4406–4407 Nucleic Acids Research, 1993, Vol. 21, No. 18

Conserved sites in the 5′-3′ exonuclease domain of *Escherichia coli* DNA polymerase

Pablo D.Gutman and Kenneth W.Minton*
Department of Pathology, F.E.Hébert School of Medicine, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814-4799, USA

Received April 27, 1993; Revised and Accepted July 27, 1993 GenBank accession no. L14581

The middle and C-terminal domains of *Escherichia coli* DNA polymerase I exhibit 3′-5′ exonuclease and DNA polymerase activity, respectively, and are both present in the large (Klenow) proteolytic fragment of this enzyme. The structural basis for these two activities has been extensively investigated using crystallographic studies and site specific mutagenesis of the Klenow fragment (rev. in 1). Little is known, however, about the structural basis for the 5′-3′ exonuclease that is characteristic of the N-terminal domain, and exhibited by the small proteolytic fragment of DNA PolI (N-terminal 323 amino acids). Sayers and Eckstein (2) compared this region with three other homologous 5′-3′ exonucleases: the phage T5 D15 exonuclease and the N-terminal domains of *Thermus aquaticus* and *Streptococcus pneumoniae* DNA Pol. Their comparison identified four conserved regions (2). The genes of three additional bacterial DNA PolI-like enzymes have now been sequenced (3–5) including one from this laboratory (*Deinococcus radiodurans*), and it has been appreciated that three additional bacteriophage have homologous exonucleases (6–8), permitting comparison of the *E.coli* DNA PolI N-terminal domain with nine related exonucleases with sufficient diversity that highly conserved regions can be better discerned. In addition, application of a computer-assisted technique for multiple sequence alignment (9) has allowed us to obtain a total alignment of the ten exonuclease sequences (Figure 1). Thus, the four regions previously detected (2) are now substantially refined (motifs A, D, E and F), and two additional conserved domains are identified (motifs B and C), one of which is supported by mutational data (Figure 1).

Within the 323 a.a. N-terminal proteolytic fragment of DNA PolI, the residues required for DNA PolI 5′-3′ exonuclease activity are within the first 297 amino acids, since the *E.coli resA1* mutation [Gln(298)→Amber] yields an N-terminal 297 a.a. polypeptide fragment that has 5′-3′ exonuclease activity (10). The region required for exonuclease activity might be even smaller, since none of the six conserved regions we have detected are beyond the first ~215 N-terminal amino acids (Figure 1). The six motifs include all 14 a.a. identities present (same residue aligned in all ten sequences). The importance of two of the six motifs is supported by the location of three of four sequenced mutations (10) in *E.coli polA* that result in defective 5′-3′ exonuclease activity: *polA107* is a substitution at an identity in motif C, while *polA480ex* and *polA214* are substitutions at two different identities within motif F (Figure 1). A fourth mutation that causes defective 5′-3′ exonucleolytic activity, *polA4113* produces a change of Gly[103] to Glu (10). This residue does not fall within a motif, but is present in five of the six bacterial polymerases and absent in the four T-phage exonucleases. This latter mutation results in a DNA PolI that is thermolabile for both 5′-3′ exonuclease and polymerase activities (10). Consequently, it may play a role in the interaction of the 5′-3′ exonuclease and polymerase domains, and accordingly might not be necessary in the T-phage 5′-3′ exonucleases, since they do not have covalently associated polymerase activities.

ACKNOWLEDGEMENTS

We are grateful to Catherine M.Joyce for invaluable discussions and careful review of this manuscript. This work was supported by USPHS grant GM39933.

REFERENCES

1. Joyce,C.M. (1991) *Curr. Opin. Struct. Biol.* **1**, 123–129.
2. Sayers,J. R. and Eckstein,F. (1991) *Nucleic Acids Res.* **19**, 4127–4132.
3. Akhmetzjanov,A. and Vakhitov,V. (1992) *Nucleic Acids Res.* **20**, 5839.
4. Uemori,T., Ishino,Y., Fujita,K., Asada,K. and Kato,I. (1993) *J. Biochem. (Tokyo)* **113**, 401–410.
5. Gutman,P.D., Fuchs,P., Ouyang,L. and Minton,K.W. (1993) *J. Bacteriol.* **175**, 3581–3590.
6. Lopez,P., Martinez,S., Diaz,A., Espinoza,M. and Lacks,S.A. (1989) *J. Biol. Chem.* **264**, 4255–4263.
7. Beck,P., Gonzalez,S., Ward,C. and Molineux,I. (1989) *J. Mol. Biol.* **210**, 687–701.
8. Hollingsworth,H.C. and Nossal,N.G. (1991) *J. Biol. Chem.* **266**, 1888–1897.
9. Schuler,G.D., Altschul,S.F. and Lipman,D.J. (1991) *Proteins: Structure, Function, and Genetics* **8**, 180–190.
10. Joyce,C.M., Fujii,D.M., Laks,H.S., Hughes,C.M. and Grindley,N.D.F. (1985) *J. Mol. Biol.* **186**, 283–293.
11. Joyce,C.M., Kelley,W.S. and Grindley,N.D.F. (1982) *J. Biol. Chem.* **257**, 1958–1964.
12. Lawyer,F.C., Stoffel,S., Salk,R.K., Myambo,K., Drummond,R. and Gelfand,D.H. (1989) *J. Biol. Chem.* **264**, 6427–6437.
13. Kaliman,A.V., Krutilina,A.I., Kryukov,V.M. and Bayev,A.A. (1986) *FEBS Lett.* **195**, 61–64.
14. Dunn,J.J. and Studier,F.W. (1983) *J. Mol. Biol.* **166**, 477–535.

* To whom correspondence should be addressed

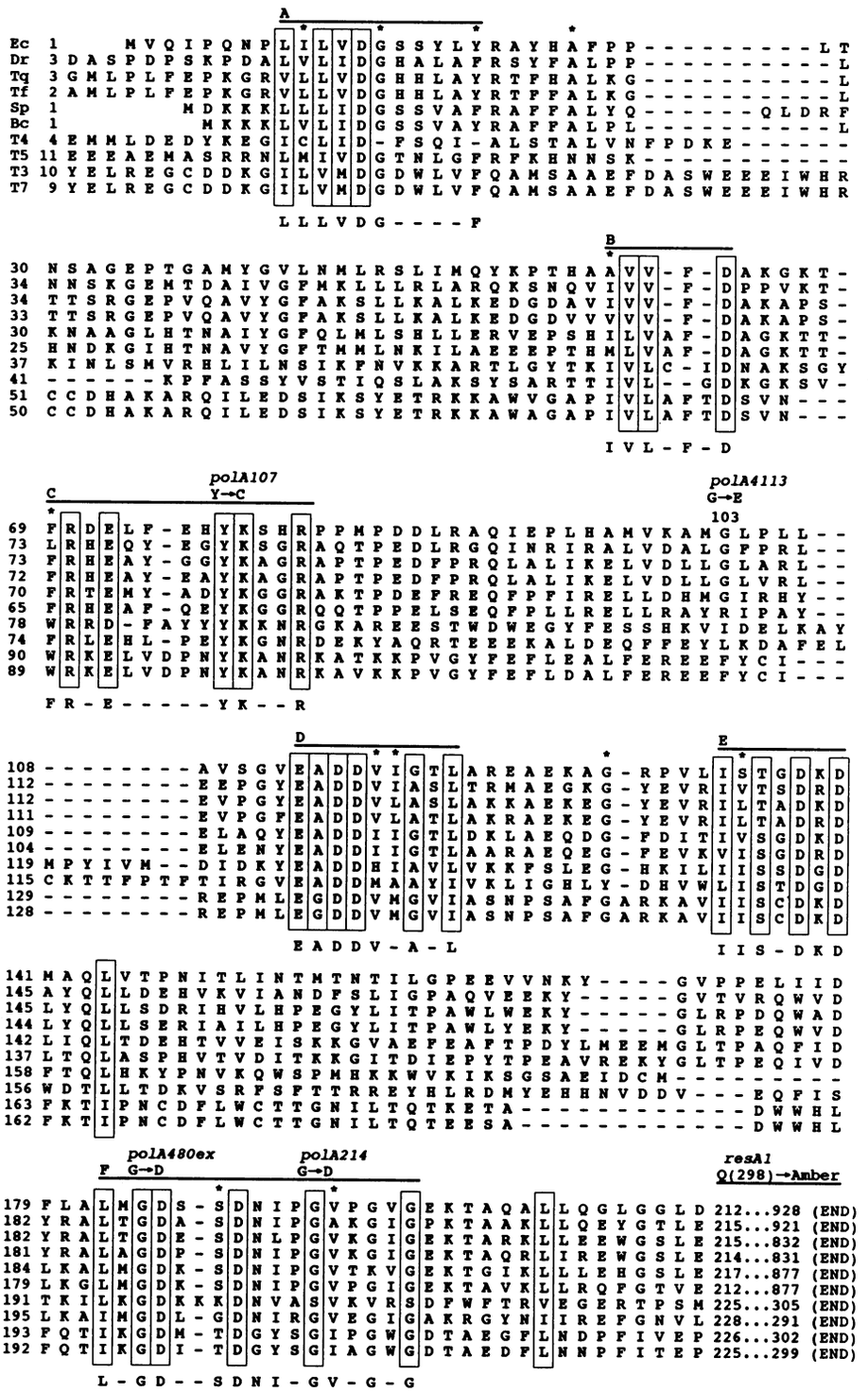

Figure 1. Conserved motifs in the 5′-3′ exonuclease in PolI family members and related phage exonucleases. Complete alignment of the N-terminal domain of DNA polymerases and homologous bacteriophage 5′-3′ exonucleases. Six conserved regions are identified, indicated by overlining, labeled A–F; the consensus sequences are indicated below these regions. Boxed aligned residues indicate identity or homology of all ten aligned residues. Asterisks indicate similarity or identity of all but one aligned residue. The following similarities were employed: acidic (D + E); basic (H + K + R); large aliphatic (I + L + M + V); aromatic (F + W + Y); hydroxyl/small aliphatic (A + S + T + G). Above the alignment four mutations are indicated that have been characterized in *E.coli* DNA polymerase I (see text). Ec — *E.coli* DNA PolI (11); Dr — *Deinococcus radiodurans* DNA Pol (5); Tq — *Thermus aquaticus* DNA Pol (12); Tf — *Thermus flavus* DNA Pol (3); Sp — *Streptococcus pneumoniae* DNA PolI (6); Bc — *Bacillus caldotenax* DNA Pol (4); T5 — phage T5 D15 exonuclease (13); T4 — phage T4 RNAse H/5′-3′ DNA exonuclease (8); T3 — phage T3 gene 6 exonuclease (7); and T7 — phage T7 gene 6 exonuclease (14). The numbering of the phage T7 gene 6 exonuclease is not that originally reported and has been corrected in accordance with subsequent findings (see 'Note in Proof' in 6). Data base searching was performed through the National Center for Biotechnology Information, Bethesda, MD. Microcomputer-assistance in alignment of a.a.'s employed the MACAW program (9).