Extremophiles (1998) 2:171–177

REVIEW

Francesca M. Pisani · Mariarita De Felice
Giuseppe Manco · Mosè Rossi

Domain organization and biochemical features of *Sulfolobus solfataricus* DNA polymerase

Received: January 22, 1998 / Accepted: February 16, 1998

Abstract DNA polymerase from *Sulfolobus solfataricus*, strain MT4 (Sso DNA pol), was one of the first archaeal DNA polymerases to be isolated and characterized. Its encoding gene was cloned and sequenced, indicating that Sso DNA pol belongs to family B of DNA polymerases. By limited proteolysis experiments carried out on the recombinant homogeneous protein, we were able to demonstrate that the enzyme has a modular organization of its associated catalytic functions (DNA polymerase and 3′-5′ exonuclease). Indeed, the synthetic function was ascribed to the enzyme C-terminal portion, whereas the N-terminal half was found to be responsible for the exonucleolytic activity. In addition, partial proteolysis studies were utilized to map conformational changes on DNA binding by comparing the cleavage map in the absence or presence of nucleic acid ligands. This analysis allowed us to identify two segments of the Sso DNA pol amino acid chain affected by structural modifications following nucleic acid binding: region 1 and region 2, in the middle and at the C-terminal end of the protein chain, respectively. Site-directed mutagenesis studies will be performed to better investigate the role of these two protein segments in DNA substrate interaction.

Key words *Sulfolobus solfataricus* · DNA replication · DNA polymerase · 3′-5′ Exonuclease · Modular organization

Introduction

Much information is available on gene transcription and translation in Archaea (Dennis 1997; Olsen and Woese 1997; Reeve et al. 1997), but very little is known about the mechanisms of DNA replication and repair. In particular, discovering how genomes can be duplicated and maintained at temperatures near or above 100°C is a great and fascinating challenge for the microbiologist working on hyperthermophilic archaeal species.

Genome structure and gene organization in Archaea is eubacterial like (Forterre et al. 1992). The chromosome of Archaea is composed of a single circular DNA molecule of 1–3 $\times$ 10^6 bp. Genes are usually tightly packed (when not overlapping), often linked in operons that sometimes share with the eubacterial counterparts an identical segmental order of their open reading frames (ORFs). Because it is likely that the structural and functional organization of a genome coevolved with the biochemical apparatus committed to its duplication and maintainance, one would expect Archaea to resemble Bacteria in terms of replication proteins. However, the recent sequencing of the *Methanococcus jannaschii* genome has revealed the eukaryotic-like features of the archaeal replication machinery (Bult et al. 1996). Undoubtedly, DNA polymerases are the most investigated biochemical component of archaebacterial DNA duplication and repair apparatus. In particular, DNA polymerases from hyperthermophilic archaeal species have attracted a special interest in view of their potential use as a tool to improve the in vitro techniques of DNA manipulation, such as sequencing reactions and enzymatic amplification of target sequences at elevated temperatures (Perler et al. 1996). This review focuses on one representative of these important group of enzymes, DNA polymerase from the thermoacidophilic archaeon *Sulfolobus solfataricus* (Sso DNA pol). A brief section on the classification of DNA polymerases and their structure–function relationships introduces this topic.

Communicated by K. Horikoshi

F. M. Pisani (✉) · M. De Felice · G. Manco · M. Rossi
Istituto di Biochimica delle Proteine ed Enzimologia, Consiglio Nazionale delle Ricerche, Via Marconi, 10. 80125-Napoli, Italy
Tel. +39-81-7257316; Fax +39-81-7257240
e-mail: pisani@dafne.ibpe.na.cnr.it

Classification of DNA polymerases and their structure–function relationships

On the basis of sequence similarities, DNA polymerases have been classified into families A, B, C, and X, which

include enzymes similar to *Escherichia coli* DNA pol I, II, and III, and eukaryotic DNA pol β, respectively (see Table 1; Braithwaite and Ito 1993).

DNA pol I from *E. coli* is the prototype of family A (Kornberg and Baker 1992). This enzyme is a monomer of 103 kDa with two associated exonuclease activities acting in the 5′-3′ and 3′-5′ direction, in addition to the synthetic function. The 3′-5′ exonuclease activity is responsible for the accuracy of DNA duplication ("proofreading" function), whereas the exonuclease that acts in the 5′-3′ direction coordinates excision of the RNA primers during replication and removal of damaged bases in DNA repair reactions. Family A includes exclusively DNA polymerases from eubacterial sources, either mesophilic (*Bacillus subtilis*, *Streptococcus pneumoniae*, T5 and T7 bacteriophages) or thermophilic (*Bacillus acidocaldarius*, *B. caldotenax*, *B. stearothermophilus*; *Thermus aquaticus*, *T. thermophilus*), the only exception being the eukaryotic mitochondrial DNA polymerase γ.

Family B is the most heterogeneous group, because it includes DNA polymerases from Eukarya (DNA polymerases α, δ, and ε and viral replicases), from various species of Archaea, and from Bacteria (*E. coli* DNA pol II, T4 and φ29 bacteriophages replicases). *E. coli* DNA pol II is a single polypeptide of 90 kDa that contains a 3′-5′ (but not a 5′-3′) exonuclease activity. The in vivo function of *E. coli* DNA pol II has not yet been established with certainity, but some genetic evidence suggests its involvement in repair reactions. Analogously, the role for each of the three eukaryotic DNA polymerases, all essential for replication, have not been defined. DNA pol α, which is associated to DNA primase and is able to use RNA primers, could be responsible for the initial elongation of nascent DNA chains. Synthesis would be then continued by DNA polymerase δ, associated with the processivity factor PCNA (proliferating-cell nuclear antigen) and replication factor C complex, or by the intrinsically processive DNA polymerase ε (Bambara et al. 1997).

Family C groups only eubacterial enzymes similar in their amino acidic sequence to *E. coli* DNA pol III, that, similarly to this latter, are involved in genome duplication as catalytic components of multisubunit complexes acting at

Table 1. Classification of DNA polymerases

Group	Enzyme	Amino acids (no.)	Exonuclease	
			5′-3′	3′-5′
Family A	Eubacterial DNA pols			
	DNA pol I from *E. coli*[a]	928	+	+
	T. aquaticus[a]	832	+	−
	S. pneumoniae	877	+	−
	Bacteriophage DNA pols			
	DNA pol from T5	829	−	+
	T7	704	−	+
	Mitochondrial DNA pols			
	DNA pol γ from *S. cerevisiae*	1254	−	+
Family B	Eubacterial DNA pols			
	DNA pol II from *E. coli*	783	−	+
	Bacteriophage DNA pols			
	DNA pol from φ29	575	−	+
	T4	898	−	+
	RB69[a]	903	−	+
	Archaeal DNA pols			
	DNA pol I from *S. solfataricus*	882	−	+
	T. litoralis	774	−	+
	P. furiosus	775	−	+
	Eukaryal cell DNA pols			
	DNA pol α from *H. sapiens*	1462	−	−
	S. cerevisiae (I)	1468	−	+
	DNA pol δ from *H. sapiens*	1107	−	+
	S. cerevisiae (III)	1097	−	+
	DNA pol ε from *S. cerevisiae* (II)	2222	−	+
	Viral DNA pols			
	DNA pol from Herpes simplex-1	1235	+	+
	Vaccinia virus	937	−	+
	Adenovirus 2	1056	−	+
Family C	Eubacterial DNA pols			
	DNA pol III from *E. coli* (α subunit)	1160	−	−
	B. subtilis	1437	−	+
Family X	Eukaryal DNA pols			
	DNA pol β from *H. sapiens*[a]	335	−	−

[a] Indicates a protein whose crystallographic structure was solved. Note that for *E. coli* DNA pol I, only the structure of the Klenow fragment was determined.

the DNA replication fork (Maki et al. 1988). Family X includes eukaryotic DNA polymerase β, shown to be implicated in DNA repair reactions (Mullen and Wilson 1997). The three-dimensional structure of this enzyme, which is the smallest known polymerase having a mass of about 37 kDa and lacking a 3′-5′ exonuclease activity, was resolved (Pelletier et al. 1994; Davies et al. 1994).

As previously discussed, DNA polymerases are often multifunctional enzymes, having one or two exonuclease activities associated with the synthetic function. Site-directed mutagenesis and crystallographic studies have demonstrated that these activities are organized in a modular fashion with the degradative activities having a N-terminal location, whereas the synthetic function has been mapped within the C-terminal portion of the enzyme molecule (Bernad et al. 1989; Joyce and Steitz 1994).

The structures of the *E. coli* Klenow fragment (Ollis et al. 1985) and *T. aquaticus* (Kim et al. 1995; Korolev et al. 1995) family A DNA polymerases, and that of the HIV-1 reverse transcriptase (Kohlstaedt et al. 1992), show a C-terminal polymerase domain possessing the shape of a hand with "thumb," "palm," and "fingers" subdomains. The palm, which possesses a β-like structure, contains the catalytic pocket; the thumb is a flexible α-helical protrusion interacting with the product duplex DNA; and the fingers on the opposite site make contact with the substrate template strand.

Analysis of the structure–function relationships of family B DNA polymerases has been mainly based on site-specific mutagenesis studies of the most conserved amino acid residues and on comparison with the Klenow fragment, since no structural model was available for this class of enzymes (Blanco and Salas 1996). Alignments of family B DNA polymerases sequences identified five conserved motifs, located with the same segmental order in the C-terminal half of each polypeptide chain. Site-specific mutagenesis studies have implicated these sequence motifs in various critical functions, such as catalysis and substrate and metal ion binding, in various systems. The three-dimensional structure of the family B DNA polymerase from *E. coli* bacteriophage RB69 has been recently solved, providing an important structural framework for this biochemical analysis (Wang et al. 1997). This study has revealed that the C-terminal half of RB69 DNA pol possesses the characteristic hand shape, but, although the palm subdomain has the same topology of reverse transcriptase and the Klenow fragment, the fingers and thumb subdomains are unrelated to all other known polymerases structures. On the other hand, the N-terminal exonuclease domain of RB69 DNA pol is structurally similar to that of the Klenow fragment. This finding was anticipated by the proposal of Bernad et al. (1989) that the 3′-5′ exonuclease domain is evolutionarily conserved in all proofreading DNA polymerases, being composed of three N-terminal amino acid segments (Exo I, Exo II, and Exo III), which contain invariant residues involved in metal ion binding, substrate interaction, and catalysis (for review, see Derbyshire et al. 1995).

DNA polymerase from *Sulfolobus solfataricus* (Sso DNA pol) belongs to family B

DNA polymerase from *Sulfolobus solfataricus*, strain MT4, was one of the first archaeal DNA polymerases to be isolated and characterized (Rossi et al. 1986). This enzyme was shown to be the most abundant (80%–85%) of two chromatographically distinguishable DNA polymerase activities. For the enzyme purification, *Sulfolobus* cells were collected during the log phase of growth, when the polymerase activity peaks. The purification procedure, based on a heparin and various ion-exchange chromatographic steps, yielded a homogeneous enzyme, whose molecular mass was about 110 kDa, as determined by gel filtration and glycerol gradient centrifugation (Rella et al. 1990). Sodium dodecyl sulfate-gel (SDS-gel) electrophoresis of the native enzyme revealed a major 100-kDa protein band and minor components of 50–55 and 35–40 kDa. These polypeptides were thought to be proteolytic fragments of the 100-kDa species because they were found to be immunologically related, and by activity gel analysis, in addition to the 100-kDa protein, the 50-kDa polypeptide was also found to retain synthetic activity. Sso DNA pol utilizes preferentially activated natural DNAs (such as genomic DNA from calf thymus, *S. solfataricus*, *Micrococcus lysodeikticus*) as substrates in the presence of divalent cations (with Mg^{2+} being a better cofactor than Mn^{2+}). Sso DNA pol is stable for several hours at 65°–70°C and has a half-life of about 35 min at 85°C and about 6 min at 90°C. The enzyme synthetic function is highly thermophilic, with an optimal temperature that was shown to depend on the melting temperature of the DNA duplex used as template/primer more than on the intrinsic protein thermal stability. Using activated calf thymus DNA (which possesses a melting temperature of about 78°C in the assay conditions), Sso DNA pol displays a maximal activity at 75°C.

Inhibition studies indicated that native Sso DNA pol was inhibited by aphidicolin, although the concentration required to obtain 50% inhibition was higher than that needed to achieve the same effect on DNA polymerase a from eukaryotes (Hubscher 1983) and on DNA polymerase from *Methanococcus vannielii* (Zabel et al. 1987). On the other hand, this drug was not effective on DNA polymerase activity isolated from *S. acidocaldarius* (Elie et al. 1989) and *Methanobacterium thermoautotrophicum* (Klimczak et al. 1986). Although initially no exonuclease activity was found to be associated to Sso DNA pol (Rossi et al. 1986), the presence of a 3′-5′ exonuclease function was subsequently detected in homogeneous native enzyme preparations using a more sensitive assay (Pisani and Rossi 1994). More recently, this activity has been biochemically characterized in the recombinant Sso DNA pol and found to possess thermostability features similar to that observed for the polymerase function; but, in contrast to this latter, it has a remarkable (if not absolute) preference for Mn^{2+} ion as a cofactor (unpublished results). The main biochemical features of Sso DNA pol and several other archaeal DNA polymerases are reported in Table 2.

Table 2. Biochemical features of some archaeal DNA polymerases

Enzyme	Source	Protein sequence[a]	Optimal temperature (°C)	Exonuclease		Size (kDa)	Half-life[b] (min at °C)
				5′-3′	3′-5′		
Sac	*S. acidocaldarius*	N.D.[c]	55–75	−	+	100	15 at 87
*Sso*I	*S. solfataricus*	Family B	75	−	+	101.3	6 at 90
*Sso*II	*S. solfataricus*	Family B	N.D.[c]	N.D.[c]	N.D.[c]	73.5	N.D.
*Poc*I	*P. occultum*	Family B	70–80	+	+	100	N.D.
*Poc*II	*P. occultum*	Family B	70–80	+	+	100	N.D.
Tli	*T. litoralis*	Family B	70–80	−	+	89.9	108 at 100
Pfu	*P. furiosus*	Family B	70–80	−	+	90.1	240 at 95
Pwo	*P. woesei*	N.A.[c]	70–80	−	+	90	>120 at 100
Mth	*M. thermoautotrophicum*	N.D.[c]	65	+	+	72	<12 at 100
Mvo	*M. voltae*	Family B	35–40	N.D.[c]	N.D.[c]	82.5	N.D.

[a] Protein sequences were determined by reverse-translation of encoding genes. Protein data bank accession numbers are as follows: *Sso*I, X64466; *Sso*II, X71587; *Poc*I, D38573; *Poc*II, D38574; *Tli*, M74198; *Pfu*, D12983; *Mvo*, L33366.
[b] Expressed as the time required to reduce the polymerase activity to 50% of the value assayed at the optimal temperature at zero time.
[c] N.D., not determined; N.A., not available.

Isolation of the Sso DNA pol encoding gene was carried out by means of two degenerate oligonucleotide probes, which were designed on the basis of the partial N-terminal sequence of the 50- and 40-kDa polypeptide species, hypothesized to be proteolytic fragments of the 100-kDa protein, as previously discussed (Pisani et al. 1992). The N-terminus of the 100-kDa protein was found to be blocked to Edman degradation. The nucleotide sequence of the gene, cloned from *S. solfataricus* genomic libraries in λ vectors, revealed the presence of a 882-codon ORF, which corresponds to a polypeptide with a predicted molecular mass of 103 kDa, in good agreement with that reported for the native Sso DNA pol. This gene did not seem to contain inteins, the intervening sequences that splice at protein and not at mRNA level. Inteins were found in DNA polymerase-encoding genes of several *Thermococcus* and certain *Pyrococcus* species (Perler et al. 1996). The 50- and 40-kDa polypeptide N-terminal sequences were found to be entirely encoded within this ORF, demonstrating the occurrence of a proteolytic degradation during the purification procedure. Computer-assisted similarity searches revealed that Sso DNA pol belongs to family B, sharing with the enzymes from this group the five highly conserved C-terminal motifs previously discussed, in addition to the three Exo sequences.

Presently, about ten new archaeal DNA polymerase sequences are available (see Table 2). A comparison among them indicates that Sso DNA pol is more similar to DNA pol I from *Pyrodictium occultum* (51% identity and 69% similarity) than to DNA polymerases from *Thermococcus* or *Pyrococcus* (about 32% identity and 54% similarity). This could result from either the phylogenetic distance among the various taxonomic groups or the difference in the biological function carried out by each enzyme in vivo.

Sulfolobus solfataricus possesses two distinct DNA polymerase genes

Sulfolobus solfataricus was the first hyperthermophile shown to possess two distinct DNA polymerase genes. The Sso DNA pol II gene was isolated from *S. solfataricus* P2 strain by means of degenerate oligonucleotides designed on the basis of the polymerase-conserved sequence "YGDTDS" (Prangishvili and Klenk 1993). The cloned gene contains an ORF that encodes a putative 626-amino-acid residue protein of 74 kDa. Prangishvili and Klenk demonstrated, by Southern hybridization experiments with oligonucleotide probes specific for Sso DNA pol I or II, that two distinct family B DNA polymerase genes exist, at least in *S. solfataricus* P2 strain. The biochemical properties of Sso DNA pol II are not known; the protein has not been produced in a recombinant form and it is not yet clear whether the corresponding gene is really expressed in *S. solfataricus.* However, quite unexpectedly, protein sequence comparison revealed that Sso DNA pol I and II have only 23% identity and 50% similarity. Two distinct family B DNA polymerase genes was also isolated from *Pyrodictium occultum* (Uemori et al. 1995), and evidence for the presence of more than one DNA polymerase activity in *Pyrococcus furiosus* was also reported by Uemori et al. (1996). These findings seem to be in contrast with the presence of only one DNA polymerase gene in the *Methanococcus jannaschii* genome (Bult et al. 1996).

Sso DNA pol possesses a modular organization of its associated catalytic activities

As previously discussed, native DNA polymerase from *S. solfataricus* has a protease-hypersensitive site that is likely to be cleaved by the action of endogenous protease during the purification procedure. As a consequence, two proteolytic fragments of about 50 and 40 kDa, in addition to the intact 100-kDa protein species, were detected on SDS-PAGE (-polyacrylamide gel electrophoresis) of highly purified Sso DNA pol. N-terminal sequence analysis enabled us to assign the 40- and 50-kDa polypeptides to the N- and C-terminal portion of the protein, respectively. To assess whether specific catalytic functions could be ascribed to the native Sso DNA pol proteolytic fragments, we utilized the

bidimensional activity gel assay procedure described by Longley and Mosbaugh (1991). This technique relies on the use of a defined ^{32}P-labeled oligonucleotide annealed to M13 single-stranded DNA, which is cast within a SDS-polyacrylamide resolving gel. Following separation of the polypeptides and in situ enzyme renaturation and reaction, radiolabeled DNA products of polymerization (or degradation) are separated according to size in a second dimension of electrophoresis through a denaturing gel. The substrate employed to detect the DNA polymerase activity consisted of a ^{32}P-5′-end-labeled 24-mer oligonucleotide fully matched to the template; on the other hand, the substrate for 3′-5′ exonuclease activity assay was a 25-mer oligonucleotide forming a 3′-terminal mispair with the same DNA template. Both the 100- and 50-kDa polypeptides retained this synthetic capability in the presence of Mg^{2+} ion. On the other hand, the 3′-5′ exonuclease assay, performed in a reaction buffer containing Mn^{2+} ion, revealed that both the 100- and 40-kDa molecular species were able to degrade the primer. On the basis of this biochemical analysis, we proposed that the Sso DNA pol consists of two distinct structural domains with different catalytic activity that are able to function in an independent fashion (Pisani and Rossi 1994). The two domains are connected by a region of the polypeptide chain sensitive to the proteolytic attack and forming quite likely an unstructured loop on the molecule surface, as the result of its elevated hydrophilicity and high content in glycine residues. Several packing interactions are likely to provide a tight linkage between the two domains, because they do not dissociate when the interconnecting loop is nicked, as indicated by the results of native PAGE and gel filtration experiments on highly purified Sso DNA pol. Quite interestingly, the N-terminal domain responsible for the exonuclease activity includes the three Exo motifs conserved among DNA polymerases of family A and B, whereas the C-terminal domain, to which the polymerization function was assigned, contains the five highly conserved sequence regions described by Blanco et al. (1991).

The model of modular organization proposed for Sso DNA pol was confirmed by a limited proteolysis analysis carried out on the recombinant enzyme, overproduced in *E. coli* under the control of a T7 RNA polymerase promoter and purified to homogeneity (see Fig. 1; Pisani et al. 1996). Partial trypsin digestion under native conditions produced relatively stable fragments corresponding to the N- and C-terminal halves of the enzyme molecule, which, by means of activity gel analysis, were found to retain 3′-5′ exonuclease and polymerase activity, respectively. In accordance with these findings, we have recently overexpressed in *E. coli* two C-terminal deletion derivatives of Sso DNA pol containing the first 438 or 508 N-terminal residues (including the three Exo motifs). The corresponding truncated proteins, purified and biochemically characterized, retain 3′-5′ exonuclease activity but are devoid of synthetic capability. Quite interestingly and unexpectedly, these deletion derivatives have been found to possess a substantially higher thermal stability in comparison to the full-sized recombinant enzyme. Furthermore, attempts at overexpressing independently in *E. coli* the Sso DNA pol C-terminal domain in a soluble and active form are in progress.

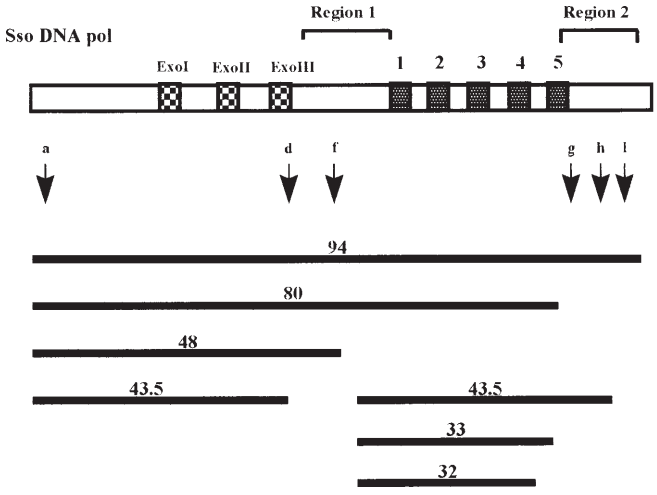

Fig. 1. Cleavage map of trypsin on Sso DNA pol. The *top diagram* indicates the position of the sequence similarity motifs of family B DNA polymerases, according to Blanco et al. (1991), within the Sso DNA pol primary structure. *Lines* represent the proteolytic fragments obtained by trypsin digestion (30 min at 37°C at a protease/polymerase ratio w/w of 1/1000). The apparent molecular size of each polypeptide is indicated in kDa. The cleavage sites at the N-terminal end were determined by protein sequencing as previously described (Pisani et al. 1996). Relative distance of cleavage sites from the protein N-terminus is indicated by alphabetical order to allow a comparison with the map depicted in Fig. 2

DNA-induced conformational changes of Sso DNA pol were mapped by limited proteolysis studies

In addition to being utilized to localize interdomain regions of multimodular proteins, the limited proteolysis technique can also be employed to detect subtle conformational changes of a protein on binding to a specific ligand. Indeed, comparison of the cleavage map in the absence or presence of the ligand allows to precisely locate where the structural changes have taken place. Partial proteolysis experiments on recombinant Sso DNA pol in native conditions by trypsin (or other proteases) revealed two main sites of protease sensitivity: within the center of the polypeptide chain (region 1), and near the the C-terminus (region 2; see Fig. 1). Limited proteolysis experiments carried out on binding to nucleic acid molecules (either activated calf thymus DNA or single-stranded M13 DNA) indicated that DNA-induced conformational changes took place mainly within region 1 and 2 for the following reasons: (i) the general susceptibilty to the proteolytic attack of both these polypepetide chain segments was noticeably increased; (ii) within these regions certain cleavage sites disappeared, whereas new others became accessible to the proteases; and (iii) a polypeptide fragment of about 8 kDa, which includes region 1, was protected from degradation only in the pres-

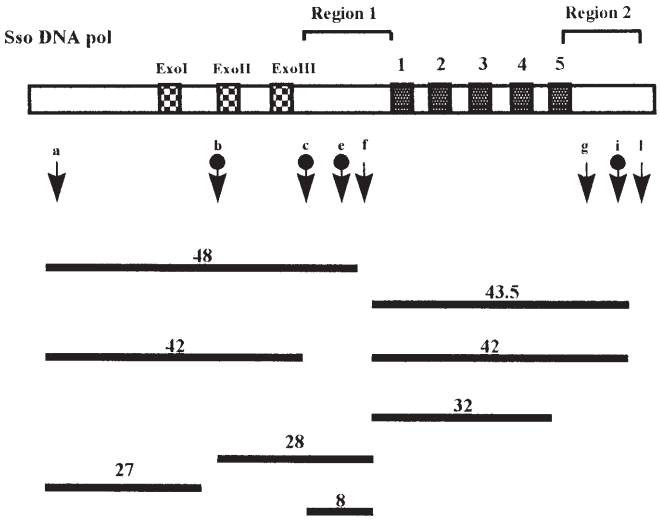

Fig. 2. Cleavage map of trypsin on Sso DNA pol following activated calf thymus DNA binding. Cleavage sites, which are present only when digestion is carried out in the presence of nucleic acid ligand, are represented by the *modified arrows.* (See legend to Fig. 1)

ence of DNA, suggesting a direct interaction with the nucleic acid ligand (see Fig. 2) (Pisani et al. 1996).

A new consensus motif, "Y-GG/A," lying between the Exo III and "D–SLYP" (motif 1 of Fig. 1 and Fig. 2) sequences, has been identified in nearly all family B DNA polymerases (Braithwaite and Ito 1993). Interestingly, in Sso DNA pol this sequence motif is located within region 1. On the basis of mutagenesis studies on DNA polymerase from T4 (Stocki et al. 1995) and ϕ29 (Truninger et al. 1995) bacteriophages, the "Y-GG/A" motif has been proposed to promote intramolecular switching of the primer terminus between the polymerase and 3′-5′ exonuclease separate active sites. More recently, the analysis of bacteriophage RB69 family B DNA polymerase crystallographic structure has revealed that the highly conserved "Y-GG/A" motif (and a portion of the polypeptide chain corresponding to region 1 of Sso DNA pol) is located within a long solvent-accessible loop between the exonuclease and polymerase domains, and its Tyr^{391} is expected to make contact with the DNA substrate primer strand (Wang et al. 1997). Our results, which indicate that region 1 of Sso DNA pol is flexible, exposed to the solvent on the protein molecule surface, and subjected to conformational changes on DNA binding, are consistent with this finding (Pisani et al. 1996).

Future prospects

One of the most interesting issues in the study of the archaeal DNA replication apparatus is to understand the physiological role of each distinct DNA polymerase, if multiple activities are present in a certain species. Development of genetic tools to specifically manipulate and transform Archaea is needed to ascertain this in vivo biological function. Of related, and not least, interest is the isolation of archaeal DNA polymerases accessory subunits, such as PCNA-like and eukaryotic replication C factor homologs (Kuryan and O'Donnell 1993; Edgell and Doolittle 1997). Another important goal of our research is the identification of Sso DNA pol subdomains or specific amino acid residues responsible for nucleic acid substrate interaction. This finding could enable us to design Sso DNA pol site-specific mutants with improved catalytic activity and processivity (the ability of a polymerase to synthesize long stretches of DNA without dissociating from the primer/template). Furthermore, attempts at obtaining Sso DNA pol crystals suitable for X-ray diffraction studies have been successful. Resolution of the protein three-dimensional structure will be invaluable to elucidate the structure–function relationships of this as well as other homologous thermostable archaeal DNA polymerases.

Note added in proofs: Edgell et al. found in *Sulfolobus solfataricus* (strain P2) genome an ORF encoding a putative third family B DNA polymerase, designated as B3 (Edgell DR, Klenk H-P, Doolittle WF, 1997, Gene duplications in evolution of archaeal family B DNA polymerases. J Bacteriol 179:2632–2640

Acknowledgments This work was supported by Progetto Finalizzato Biotecnologie (C. N. R., Italy) and partly by European Union Contract B104-CT96-0488.

References

Bambara RA, Murante RS, Henricksen LA (1997) Enzymes and reactions at the eukaryotic DNA replication fork. J Biol Chem 272:4647–4650

Bernad A, Blanco L, Lazaro JM, Martin G, Salas M (1989) A conserved 3′-5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. Cell 59:219–228

Blanco L, Salas M (1996) Relating structurre to function in ϕ29 DNA polymerase. J Biol Chem 271:8509–8512

Blanco L, Bernad A, Blasco MA, Salas M (1991) A general structure for DNA-dependent DNA polymerases. Gene 100:27–38

Braithwaite DK, Ito J (1993) Compilation, alignment and phylogenetic relationships of DNA polymerases. Nucleic Acids Res 21:787–802

Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb J-F, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JL, Presley EA, Nguyen D, Utterback TR, Kelley JM, Peterson JD, Sadow PW, Hanna MC, Cotton MD, Hurst MA, Roberts KM, Kaine BP, Borodovsky M, Klenk H-P, Fraser CM, Smith HO, Woese CR, Venter JC (1996) Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii.* Science 273:1058–1073

Davies JFN, Almassy RJ, Hostomska Z, Ferre RA, Hostomski Z (1994) 2.3 Å crystal structure of the catalytic domain of DNA polymerase b. Cell 76:1123–1133

Dennis PP (1997) Ancient ciphers: translation in Archaea. Cell 89:1007–1010

Derbyshire V, Pinsonneault JK, Joyce CM (1995) Structure-function analysis of 3′ → 5′-exonuclease of DNA polymerases. Methods Enzymol 262:363–385

Edgell DR, Doolittle WF (1997) Archaea and the origin(s) of DNA replication proteins. Cell 89:995–998

Elie C, De Recondo AM, Forterre P (1989) Thermostable DNA polymerase isolated from the archaebacterium *Sulfolobus acidocaldarius*. Purification, characterization and immunological properties. Eur J Biochem 178:619–626

Forterre P, Charbonnier F, Marguet E, Harper F, Henckes G (1992) Chromosome structure and DNA topology in extremely thermophilic Archaebacteria. In: Danson MJ, Hough DW, Lunt GG (eds) The Archaebacteria: biochemistry and biotechnology. Biochemical Society, London, pp 99–112

Hubscher U (1983) DNA polymerase in eukaryotes and prokaryotes. Experientia (Basel) 39:1–26

Joyce CM, Steitz TA (1994) Function and structure relationships in DNA polymerases. Annu Rev Biochem 63:777–822

Kim Y, Eom SH, Wang J, Lee DS, Suh SW, Steitz TA (1995) Crystal structure of *Thermus aquaticus* DNA polymerase. Nature (Lond) 376:612–616

Klimczak LJ, Grummt F, Burger KJ (1986) Purification and characterization of a DNA polymerase from the archaebacterium *Methanobacterium thermoautotrophicum*. Biochemistry 25:4850–4855

Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 256:1783–1790

Kornberg A, Baker TA (1992) DNA polymerase I of *E. coli*. In: Kornberg A, Baker TA (eds) DNA replication. Freeman, New York, pp 113–164

Korolev S, Nayal M, Barnes WM, DiCera E, Waksman G (1995) Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5 Å resolution: structural basis for thermostability. Proc Natl Acad Sci USA 90:4299–4303

Kuryan J, O'Donnell M (1993) Sliding clamps of DNA polymerases. J Mol Biol 234:915–925

Longley MJ, Mosbaugh DW (1991) Characterization of DNA metabolizing enzymes *in situ* following polyacrylamide gel electrophoresis. Biochemistry 30:2655–2664

Maki H, Maki S, Kornberg A (1988) DNA polymerase III holoenzyme of *Escherichia coli.* Distinctive processive polymerase reconstituted from purified subunits. J Biol Chem 263:6570–6579

Mullen GP, Wilson SH (1997) DNA polymerase b in abasic site repair: a structurally conserved helix-hairpin-helix motif in lesion detection by base excision repair enzymes. Biochemistry 36:4713–4717

Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA (1985) Structure of large fragment of *Escherichia coli* DNA polymerase I. Nature (Lond) 313:762–766

Olsen GJ, Woese CR (1997) Archaeal genomics: an overview. Cell 89:991–994

Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J (1994) Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. Science 264:1891–1903

Perler FB, Kumar S, Kong H (1996) Thermostable DNA polymerases. Adv Protein Chem 48:377–435

Pisani FM, Rossi M (1994) Evidence that an archaeal α-like DNA polymerase has a modular organization of its associated catalytic functions. J Biol Chem 269:7887–7892

Pisani FM, De Martino C, Rossi M (1992) A DNA polymerase from the archaeon *Sulfolobus solfataricus* shows sequence similarity to family B DNA polymerases. Nucleic Acids Res 20:2711–2716

Pisani FM, Manco G, Carratore V, Rossi M (1996) Domain organization and DNA-induced conformational changes of an archaeal family B DNA polymerase. Biochemistry 35:9155–9166

Prangishvili D, Klenk HP (1993) Nucleotide sequence of the gene for a 74 kDa DNA polymerase from the archaeon *Sulfolobus solfataricus*. Nucleic Acids Res 21:2768

Reeve JN, Sandman K, Daniels CJ (1997) Archaeal histones, nucleosomes, and transcription initiation. Cell 89:999–1002

Rella R, Raia CA, Pisani FM, D'Auria S, Nucci R, Gambacorta A, De Rosa M, Rossi M (1990) Purification and properties of a thermophilic and thermostable DNA polymerase from the archaebacterium *Sulfolobus solfataricus*. Ital J Biochem 39:83–99

Rossi M, Rella R, Pensa M, Bartolucci S, De Rosa M, Gambacorta A, Raia CA, Dell'Aversano Orabona N (1986) Structure and properties of a thermophililc and thermostable DNA polymerase isolated from *Sulfolobus solfataricus*. Syst Appl Biochem 7:337–341

Stocki SA, Nonay RL, Reha-Krantz LJ (1995) Dynamics of bacteriophage T4 DNA polymerase function: identification of amino acid residues that affect switching between polymerase and 3′-5′ exonuclease activties. J Mol Biol 254:15–28

Truninger V, Lazaro JM, Salas M, Blanco L (1995) A DNA binding motif coordinating synthesis and degradation in proofreading DNA polymerases. EMBO J 15:3430–3441

Uemori T, Ishino Y, Doi H, Kato I (1995) The hyperthermophilic archaeon *Pyrodictium occultum* has two α-like DNA polymerases. J Bacteriol 177:2164–2177

Uemori T, Satoh Y, Kato I, Doi H, Ishino Y (1996) Organization of a novel DNA polymerase supposed to be a replicational enzyme from *Pyrococcus furiosus* (abstract). In: "Thermophiles '96" international conference, Athens, GA, USA, 4–9 September 1996, p 2.I.2

Wang J, Sattar AKM, Wang CC, Karam JD, Konigsberg WH, Steitz TA (1997) Crystal structure of a pol, a family replication DNA polymerase from bacteriophage RB69. Cell 89:1087–1089

Zabel H-P, Holler E, Winter J (1987) Mode of inhibition of the DNA polymerase of *Methanococcus vannielii* by aphidicolin. Eur J Biochem 165:171–175