competition with dATP showed that three regions of β-pol in solution combine to form the dNTP binding pocket as follows: residues 4−40 of the 8 kDa domain; residues 142−206 of the 10 kDa subdomain; and residues 263−280 of the 12 kDa subdomain (α-helices M and N). These results are discussed in light of the recent crystal structure of dATP bound to rat β-pol.

DNA polymerase β (β-pol) is a constitutively expressed "housekeeping" enzyme in vertebrates involved in short gap filling synthesis during base excision repair [for discussion see Singhal et al. (1995) and Sobol et al. (1996)]. Mammalian β-pols characterized to date are highly conserved at the primary structure level (Wilson, 1990), and human and rat β-pols have been overexpressed in *Escherichia coli* (Abbotts et al., 1988; Date et al., 1988; Kumar et al., 1990b). These recombinant β-pols are fully active in DNA synthesis and possess template·primer specificity and other catalytic properties similar to those of the natural homologue (Abbotts et al., 1988). The purified recombinant human and rat β-pols are monomeric proteins at physiological ionic strength and are polypeptides of 334 amino acids (SenGupta et al., 1986; Zmudzka et al., 1986; Abbotts et al., 1988; Kim et al., 1994), differing from the natural enzymes only by the absence of the natural N-terminal Met, which is removed when expressed in *E. coli*.

Controlled proteolysis and chemical cleavage of rat and human β-pol revealed that the enzymes in solution have similar and distinctive domain structure: an N-terminal "template binding" domain of 8 kDa (residues 4 to ~80) and a C-terminal "catalytic" domain of 31 kDa (residues 88−335) (Kumar et al., 1990a,b). Chymotryptic removal of the N-terminus of the 31 kDa domain produced a 27 kDa fragment (residues 142−335) that was devoid of enzymatic activity, suggesting that residues 88−141 are required for activity (Kumar et al., 1990a). Lys^{72}, the Gly^{188}−Leu^{217} peptide fragment, and some additional portions of the 31 kDa domain have been implicated in binding a substrate (dNTP) or a metal activator (Mg^{2+}) (Basu et al., 1989; Evans & Coleman, 1989; Kumar et al., 1990a; Recupero et al., 1992).

The crystal structure determinations of Klenow fragment of *E. coli* DNA Pol I, HIV-1 reverse transcriptase, RNA polymerase from bacteriophage T7 revealed a common groove-like configuration and three distinct structural modules or subdomains (designated as fingers, palm, and thumb because of their resemblance to features of the right-hand) forming an obvious DNA binding channel. These observations led to the hypothesis that perhaps polymerases may share a common nucleotidyl transfer mechanism centered around the highly conserved carboxylic acid residues in the palm (Kohlstaedt et al., 1992; Jacobo-Molina et al., 1993; Sousa et al., 1993; Ollis et al., 1985). The crystal structure of rat β-pol nevertheless revealed a polymerase configuration (Davies et al., 1994; Pelletier et al., 1994; Sawaya et al., 1994) consisting of palm (with conserved carboxylic acid residues), fingers, and thumb (along with an additional 8 kDa domain attached to the fingers). Two other mammalian polymerases, terminal deoxynucleotidyltransferase and polyadenylate polymerase, share some similarities in primary structure with β-pol (Anderson et al., 1987; Matsukage et al., 1987; Kumar et al., 1990b; Raabe et al., 1991).

The combination of a relatively simple primary structure, crystal structures, and large amounts of homogeneous enzyme make β-pol a most attractive model for DNA polymerase structure-function studies. Photoaffinity labeling with azidonucleotide analogues, along with dNTP competition, can provide a direct approach for elucidating nucleotide binding regions in a nucleotide-binding protein in solution. Azidonucleotides, which serve as substrates for the enzymatic reaction, bind at the nucleotide binding pocket of many enzymes. Subsequent photolysis with UV light activates the azido moiety, such that it reacts with nearby amino acid residues in the nucleotide-binding pocket. The protein region containing the covalently attached nucleotide can be identi-

† This work was supported in part by grants to S.H.W. from the NIH (ES06839) and the Robert A. Welch Foundation (H-1265) and by NIH Grant ES06676.

* To whom correspondence should be addressed.
‡ University of Texas Medical Branch.
§ Williams College.
|| Deceased.

⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1996.

0006-2960/96/0435-3728$12.00/0

fied after protein fragmentation and sequencing (Potter & Haley, 1983). In the present study, we used this approach to map the regions of the β-pol involved in nucleotide binding. Photolabeling with 5-azido-dUTP, in the presence and absence of dNTP competitor revealed that three different segments of intact β-pol form the dNTP binding pocket.

MATERIALS AND METHODS

Isolation of DNA Polymerase β. Recombinant rat β-pol was purified from *E. coli* RR1 (pRK 248cIts, pRC-Rβ1) as described by Kumar et al. (1990b). The concentration of stock solutions of β-pol was calculated from its ultraviolet absorbance using the following molar absorption coefficient $\epsilon_{280} = 2.12 \times 10^4\ M^{-1}\ cm^{-1}$.

Controlled Proteolysis with Trypsin. Large scale digestion of β-pol with trypsin was carried out in 25 mM Tris-HCl, pH 7.5, 25 mM NaCl, and 1 mM EDTA at 25 °C as described by Kumar et al. (1990a,b). Treatment with a low concentration of trypsin (mild digestion, substrate to enzyme ratio of 1000:1, w/w) generated 8 and 31 kDa fragments after a 90 min digestion of β-pol at 25 °C. Increasing concentrations of trypsin produced different molecular size fragments of β-pol. Three stable fragments of 10, 12, and 27 kDa were generated after a 60 min digestion at a substrate to enzyme ratio 10:1 (w/w, harsh digestion). The reactions were terminated by boiling the samples in SDS−PAGE sample buffer (62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 5% β-mercaptoethanol, and 0.001% bromophenol blue) and fragments were purified as described by Kumar et al. (1990b).

Protein Microsequencing of 27, 12, and 10 kDa Fragments. After β-pol was digested with trypsin as described above under mild or harsh digestion conditions, the resulting fragments were separated by electrophoresis on a 15% SDS−polyacrylamide gel (Laemmli, 1970) and electrotransferred to a 0.45-μm polyvinylidene fluoride (PVDF) membrane (Immobilon P; Millipore) in 10 mM 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS), pH 11.0, containing 10% (v/v) methanol (Towbin et al., 1979). Membrane was stained with Coomassie Brilliant Blue, and the discrete bands (27, 12, and 10 kDa) were excised and were subjected to microsequence analysis using an Applied Biosystems Inc. model 477A protein sequencer.

Photoaffinity Labeling of β-pol with 8-Azido-ATP and 5-Azido-dUTP. 8-Azido-ATP was purchased from Sigma, whereas 5-azido-dUTP was synthesized using a published procedure (Evans & Haley, 1987). [γ-^{32}P]-8-Azido-Rβ1 and [γ-^{32}P]-5-azido-dUTP were synthesized according to a standard procedure (Glynn & Chappell, 1964; Evans & Haley, 1987). The reaction mixture contained 50 mM Tris-HCl, pH 7.1, 25 mM NaCl, and 8 mM $MgCl_2$, along with 10−20 μM of [γ-^{32}P]-8-azido-ATP or [γ-^{32}P]-5-azido-dUTP (4.7 mCi/μmol) and 10 μM homogeneous β-pol. The reactions were performed in the presence or absence of 2 mM dATP. Each reaction was photolyzed for 1 min with a 302-nm UV lamp (Spectronics Corp., model EB-28) from a distance of 0.5 cm (5 mW/cm^2). After photolysis the reaction mixture was subjected directly to SDS−PAGE or after trypsin digestion as described below. All control experiments were exposed to the same conditions and treatments as the photolyzed samples. Labeling of β-pol was not observed when the above reaction was not photolyzed.

Proteolysis of ATP or dUTP Cross-Linked β-pol and Analysis of Fragments

Mild Trypsin Digestion. The β-pol that cross-linked with [γ-^{32}P]-8-azido-ATP or [γ-^{32}P]-5-azido-dUTP was digested with trypsin (substrate/enzyme ratio 1000:1, w/w). Under this reaction condition most of the cross-linked β-pol remained undigested even after 120 min of reaction. Hence, in subsequent experiments cross-linked β-pol was digested with trypsin at substrate/enzyme ratio 500:1 (w/w) at 25 °C for different time periods from 30 to 90 min.

Harsh Trypsin Digestion. 8-Azido-ATP or 5-azido-dUTP cross-linked β-pol (1.2 μg) was digested with trypsin (substrate to enzyme ratio 10:1, w/w) in a final volume of 15 μL in 25 mM Tris-HCl, pH 7.5, 25 mM NaCl, 4 mM $MgCl_2$, and 1 mM EDTA. The reaction was carried out at room temperature, and aliquots removed at 0, 30, 60, and 90 min were mixed with gel loading buffer and boiled for 5 min. The cross-linked β-pol and its digestion products were separated on a 15% SDS−polyacrylamide gel. The dried gels were subjected to autoradiography at −70 °C with X-omat film (Kodak).

Protein Microsequencing of Labeled Fragments. 8-Azido-ATP or 5-azido-dUTP cross-linked β-pol (200 μg) were digested with trypsin under mild and harsh digestion conditions. After resolution on a 15% SDS−polyacrylamide gel, the various fragments were transferred to PVDF membranes as described above. To identify the labeled β-pol fragments, the PVDF membranes were stained with Coomassie Brilliant Blue and subjected to autoradiography at −70 °C with X-omat film. The areas corresponding to the bands on X-ray film were excised and subjected to microsequence analysis as described above.

RESULTS

Solution Structure of Rat β-pol. Mild trypsin proteolysis of β-pol results in release of two major fragments corresponding to an N-terminal 8 kDa domain (residues 4 to ~80) and a C-terminal 31 kDa domain (residues 88−335). Chymotryptic removal of 54 residues from the N-terminus of the 31 kDa domain produces the 27 kDa fragment (residues 142−335) (Kumar et al., 1990a,b). In this study, we found that with higher concentrations of trypsin, three subdomains of the 31 kDa domain are produced with molecular masses of 6, 10, and 12 kDa. In the experiment shown in Figure 1A, intact β-pol was digested into the 31 kDa domain and 27 kDa fragment within 5 min (lanes 2 and 3). The 31 kDa domain disappeared after 15 min of incubation, and the 27 kDa fragment was the major product. When digestion was continued for up to 120 min, the amount of 27 kDa fragment decreased, with concomitant increase in the 10 and 12 kDa fragments (lanes 3−8). These results suggested that the 10 and 12 kDa fragments are derived from the 27 kDa fragment. The positions of the cleavage sites for production of the 10 and 12 kDa fragments were determined by isolating each fragment and sequencing the first few residues from the N-terminus (Figure 1B). The fact that the 10 kDa fragment has the same N-terminal sequence (YFEDF) as the 27 kDa fragment confirmed the formation of the 10 kDa fragment from the 27 kDa fragment and mapped the 10 kDa fragment boundaries (residues 142−222). The C-terminal boundary of the 10 kDa fragment was

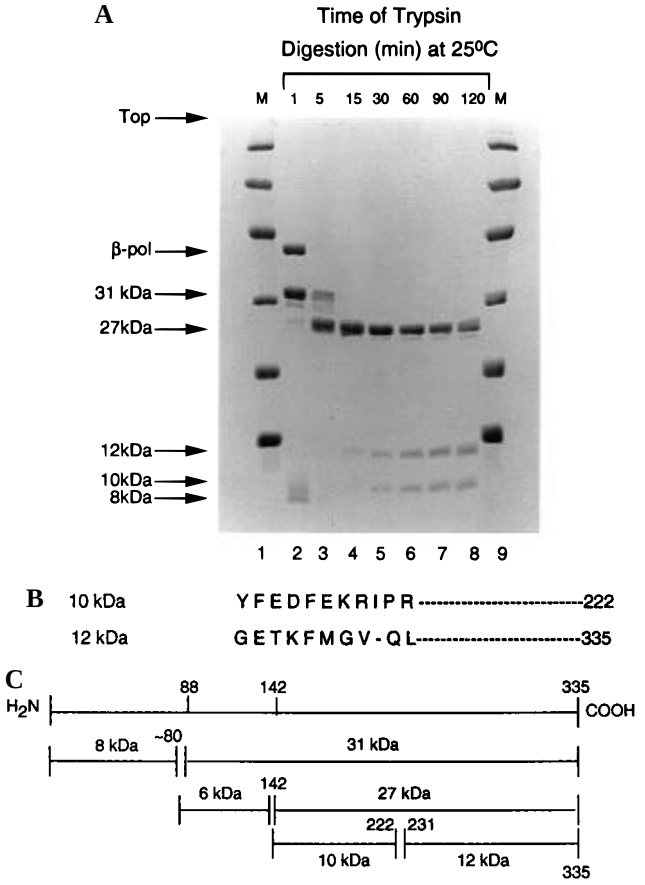


FIGURE 1: Domain mapping of β-pol. Panel A: Photograph of a Coomassie Brilliant Blue stained SDS−polyacrylamide gel illustrating the controlled proteolysis of purified rat β-pol with trypsin (substrate/enzyme ratio = 10:1, w/w). Lanes 2−7, cleavage of β-pol into 31 and 27 kDa fragments, which on prolonged digestion produces two additional fragments of 10 and 12 kDa, time (min) of trypsin digestion is shown at the top of each lane. Lanes 1 and 9 are protein molecular weight markers of 97, 69, 46, 31, 20, and 14 kDa (top to bottom). Panel B: Amino-terminal sequence analysis of the 10 and 12 kDa fragments. The carboxy terminus of each fragment was deduced on the basis of size of each fragment on SDS−PAGE and specificity of trypsin. The 10 kDa fragment begins at Tyr^{142} and spans residues 142−222 of β-pol, whereas the 12 kDa fragment starts at residue Gly^{231} and extends up to the C-terminus of β-pol. Panel C: Schematic representation of various domains and subdomains obtained from tryptic digestion of rat β-pol.

deduced from its apparent mass and the presence of a cleavage site for trypsin at residue 222.

The 12 kDa fragment had the N-terminal sequence GETKFM where the G is residue 231 of β-pol. On the basis of apparent molecular mass, we estimate that the 12 kDa fragment extends to the C-terminus of β-pol, i.e., the 12 kDa fragment contains 105 amino acids (residues 231−335; Figure 1B,C). Hence, proteolysis of β-pol produces three fragments from the 31 kDa domain (Figure 1C), designated as the 6, 10, and 12 kDa subdomains, respectively. Note that the 8 kDa domain (residues 4 to ~80) and 6 kDa (residues ~88 to 141) subdomain were not observed by SDS−PAGE analysis under the relatively harsh proteolytic conditions used here (Figure 1A).

Photoaffinity Labeling of β-pol with [γ-^{32}P]-8-Azido-ATP. After conditions for producing the domains and subdomains of β-pol had been established, we used this proteolytic division of the enzyme to identify regions involved in dNTP binding. β-Pol was photoaffinity labeled with azido-ATP

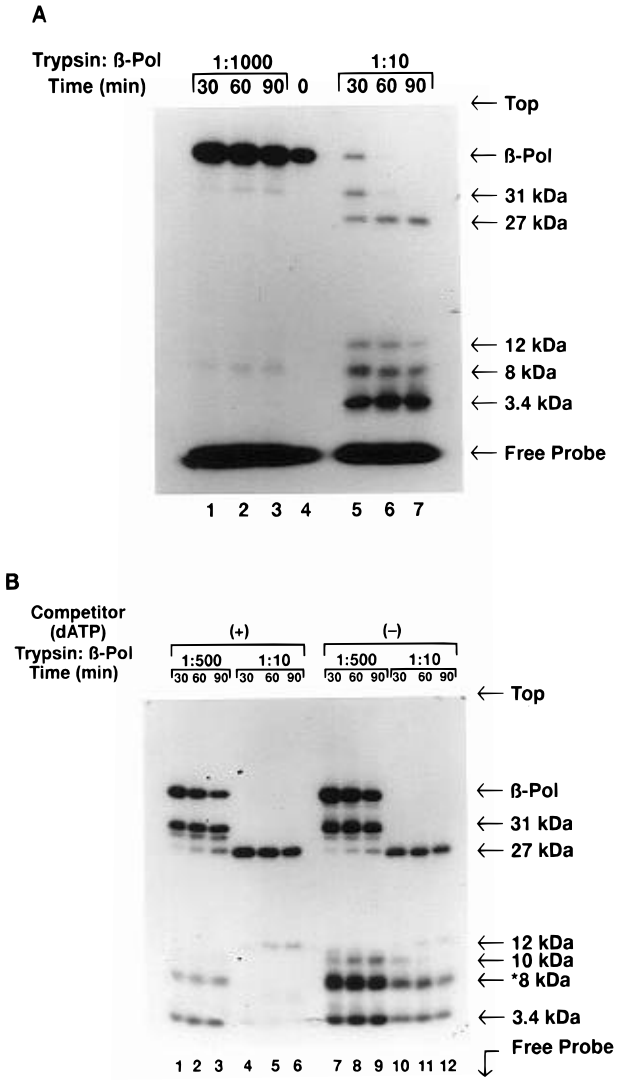


FIGURE 2: Proteolysis of photoaffinity-labeled β-pol. β-Pol was photoaffinity labeled with [γ-^{32}P]-8-azido-ATP and digested with trypsin. Panel A: β-pol was cross-linked with [γ-^{32}P]-8-azido-ATP in the absence of competitor dATP; lanes 1−3, mild tryptic digestion of cross-linked β-pol from 30 to 90 min; lanes 5−7, harsh trypsin digestion of cross-linked β-pol from 30 to 90 min; and lane 4 represents undigested photolabeled β-pol. Panel B: Mild (lanes 1−3 and 7−9) and harsh (lanes 4−6 and 10−12) tryptic digestion of photolabeled β-pol. The β-pol was photolabeled with [γ-^{32}P]-8-azido-ATP in the presence (lanes 1−6) or absence (lanes 7−12) of 2 mM competitor dATP. Reaction conditions are indicated at the top of each lane. Products were separated on a 15% SDS−polyacrylamide gel. The *8 kDa represents two different 8 kDa fragments, one generated from the N-terminus of β-pol (lanes 7−9) and the other from the N-terminus of the 10 kDa fragment (lanes 10−12) upon trypsin digestion.

(molecular ratio of 1:1). Cross-linked β-pol was digested with trypsin under both mild and harsh conditions (Figure 2A). Mild digestion (substrate:enzyme ratio, 1000:1, w/w) resulted in only partial degradation of intact β-pol. Most of the cross-linked β-pol remained undigested even after 90 min of incubation (Figure 2A, lanes 1−3), indicating that the "hinge region" connecting the 8 and 31 kDa domains of β-pol was less sensitive to protease digestion after the cross-linking procedure. Therefore, we increased the amount of trypsin in the digestion. At a 500:1 β-pol:trypsin ratio (w/w), labeled proteolytic fragments of 31, 27, 10, 8, and ~3.4 kDa were generated from labeled β-pol (Figure 2B, lanes 7−9). To determine if these fragments were cross-linked at a binding

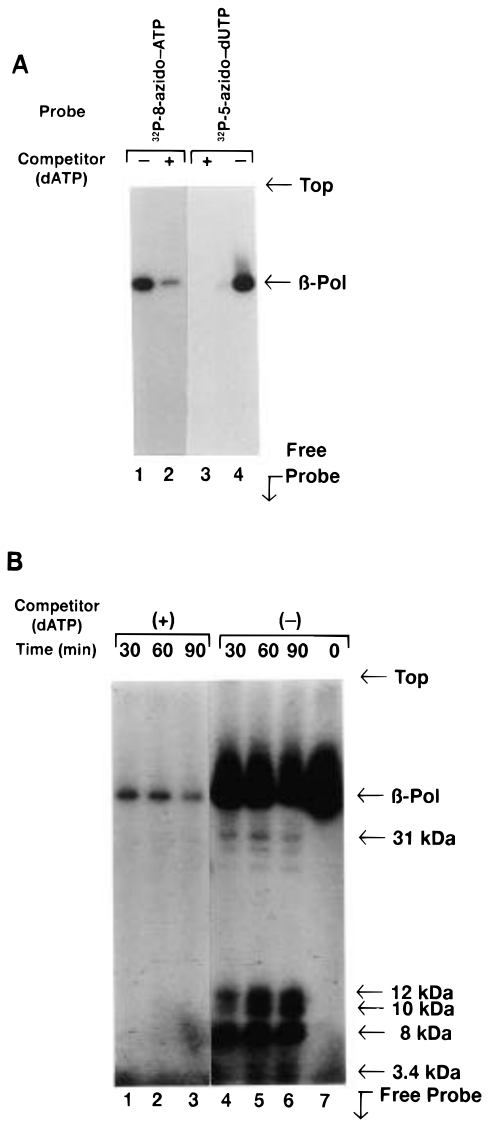
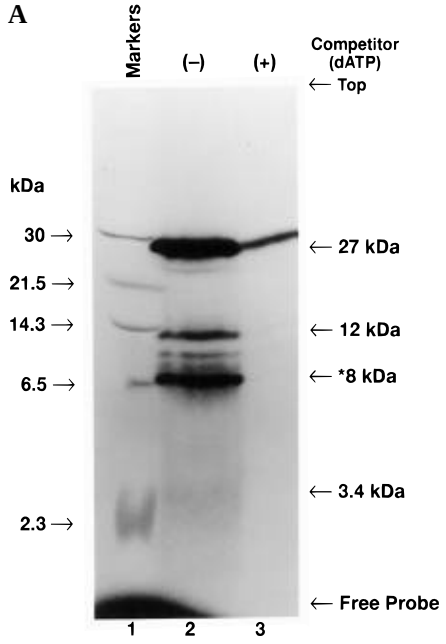


FIGURE 3: Photolabeling of β-pol-dNTP binding pocket and proteolysis of the photolabeled complex. Panel A: Photograph of autoradiogram showing the comparison of β-pol specific labeling with [γ-^{32}P]-8-azido-ATP (lanes 1 and 2) or [γ-^{32}P]-5-azido-UTP (lanes 3 and 4). The β-pol was photolabeled in the presence (lanes 2 and 3) or absence (lane 1 and 4) of 2 mM competitor dATP. The products were separated on a 15% SDS−polyacrylamide gel and visualized by autoradiography. Panel B: Photograph of autoradiogram illustrating results of mild tryptic digestion of [γ-^{32}P]-5-azido-dUTP-linked β-pol. Photolabeling was performed in the presence (lanes 1−3) or absence (lanes 4−7) of dATP. The two minor bands visible just below the 31 kDa band at ~27 kDa represent partial degradation of the 31 kDa domain as revealed by sequencing; these two peptides are cleaved to the 10 and 12 kDa fragments. Time (min) of trypsin digestion is indicated at the top of each lane.

site specific for dNTPs, β-pol was labeled with [γ-^{32}P]-8-azido-ATP in the presence of 2 mM dATP. As shown in Figure 2B the smaller labeled fragments of 10, 8, and ~3.4 kDa failed to appear in the digest of β-pol that had been cross-linked in the presence of competitor dATP, confirming the ability of 8-azido-ATP and dATP to bind at the same site under these conditions. The 27 kDa fragment and 31 kDa domain were partially cross-linked with azido-ATP in the presence of competitor dATP (Figure 2B, cf. lanes 7−9 to lanes 1−3), suggesting the presence of some modest nonspecific labeling. In summary, harsh trypsin digestion of 8-azido-ATP cross-linked β-pol produced five labeled



B Sequence of cross-linked fragment.

Mass Cross-linked Fragment	Sequence	Corresponding domain of ß-Pol
27 kDa	^{142}YFEDFE – – – –	31 kDa
12 kDa	^{231}GETKFM – – – –	12 kDa
*8 kDa	^{142}YFEDFE – – – –	10 kDa
3.4 kDa	^{4}RKAPQ – – – –	8 kDa

FIGURE 4: Identification of specifically photolabeled fragments of β-pol. Panel A: Photograph of an autoradiogram illustrating the harsh digestion of β-pol cross-linked to 5-azido-dUTP in the presence (lane 3) or absence (lane 2) of 2 mM dATP. Digestion products were separated on a 17.5% SDS−polyacrylamide gel, transferred onto a PVDF membrane, and exposed to X-ray film. Molecular weights on the right-hand side of the figure represent the size of each fragment, and protein molecular weight standards (lane 1) are shown on the left-hand side. The fragments between *8 and 12 kDa represent the 10 kDa fragment and partial degradation of the 12 kDa fragment. Panel B: Various bands observed by autoradiography (Panel A) were localized on PVDF membrane and excised. On the basis of the N-terminal sequence of each cross-linked fragment the domain of origin was assigned.

fragments with molecular masses of 27, 12, 10, 8, and 3.4 kDa (lanes 10−12). When competitor dATP was present, 8-azido-ATP cross-linking to the 10, 8, and 3.4 kDa fragments was strongly reduced (Figure 2B, cf. lanes 4−6 to lanes 10−12).

Photoaffinity Labeling of β-pol with [γ-^{32}P]-5-Azido-dUTP. To confirm and extend the azido-ATP labeling results we also conducted studies with the 5-azido derivative of deoxyribonucleotide, dUTP. The β-pol was specifically cross-linked with 5-azido-dUTP, as the addition of dATP markedly decreased cross-linking (Figure 3A, cf. lanes 1 and 2 to lanes 3 and 4). Cross-linked β-pol was digested under both mild and harsh conditions: With mild conditions, labeled fragments of 12, 8, and 3.4 kDa were observed; labeling of these fragments was blocked by competition with dATP (Figure 3B, cf. lanes 1−3 to lanes 4−6); with harsh conditions, fragments of 3.4, *8.0, 12.0, and 27 kDa were observed; except for the 27 kDa fragment, labeling was blocked by competition with dATP (Figure 4A). The *8 kDa fragment is distinct from the 8 kDa domain (see below). The identity of these labeled fragments was verified by

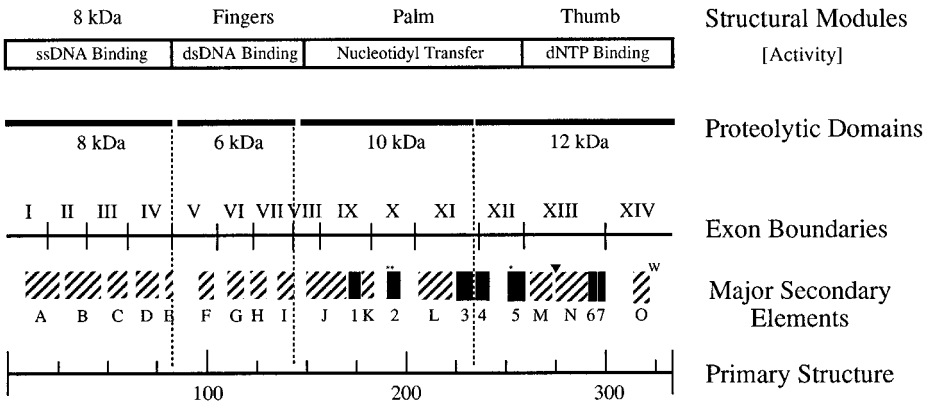


FIGURE 5: Relationship between gene and protein structures of DNA polymerase β. DNA polymerase β has four structural modules and four proteolytic domains associated with primary dedicated activities (Kumar et al., 1990a,b; Sawaya et al., 1994): DNA recognition, nucleotidyl transferase, and dNTP binding. Locations of structural modules, proteolytic domains, and the 14 known exons and positions of α-helices (stripped boxes designated A−O) and β-sheets (filled boxes designated 1−7) are compared with the primary structure of β-pol along the bottom. Assignment of secondary structural elements was determined by the DSSP procedure (Kabsch & Sander, 1983). Symbols at the top of major secondary elements are as follows: *, DNA polymerase conserved carboxylic acids; ▼, cis-peptide bond; W, tryptophan.

Table 1: Relative Photochemical Labeling of Regions of DNA Polymerase β with [γ-^{32}P]-5-Azido-dUTP[a]

	proteolytic fragment								
conditions	8 kDa domain			10 kDa subdomain			12 kDa subdomain		
competitor (dATP)	−	+		−	+		−	+	
mild digestion	2.3	0.2	(2.1)[b]	0.5	0.0	(0.5)	0.8	0.0	(0.8)
harsh digestion	0.1	0.0	(0.1)	4.4	0.0	(4.4)	2.1	0.1	(2.0)

[a] DNA polymerase β was photolabeled with 5-azido-dUTP in the presence (+) or absence (−) of 2 mM dATP and digested with trypsin as described under Materials and Methods. Relative labeling corresponds to integrated densitometric signals [(−)/(+)]. [b] Values in parenthesis are the difference in labeling in the absence and presence of competitor.

protein microsequencing after transfer to a PVDF membrane. The 8 and 3.4 kDa fragments produced under mild conditions shared the N-terminal sequence RKAPQ, corresponding to N-terminal residues 4−8 of β-pol. Hence, since the 3.4 kDa fragment is from the N-terminus of β-pol, the results clearly indicate that a portion of the 8 kDa domain participates in 5-azido-dUTP binding. In experiments not shown, we found that the N-terminal 8 kDa domain is not labeled in the absence of the 31 kDa domain.

Microsequencing of the 3.4, *8, 12, and 27 kDa fragments, obtained with harsh conditions, also was performed (Figure 4B). The sequence of the 3.4 kDa fragment was RKAPQE corresponding to the N-terminal residues of β-pol. The sequence of the *8 kDa fragment (from harsh digestion) was YFEDFEKRIP, corresponding to the N-terminus of the 10 kDa subdomain (residues 142−222), indicating that amino acid(s) in the 10 kDa subdomain also participate in dATP binding. The N-termini of the remaining two fragments (12 and 27 kDa) had the sequences GETKFM and YFEDFE, respectively, representing the 12 kDa subdomain and 27 kDa fragment (Figures 4B and 1).

Specificity of Labeling β-Polymerase. To quantify photoaffinity labeling in regions of β-pol, the enzyme was labeled with [^{32}P]-5-azido-dUTP in the presence and absence of 2 mM dATP. Following either mild or harsh trypsin digestion and SDS−PAGE, the decrease in ^{32}P incorporation resulting from dATP protection was determined for the N-terminal 8 kDa domain and the 10 and 12 kDa subdomains. The results indicate that specific photolabeling occurred (i.e., competed by dATP) in the 8 kDa domain and in the 10 and 12 kDa subdomains, Table 1. These regions of β-pol do not appear to independently possess dNTP binding capacity, as the harsh trypsin digestion products of β-pol in solution did not show any cross-linking with 5-azido-dUTP (data not shown).

DISCUSSION

DNA polymerase β, a 39 kDa monomer in solution at physiologic ionic strength (Abbotts et al., 1988; Kim et al., 1994), is considered the simplest naturally occurring prokaryotic or eukaryotic DNA polymerase. The enzyme has been extensively studied by physical biochemical approaches in solution (Kim et al., 1994) and in crystals as various binary and ternary complexes with DNA and ddNTP (Pelletier et al., 1994; Sawaya et al., 1994). In the present study, we show that β-pol can be divided into four proteolytic domains or subdomains, of approximately 60−100 amino acids each as follows: N-terminal 8 kDa domain (residues 4 to ~80); 6 kDa subdomain (residues ~88 to 141); 10 kDa subdomain (residues 142−222); and a C-terminal 12 kDa subdomain (residues 231−335). These results from proteolytic domain mapping are in agreement with the recently published crystal structure of β-pol (Davies et al., 1994; Sawaya et al., 1994), where the three subdomains in the 31 kDa domain were represented as finger, palm, and thumb regions or modules (Figure 5).

With the results reported here on the four proteolytic domains of β-pol, along with the crystal structure of β-pol, it is interesting to consider structural features of the protein in comparison with the organization of the gene (Chyan et al., 1994). There have been two popular theories on the emergence of structural features of eukaryotic genes. The "intron early" theory suggests that genes are assembled from exons encoding protein with a dedicated role by intron-mediated recombination. In contrast, the "intron late" model suggests that introns are inserted late evolutionarily, such

that longer continuous exons are invaded by introns. For the intron early model, as first suggested by Blake (1978), exons encode discrete protein domains suitable for transposition as a unit. Thus, correspondence of exons with globular, folded domains that were functional either structurally or catalytically was anticipated. However, it became clear that exons are generally too small to encode full protein domains. Blake (1983) later proposed that domains are encoded by groups of exons and that individual exons encode secondary structure elements. If genes evolved in this fashion, the positions of introns relative to globular domain secondary structural elements might not be random. More recently, Stolzfus et al. (1994) tested this theory by comparing a number of protein and gene structures of "ancient" proteins such as alcohol dehydrogenase, globins, pyruvate kinase, and triosephosphate isomerase. Using several lines of evidence in favor of the intron late model, they suggest there is no significant correspondence between exons and units of protein secondary structure and concluded that the intron early model is not logical.

The human gene encoding β-pol spans 33 kb and contains 14 exons (Chyan et al., 1994). As might be expected, all the domain boundaries occur between elements of secondary structure with the boundaries of the 8 kDa domain and 6 kDa subdomain occurring at exon-intron junctions at the end of a secondary structural element (Figure 5). In contrast, the 10 and 12 kDa subdomains end and begin with exon XI, respectively. More than half (i.e., seven) of the 13 introns occur at the ends of secondary structural elements, α-helices or β-strands (i.e., 147 residues). The probability of selecting such a residue at random is 147/335 or 0.44 for β-pol (335 residues). If the 13 introns were inserted randomly, one would expect to hit $13P$ or 5.7 ends of secondary structural elements with a standard deviation of about 1.8 [SD = $13P(1 - P)^{0.5}$]. Since seven introns were observed to coincide with the end of a secondary structural element, the introns *do not* appear to occur preferentially at the ends of these structural elements. This limited analysis is consistent with the idea that introns for human DNA polymerase β may have been acquired in a random fashion, i.e., in a manner unrelated to protein secondary structure.

Photoaffinity labeling with nucleotides has been an important technique in determining protein nucleotide interactions (Havron & Sperling, 1977; Salvucci et al., 1992; Prasad et al., 1993), and we exploited the approach here to study the involvement of various regions of β-pol in nucleotide binding. We found that parts of the 8 kDa domain, and 10 and 12 kDa subdomains bind dNTP (Table 1). The label in the 8 kDa domain was found in a 3.4 kDa fragment, corresponding to the first 40 amino acids (residues 4−40) of β-pol. The 8 kDa domain is also known to possess single-strand nucleic acid-binding capacity (Kumar et al., 1990b), and Prasad et al. (1993) showed proximity of Ser^{30} and His^{34} to the single-strand DNA binding site, as revealed by UV cross-linking experiments with oligo$(dT)_{16}$. Additionally, this domain has been shown to "direct" β-pol to the 5′-phosphate in gapped DNA substrates (Prasad et al., 1994) and to have deoxyribose phosphosdiesterase activity important in base excision repair (Matsumoto & Kim, 1995).

Label in the 10 kDa subdomain was localized to the *8 kDa N-terminal region (residues 142−206) (Figures 4 and 5). Hence, this N-terminal part of the 10 kDa subdomain appears to undergo a structural alteration upon substrate binding, because the 142−206 fragment is not observed when β-pol is digested in the absence of dNTP. The location of dATP in the β-pol crystal structure of the β-pol/dATP complex also suggests the involvement of Asp^{190} and Asp^{192} in catalytic function, and other amino acids of the 142−206 fragment, as revealed in the crystal structure, are in close proximity to the active site (Figure 6). The involvement of the 10 kDa subdomain in dNTP substrate binding is also in general agreement with the idea that this region is conserved among other DNA polymerases. The ^{189}GDMD^{192} region of rat β-pol resembles the GDTD motif believed to be involved in substrate/dNTP binding of several replicative polymerases (Chen & Horwitz, 1989; Evans et al., 1989; Wang et al., 1989).

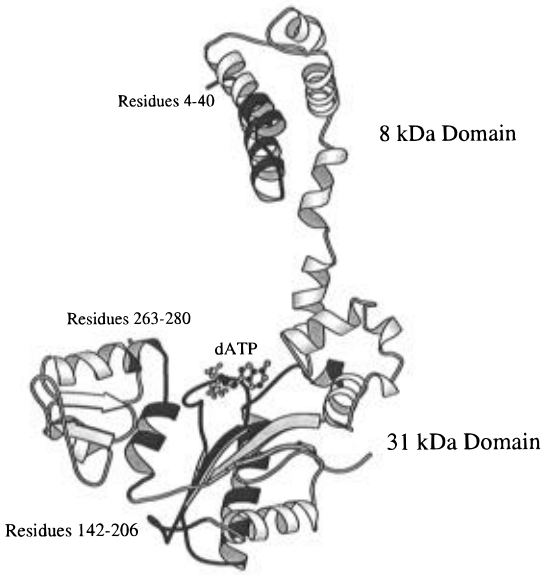


FIGURE 6: Rat β-pol binary complex. Ribbon representation of β-pol complexed with dATP. The regions of primary structure cross-linked with azidonucleotide are shaded darker, and the boundaries of the isolated cross-linked fragments are indicated. Residues 4−8 are not depicted since they are not observed in the crystallographic structure. This figure was made with MOLSCRIPT (Kraulis, 1991).

Photoaffinity labeling in the 12 kDa subdomain of β-pol also was competed by excess dATP; the degree of competition was comparable to that in the 8 kDa domain and 10 kDa subdomain (Table 1). By amino acid sequencing, we identified the peptide, Asp^{263}−Lys^{280}, of the 12 kDa subdomain as the site of cross-linking (data not shown). Thus, the 12 kDa subdomain also participates in the dNTP binding pocket. In the binary and ternary complex crystal structures (Sawaya et al., 1994; Pelletier et al., 1994), α-helices M and N correspond to the peptide Asp^{263}−Lys^{280}, and this region is in the binding site of the incoming nucleotide substrate (Figure 6). We note that Recupero et al. (1992) proposed a dNTP binding motif in this region on the basis of a monoclonal antibody (mAb) study and from a comparison of the sequence of β-pol with other NTP binding enzymes. Overall, these results are in accord with earlier studies, such as those of pyridoxal phosphate inhibition (Basu et al., 1989) and mAb studies (Recupero et al., 1992), that predicted involvement of different parts of β-pol in dNTP binding. We propose involvement of amino acids in three regions of β-pol (8, 10, and 12 kDa) in the dNTP binding site under natural catalytic conditions. The involvement of more than one domain in constituting the functional dNTP binding pocket was further supported by our observation that none

of the domains (8, 10, or 12 kDa) independently bound and cross-linked azidonucleotides.

Since our studies were conducted in the absence of template·primer, the question of the dNTP substrate binding site under catalytic conditions arises. An earlier report (Evans et al., 1989) and the crystal structure of β-pol (Pelletier et al., 1994; Sawaya et al., 1994) support our studies and show the involvement of parts of the 10 and 12 kDa subdomains. For example, in the 12 kDa domain α-helices M and N are in close contact with the incoming nucleotide in the ternary complex crystal structure; yet α-helices M and N are not in close contact with dATP in the binary complex crystal structure (see Figure 6). Further, under the crystallization conditions used, involvement of the 8 kDa domain in substrate binding in both the ternary and binary complex crystal structures was not observed. The 8 kDa domain moves much closer to the bound ddCTP in the ternary complex and to the 10 kDa and 12 kDa subdomains of β-pol in both crystal forms of the ternary complex. Yet the 8 kDa domain is not close enough to contribute to the dNTP binding site (Pelletier et al., 1994). As we observed strong labeling of the 3.4 kDa fragment (N-terminal part of the 8 kDa domain), along with complete competition in the presence of dATP, the 8 kDa domain of β-pol clearly is close to the active site and forms part of the dNTP binding pocket in solution. The low level of labeling dNTP used in our experiments (ratio of enzyme:labeling dNTP, ∼1:1) and the lack of labeling in the absence of intact enzyme indicate the specificity of the labeling for the dNTP binding site.

ACKNOWLEDGMENT

This communication is dedicated to the memory of Dr. Amalendra Kumar who was involved in the preliminary studies and conceptual framework of this manuscript. We thank Drs. Sankar Mitra, Ben Van Houten, and Phyllis Strauss for helpful discussion and critical reading of the manuscript, Stefan Serabyn for oligopeptide sequencing, Dr. Michael R. Sawaya for the coordinates to produce Figure 6, and Kay Miller for typing the manuscript.

REFERENCES

Abbotts, J., SenGupta, D. N., Zmudzka, B., Widen, S. G., Notario, V., & Wilson, S. H. (1988) *Biochemistry 27*, 901−909.
Anderson, R. S., Lawrence, C. B., Wilson, S. H., & Beattie, K. L. (1987) *Gene* (Amst.) *60*, 163−173.
Basu, A., Kedar, P., Wilson, S. H., & Modak, M. J. (1989) *Biochemistry 28*, 6305−6309.
Blake, C. C. F. (1978) *Nature 273*, 267.
Blake, C. (1983) *Nature 306*, 535−537.
Chen, M., & Horwitz, M. S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6116−6120.
Chyan, Y. J., Ackerman, S., Shepherd, N. S., McBride, O. W., Widen, S. G., Wilson, S. H., & Wood, T. G. (1994) *Nucleic Acids Res. 22*, 2719−2725.
Date, T., Yamaguchi, M., Hirose, F., Nishimoto, Y., Tanihara, K., & Matsukage, A. (1988) *Biochemistry 27*, 2983−2990.
Davies, J. F., Almassy, R. J., Hostomska, Z., Ferre, R. A., & Hostomsky, Z. (1994) *Cell 76*, 1123−1133.
Evans, R. K., & Haley, B. E. (1987) *Biochemistry 26*, 269−276.
Evans, R. K., & Coleman, M. S. (1989) *Biochemistry 28*, 707−712.
Evans, R. K., Beach, C. M., & Coleman, M. S. (1989) *Biochemistry 28*, 713−720.
Glynn, I. M., & Chappell, J. B. (1964) *Biochem. J. 90*, 147−149.
Havron, A., & Sperling, J. (1977) *Biochemistry 16*, 5631−5635.
Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Lu, X., Tantillo, C., Williams, R. L., Kramer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., & Arnold, E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6320−6324.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.
Kim S.-J., Lewis, M. S., Knutson, J. R., Porter, D. K., Kumar, A., & Wilson, S. H. (1994) *J. Mol. Biol. 244*, 224−235.
Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., & Steitz, T. A. (1992) *Science 256*, 1783−1790.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kumar, A., Abbotts, J., Karawya, E. M., & Wilson, S. H. (1990a) *Biochemistry 29*, 7156−7159.
Kumar, A., Widen, S. G., Williams, K. R., Kedar, P., Karpel, R. L., & Wilson, S. H. (1990b) *J. Biol. Chem. 265*, 2124−2131.
Laemmli, U. K. (1970) *Nature 227*, 680−685.
Matsukage, A., Nishikawa, K., Ooi, T., Seto, Y., & Yamaguchi, M. (1987) *J. Biol. Chem. 262*, 8960−8962.
Matsumoto, Y., & Kim, K. (1995) *Science 269*, 699−702.
Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., & Steitz, T. A. (1985) *Nature 313*, 762−766.
Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., & Kraut, J. (1994) *Science 264*, 1891−1903.
Potter, R. L., & Haley, B. E. (1983) *Methods Enzymol. 91*, 613−633.
Prasad, R., Kumar, A., Widen, S. G., Casas-Finet, J. R., & Wilson, S. H. (1993) *J. Biol. Chem. 268*, 22746−22755.
Prasad, R., Beard, W. A., & Wilson, S. H. (1994) *J. Biol. Chem. 269*, 18096−180101.
Raabe, T., Bollum, F. J., & Manley, J. L. (1991) *Nature 353*, 229.
Recupero, A. J., Rein, D. C., & Meyer, R. R. (1992) *Biochemistry 31*, 7989−7997.
Salvucci, M. E., Chavan, A. J., & Haley, B. E. (1992) *Biochemistry 31*, 4479−4487.
Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H., & Kraut, J. (1994) *Science 264*, 1930−1935.
SenGupta, D. N., Zmudzka, B. Z., Kumar, P., Cobianchi, F., Skowronski, J., & Wilson, S. H. (1986) *Biochem. Biophys. Res. Commun. 136*, 341−347.
Singhal, R. K., Prasad, R., & Wilson, S. H. (1995) *J. Biol. Chem. 270*, 949−957.
Sobol, R. W., Horton, J. K., Singhal, R. K., Prasad, R., Kühn, R., Gu, H., Rajewsky, K., & Wilson, S. H. (1996) *Nature 379*, 183−186.
Sousa, R., Chung, Y. J., Rose, J. P., & Wang, B. (1993) *Nature 364*, 593−599.
Stoltzfus, A., Spencer, D. F., Zuker, M., Logsdon, J. M., Jr., & Doolittle, W. F. (1994) *Science 265*, 202−207.
Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350−4354.
Wilson, S. H. (1990) in *The Eukaryotic Nucleus* (Strauss, P. R., & Wilson, S. H., Eds.) Vol. 1, pp 199−234, The Telford Press, Inc., Caldwell, NJ.
Wang, T. S., Wong, S. W., & Korn, D. (1989) *FASEB J. 3*, 14−21.
Zmudzka, B. Z., SenGupta, D., Matsukage, A., Cobianchi, F., Kumar, P., & Wilson, S. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5106−5110.

BI952632H