Re-engineering the polymerase domain of Klenow fragment and evaluation of overproduction and purification strategies

Victoria Derbyshire⁺, Mekbib Astatke and Catherine M. Joyce^{*}

Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520, USA

Received August 10, 1993, Revised and Accepted October 21, 1993

ABSTRACT

We describe experiments to produce large quantities of the polymerase domain of *E. coli* DNA polymerase I for biochemical and biophysical studies. The polymerase domain derivative used in previous studies was insoluble when overproduced and tended to aggregate during purification. These problems were solved by a combination of two distinct strategies. By changing the expression system, we were able to obtain the overproduced protein in a soluble form, a necessary first step since attempts to purify the polymerase domain from the insoluble pellet were unsuccessful. The tendency of the polymerase domain to aggregate was eliminated by re-engineering the protein so as to remove both a solvent-exposed hydrophobic patch and a potentially unstructured region at the extreme N-terminus. Unlike the original construct, the re-engineered derivatives chromatographed as a single species and could be purified to homogeneity in good yield. Our experience in this study emphasizes the level of ignorance of the factors that influence protein overproduction and the need, in difficult cases, to evaluate many strategies in a semi-empirical manner.

INTRODUCTION

DNA polymerase I of *Escherichia coli* is a multifunctional, single-subunit, protein of molecular weight 103 kDa. It has three enzymatic activities: DNA polymerase, 5'–3' exonuclease and 3'–5' exonuclease (1). The protein can be proteolytically cleaved to give a 68 kDa carboxy-terminal fragment, the Klenow fragment, which retains the polymerase and 3'–5' exonuclease activities of the parent molecule, but no longer has the 5'–3' exonuclease activity (2, 3). The DNA encoding Klenow fragment has been cloned into various expression vectors making it possible to purify large quantities of protein for biophysical and biochemical studies (4, 5). Because of its small size and the

availability of structural data, Klenow fragment provides an excellent simple model for the study of the enzymology of DNA replication.

The three-dimensional crystal structure of Klenow fragment has been solved to high resolution (6). The enzyme is folded into two distinct structural domains, an amino-terminal domain containing the 3'–5' exonuclease active site and a larger carboxy-terminal polymerase domain. Genetic and biochemical studies of the wild-type Klenow fragment and many mutant derivatives have yielded considerable information about the molecular details of both enzymatic reactions (5, 7–9). Our understanding of the 3'–5' exonuclease reaction is more complete because the mutagenesis and structural studies have proceeded in parallel, largely due to the availability of complexes having the substrate (single-stranded DNA) or the product (dNMP) bound at the 3'–5' exonuclease active site (6, 10, 11).

By contrast, structural information on the binding of substrates to the polymerase active site has been much harder to obtain. Consequently, although extensive mutagenesis experiments have identified amino acid residues involved in various steps of the polymerase reaction (5, 9), there is as yet no defined picture of the active site nor a detailed mechanism for the reaction. In an effort to get crystals with DNA bound at the polymerase active site, many different crystallization conditions have been tried with a variety of DNA substrates. Even when duplex DNA substrates were used, the complexes seen in co-crystals obtained at high ionic strength contained single-stranded DNA bound to the 3'–5' exonuclease active site (10, 12). Whereas, under low salt aqueous conditions, a duplex DNA terminus is bound predominantly at the polymerase site (13), crystallization at high salt may have favored interactions with the exonuclease site which provides a series of hydrophobic contacts with the DNA substrate (11, 14). In at least some of the co-crystals studied, it seemed likely that the duplex DNA, if present, was disordered, since, although changes were seen in the protein structure of the polymerase domain, no associated DNA was visible in the electron-density maps (10, 12). More recently, the structure of an editing complex

^{*} To whom correspondence should be addressed

⁺ Present address: The David Axelrod Institute for Public Health, New York State Department of Health, Wadsworth Center for Laboratories and Research, Empire State Plaza, PO Box 509, Albany, NY 12201, USA

in which the duplex portion of the DNA can be seen, has been solved at high-resolution and this has led to a proposal for the mode of DNA binding at the polymerase active site (12).

An alternative approach to obtain co-crystals having a duplex DNA terminus at the polymerase active site would be to use the isolated polymerase domain of Klenow fragment. This would eliminate competition from the alternative DNA binding region at the 3'-5' exonuclease active site, and might provide an opportunity to explore other crystal forms and crystallization conditions in the hope of finding conditions more conducive to DNA binding at the polymerase active site. The carboxy-terminal polymerase domain (residues 515-928 of DNA polymerase I) had previously been purified from an overproducing strain (15). However, the protein obtained in this way was not appropriate for structural studies. When expressed at high levels, the polymerase domain was essentially completely insoluble and could be solubilized from the cell pellet only by denaturation. Even when renatured in aqueous solution the protein had a tendency to aggregate, resulting in massive losses during the subsequent chromatographic steps. Although much lower levels of expression gave some polymerase domain that was soluble in the cell extract, this strategy also was clearly impractical for preparing the large amounts of protein necessary for structural studies. The enzymatic activity of the polymerase domain purified in these studies could be increased by the addition of glycerol to the reaction mix, consistent with the structure being in some way compromised.

To facilitate purification of the polymerase domain in sufficient quantities for structural studies, we adopted two approaches. One was to modify the polymerase domain clone so as to remove two potential sources of structural instability, which might have contributed to the tendency of the protein to aggregate. These were an extra seven amino acids preceding the start of helix G (Fig. 1) at the N terminus of the protein, which could potentially assume an unstructured conformation, and a hydrophobic region on the face of helix G that is normally buried in Klenow fragment but becomes exposed when the 3'-5' exonuclease domain is absent. The second approach was to evaluate some of the strategies that have been developed more recently for the purification of active proteins from an insoluble cell pellet (e.g. ref. 16).

MATERIALS AND METHODS

Materials

Oligodeoxyribonucleotides were synthesized on a Biosearch 8600 DNA synthesizer and were purified by gel electrophoresis as described previously (5). Restriction endonucleases, T4 polynucleotide kinase and DNA ligase were from New England Biolabs, Boehringer Mannheim or Collaborative Research and were used according to the manufacturers' instructions. *Bst* polymerase was from Bio-Rad.

General techniques

SDS-polyacrylamide gel electrophoresis was carried out according to Laemmli (17). Where necessary (see text), the polymerase domain was detected on gels by Western blotting (18), using rabbit antiserum directed against Klenow fragment, and enhanced chemiluminescent detection (Amersham). The protein concentration of the purified polymerase domain was determined by the Bradford colorimetric assay (19), using the reagent supplied by Bio-Rad. Bovine serum albumin was used

as a standard, since we had previously determined that it gave the same results as a Klenow fragment standard. The concentration of the polymerase domain in partially purified samples was determined by densitometric scanning of samples run on SDS-polyacrylamide gels, stained with Coomassie Blue, in comparison with a Klenow fragment standard of known concentration. DNA manipulations were carried out as described by Maniatis *et al.* (20).

Mutagenesis of the polymerase domain

The mixed oligonucleotides, A and B, and the partially complementary oligonucleotide C are described in Fig. 2. Oligonucleotide A or B was annealed to oligonucleotide C in a 30 μ l reaction mix containing 6.8 pmol of each oligonucleotide in 10 mM Tris-HCl, pH 8.0, 5 mM MgCl₂. Trace amounts of the same oligonucleotides, 5'-labeled with ³²P, were added to each reaction to facilitate subsequent gel purification. To minimize problems due to self-priming by the oligonucleotides, the annealing and primer extension reactions were carried out at high temperatures. The annealing reactions were incubated at 80°C for 10 min, quickly frozen on dry ice and then incubated at 55°C for 30 min. Extension reactions were then initiated by the addition of 20 μ l of a solution containing 0.25 mM of each dNTP, 7 μ l of 60 mM Tris-HCl, pH 7.5, 7 μ l of 100 mM NaCl, 7 μ l of 1 mg/ml bovine serum albumin and 3 units of *Bst* polymerase. After incubation at 55°C for 20 min the reaction was quenched by the addition of 5 μ l of 0.5 M EDTA. The DNA was precipitated with ethanol and the reaction products were fractionated on 6% polyacrylamide-urea gels with appropriate size markers. The DNA was visualized by autoradiography and fragments of the correct size (100 bp) were excised. Even though the primer extension reactions were carried out at high temperature, the reactions contained substantial amounts of other by-products. After isolation from the gel slice, the DNA was digested with *Nde*I and *Cla*I. Approximately 0.02 pmol of each fragment library was ligated to an equivalent amount of the large *Nde*I-*Cla*I fragment of pVD19 (see below). Portions of the ligation mix were introduced into strain LE392 (*lacY galK galT metB trpR supE supF hsdR*) by electroporation. Plasmids having the correct structure, as judged by restriction mapping, were further characterized by DNA sequencing.

Plasmids

The recipient plasmid for the oligonucleotide-derived mutant library was pVD19. It was derived from a plasmid containing a fragment from the previous polymerase domain overproducer (15), spanning the amino-terminus of the wild-type polymerase domain from the *Ava*I site preceding the coding sequence to the *Sac*I site (nucleotide 1672 on the *polA* sequence; ref. 21), cloned into pUC118 (22). To facilitate gel purification of the large *Nde*I-*Cla*I fragment for cloning the mutant library, the 120 bp *Nde*I-*Cla*I fragment at the start of the polymerase domain coding sequence was replaced with a 1200 bp *Nde*I-*Cla*I fragment from pNG16 (23), giving pVD19. Because the *Cla*I site in this region of the *polA* gene is methylated, the plasmid DNAs used in these manipulations were prepared from a *dam*⁻ strain.

The mutant sequences characterized in the pVD19 recipient plasmid were transferred to an overproducer plasmid derived from the pAS1 expression vector (24), such that the coding sequence of the polymerase domain was placed under the control of the λ P_L promoter. Each mutant sequence was excised on a fragment extending from the unique *Bst*XI site in control

sequences upstream of the translational start, to the unique *SacI* site (nucleotide 1672 on the *polA* sequence) which was ligated so as to reconstruct the entire polymerase domain coding sequence. Expression plasmids using a T7 promoter were derived from pGEM3 (Promega), having T7 transcriptional signals, but modified so as to have the translational control signals (and the *BstXI* site) of the λ *cII* gene, as in the pAS1 vector. First, wild-type Klenow fragment was cloned using the *BstXI*-*EcoRI* fragment from a Klenow fragment expression plasmid (5). Polymerase domain derivatives were then cloned by replacing the *BstXI*-*SacI* fragment encoding the Klenow fragment N-terminus with a series of *BstXI*-*SacI* fragments from the λ P_L expression plasmids described above.

Expression of polymerase domain derivatives from the λ P_L promoter, and tests of solubility

The series of overproducer plasmids was introduced into the bacterial host AR120, and expression of the polymerase domain was induced by addition of nalidixic acid to 40 μ g/ml, as previously described (15). In an effort to improve solubility, inductions were carried out at room temperature using either nalidixic acid (40 μ g/ml) or mitomycin C (100 μ g/ml). The level of expression was the same with either inducing treatment; the mitomycin C-induced samples were processed for determining solubility.

To test solubility of the overproduced protein, the cell pellet from a 1.5 ml induced culture was resuspended in 1 ml of 50 mM Tris-HCl, pH 7.5, containing 2 mM EDTA, 1 mM DTT, 0.02 mM PMSF and 2 mg/ml lysozyme, and was lysed by sonication as previously described (4). Alternatively, the cells were resuspended in a small volume (300 μ l) of the same buffer and were lysed by four cycles of freezing and thawing. Samples of the extract before centrifugation, and of the supernatant and pellet obtained after spinning for 5 min in a microfuge at 4°C, were examined by SDS-polyacrylamide gel electrophoresis.

Attempts at solubilization using ion-exchange resins

Using the λ P_L expression system, the cell pellet from a 20 ml nalidixic acid-induced culture was lysed by sonication (4), in a total volume of 10 ml. Portions (1 ml) were centrifuged for 5 min in a microfuge at 4°C, and the pellets were resuspended in 100 μ l of a series of 50 mM buffers at pH values 5.5 (Mes), 6.5 (Pipes), 7.5 (Hepes), 8.5 (Hepes) and 9.6 (Caps). (In each case the acidic form of the buffer was titrated to the desired pH with NaOH.) There were two extracts for each pH value; one was treated with Q-Sepharose Fast Flow (Pharmacia) and one with S-Sepharose Fast Flow. In each case approx. 100 μ l of the ion-exchange resin, equilibrated with the appropriate buffer, was added, and the mixture was mixed gently for 2 h at 4°C. After centrifugation and removal of unbound proteins in the supernatant, any bound proteins were eluted by gentle mixing with 200 μ l of the corresponding buffer containing 1 M NaCl, for 16 h at 4°C.

Purification of polymerase domain derivatives from the insoluble cell pellet

Following the method of Nagai and Thøgersen (25), cells (approx. 0.5 g) from a 100 ml nalidixic acid induction were lysed and inclusion bodies were purified by a series of detergent washes. The pellet was resuspended in 25 ml of 50 mM Tris-HCl, pH 8.5, 1 mM DTT, containing 8 M urea, and mixed gently for 20 min with 3 ml of Q Sepharose Fast Flow equilibrated with

the same buffer. After 2 washes with this buffer (30 ml), bound proteins were eluted with sequential washes (10 ml) of the same buffer containing NaCl at 50 mM, 100 mM, 150 mM, 200 mM, 250 mM and 500 mM. In each case, the 100 mM elution was the purest and was estimated to contain 2 to 5 mg/ml of the polymerase domain. Two methods for renaturation of the protein in these fractions were tested. In one, a portion (2 ml) of the eluate was dialyzed against 50 mM Tris-HCl, pH 7.5, 1 mM DTT, containing 8 M urea, to remove the NaCl, and was loaded onto a Q Sepharose Fast Flow column (0.4 ml) in the same buffer. The column was washed and then the urea was removed by sequential washes (6 ml) of 50 mM Tris-HCl, pH 7.5, 1 mM DTT, containing 7 M, 6 M, 5 M, 4 M, 3 M, 2 M and 1 M urea. After washing the column with 50 mM Tris-HCl, pH 7.5, 1 mM DTT, the bound protein was eluted with a series of washes (1 ml) of the same buffer containing NaCl at 50 mM, 100 mM, 150 mM, 200 mM, 250 mM, 500 mM and 1 M. The polymerase domain eluted in the 500 mM and 1 M NaCl fractions. In the second method, a portion (0.4 ml) of the original 100 mM NaCl eluate was diluted to approx. 40 μ g/ml protein with 50 mM Tris-HCl, pH 7.5, 1 mM DTT, containing 8 M urea, and then dialyzed against a series of buffers to remove the urea. For each dialysis, approx. 50 ml protein solution was dialyzed against 4 l of buffer for 3 h or overnight. The buffers were 50 mM Tris-HCl, pH 7.5, 1 mM DTT, containing 4 M urea, 2 M urea with 0.5 M NaCl, 1 M urea with 0.5 M NaCl, 0.5 M urea with 0.5 M NaCl, 0.5 M NaCl alone, and finally the same buffer with no additions. The dialysate was then concentrated by extraction onto Q Sepharose Fast Flow resin (1 ml). After mixing gently for 16 h the resin was washed twice with 50 mM Tris-HCl, pH 7.5, 1 mM DTT (10 ml), and then eluted with 1 ml aliquots of the same buffer containing NaCl at 50 mM, 100 mM, 150 mM, 200 mM, 250 mM, 500 mM, 1 M and 2 M. The polymerase domain eluted in the 500 mM, 1 M and 2 M NaCl fractions.

Expression of polymerase domain derivatives from the T7 promoter

The appropriate overproducer plasmid was introduced into the expression host BL21(DE3)pLysS (26,27). Fresh transformants were used in every case as a precaution against plasmid instability, which was an extremely serious problem with the Klenow fragment overproducer but much less so (if at all) for the polymerase domain clones. Cells were grown to mid-log at 37°C in LB containing carbenicillin (50 μ g/ml). They were transferred to 30°C for 20 min and then protein expression was induced by the addition of IPTG to 0.4 mM. Cells were harvested after 3 h. A crude cell extract was obtained as described previously for Klenow fragment (4), except that no lysozyme was added to the lysis buffer since the pLysS plasmid produced adequate amounts intracellularly. This crude extract could be assayed for polymerase activity without further purification.

Purification of polymerase domain derivatives

The polymerase domain derivative having asparagine substitutions at the four mutated positions was purified to homogeneity by a combination of FPLC and conventional chromatography. All steps were carried out at 0–4°C. Approx. 5–6 g of cells were processed as described above. Solid ammonium sulfate was added to the crude cell lysate to 40% saturation (22.6 g per 100 ml extract). The precipitate was discarded and solid ammonium sulfate was added to 85% saturation (a further 29.6 g per 100 ml extract). The precipitate was dissolved in 50 ml of 50 mM

Tris-HCl, pH 7.5, 1 mM DTT and dialyzed against this buffer. Because the dialysate was slightly turbid it was spun at $149,000 \times g$ for 30 min, before being filtered through a Millipore 0.22 μm filter unit and applied at 1 ml/min to a Mono Q HR 16/10 column equilibrated with 50 mM Tris-HCl, pH 7.5, 1 mM DTT. The column was washed with 150 ml of the same buffer and then the polymerase domain was eluted at 4 ml/min with a 250 ml linear gradient of 0–0.5 M NaCl in the same buffer. Peak fractions of the polymerase domain (at 120–150 mM NaCl) were pooled and dialyzed against 50 mM Pipes, pH 7.0, before being loaded on a 15 ml BioRex 70 column equilibrated in the same buffer. The column was washed with 75 ml 50 mM Pipes, pH 7.0, and then the polymerase domain was eluted at 0.7 ml/min with a 120 ml linear gradient of 0–1.5 M NaCl in the same buffer. The peak fractions (at 130–240 mM NaCl) were pooled and dialyzed against 50 mM Tris-HCl, pH 7.5, 1 mM DTT, containing 1.7 M ammonium sulfate. The dialyzed protein was applied to a phenyl Superose HR 5/5 column equilibrated with the same buffer. The column was washed with 5 ml of this buffer and the polymerase domain was eluted at 0.5 ml/min with a reverse ammonium sulfate gradient (1.7–0 M) in the same buffer. The peak fractions (at 0.7–1.0 M ammonium sulfate) were pooled and, assuming the pool to be 15% saturated in ammonium sulfate, solid ammonium sulfate was added to 85% saturation (46.0 g per 100 ml). The ammonium sulfate precipitate was dissolved in 150–200 μl of 100 mM Tris-HCl, pH 7.5, 1 mM DTT. This was clarified by centrifugation and 200 μl was applied to a Superose 12 HR 10/30 column equilibrated in the same buffer. The column was developed with 30 ml of this buffer at 0.5 ml/min. The polymerase domain eluted after approx. 15 ml. Peak fractions (approx. 1 mg/ml polymerase domain) were diluted with an equal volume of glycerol and stored at -20°C .

For the partial purification of polymerase domain derivatives from 100 ml cultures of induced cells, the BioRex 70 and phenyl Superose chromatography steps were omitted and chromatography on Mono Q was carried out on the smaller Mono Q HR 5/5 column.

Measurement of polymerase activity

The polymerase domains and Klenow fragment were assayed by the standard DNA polymerase assay (28), with poly[d(A-T)] as the template. One unit of polymerase activity is defined as the amount of enzyme catalyzing the incorporation of 3.3 nmol of dNTP into trichloroacetic acid-insoluble material in 10 min at 37°C . The steady-state kinetic constants, $K_m(\text{dNTP})$ and k_{cat} , for the homogeneous polymerase domain were measured on the homopolymer substrate, poly(dA) primed with oligo(dT)₁₆, as previously described (5). The dissociation constant, $K_D(\text{DNA})$, was measured by a DNase I footprinting titration as previously described (5), except that the DNA substrate was a 5' end labeled 68-mer 'hairpin' oligonucleotide with a 12-nucleotide 5' extension, having the identical DNA sequence around the primer terminus as in the previous M13-derived substrate. All values were the average of at least two determinations.

RESULTS

Mutagenesis to generate polymerase domain derivatives

In an attempt to improve the solubility and chromatographic behavior of the polymerase domain, the original expression plasmid (15) was modified in two ways. In the original clone, the initiator methionine preceded Asp515 of DNA polymerase

Figure 1. Polymerase domain constructs in relation to the three-dimensional structure of Klenow fragment. (A.) Structure of the entire Klenow fragment molecule (6). The approximate locations of the polymerase catalytic site, on the larger C-terminal domain, and the 3'-5' exonuclease catalytic site, on the N-terminal domain, are indicated. (B.) The polymerase domain. The construct used in this work has its N-terminus at the start of helix G, and mutations at residues Leu523, Phe526, Met531 and Val534 (whose positions are shown schematically) to make the face of helix G more hydrophilic. The shaded region indicates the location on the Klenow fragment structure of the additional six amino acids at the N-terminus of the 'wild-type' polymerase domain derivative originally purified by Freemont *et al.* (15).

I, close to the C-terminus of helix F (Fig. 1), such that the expressed protein contained an extra seven amino acids preceding the start of helix G. We deleted the corresponding coding sequence so as to place the initiation codon next to Gly521, the N-terminal residue of helix G. Additionally, it seemed possible that the physical properties of the original 'wild-type' polymerase domain could be the result of exposing to solution a hydrophobic region of the protein that is normally buried. By studying the refined three-dimensional structure (L.S. Beese and T.A. Steitz, personal communication) for the face of helix G that becomes exposed when the 3'-5' exonuclease domain is removed, we identified four amino acids, Leu523, Phe526, Met531 and Val534, as providing a solvent-exposed hydrophobic face to this helix. We therefore planned to substitute hydrophilic amino acids

Figure 2. Strategy for cloning and mutagenesis of the C-terminal polymerase domain of Klenow fragment. Panel A shows schematically how primed synthesis and restriction enzyme digestion were used to generate a library of DNA fragments for cloning. The oligonucleotides A and B spanned the amino terminal region of the original polymerase domain clone (15), but had the six codons following the initiation codon deleted, so as to place the initiator methionine at the start of helix G. The strategy for mixing the nucleotide precursors during synthesis so as to generate a library of mutations at codons 523, 526, 531 and 534 is detailed in panel B. Panel C shows the entire sequence of the oligonucleotides, indicating the locations of the initiation codon and the mutated codons in relation to the restriction sites used in cloning.

for these four residues. Because we could not predict which particular hydrophilic amino acid substitutions would improve the solubility properties of the polymerase domain, we made a library of derivatives with different combinations of amino acid substitutions. To maximize the chances that these substitutions would be compatible with the overall protein structure, we were guided by a few common-sense considerations in designing our mutagenesis strategy. First we aimed to have glutamine and asparagine substitutions highly represented in our library, since these residues would be unlikely to generate unfavorable charge interactions. In addition, we avoided arginine or lysine at positions 523 and 531, because these residues could form salt bridges with Glu527. We also aimed to get some derivatives having both a positively and a negatively charged substitution in case maintenance of the isoelectric point was important.

We used a mixed oligonucleotide cloning strategy (Fig. 2) to generate derivatives of the polymerase domain of Klenow fragment having a re-engineered N-terminus and a variety of hydrophilic side chains on the exposed face of helix G. To satisfy the somewhat conflicting goals of generating many different domain derivatives and yet obtaining proteins containing glutamine and asparagine with ease, we made two mixed oligonucleotides (A and B). Oligonucleotide A was designed to give just 4 combinations of those residues that we thought were most likely to be successful. Oligonucleotide B gave 256 possible combinations of amino acid substitutions. After annealing to the partially complementary oligonucleotide C (Fig. 2), primed synthesis was used to generate double-stranded DNA fragments for cloning, as described in Methods. Plasmids were examined by restriction analysis to identify those that had insertions of the

correct size and had maintained the restriction sites used in the cloning. Of 93 clones analyzed from both mixed oligonucleotide cloning experiments, 74 were correct by restriction digestion. These 74 were sequenced to determine the nature of the amino acid substitutions (Table I). It is clear that some aspect of the primer extension and subcloning procedures was mutagenic since we found five additional nucleotide substitutions outside the codons mutagenized and we obtained one unexpected substitution within the target codons in the oligonucleotide A experiment. Comparison of Fig. 2 (panel B) and Table I shows that, although all possible substitutions were obtained at each of the mutagenized positions, the distribution of amino acid sequences in the final pool was not completely random, presumably due to different efficiencies of incorporation of nucleotides in the oligonucleotide synthesis. By sequencing 74 clones we obtained 28 unique combinations of amino acid changes that had the wild-type protein sequence throughout the rest of the region that was to be subcloned into the expression vector; we felt that these 28 sequences (listed in Table I) provided sufficient variety for subsequent tests. (Throughout this paper, each derivative is identified by the four amino acid residues (in the one-letter code) at positions 523, 526, 531 and 534; thus NNNN denotes the mutant protein having Asn at all four positions.)

Expression of mutant derivatives of the polymerase domain from the λ P_L promoter

We subcloned the 28 derivatives into an overproducer plasmid where expression was under the control of the inducible strong leftward promoter (P_L) of phage λ . This is the expression system currently used for Klenow fragment (5,7), and used in

Table I. Library of polymerase domain derivatives

Position ^a	523	526	531	534	Frequency
Oligo A ^b	L→Q	F→N or D	M→Q	V→N or K	
	Q	N	Q	N	6
	Q ^c	N	Q	K	7
	Q	D	Q	N	1
	Q	D	Q	K	3
	Q ^c	H*	Q	N	1
Oligo B ^b	L→N,Q,H or K	F→N,K,D or E	M→N,Q,H or K	V→N,S,K or R	
	N ^c	N	N	N	1
	N ^c	N	N	S	1
	N ^c	N	N	K	1
	N ^c	N	H	N	7
	N ^c	N	H	K	1
	N ^c	N	K	N	1
	N	D	H	N	3
	N ^c	D	H	R	1
	Q	N	N	K	1
	Q	N	H	K	1
	Q ^c	K	H	N	1
	Q ^c	D	H	N	1
	Q	E	H	S	1
	H	N	N	N	3
	H ^c	N	N	S	1
	H	N	N	K	1
	H	N	H	N	7
	H	D	H	N	1
	H	E	H	N	1
	K	N	Q	N	1
	K	N	H	N	1
	K	D	H	K	1
	K	D	N	K	1

^a The initiator methionine of each protein precedes residue 521 of the DNA polymerase I sequence (21); amino acids 523, 526, 531 and 534 were mutated as indicated.

^b The possible mutations encoded by each mixed oligonucleotide are listed. The * indicates an unexpected mutation that was obtained.

^c These domains were also examined using a T7 expression system (26).

the previous study of the wild-type polymerase domain (15). Protein synthesis was induced from each expression plasmid by SOS-induction with nalidixic acid, as in the previous study (15). After 16 hours at 37°C, all of the polymerase domain derivatives were expressed at high levels but fractionation of cell extracts into soluble and insoluble material showed that they were all essentially insoluble (Fig. 3). Although re-engineering clearly had not made the polymerase domain soluble when expressed in this system, we reasoned that it might have stabilized the protein structure sufficiently that the domain would not aggregate when extracted from the insoluble pellet. Moreover, the initial fractionation of the polymerase domain into the cell pellet offered a potential advantage by providing a substantial purification step which could be useful when purifying protein on a large scale for structural work. We therefore decided to evaluate methods for purification of the polymerase domain from the insoluble material in the cell pellet.

Methods for purification of polymerase domain derivatives from the insoluble cell pellet

Initially we tried two non-denaturing methods to solubilize the polymerase domain. None of the 28 derivatives were solubilized by extraction with buffers containing 1 M NaCl. Since we had so far detected no difference in behavior between the clones, the NNNS and NNNK derivatives were chosen for more detailed study. Following the observation that some insoluble proteins can be solubilized from bacterial lysates by treatment with ion-exchange resins (29), we tested extraction with a series of buffers from pH 5.5 to pH 9.6 in the presence of S Sepharose Fast Flow

or Q Sepharose Fast Flow (Pharmacia-LKB). Neither derivative was solubilized by this treatment. We therefore concluded that denaturation would be necessary to release the polymerase domain from the insoluble cell pellet.

Although, in the previous study (15), the polymerase domain had been purified by guanidinium hydrochloride denaturation and subsequent renaturation, the yield had been very low and the purified protein aggregated. We hoped that the combination of an improved denaturation-renaturation protocol and the use of re-engineered polymerase domain derivatives would be more successful. We chose the re-engineered NNNS and NNNK derivatives for these experiments, and compared them with the original clone, LFMV (which has the additional six amino acids at the N-terminus).

We used a detergent extraction protocol (25) as an initial purification step in order to purify the inclusion bodies of insoluble protein away from the membrane porin proteins which are the major contaminants in the cell pellet. This left the polymerase domain >90% pure as judged by SDS-polyacrylamide gel electrophoresis. We then solubilized the inclusion bodies in 8 M urea and carried out batch anion exchange chromatography in the presence of urea. Following this, the extracted proteins were renatured using two different approaches to avoid the potential problem that proteins, when refolded in concentrated solution, may aggregate as denaturant is removed rather than folding into the native structure. The first method involved renaturing protein bound to an ion-exchange column. This should discourage aggregation because individual protein molecules would be isolated on the column and would not have

Figure 3. Insolubility of re-engineered polymerase domain derivatives when expressed from the λ P_L promoter. A whole cell extract ('slurry') from cells over-expressing a re-engineered polymerase domain shows high level expression. Fractionation of this extract into the insoluble pellet and the supernatant (see Methods) shows that there is no detectable polymerase domain in the soluble fraction.

the opportunity to interact with others while refolding. The column method would also eliminate the need for a high dilution step (as in the second method, see below) and so would be more convenient when purifying large amounts of protein. The inclusion body proteins were applied to an anion exchange column in the presence of 8 M urea. The urea was removed stepwise and then the polymerase domain was eluted with NaCl. Although the yield from this method was quite high (approx. 60%) and the polymerase domain derivatives isolated in this way were >90% pure, they had no detectable activity. In the second method, the denatured protein solution was substantially diluted and then renatured by a series of dialysis steps to remove the urea. The renatured protein was then recovered and concentrated by batch anion exchange chromatography. The polymerase domain derivatives isolated by this second approach were >90% pure but the yield (approx. 10%) was lower than in the first method, presumably as a result of the dilution step; the specific activities in the standard polymerase assay were 0.2% of that of wild-type Klenow fragment. Comparing our denaturation-renaturation results with those of Freemont *et al.* (15); one can see that the yields were much greater in the present study, approximately 10% or 60% (depending on the protocol), compared with 2%. However, the specific activities of the polymerase domain derivatives purified in our study were at least ten-fold lower. As pointed out by Freemont *et al.* (15), it is possible that the column chromatography step in the previous method effectively filtered out the most aggregated (and probably least active) material leaving the final preparation enriched for the least aggregated, most active enzyme.

The low specific activity of the product obtained by two different denaturation/renaturation procedures raised the possibility that, despite re-engineering, the polymerase domain might not be amenable to this approach. We therefore decided against further investigation of the parameters involved in the refolding process (e.g. ref. 16) and instead concentrated our

Figure 4. Effect of temperature and time of induction on the solubility of Klenow fragment and the polymerase domain when expressed from the T7 promoter. The first set of five lanes shows an experiment using an overproducer for the NNHK polymerase domain, whose behavior was typical of the domain derivatives tested. The second set of five lanes is from an overproducer which expresses wild-type Klenow fragment from the T7 promoter. In each case the first lane (marked '0') is the soluble fraction from an uninduced cell extract. The other four lanes show the soluble fraction from cells induced for the indicated times at 20°C or 37°C. The positions of molecular weight standards (in kDa) and of Klenow fragment ('K') and the polymerase domain ('P') are shown.

efforts on expressing one or more of the new polymerase domain derivatives in a soluble form.

Effect on solubility of changing the expression system

Our preliminary results with the polymerase domain derivatives expressed from the λ P_L promoter encouraged us to vary expression conditions to maximize the yield of soluble protein. Western blot analysis of SDS-polyacrylamide gels of the soluble cell extracts from cultures induced at 37°C showed that several of the polymerase domain derivatives were slightly soluble (<1%). In other systems, solubility has been increased by lowering the temperature at which inductions were carried out (e.g. refs. 30,31). When we induced cells at room temperature the level of expression of the polymerase domain was <1% of that obtained at 37°C, so that Western blot analysis was required even for detection of the polymerase domain in the total cell protein; however, several of the proteins were now almost completely soluble. We therefore re-cloned a subset of the polymerase domain derivatives into an expression vector based on a bacteriophage T7 promoter (26), from which high level expression could be achieved at room temperature. We choose 11 derivatives (indicated in Table I) that appeared to be the most soluble in the initial Western blot analysis of supernatants from the room temperature and 37°C inductions and, for comparison, one (NNKN) that appeared to be totally insoluble. We also cloned the original polymerase domain construct and whole Klenow fragment into the same expression system. Cells carrying the resulting plasmids were induced at various temperatures for a range of times to establish conditions that optimized expression level and solubility.

Each of the 14 proteins tested showed decreased solubility as either the temperature or the time of induction was increased (Fig. 4). It did not seem that the level of solubility of the overproduced protein was related in a simple way to the concentration of protein

Figure 5. Purification of the NNNN polymerase domain. Lane A, soluble cell extract; lane B, pooled fractions from Mono Q; lane C, pooled fractions from Bio-Rex 70; lane D, pooled fractions from phenyl Superose; lane E, peak fraction from Superose 12.

in the cell, nor was there a temperature that represented a cut-off between solubility and insolubility. Rather, at any given temperature, the expressed protein was completely soluble early in the induction period, and then, as the intracellular concentration increased, a gradually increasing fraction became insoluble, until, at a critical concentration, the protein became completely insoluble. For example, when cells were induced at 37°C (Fig. 4) there were substantial amounts of soluble protein after 3 hours of induction but the induced protein was no longer detectable in the soluble fraction by 24 hours, even though the concentration of protein in the cells was still high (not shown). At a lower induction temperature (20°C), the same trend in induced protein solubility was seen, but on a longer time scale. The behavior of all of the re-engineered domain derivatives was the same, regardless of the level of solubility detected in the earlier experiments. Moreover, Klenow fragment and the original polymerase domain clone behaved the same as the re-engineered derivatives. This contrasts sharply with the situation when these same proteins are expressed at high levels from the λ P_L promoter; in this case Klenow fragment is soluble and all of the polymerase domain derivatives are essentially insoluble.

Purification of the soluble polymerase domain derivatives

Initially, we chose 5 re-engineered polymerase domain derivatives (listed in Table II) and partially purified them on a small scale using FPLC (see Methods). The peaks obtained after two columns were approximately 95% pure as judged by SDS-polyacrylamide gel electrophoresis. Moreover, the elution position on gel filtration was as expected for polymerase domain monomers; no polymerase domain was seen in the void volume, in contrast to the results of Freemont *et al.* (15). The original wild-type polymerase domain behaved differently from the re-engineered

Table II. Polymerase activity of polymerase domain derivatives.

Protein	units/mg ^a	% of Klenow fragment activity ^b
<i>Homogeneous protein</i>		
NNNN	251	2.3
<i>Partially purified fractions</i>		
QHQN	197	2.6
NNKN	190	2.5
NNNS	177	2.3
NNNN	150	2.0
NNNK	174	2.3
<i>Crude extracts^c</i>		
QNQK	221	2.9
QDHN	131	1.7
NDHR	108	1.4
NNHK	179	2.3
HNNS	350	4.6
QKHN	118	1.5
LFMV ^d	122	1.6

^a Specific activities in units/mg were measured under the standard assay conditions (28) with poly[d(AT)] as substrate. Each value is the average of at least two determinations.

^b A sample of pure Klenow fragment assayed at the same time as the partially purified and crude samples had a specific activity of 7,670 units/mg. The homogeneous protein was assayed in a separate experiment, when the standard had a value of 10,800 units/mg.

^c These assays were more variable than those using the purer fractions, most probably due to the difficulty of pipetting the viscous lysate.

^d The original polymerase domain derivative described by Freemont *et al.* (15).

derivatives in our purification procedure, even though we had established conditions under which this protein was soluble in the overproducer cell extract. Only a small fraction (approx. 5%) of the wild-type domain eluted from the Mono Q column at a similar position to the re-engineered domains, while the remaining 95% could only be removed by stripping the column with acid and base. (The small amount that eluted normally from the Mono Q column chromatographed as a monomer on gel filtration.) Since we observed no difference between the five clones tested in the initial purification studies, we chose a single derivative (NNNN) and purified it to homogeneity on a larger scale, as described in Methods (Fig. 5). Approximately 4 mg of pure polymerase domain were obtained from 5 g induced cells.

The re-engineered polymerase domain derivatives, purified by either of the procedures described above, were assayed for polymerase activity in the standard poly[d(A-T)] assay (28). In each case the polymerase specific activity was 2–3% of that of wild-type Klenow fragment under the same assay conditions (Table II). Unfortunately, as described above, the original wild-type polymerase domain was not recovered from the purification in sufficient yield to allow a comparison with the re-engineered polymerase domain derivatives. However, a rough comparison of polymerase activity in crude cell extracts (Table II) indicated that both the original and the re-engineered constructs have similar activity, and ruled out the possibility that the mutations on the newly exposed face of helix G may have affected the polymerase activity. Freemont *et al.* (15) had reported a specific activity around 10% of that of Klenow fragment for the original wild-type polymerase domain fractionated from a soluble extract. Given the variability we have observed from one batch to another of poly[d(A-T)] and the difficulty of making reliable estimates of protein concentration in impure samples, we feel it is unlikely that this apparent difference in specific activity in the two studies

represents a significant inconsistency. None of the five partially purified domain derivatives showed any significant increase in polymerase activity in the presence of 12% (v/v) glycerol, unlike the previous derivative whose activity increased fivefold under these conditions (15). Although this result is subject to the *caveat* that we were not able to test a comparable extract of the original polymerase domain derivative, it could indicate that re-engineering had alleviated the structural problems of the original construct.

Determination of the steady-state kinetic parameters for the polymerase reaction catalyzed by the homogeneous NNNN polymerase domain on the homopolymer substrate, poly(dA)-oligo(dT), gave a $K_{m(dNTP)}$ of 1.4 μM , very close to the previously determined value of 2.8 μM for wild-type Klenow fragment (9). By contrast k_{cat} was about three orders of magnitude lower than the wild type value: $1.2 \times 10^{-3} \text{ s}^{-1}$, compared with 2.4 s^{-1} . The dissociation constant for binding of the polymerase domain to a small double-stranded DNA oligonucleotide was 13 nM, compared with 0.09 nM for the D424A (exonuclease-deficient) mutant of Klenow fragment on the same DNA substrate.

DISCUSSION

The aim of this study was to provide a reliable means for generating large quantities of the polymerase domain of Klenow fragment for use in structural and other studies. We have succeeded in this objective, as shown by the purification to homogeneity of the NNNN domain derivative described above. From the results of this study, we now understand that the difficulties encountered in previous attempts to overproduce the polymerase domain (15) are attributable to two apparently distinct causes. The first was insolubility of the overproduced protein, a problem that we have solved by altering the expression vector and induction protocol. The second problem was the unsatisfactory chromatographic behavior of the original overproduced domain, which we attributed to aggregation, and have solved by re-engineering the immediate N-terminal region of the protein.

Structural considerations suggested that the tendency of the original domain to aggregate could be due either to the six potentially disordered amino acid residues at the N-terminus, or to a series of hydrophobic side chains on the newly-exposed face of helix G. We therefore made a library of derivatives having a variety of hydrophilic substitutions on the face of helix G and lacking the six N-terminal amino acids that preceded the start of helix G. In our studies of this library, there was every indication that all the derivatives behaved identically to one another. Of the twelve derivatives that were investigated in some detail, all had the same solubility properties, being insoluble when expressed from the λ P_L promoter and soluble when expressed under appropriate conditions from a T7 promoter. Moreover, the five derivatives tested behaved identically through several fractionation steps, chromatographing as monomeric non-aggregating proteins. By contrast, the original domain derivative did not chromatograph as a single species, resulting in huge losses on Mono Q. It appears that only a small fraction of this derivative is present in a monomeric form in solution. Thus, although the solubility of the original domain derivative could be improved by changing the expression system, re-engineering the N-terminal region was necessary in order to obtain a chromatographically well-behaved product.

Our observation that all of the re-engineered derivatives tested, regardless of the exact nature of their hydrophilic substitutions, were successful in alleviating the presumed aggregation problems of the original clone is perhaps not surprising, for several reasons. First, the mutations had been chosen so as to be compatible with the local protein structure. Additionally, it is likely that the protein structure can tolerate a wide range of polar substitutions in surface residues, as has been found in other systems (e.g. refs. 32–35). It is also possible that the hydrophilic substitutions were irrelevant and that it was the removal of the six N-terminal amino acids that was responsible for the improved chromatographic behavior of the re-engineered polymerase domain derivatives. Since the constructs described above allowed us to achieve our objective of overproducing a soluble, non-aggregating polymerase domain, we did not make the related construct having wild-type sequence but lacking the six N-terminal amino acids. Thus the available data do not rule out the latter explanation.

Although we have attributed the problems with the original polymerase domain construct to aggregation, this is not the only possible explanation. Two recent studies have demonstrated an increase in protein stability when a hydrophobic surface residue is replaced by a polar residue (36, 37). In both cases the authors argue persuasively that their observations are best explained by the amino acid substitutions affecting the equilibrium between native and denatured states of the protein. In a similar way, the chromatographic behavior of the original polymerase domain construct, which contains the group of exposed hydrophobic residues, could be explained by a portion of the protein adopting an unfolded conformation. Any explanation for the properties of this protein, whether invoking aggregation or denaturation, should take into account the fact that the polymerase specific activities of the original domain and its re-engineered derivatives are essentially the same. This observation is probably compatible with aggregation or rather limited unfolding of the original construct, but not with extensive denaturation.

Several of the results of this study serve to illustrate the current lack of understanding of the factors involved in protein overproduction and purification. There are many examples of proteins that are insoluble when over-expressed but can be purified by denaturation and renaturation to give soluble, active proteins (38). However, in the present study, despite using strategies designed to maximize the chances of success, isolation of the polymerase domain from the insoluble cell pellet was not a successful approach, since the final product had little or no enzymatic activity. Perhaps 'unnatural' proteins, such as protein fragments or domains, are less good subjects for this technique, since the refolding process may be more difficult than for the native molecule. Even more surprising was the dramatic influence of the expression system on the solubility of the protein product. We had predicted that the T7 system, allowing lower induction temperatures than the λ P_L system, might improve the solubility of the polymerase domain derivatives, as indeed was the case. However, we were surprised to find that Klenow fragment, which is completely soluble when expressed at 37°C or 42°C from the λ P_L promoter, was much less soluble when using a T7 expression plasmid. In contrast with the λ P_L system, all of the proteins studied showed identical solubility behavior in the crude extract when using the T7 expression system. There was no difference when comparing the intact parent molecule (Klenow fragment) with the domains, or when comparing the original domain construct, which tended to aggregate, with the re-engineered monomeric derivatives. We also do not have a

definitive explanation for the peculiar time course we have observed for protein solubility in the T7 system. Early in the induction period there are substantial amounts of soluble protein until, it seems, a threshold level is reached and the overproduced protein disappears from the soluble fraction. A possible explanation is that the cytoplasm initially becomes supersaturated with the polymerase domain. Alternatively, following the suggestion of Darby and Creighton (39), insolubility may be a consequence of binding to another cellular component, presumably one that accumulates relatively late in the induction period. It is also conceivable that this phenomenon may be more general than we had suspected, but we are unaware of any data that address this point. In conclusion, it is very clear from the present study that, when attempting to solve a difficult overproduction problem, one should be prepared to test a range of expression systems, induction protocols and purification strategies.

The homogeneous polymerase domain had a specific activity about 2% of that of Klenow fragment in the standard poly[d(A-T)] assay, consistent with previous data (15). The K_m for dNTP utilization for the homogeneous NNNN derivative was very similar to the value for wild-type Klenow fragment, as might be expected since the dNTP binding region is likely to be completely contained within the polymerase domain. The value of $K_{m(dNTP)}$ and the lack of an effect of glycerol on the polymerase activity contrast with the observations of Freemont *et al.* (15) and may suggest that the re-engineered derivatives have a more stable structure than the previous polymerase domain construct. Although $K_{m(dNTP)}$ for the NNNN polymerase domain was essentially wild-type, the value of k_{cat} was about 10^3 -fold lower than the wild-type value. We suspect that this decrease may be related to the impaired DNA binding of the isolated domain, just as we have observed for Klenow fragment point mutants that affect DNA binding (5), since we would not expect the removal of the exonuclease domain to have a particularly significant effect on the rate of chemical catalysis at the polymerase active site. The decreased DNA binding resulting from loss of the exonuclease domain implies, not unexpectedly, that the exonuclease domain provides some of the binding contacts to a DNA substrate at the polymerase active site, but does not of itself discriminate between models that have been put forward for DNA binding (12, 40). A detailed examination of the geometry of the polymerase domain footprint should permit this distinction to be made.

ACKNOWLEDGEMENTS

We are grateful to Lorena Beese and Tom Steitz for their advice on structural aspects of the mutagenesis of surface residues of the polymerase domain, to Nigel Grindley for his interest and insightful comments on this work, to Keith Derbyshire for advice and helpful discussions, to Xiaojun Chen Sun for expert technical assistance, and to Kate Tatham for her help in preparation of the manuscript. This work was supported by Grant GM-28550 from the National Institutes of Health to Nigel D.F. Grindley

REFERENCES

- Kornberg, A. and Baker, T.A. (1992) *DNA Replication*, 2nd Ed., pp. 113–164, Freeman Publications, San Francisco.
- Brutlag, D., Atkinson, M.R., Setlow, P. and Kornberg, A. (1969) *Biochem. Biophys. Res. Commun.* **37**, 982–989.
- Klenow, H. and Henningsen, I. (1970) *Proc. Natl. Acad. Sci. U.S.A.* **65**, 168–175.
- Joyce, C.M. and Grindley, N.D.F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1830–1834.
- Polesky, A.H., Steitz, T.A., Grindley, N.D.F. and Joyce, C.M. (1990) *J. Biol. Chem.* **265**, 14579–14591.
- Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. and Steitz, T.A. (1985) *Nature (London)* **313**, 762–766.
- Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L., Friedman, J.M., Joyce, C.M. and Steitz, T.A. (1988) *Science* **240**, 199–201.
- Derbyshire, V., Grindley, N.D.F. and Joyce, C.M. (1991) *EMBO J.* **10**, 17–24.
- Polesky, A.H., Dahlberg, M.E., Benkovic, S.J., Grindley, N.D.F. and Joyce, C.M. (1992) *J. Biol. Chem.* **267**, 8417–8428.
- Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R. and Steitz, T.A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8924–8928.
- Beese, L.S. and Steitz, T.A. (1991) *EMBO J.* **10**, 25–33.
- Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) *Science* **260**, 352–355.
- Guest, C.R., Hochstrasser, R.A., Dupuy, C.G., Allen, D.J., Benkovic, S.J. and Millar, D.P. (1992) *Biochemistry* **30**, 8759–8770.
- Steitz, T.A., Beese, L., Freemont, P.S., Friedman, J. and Sanderson, M.R. (1987) *Cold Spring Harbor Symp. Quant. Biol.*, **52**, 465–471.
- Freemont, P.S., Ollis, D.L., Steitz, T.A. and Joyce, C.M. (1986) *Proteins* **1**, 66–73.
- Claassen, L.A., Byungchan, A., Koo, H.-S. and Grossman, L. (1991) *J. Biol. Chem.* **266**, 11380–11387.
- Laemmli, U. (1970) *Nature* **227**, 680–685.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248–254.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Joyce, C.M., Kelley, W.S. and Grindley, N.D.F. (1982) *J. Biol. Chem.* **257**, 1958–1964.
- Vieira, J. and Messing, J. (1987) *Methods Enzymol.* **153**, 3–11.
- Grindley, N.D.F. and Joyce, C.M. (1980) *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 7176–7180.
- Rosenberg, M., Ho, Y. and Shatzman, A. (1983) *Methods Enzymol.* **101**, 123–138.
- Nagai, K. and Thøgersen, H.C. (1987) *Methods Enzymol.* **153**, 461–481.
- Studier, F.W., Rosenberg, A.H. and Dunn, J.J. (1990) *Method Enzymol.* **185**, 60–89.
- Studier, F.W. (1991) *J. Mol. Biol.* **219**, 37–44.
- Setlow, P. (1974) *Methods Enzymol.* **29**, 3–12.
- Hoess, A., Arthur, A.K., Wanner, G. and Fanning, E. (1988) *Bio/Technology* **6**, 1214.
- Lin, T.-C., Rush, J., Spicer, E.K. and Konigsberg, W.H. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7000–7004.
- Pan, H., Clary, D. and Sadowski, P.D. (1991) *J. Biol. Chem.* **266**, 11347–11354.
- Perutz, M.F., Kendrew, J.C. and Watson, H.C. (1965) *J. Mol. Biol.* **13**, 669–678.
- Bashford, D., Chothia, C. and Lesk, A.M. (1987) *J. Mol. Biol.* **196**, 199–216.
- Reidhaar-Olson, J.R. and Sauer, R.T. (1988) *Science* **241**, 53–57.
- Reidhaar-Olson, J.R. and Sauer, R.T. (1990) *Proteins* **7**, 306–316.
- Pakula, A.A. and Sauer, R.T. (1990) *Nature* **344**, 363–364.
- Bowler, B.E., May, K., Zaragoza, T., York, P., Dong, A. and Caughey, W.S. (1993) *Biochemistry* **32**, 183–190.
- Marston, F.A.O. (1986) *Biochem. J.* **240**, 1–12.
- Darby, N.J. and Creighton, T.E. (1990) *Nature* **344**, 715–716.
- Joyce, C.M. and Steitz, T.A. (1987) *Trends Biochem. Sci.* **12**, 288–292.