Biochemical Pharmacology, Vol. 40, No. 1, pp. 69–81, 1990.
Printed in Great Britain.

0006-2952/90 $3.00 + 0.00

NMR STUDIES OF THE ACTIVE SITE OF DNA POLYMERASE I AND OF A 50-RESIDUE PEPTIDE FRAGMENT OF THE ENZYME

GREGORY P. MULLEN, JOSEPH B. VAUGHN, JR., PONNIAH SHENBAGAMURTHI and ALBERT S. MILDVAN*

Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, U.S.A.

Abstract—Transferred nuclear Overhauser effects (NOEs) and selective T_1 measurements were used to determine interproton distances in the substrates Mg^{2+}dATP and Mg^{2+}TTP bound to the large fragment of DNA polymerase I (Pol I). The distances are consistent with high *anti*, O1′ endo conformations for the enzyme-bound substrates, similar to nucleotides of B-DNA. These substrate conformations show little or no change when the complementary RNA templates $(rU)_{57}$ or $(rA)_{50}$ are bound. In contrast, multiple conformations, including *syn* and *anti* species, are required to fit the interproton distances measured on the enzyme-bound guanine nucleotide substrates Mg^{2+}dGTP and Mg^{2+}ddGTP. These multiple substrate conformations simplify to a single high *anti*, O1′ endo conformation when the complementary template $(rC)_{37}$ is bound, possibly due to base-pairing with the template, as in the active complex. In the presence of both template and primer, enzyme-bound Mg^{2+}ddGTP reverts to multiple conformations. This ability of Pol I to decrease the fraction of bound substrate which is appropriate for primer elongation may be an error-preventing mechanism. In all cases, the conformations of the average nucleotide of the enzyme-bound RNA templates are also B-like. Transferred NOEs from protons of the enzyme to those of bound dNTP substrates suggest hydrophobic (Ile, Leu) and an aromatic amino acid (Tyr) at the substrate binding site. Peptide I, a synthetic 50-residue peptide based on residues 728 to 777 of the Pol I sequence, containing the conserved sequence L-I-Y-G, retains significant secondary and tertiary structure in solution as found by circular dichroism (CD) and 2D NMR. While the X-ray structure shows 48% helix in this region, the sequence specific NOESY analysis suggests 18% helix, and the preservation of two of the three β turns. Peptide I shows tight binding of dNTP substrates, the substrate analog 2′,3′-trinitrophenyl-ATP, and duplex DNA, providing direct evidence that the active site for polymerization lies in this region of the enzyme, with the substrate binding along the O-helix near Leu-764, Ile-765, and Tyr-766. Another synthetic peptide, peptide II, based on residues 840 to 888 of the Pol I sequence also retains much secondary structure as detected by CD but does not bind the substrate analog TNP-ATP.

DNA polymerase I (Pol I) from *Escherichia coli*, a large protein 928 amino acids in length, is a tripartite enzyme consisting of three catalytic domains (Fig. 1). The amino terminal domain (residues 1 to 325) catalyzes a 5′,3′-exonuclease reaction, the middle domain (residues 326 to 542) catalyzes a 3′,5′-exonuclease reaction, and the carboxy-terminal domain (residues 543 to 928) catalyzes the DNA polymerase reaction [1–4]. The middle and carboxy-terminal domains (residues 326 to 928) together constitute the Klenow or large fragment of Pol I which has been cloned, sequenced [5, 6], and for which a 2.8 Å X-ray structure has been determined [2]. The binding site for the deoxynucleoside monophosphate (dNMP) product of the 3′,5′-exonuclease reaction has been located by X-ray analysis [4] (Fig. 1) and has been confirmed by site directed mutations of nearby metal-liganding residues. Thus, the D424A and D335A + E357A mutants which have lost one and two metal-binding sites, respectively, have profoundly decreased exonuclease activity, but normal polymerase activity [7].

Less is known about the polymerase active site. We have been using NMR methods to determine the conformations and amino acid environments of enzyme-bound deoxynucleoside triphosphate (dNTP) substrates and templates bound to the large fragment of Pol I [8–11] in order to apply the *NMR docking* procedure [12–14] to locate their binding sites on the enzyme. We have confirmed our proposed substrate binding site on the enzyme by the synthesis of a 50-residue peptide based on the local sequence at this site and have shown that this peptide tightly binds dNTP substrates as well as duplex DNA [15]. This paper reviews our progress in these studies.

METHODS

As reviewed in detail elsewhere [14, 16–18], the rates of longitudinal $(1/T_1)$, transverse $(1/T_2)$, and cross-relaxation (σ) of magnetic nuclei are NMR parameters especially valuable for the study of enzyme–substrate complexes because they yield both structural and kinetic information. Thus, paramagnetic effects on longitudinal relaxation rates $(1/T_{1P})$ may be used to derive distances ≤ 20 Å from paramagnetic centers such as Mn^{2+} to magnetic nuclei such as protons [16]. The cross-relaxation rate between two protons H_A and H_B, as evaluated from the product of the longitudinal relaxation rate of H_B and the nuclear Overhauser effect from H_A to H_B, may be used to obtain interproton distances ≤ 4 Å between H_A and H_B [17, 18]. The transverse relaxation rates of resonances of an enzyme-bound substrate are often dominated by the rate constant (k_{off})

* To whom all correspondence should be sent.

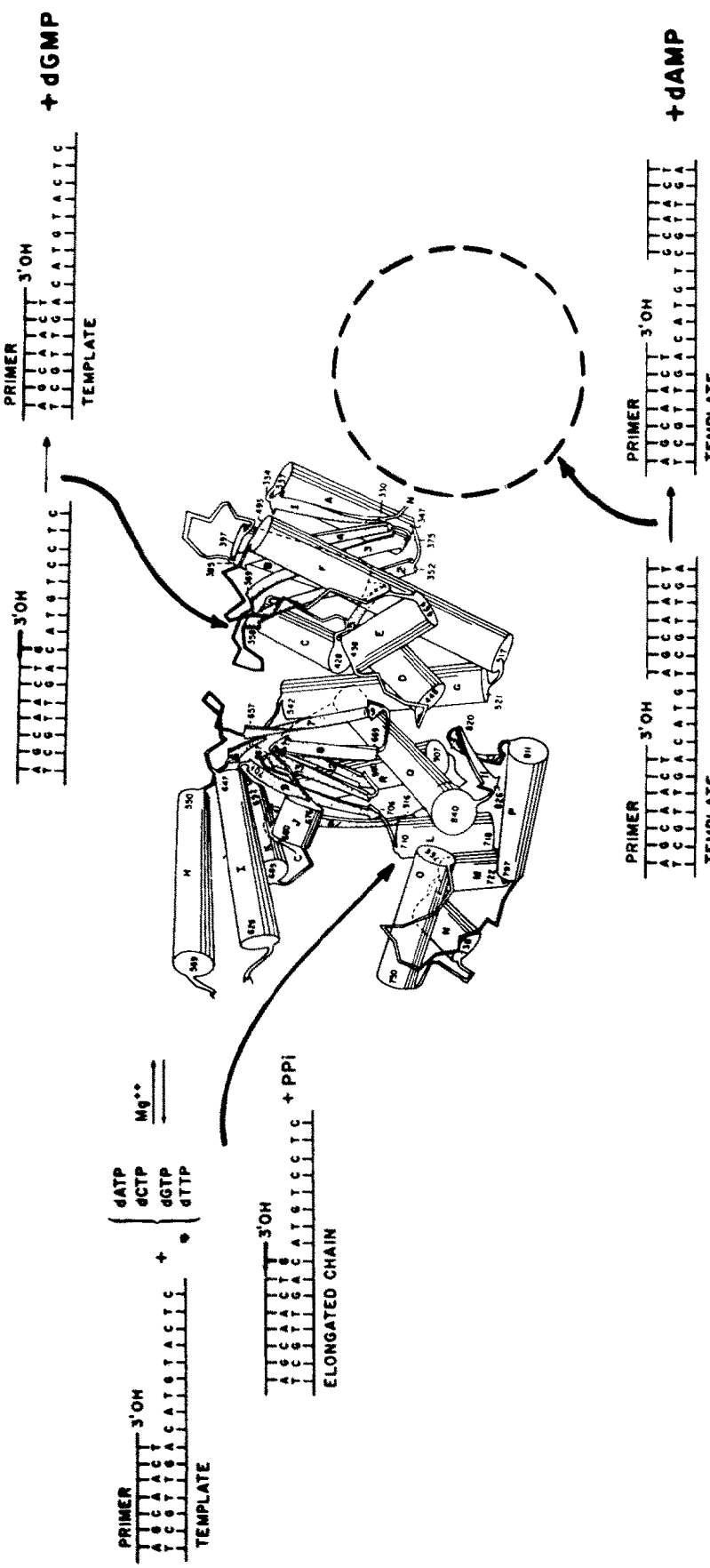

Fig. 1. Reactions catalyzed by DNA polymerase I from *E. coli*. *Upper left*: polymerase reaction catalyzed by the carboxy-terminal domain. *Upper right*: 3′,5′-exonuclease reaction catalyzed by the middle domain. *Lower right*: 5′,3′-exonuclease reaction catalyzed by the amino terminal domain. The X-ray structure of the polymerase and 3′,5′-exonuclease domains, or Klenow fragment, is shown [2].

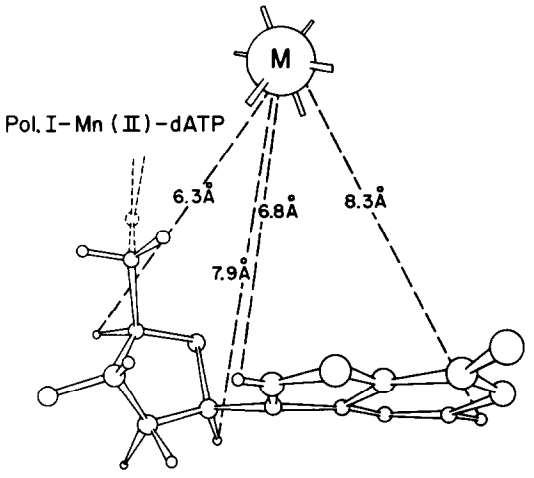

Fig. 2. Conformation of Mn^{2+}dATP bound to Pol I based on distances from Mn^{2+} bound at the polymerase site [8].

for dissociation of the enzyme–substrate complex and may therefore be used to evaluate k_{off} or its lower limit [16].

In our studies of DNA polymerase, measurements of $1/T_{1P}$ were made by the nonselective saturation recovery method at 100 and 220 MHz using Varian NMR spectrometers [8]. Cross-relaxation rates were determined at 250 MHz, using the Bruker AM 250 NMR Spectrometer, by fitting the time course of the development of the transferred NOE to the two spin equation. The T_1 values of the recipient protons were measured independently by the selective saturation recovery method [9–11]. For interproton distances in dNTP substrates, the invariant distance of 2.37 ± 0.1 Å between the deoxyribose H1′ and H2″ protons was used as an internal reference. The distance of 2.9 ± 0.2 Å from H1′ to H2′ served as a secondary internal reference. Interproton distances were used to construct molecular models of dNTP substrates from which the conformational angles were obtained [9–11, 18].

When the interproton distances in a dNTP substrate could not be accommodated by a single conformation without unrealistic distortions of the base or sugar, an alternative approach was used. The distances were fit by assuming the averaging of the minimum number of low energy nucleotide conformations [18, 19]. The latter were determined by the observed clustering of low energy conformations in X-ray structures of many nucleotides [20] and by theoretical calculations [21].

The NMR docking procedure [12–14] makes use of the structural parameters $1/T_{1P}$ and σ to position a substrate into the X-ray structure of an enzyme. It consists of four steps: (1) Determine the conformation of the enzyme-bound substrate (MgdNTP) by interproton distances from intramolecular cross-relaxation rates. (2) Measure distances from a paramagnetic substrate analog to assigned proton resonances of the enzyme by paramagnetic effects on $1/T_1$. (3) Identify amino acids within 4 Å of bound substrates by intermolecular cross-relaxation rates, from protons of the enzyme to those of the substrate. (4) Use the distances from (2) to (3) to position the proper conformation of the substrate into the X-ray structure of the enzyme, with a computer graphics system.

With DNA polymerase we have been able to carry out only steps (1) and (3) of the NMR docking procedure since an appropriate paramagnetic analog of MgdATP has not yet been found, and the complete atomic coordinates of the enzyme are not yet available. Nevertheless, as will be shown, we have been able to locate the approximate binding site of the dNTP substrate.

An independent procedure for locating the substrate binding site on an enzyme is to synthesize a small peptide, 40 to 50 residues in length, which is believed to constitute or include the substrate binding site, and to show that this peptide binds the substrate with an affinity and in a conformation comparable to that found with the complete enzyme. This method was first used to determine the ATP binding site on adenylate kinase [12]. In the present case two candidate peptides, 49 and 50 residues in length, were synthesized by the solid phase method of Merrifield [22] using an Applied Biosystems model 430A peptide synthesizer, and purified by reversed phase HPLC as previously described [15]. Peptide I consisted of residues 728 to 777 and peptide II con-

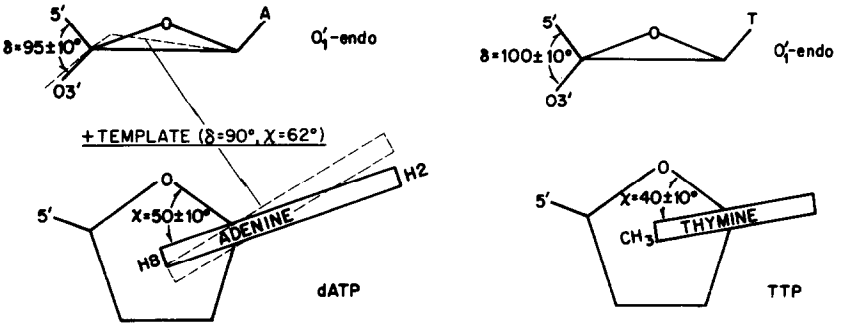

Fig. 3. Conformations of Mg^{2+}dATP and Mg^{2+}TTP bound to the large fragment of Pol I based on interproton distances determined by cross-relaxation rates [9, 10]. The conformation of Mg^{2+}dATP is shown both in the absence (solid lines) and in the presence (dotted lines) of the template $(rU)_{54}$. The conformations of MgTTP in the absence and presence of the template $(rA)_{50}$ were indistinguishable.

Table 1. Substrate conformations in complexes of the large fragment of Pol I*

Complex	Percentage contribution	Conformation	χ	δ (degrees)	γ
E-*dATP*-Mg^{2+}†	100	High *anti*, O1′ endo, C3′ endo	50 ± 10	95 ± 10	
E-$(rU)_{54}$-*dATP*-Mg^{2+}‡	100	High *anti*, O1′ endo, C3′ endo	62 ± 10	90 ± 10	
E-*TTP*-Mg^{2+}† and E-$(rA)_{50}$-*TTP*-Mg^{2+}‡	100	High *anti*, O1′ endo	40 ± 10	100 ± 10	
E-*dGTP*-Mg^{2+}§	60 ± 10	High *anti*, O1′ endo, C3′ endo	45	90	
	40 ± 10	*Syn*, O1′ endo, C2′ endo	212	135	
E-$(rC)_{37}$-*dGTP*-Mg^{2+}§	100	High *anti*, O1′ endo, C3′ endo	50 (35 to 60)	90 (125 to 85)	
E-*ddGTP*-Mg^{2+}§	30 ± 10	Low *anti*, C3′ endo	30	95	180‖
and					
E-$(rC)_{37}$-$(rI)_{14}$-ddG-*ddGTP*-Mg^{2+}§	30 ± 10	High *anti*, O1′ endo, C2′ endo	55	135	180‖
	40 ± 10	*Syn*, O1′ endo, C2′ endo	212	135	180‖
E-$(rC)_{37}$-*ddGTP*-Mg^{2+}§	100	High *anti*, O1′ endo, C2′ endo	45 ± 10	135 ± 10	180 ± 10‖

* When the interproton distances yielded a single substrate conformation, errors in the conformational angles are given. When more than one conformation was required by the distances, the errors are expressed in the percentage contribution of each well defined member of the basis set.

† From Ref. 9.

‡ From Ref. 10.

§ From Ref. 11.

‖ A γ value of 180° indicates a gauche, *trans* torsional angle about the C4′—C5′ bond.

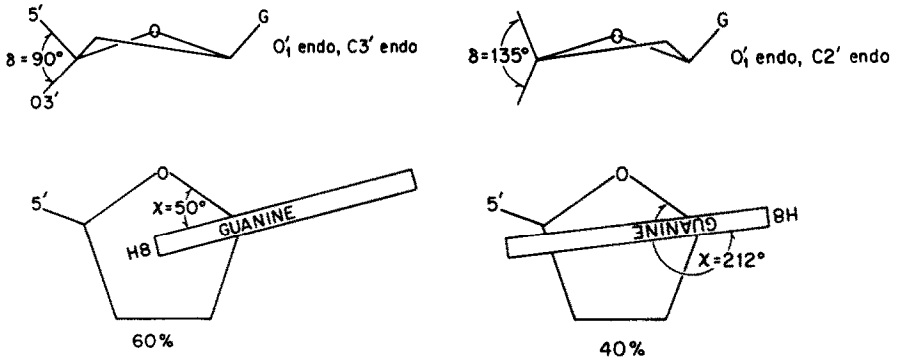

Fig. 4. Conformations of Mg^{2+}dGTP bound to the large fragment of Pol I. The averaging of 60% *anti* (left) and 40% *syn* (right) conformers was assumed in order to fit the interproton distances in the absence of template. In the presence of the template $(rC)_{37}$, only the *anti* conformation was detected [10, 11].

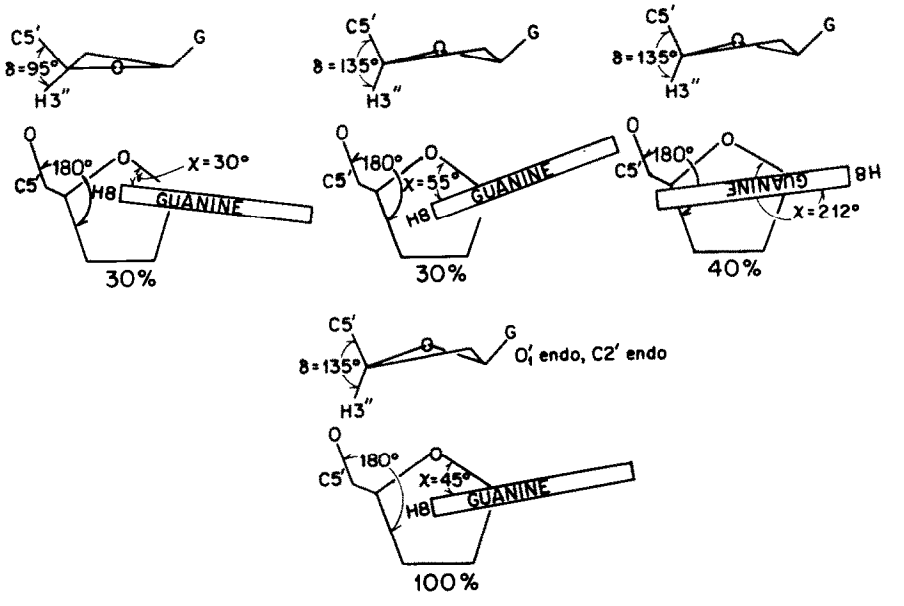

Fig. 5. Conformations of Mg^{2+}ddGTP bound to the large fragment of Pol I. The averaging of 30, 30 and 40% of the indicated conformers was assumed in order to fit the interproton distances, both in the absence of template and in the complete, chain terminated system, E-$(rC)_{37}$-$(rI)_{14}$-ddG-ddGTP-Mg^{2+}. A single conformation (lower) was found in the E-$(rC)_{37}$-ddGTP-Mg^{2+} complex [11].

sisted of residues 840 to 888 of the Pol I sequence. The binding of the fluorescent substrate analog 2',3'-trinitrophenyl-ATP (TNP-ATP) to the enzyme and to peptides I and II was studied by fluorescence enhancement by exciting at 412 nm and observing the emission at 540 nm, using an Aminco–Bowman spectrophotofluorimeter [15]. At neutral pH, the binding of authentic dNTP substrates was studied by competition with TNP-ATP. At pH 4.0 dATP is partially protonated at N1 and becomes fluorescent. The binding of dATP to peptide I at pH 3.9 was studied directly by the quenching of the fluorescence of dATP, by exciting at 240 nm and measuring the emission at 390 nm.

* Abbreviations: COSY, correlated spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; and TOCSY, total correlation spectroscopy.

The secondary structures of peptides I and II were studied by circular dichroism (CD) spectroscopy using an Aviv 60DS spectropolarimeter [15]. The solution structure of peptide I was studied by high resolution 2D NMR at 600 MHz using the Bruker AM 600 NMR spectrometer. The sequential assignments were made by phase sensitive, double quantum filtered COSY*, NOESY, and TOCSY experiments in H_2O and ^2H_2O [23–25]. In NOESY spectra in H_2O, the jump-return method was used to suppress the strong water signal [26, 27].

RESULTS AND DISCUSSION

Substrate binding and fidelity of DNA polymerase I

Both Pol I [28] and its large fragment [9] are known to bind all four dNTP substrates tightly and

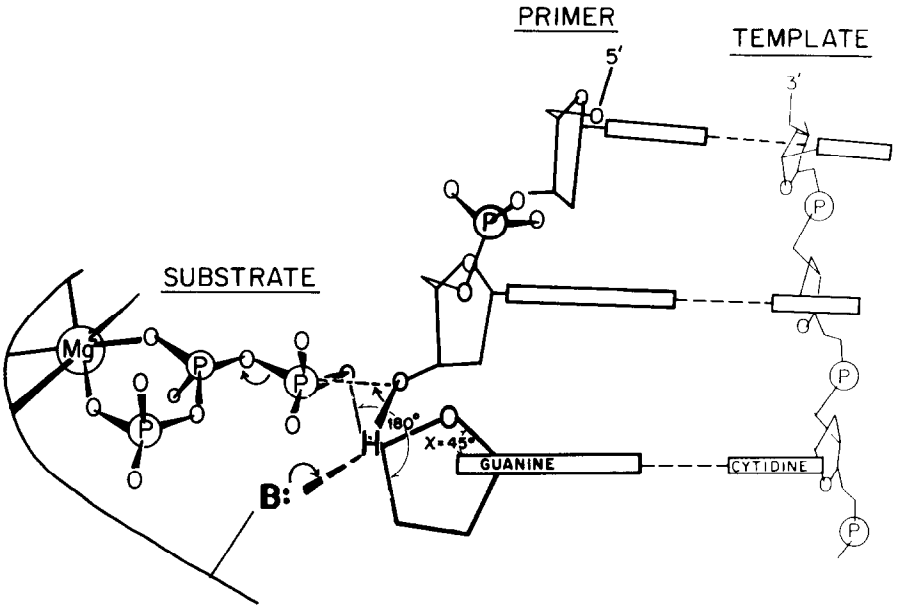

Fig. 6. Positioning of the active B-like conformation of Mg^{2+}-ddGTP into a B–DNA duplex showing a mechanism of Pol I [11].

Table 2. Conformation of average nucleotides of RNA templates bound to the large fragment of Pol I

Complex	Nucleotide unit	Conformation	χ (degrees)	δ (degrees)
E-$(rU)_{54}$*	Uridyl	High *anti*, O1′ endo	60 ± 10	105 ± 10
E-$(rA)_{50}$*	Adenyl	*Anti*	37 ± 17	
E-$(rC)_{37}$-dGTP-Mg^{2+}† and E-$(rC)_{37}$-$(rI)_{14}$-ddG-dGTP-Mg^{2+}†,‡	Cytidyl	High *anti*, O1′ endo	40 ± 10	105 ± 20

* From Ref. 10.
† From Ref. 11.
‡ Although the presence of the primer $(rI)_{14}$-ddG did not alter the conformation of the average rC unit, significant decreases in internucleotide distances from H5′5″ to H1′ and from H6 to H1′ were detected, consistent with a more B-like structure.

competitively at a single site on the enzyme which probably constitutes the polymerase site. Under most conditions, such binding is tightened by Mg^{2+}. Interestingly, the binding of (RNA) templates does not alter the relative affinities of the polymerase site for complementary versus non-complementary substrates [10, 11]. This indicates that the unusually high fidelity of Pol I in template replication, less than 1 error per 10^6 correct nucleotides incorporated [29, 30], does not result from the tighter binding of the correct substrate or from the weaker binding of the incorrect substrate in the initial complex. Rather, a subsequent verification step is necessary, prior to covalent elongation of the primer, which increases the fidelity of template copying by a factor of 10^3 [10, 11]. Evidence for such a subsequent step has been provided by rapid quench kinetics [31] but its precise structural basis is unclear. That it cannot be proofreading by the 3′,5′-exonuclease site is shown by the observation that inhibition of proofreading by the use of the substrate α-thio TTP decreased fidelity by only one order of magnitude to an error rate of 10^{-5}. This low error rate remains 10^3-fold lower than predicted by Watson–Crick base pairing alone [30].

Conformations of enzyme-bound dNTP substrates

Our early studies of the conformations of TTP and dATP bound to Pol I were based on paramagnetic effects of Mn^{2+} bound at the polymerase site on $1/T_1$ of protons of these substrates [8]. As shown for dATP (Fig. 2) a high *anti* glycosidic conformation is consistent with these distances. However, alternatives such as a *syn* conformation are also possible, and the deoxyribose pucker was undefined. More recent measurements of interproton distances in dNTP substrates bound to the large fragment of Pol I, based on NOE and selective $1/T_1$ measurements, reveal high-*anti* glycosidic conformations for enzyme-bound dATP and TTP substrates, with O1′ endo deoxyribose puckers (Fig. 3, Table 1). Such

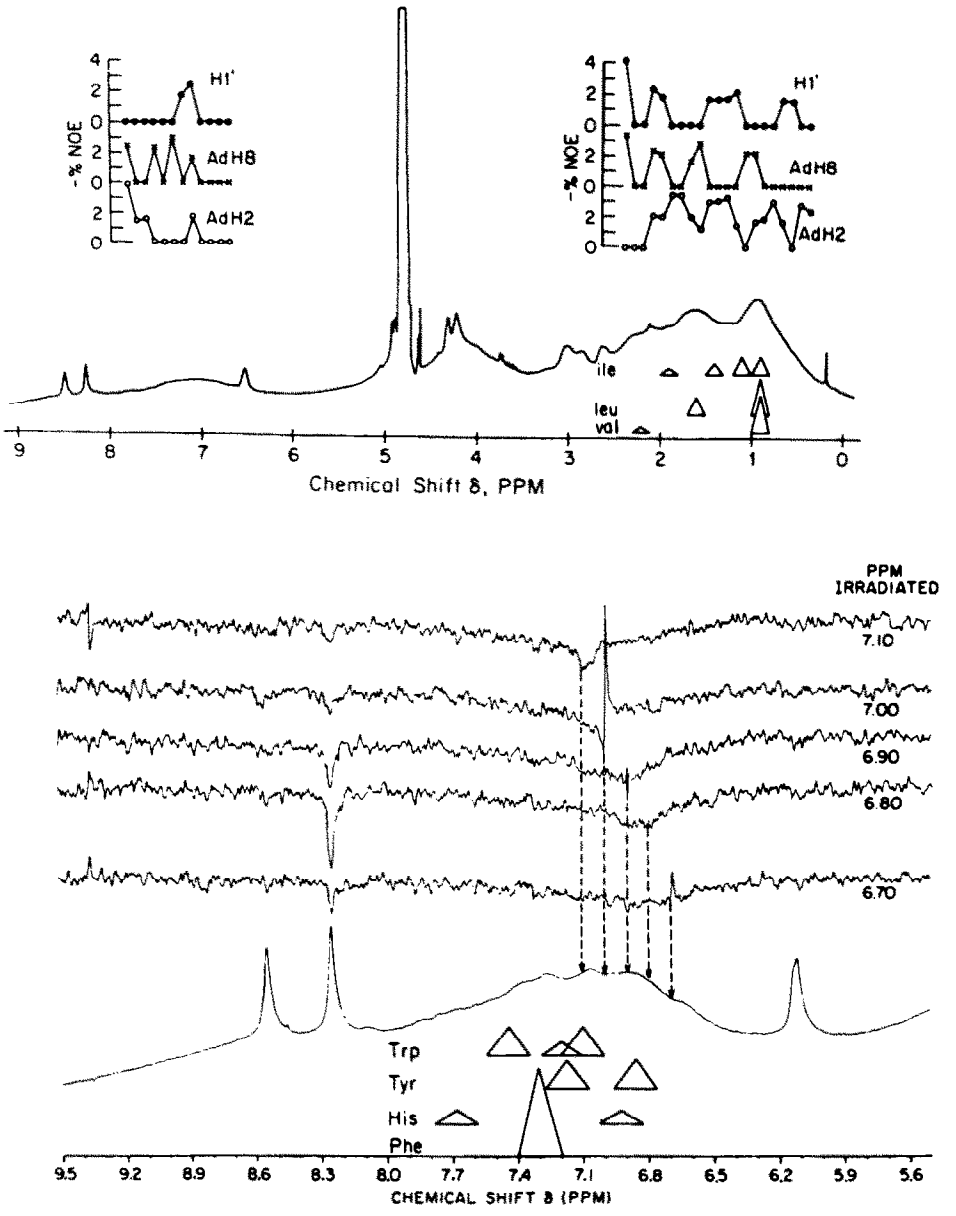

Fig. 7. NOE action spectra from protons of the large fragment of Pol I to protons of the substrate Mg^{2+}dATP (upper) and to protons of the substrate analog Mg^{2+}AMPCPP, (lower). Also shown for comparison are chemical shifts and widths of side chain resonances of hydrophobic and aromatic amino acids [9, 10].

conformations are typical of nucleotide units of B-DNA and differ significantly from those of A- or Z-DNA [9]. The binding of the complementary template $(rU)_{54}$ produced only a very small change in the conformation of bound dATP. Similarly the binding of $(rA)_{50}$ produced no detectable change in the conformation of TTP [10].

In all cases, as determined by $1/T_2$ measurements based on NMR line broadenings, the exchange rates of the bound substrates and templates onto and off of the enzyme were rapid compared to the $1/T_1$ and σ values, indicating that these parameters are in fast exchange.

Because of its greater tendency to form *syn* conformations, the substrate dGTP showed more complicated behavior, and thereby provided additional information [10, 11]. In the absence of templates, the interproton distances in enzyme-bound dGTP could not be fit by a single substrate conformation without unprecedented distortion of the deoxyribose ring. It could, however, be fit most simply by the averaging of two low energy conformations, a high *anti* O1′ endo, C3′ endo species contributing 60 ± 10% and a *syn* O1′ endo, C2′ endo species contributing 40 ± 10% (Fig. 4, Table 1). The binding of the complementary template $(rC)_{37}$ simplified the

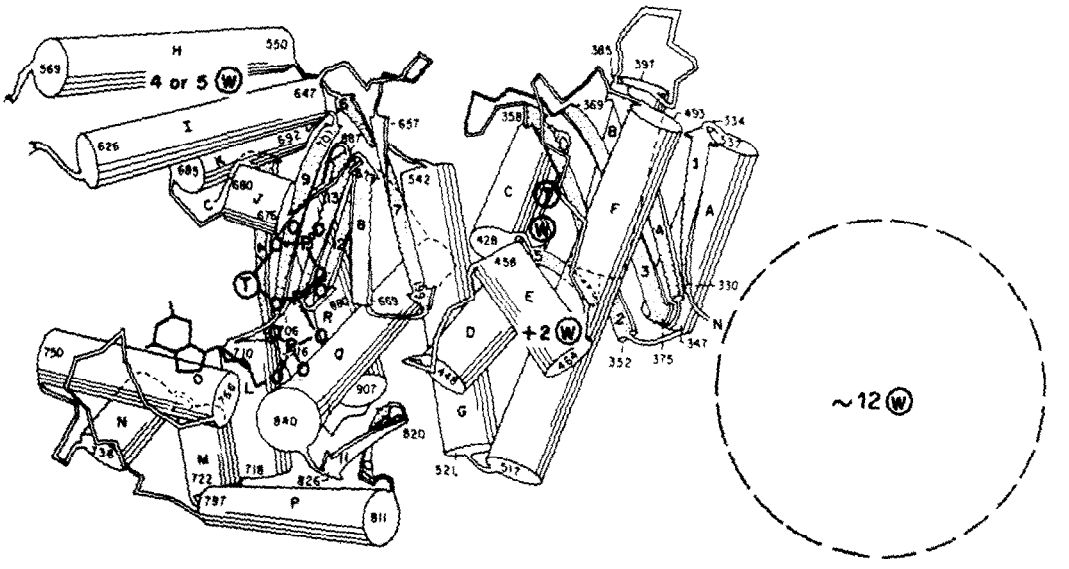

Fig. 8. Structure of Pol I [2] showing the approximate location of the dNTP binding site based on transferred NOEs [9–11], photoaffinity and affinity labeling [32, 33] and on the substrate binding properties of a peptide consisting of residues 728 to 777 [15]. Also shown are the approximate locations of tight metal binding sites Ⓣ, ($K_D \leq 10\,\mu M$) and weak metal binding sites Ⓦ, ($K_D \sim 1\,mM$), based on Mn^{2+} binding studies [34] and on the X-ray structure and mutational studies of the exonuclease site [3, 7].

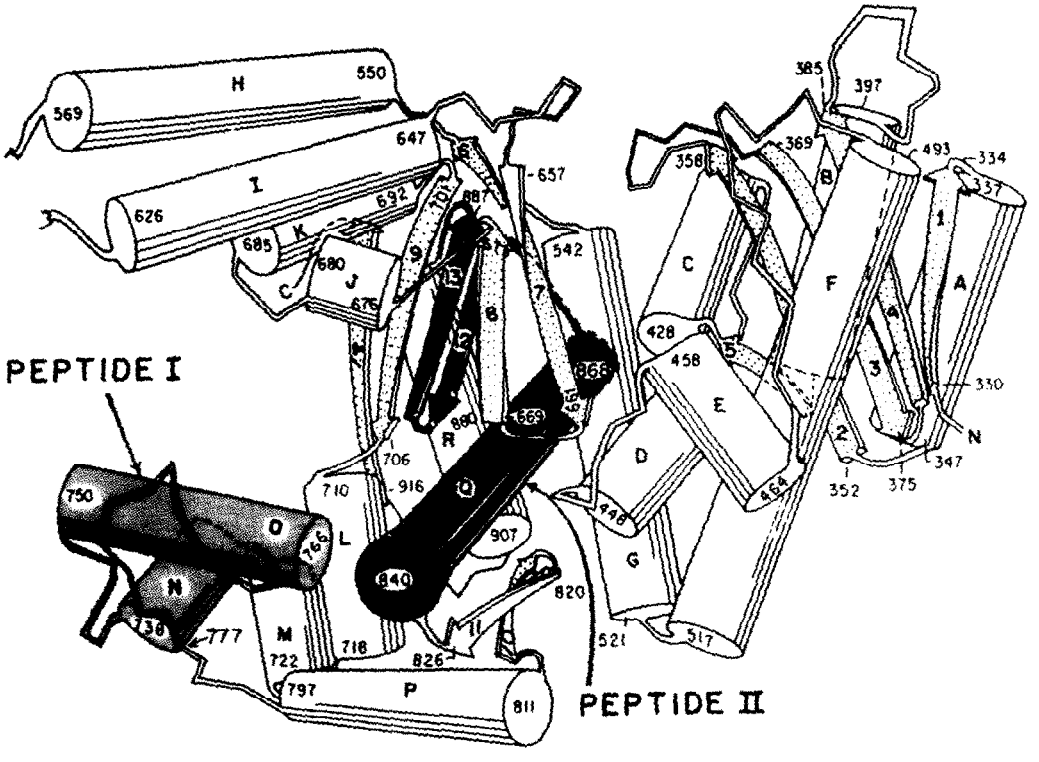

Fig. 9. Structure of the large fragment of Pol I modified from [2] to show the regions corresponding to peptides I and II.

Table 3. Secondary structures (%) of peptides I and II and comparison with the corresponding regions in the X-ray structure of the large fragment of Pol I*

	α-Helix	β-Structure	Turns	Coil
Enzyme (residues 728–777)	48	0	24	28
Peptide I (H_2O, 24°)	37	18	0	45
Peptide I (50% MeOH, 2°)	29	30	12	29
Peptide I + TNP-ATP (50% MeOH, 2°)	29	24	15	32
Enzyme (residues 840–888)	59	21	8	12
Peptide II (H_2O, 24°)	46	27	0	27
Peptide II (50% MeOH, 3°)	26	27	14	33

* The secondary structures of peptides I and II were estimated by fitting theoretical spectra to the CD spectra using the basis set of Stone *et al.* [36]. The secondary structures of the corresponding regions of the enzyme are from a computer graphics study, using the coordinates of the α-carbon atoms in the X-ray structure [2]. The uncertainties in the contributions are ≤6% in H_2O and <10% in 50% methanol.

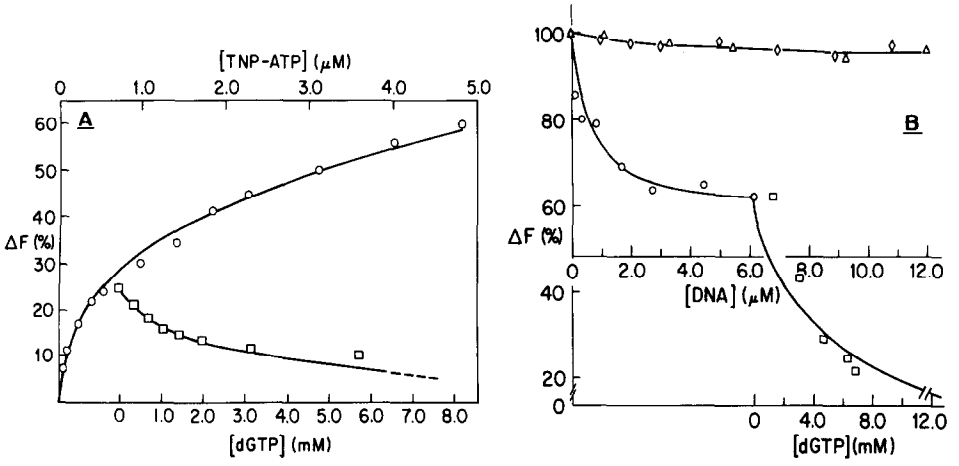

Fig. 10. Titrations of peptide I with substrates and DNA using 2′,3′-trinitrophenyl-ATP (TNP-ATP) as a fluorescent probe. (A) Fluorescence titration of peptide I (0.6 μM) with TNP-ATP (○) and displacement by dGTP (□). The solid curves were calculated assuming biphasic binding of TNP-ATP ($K_1 = 0.09$ μM, $n_1 = 0.16$; $K_2 = 5.0$ μM, $n_2 = 0.84$) and simple competitive displacement by dGTP (K_D = 336 μM). (B) Fluorescence titrations of peptide I-TNP-ATP (each 2.8 μM) with double-stranded DNA ($p(dA)_{16} \cdot p(dT)_{12}$) (○) ($K_D = 0.35$ μM, $n = 1.0$), and with single-stranded DNA $p(dA)_{12}$ (◇) and $p(dT)_{12}$ (△). Displacement of TNP-ATP from peptide I–$p(dA)_{16} \cdot p(dT)_{12}$ by dGTP (□) ($K_D = 2.1$ mM). Conditions: 10 mM potassium piperazine-*N*′,*N*′-bis(2-ethanesulfonic acid) (K^+PIPES), pH 6.9, 17 mM KCl in 50% methanol at 3° [15].

conformation of enzyme-bound dGTP to a single B-like species with a high *anti* glycosidic torsional angle and an O1′ endo, C3′ endo deoxyribose pucker, presumably due to Watson–Crick base pairing [11]. The same simplification of the dGTP conformation was observed on binding of the template $(rU)_{43}$ to the enzyme, probably due to the formation of a G-U wobble base pair [10]. Despite this effect of $(rU)_{43}$, no significant misincorporation of dGTP opposite a poly(rU) template was observed in kinetic studies [10].

To permit the observation of more complete complexes containing primer as well as template, the chain terminating substrate 2′,3′-ddGTP was studied [11]. In the enzyme–ddGTP–Mg^{2+} complex, no less than three substrate conformations were required to fit the interproton distances. These consisted of a 30 ± 10% contribution from a low *anti*, C3′ endo conformer, a 30 ± 10% contribution from a high *anti*, O1′ endo, C2′ endo conformer, and a 40 ± 10% contribution from a *syn*, O1′ endo, C2′ endo conformer (Fig. 5, Table 1). In all of these conformations the torsional angle γ about the C4′–C5′ bond could also be determined due to the absolute assignment of the 5′ and 5″ proton resonances and was found to be 180°, differing from the values of 55 ± 20° found in both B- and A-DNA. On binding of the template $(rC)_{37}$, the conformation of bound ddGTP simplified to a single high *anti*, O1′ endo, C2′ endo species compatible with B-DNA, suggesting G-C base pairing. However, the γ angle remained at 180° (Fig. 5, Table 1). Model building of this ddGTP conformation into B-DNA (Fig. 6) reveals that the γ value of 180° would permit the α-phosphorus of the

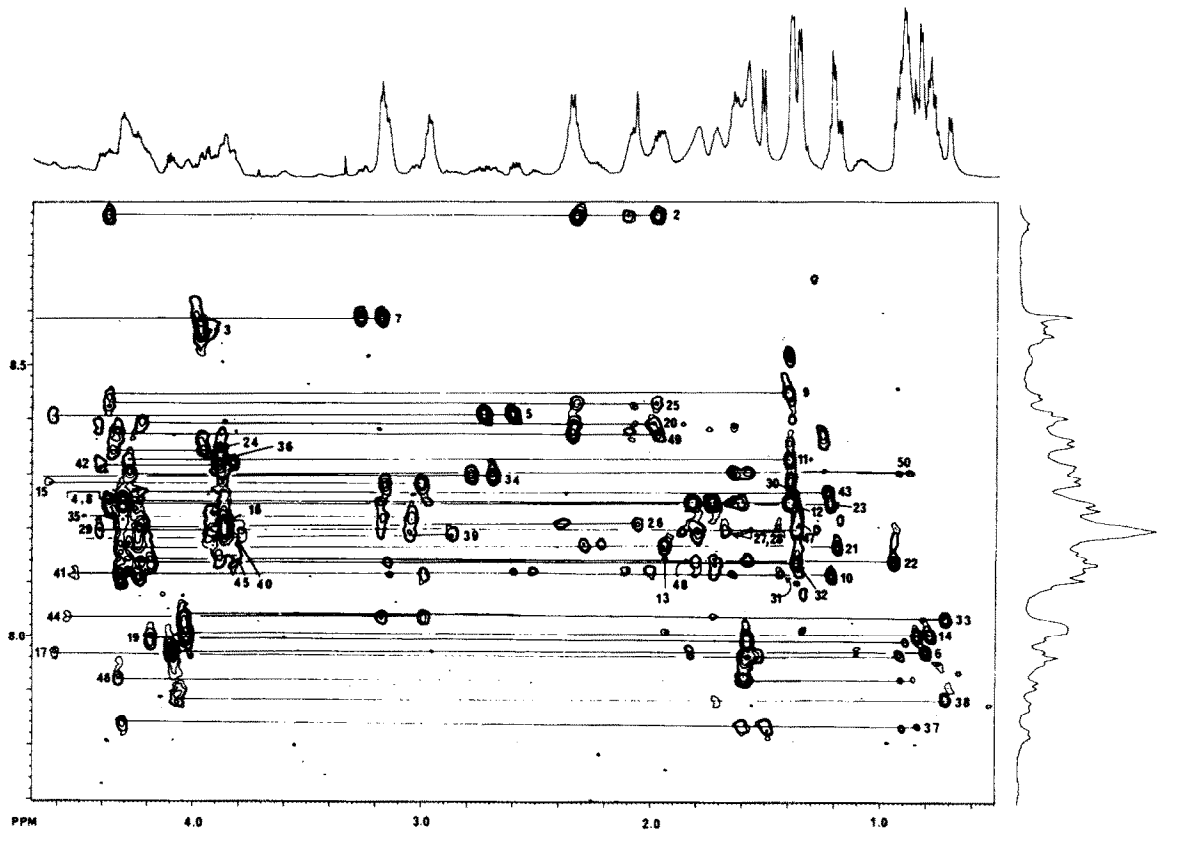

Fig. 11. TOCSY and one-dimensional spectra of the amide-aliphatic region of peptide I at pH 3.9 in H_2O containing 10% ^2H_2O at 27° acquired on a Bruker AM 600 spectrometer. For the TOCSY, the pulse sequence, presaturate-90°-t_1-MLEV-17-acquire [24], was employed using the decoupler amplifier (inverse mode) and fast switching between decoupler powers. Presaturation was for 0.7 sec at 19 dB below the low power output. The high power output with 10 dB attenuation was used in the preparation pulse, and in the MLEV-17 sequence, which had a total mixing time of 69.3 msec. The TPPI method [23] was used to yield a phase sensitive spectrum with a resolution of 8K in F2 and 4K in F1 which was obtained from 800 experiments of 4K time domain data points.

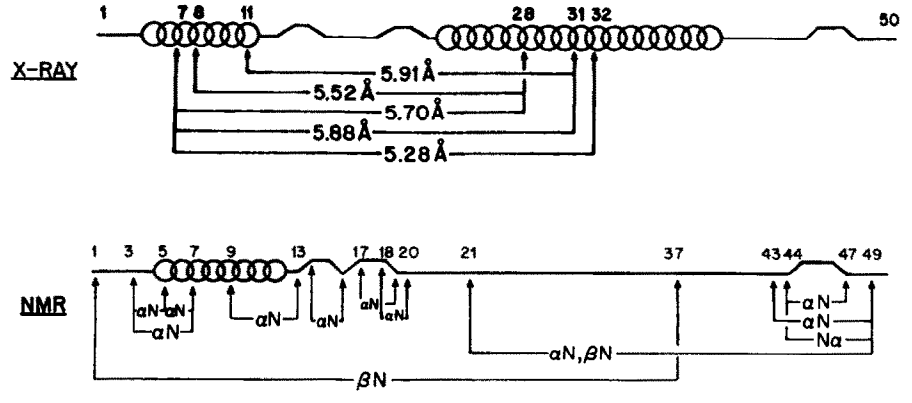

Fig. 12. Comparison of the secondary structure of peptide I in solution, as determined by 2D NMR studies at 600 MHz using the sequential assignment procedure [25], with the secondary structure of the corresponding 50 residues (728 to 777) in the crystal structure of the large fragment of Pol I [2]. Symbols: (ⒸⒸⒸⒸ) helix; (—⁀—) β-turn; and (———) coil. Also compared are tertiary interactions in peptide I in solution, as detected by the indicated remote NOEs, with tertiary interactions in residues 728 to 777 in the crystal structure [2]. The distances shown in the crystal structure are between α-carbon atoms which are less than 6 Å.

substrate to approach the 3′-OH group of the primer terminus, and the leaving β,γ pyrophosphate group to depart with inversion, suggesting this to be the active substrate conformation. In the complete, chain terminated complex, enzyme-$(rC)_{37}$-$(rI)_{14}$-ddG-ddGTP-Mg^{2+}, the bound ddGTP substrate reverts to multiple conformations only 30% of which are the active species. This ability of the enzyme to decrease the fraction of bound substrate which is appropriate for primer elongation may be an error-preventing mechanism. Such an error-preventing mechanism is necessary, since the complementary substrate ddGTP can easily be displaced from the complete system by the non-complementary substrate TTP. This rules out binding discrimination in the initial complex as the basis for the high fidelity of template copying by Pol I [10–12].

Conformations of enzyme-bound ribonucleotide templates

In these experiments RNA templates were used to avoid hydrolysis by the 3′,5′-exonuclease domain of the large fragment. Although the catalytic activity is significantly lower with RNA templates, the fidelity of replication remains high and comparable to that found with DNA templates [10, 29].

Interproton distances in the enzyme-bound RNA templates $(rU)_{54}$, $(rA)_{50}$, and $(rC)_{37}$ were determined [10, 11]. In all cases, the conformation of the average nucleotide of each template was *anti*, ruling out the Z-conformation (Table 2). With $(rU)_{54}$ and $(rC)_{37}$ the average ribose pucker could also be determined and was found to be O1′ endo, C2′ endo consistent with a B-like conformation (Table 2). Hence the polymerase site has a strong preference for the B-conformation, since RNA templates are so held when bound.

Amino acid environment of the enzyme-bound dNTP substrates

NOE action spectra, exemplified in Fig. 7, reveal the proximity of hydrophobic and aromatic amino acid residues to the H2, H8 and H1′ protons of the enzyme-bound substrate dATP, and the substrate analog AMPCPP [9, 10]. Similar action spectra were found with the substrates TTP and ddGTP [9, 11]. With ddGTP in H_2O intermolecular NOEs to the exchangeable 6-NH_2 protons were also detected [11]. The action spectra could be explained most simply by assuming the presence of two hydrophobic amino acids, one of which is Ile, and one aromatic amino acid, most likely Tyr, at the binding site of the dNTP substrate. The NMR data were consistent with photoaffinity labeling studies which revealed that 8-azido-d-ATP modified Tyr-766 which was located in the conserved amino acid sequence L-I-Y-G [32]. Other studies showed that Lys-758 was modified by pyridoxal phosphate acting as an affinity label [33]. The NMR and affinity labeling studied strongly suggested to us that the binding site for dNTP substrates might lie in part along the O-helix in the X-ray structure (Fig. 8) [2], although the complete X-ray coordinates are not yet available for detailed NMR docking studies. Figure 8 also shows the locations of tight and weak Mn^{2+}-binding sites on Pol I as determined by metal binding studies to the complete enzyme, the large fragment, and the D424A, and D355A + E357A exonuclease deficient mutants of the large fragment [34].

Synthesis and properties of active site peptides: substrate and DNA-binding studies

To test more directly whether the dNTP substrate might bind along the O-helix, a 50-residue peptide, peptide I, based on amino acids 728 to 777 of the Pol I sequence was synthesized. Peptide I was designed to include not only the entire O-helix, but also the interacting N-helix for structural support (Fig. 9) [15]. Additionally, another 49-residue peptide, peptide II, based on residues 840 to 888 of the Pol I sequence was synthesized [15], since His 881 in the homologous sequence V-H-D-E was modified by photoaffinity labeling with TTP [35]. Peptide II included the Q helix and β strands 12 and 13 in the X-ray structure (Fig. 9).

Based on CD studies, both peptides I and II retained significant secondary structure when compared with the corresponding regions of the crystalline enzyme (Table 3). However, only peptide I was found to bind substrates, as revealed by fluorescence titrations using the substrate analog TNP-ATP [15]. In aqueous buffers, peptide I was found to bind TNP-ATP tightly ($K_D = 0.03\ \mu M$) with a dissociation constant comparable to that obtained with the enzyme ($K_D = 0.07\ \mu M$). However, the stoichiometry of TNP-ATP binding by peptide I was low (n = 0.2 nucleotides/peptide) probably due to partial aggregation of the peptide [15]. When the pH values of solutions of peptide I were lowered from 7.5 to 3.9 to minimize isoelectric precipitation, and for consistency with subsequent 2D NMR studies, direct binding studies with dATP itself became possible, since protonation of the N-1 of dATP ($pK_a = 4.8$) induces a strong fluorescence of this nucleotide. A titration of dATP (8.6 μM) with peptide I revealed a 19% quenching of dATP fluorescence, a high binding affinity ($K_D = 0.5\ \mu M$) which was an order of magnitude tighter than that of the enzyme at pH 7.5, and a 1:1 peptide I–dATP complex.

At pH 7.5, in 50% aqueous methanol, a solvent better approximating a protein environment, the stoichiometry for TNP-ATP binding to peptide I increased to a value approaching 1.0 (Fig. 10A). However, the binding of TNP-ATP was biphasic with 16% of the peptide showing a high affinity ($K_D = 0.09\ \mu M$) and the remainder (84%) showing a lower affinity ($K_D = 5\ \mu M$) (Fig. 10A). Size exclusion HPLC studies under these conditions indicated that these two forms of the peptide are a monomer and a dimer respectively. Displacement of TNP-ATP from peptide I with the authentic DNA polymerase substrates dATP, TTP and dGTP yielded classical titration curves and K_D values ranging from 231 to 570 μM (Fig. 10A). These K_D values, although low, reflected binding which was two orders of magnitude weaker than that found with the enzyme, in H_2O under otherwise similar conditions ($K_D = 5.7\ \mu M$).

Most interestingly, the peptide I–TNP-ATP complex tightly and stoichiometrically binds duplex DNA ($K_D = 0.35\ \mu M$) to form a ternary complex much like the enzyme does, albeit with 100-fold lower affinity (Fig. 10B). Similarly, the peptide I–duplex DNA

complex stoichiometrically binds TNP-ATP, as well as authentic dNTP substrates in competition with TNP-ATP. The ability of peptide I, which represents only 8% of the Klenow fragment of Pol I, to bind substrates and duplex DNA indicates that residues 728 to 777 constitute a major portion of the substrate binding site of Pol I (Fig. 8).

2D *NMR studies of the solution structure of peptide I*

Because of its interesting substrate and DNA binding properties, and because significant secondary structure was retained in solution as judged by CD measurements (Table 3), peptide I was also studied by 2D NMR spectroscopy to locate the precise regions of secondary and tertiary structure. The NMR studies were carried out at pH 3.9 in H_2O and ^2H_2O, conditions under which the peptide tightly binds dATP, and from its high resolution NMR spectrum was probably monomeric. Double quantum filtered COSY and TOCSY spectra at 600 MHz were used to assign resonances to amino acid types. The data are exemplified by the TOCSY spectrum (Fig. 11). NOESY spectra were used to make sequence specific assignments. A preliminary comparison of the linear secondary structure of peptide I with that of residues 728 to 777 of the crystalline enzyme (Fig. 12) reveals a flexible helix from residues 4 to 12 closely corresponding to the N-helix of the crystalline enzyme. Thus, while the X-ray structure shows 48% helix, the sequential NOESY analysis suggests 18% helix. Two of the three β-turns found in the crystalline state were preserved as Type II and I β-turns in solution (Fig. 12). Further, as revealed by long range NOEs, the tertiary fold of peptide I in solution differs significantly from that of the corresponding residues of the crystalline enzyme (Fig. 12), undoubtedly due to the constraints provided by the remainder of the protein. We are currently refining the solution conformation of peptide I in order to learn its structure in greater detail as well as the structure of its complexes with DNA and substrates.

CONCLUSIONS

(1) Bound dNTP substrates and RNA templates are held in B-like conformations on Pol I. With dGTP and ddGTP, B-conformations are found only in the presence of template. In the absence of template, multiple conformations of enzyme-bound guanine nucleoside triphosphates are detected.

(2) A thermodynamically and kinetically unfavorable step following substrate binding is required to properly position and activate the substrate for nucleophilic attack, and to increase fidelity in template copying by Pol I.

(3) Substrates of the polymerase reaction are bound near two hydrophobic amino acids and an aromatic amino acid, probably Leu-764, Ile-765, and Tyr-766.

(4) A synthetic 50-residue peptide based on residues 728 to 777 of the Pol I sequence retains significant secondary and tertiary structure in solution and shows tight binding of dNTP substrates, the substrate analog TNP-ATP, and duplex DNA, confirming the location of the dNTP binding site as being, in part, along the O-helix.

(5) A synthetic 49-residue peptide based on residues 840 to 888 of the Pol I sequence also retains much secondary structure but does not bind the substrate analog TNP-ATP.

Acknowledgements—We are grateful to Dr. C. M. Joyce for providing us with strains of *E. coli* which overproduce the large fragment of Pol I, to Dr. T. A. Steitz for permitting us to make use of his figures of the X-ray structure of this enzyme and to Dr. G. M. Clore for valuable advice on the jump-return method. This work was supported in part by Grants DK28616 from the National Institutes of Health and NP-645 from the American Cancer Society to A. S. M. The 600 MHz NMR Spectrometer was funded by NIH Grant RR03518, NSF Grant DMB8612318, and the Johns Hopkins School of Medicine. G. P. M. holds N.I.H. Post-Doctoral Fellowship GM12058.

REFERENCES

1. Kornberg A, *DNA Replication*. WH Freeman, San Francisco, 1980.
2. Ollis DL, Brick P, Hamlin KM, Xuong NG and Steitz TA, Structure of the large fragment of *E. coli* DNA polymerase I complexed with TMP. *Nature* **313**: 762–766, 1985.
3. Freemont PS, Ollis DL, Steitz TA and Joyce CM, A domain of the Klenow fragment of *E. coli* DNA polymerase I has polymerase but no exonuclease activity. *Proteins: Structure Function Genetics* **1**: 66–73, 1986.
4. Steitz RA, Beese L, Freemont PS, Friedman JM and Sanderson MR, Structural studies of Klenow fragment: An enzyme with two active sites. In: *Cold Spring Harbor Symp Quant Biol* **LII**: 465–471, 1987.
5. Joyce CM and Grindley NDF, Construction of a plasmid that overproduces the large proteolytic fragment (Klenow fragment) of DNA polymerase I. *Proc Natl Acad Sci USA* **80**: 1830–1834, 1983.
6. Joyce CM, Kelley WS and Grindley NDF, Nucleotide sequence of *E. coli* pol A gene and primary structure of DNA polymerase I. *J Biol Chem* **257**: 1958–1964, 1982.
7. Joyce CM and Steitz TA, DNA polymerase I: From crystal structure to function via genetics. *Trends Biol Sci* **12**: 288–292, 1987.
8. Sloan DL, Loeb LA, Mildvan AS and Feldmann RJ, Conformation of dNTP substrates on DNA polymerase I from *E. coli* as determined by nuclear magnetic relaxation. *J Biol Chem* **250**: 8913–8920, 1975.
9. Ferrin LJ and Mildvan AS, NOE studies of the conformations and binding site environments of dNTP substrates bound to DNA polymerase I and its large fragment. *Biochemistry* **24**: 6904–6913, 1985.
10. Ferrin LJ and Mildvan AS, NMR studies of conformations and interactions of substrates and ribonucleotide templates bound to the large fragment of DNA polymerase I. *Biochemistry* **25**: 5131–5145, 1986.
11. Mullen GP and Mildvan AS, NOE studies of the conformations of dGTP and ddGTP bound to the large fragment of DNA polymerase I. *FASEB J* **2**: A588, 1988.
12. Fry DC, Kuby SA and Mildvan AS, NMR studies of the MgATP binding site of adenylate kinase and of a 45-residue peptide fragment of the enzyme. *Biochemistry* **24**: 4680–4694, 1985.
13. Kuliopulos A, Westbrook EM, Talalay P and Mildvan AS, Positioning of a spin-labeled substrate analog into the structure of Δ^5-3-ketosteroid isomerase by combined kinetic, magnetic resonance and X-ray diffraction methods. *Biochemistry* **26**: 3927–3937, 1987.

14. Mildvan AS, NMR studies of the interactions of substrates with enzymes and their peptide fragments. *FASEB J* **3**: 1705–1714, 1989.
15. Mullen GP, Shenbagamurthi P and Mildvan AS, Substrate and DNA binding to a 50-residue peptide fragment of DNA polymerase I. Comparison with the enzyme. *J Biol Chem* **264**: 19637–19647, 1989.
16. Mildvan AS and Gupta RK, Nuclear relaxation measurements of the geometry of enzyme-bound substrates and analogs. *Methods Enzymol* **49G**: 322–359, 1978.
17. Mildvan AS, Rosevear PR, Granot J, O'Brian C, Bramson HN and Kaiser ET, Use of NMR and EPR to study cAMP-dependent protein kinase. *Methods Enzymol* **99**: 93–119, 1983.
18. Rosevear PR and Mildvan AS, Ligand conformations and ligand–enzyme interactions as studied by the nuclear Overhauser effect. *Methods Enzymol* **177**: 333–358, 1989.
19. Rosevear PR, Bramson HN, O'Brian C, Kaiser ET and Mildvan AS, Nuclear Overhauser effect studies of the conformations of tetraamminecobalt(III)-adenosine 5′-triphosphate free and bound to bovine heart protein kinase. *Biochemistry* **22**: 3439–3447, 1983.
20. deLeeuw HPM, Hasnoot CAG and Altona C, Empirical correlations between conformational parameters in β-D-furanoside fragments derived from a statistical survey of crystal structure of nucleic acid constituents. *Isr J Chem* **20**: 108–126, 1980.
21. Levitt M and Warshel A, Extreme conformational flexibility of the furanose ring in DNA and RNA. *J Am Chem Soc* **100**: 2607–2613, 1978.
22. Merrifield RB, Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J Am Chem Soc* **85**: 2149–2154, 1963.
23. Marion D and Wüthrich K, Applications of phase sensitive two dimensional correlated spectroscopy (COSY) for measurements of ^1H–^1H spin–spin coupling constants in proteins. *Biochem Biophys Res Commun* **113**: 967–974, 1983.
24. Bax A and Davis DG, MLEV-17-based two dimensional homonuclear magnetization transfer spectroscopy. *J Magn Reson* **65**: 355–360, 1985.
25. Wider G, Macura S, Kumar A, Ernst RR and Wüthrich K, Homonuclear two-dimensional ^1H NMR of proteins. Experimental procedures. *J Magn Reson* **56**: 207–234, 1984.
26. Plateau P and Gueron M, Exchangeable proton NMR without baseline distortion using new strong-pulse sequences. *J Am Chem Soc* **104**: 7310–7311, 1982.
27. Driscoll PC, Clore GM, Beress L and Gronenborn AM, A proton nuclear magnetic resonance study of the antihypertensive and antiviral protein BDS-1 from the sea anemone *Anemonia sulcata:* Sequential and stereospecific resonance assignment and secondary structure. *Biochemistry* **28**: 2178–2187, 1989.
28. Englund PT, Huberman JA, Jovin TM and Kornberg A, Enzymatic synthesis of DNA. Binding of triphosphates to DNA polymerase. *J Biol Chem* **244**: 3038–3044, 1969.
29. Loeb LA and Kunkel TA, Fidelity of DNA synthesis. *Annu Rev Biochem* **51**: 429–457, 1982.
30. Kunkel TA, Eckstein F, Mildvan AS, Koplitz RM and Loeb LA, Deoxynucleoside 1-thiotriphosphates prevent proofreading during *in vitro* DNA synthesis. *Proc Natl Acad Sci USA* **78**: 6734–6738, 1981.
31. Kuchta RD, Benkovic P and Benkovic SJ, Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. *Biochemistry* **27**: 6716–6725, 1988.
32. Joyce CM, Ollis DL, Rush J, Steitz TA, Konigsberg WH and Grindley NDF, Relating structure to function for DNA polymerase I of *E. coli*. In: *Protein Structure, Function and Design, UCLA Symposia on Molecular and Cellular Biology* (Ed. Oxender D), Vol. 32, pp. 197–205. Alan R. Liss, New York, 1985.
33. Basu A and Modak MJ, Identification and amino acid sequence of the dNTP binding site in *E. coli* DNA polymerase I. *Biochemistry* **26**: 1704–1709, 1987.
34. Mullen GP, Serpersu EH, Ferrin LJ, Joyce CM, Derbyshire V, Loeb LA and Mildvan AS, The binding of Mn^{2+} to DNA polymerase I and to the wild type and exonuclease deficient mutants of the Klenow fragment. *J Cell Biol* **107**: 402A, 1988.
35. Pandey VN, Williams KR, Stone KL and Modak M, Photoaffinity labeling of the TTP binding domain in *E. coli* DNA polymerase I: Identification of histidine-881 as the site of cross-linking. *Biochemistry* **26**: 7744–7748, 1987.
36. Stone AL, Park JY and Martenson RE, Low ultraviolet circular dichroism spectroscopy of oligopeptides 1–95 and 96–168 derived from myelin basic protein of rabbit. *Biochemistry* **24**: 6666–6673, 1985.