

Characterization of the *Escherichia coli thyA* gene and its amplified thymidylate synthetase product

(λ thyA⁺ transducing phage/subcloning by random resection/thymidylate synthetase purification)

MARLENE BELFORT, GLADYS F. MALEY, AND FRANK MALEY

The Center for Laboratories and Research, New York State Department of Health, Empire State Plaza, Albany, New York 12201

Communicated by Phillips W. Robbins, December 27, 1982

ABSTRACT The 7.8-kilobase *Hind*III insert in phage λ NM589*thyA* [Borck, K., Beggs, J. D., Brammar, W. J., Hopkins, A. S. & Murray, N. (1976) *Mol. Gen. Genet.* 146, 199] was confirmed as originating from *Escherichia coli* by hybridization analysis and was shown to encode the thymidylate synthetase (5,10-methylenetetrahydrofolate:dUMP C-methyltransferase EC 2.1.1.45) of *E. coli* K-12 by using biochemical, structural, and immunologic criteria. The 7.8-kilobase insert was reduced in size to a quasi-random population of DNA subfragments by partial digestion with the 4-base-pair recognition enzymes *Alu* I and *Hae* III. A clone containing a 1.1- to 1.2-kilobase fragment that encompassed the gene was obtained from this mixture by selecting for Thy⁺ recombinants. Fusion of this DNA fragment to the phage λ p_L promoter in plasmid pKC30 revealed the direction of transcription of the *thyA* gene, and, in a phage λ lysogen containing a thermolabile repressor, intracellular synthetase levels were increased about 700-fold. The enzyme was purified to homogeneity from this source by affinity chromatography, and some of its properties are described.

Thymidylate synthase (5,10-methylenetetrahydrofolate:dUMP C-methyltransferase; EC 2.1.1.45) was first described by Friedkin and Kornberg (1) in extracts of *Escherichia coli* as an activity that catalyzes the formation of dTMP from dUMP in the presence of 5,10-methylenetetrahydrofolate (5,10-CH₂-H₄PteGlu). Because dTMP is an essential precursor of DNA synthesis, the synthetase has become a target for attack by various chemotherapeutic agents, such as 5-fluorouracil and 5-fluorodeoxyuridine, which exert their action by being converted to fluorodeoxyuridylate (FdUMP) (2). To understand the mechanism of action of these agents, in addition to the unique catalytic reaction promoted by this enzyme (3), an extensive investigation of the synthetase structure has been undertaken. Most of these studies have been conducted on the *Lactobacillus casei* enzyme because of its relative abundance (4, 5) and have led to a clarification of its complete amino acid sequence and the location of its active site (6).

New approaches to understanding the regulation and mode of action of the enzyme are provided by studying thymidylate synthetase genes, which have been cloned from a variety of viral (7–9) and cellular (10–12) sources. Aside from the potential for providing large quantities of enzyme (13), these cloned genes afford the possibility of both mutational analysis and phylogenetic comparison (13–15). The *E. coli thyA* gene was selected for this study because of the relatively facile nature of gene manipulation in *E. coli* and the ability to select *thy*[−] mutants. Phylogenetic comparison also is facilitated in *E. coli* because the thymidylate synthetase gene from organisms as distantly related as yeast are known to be expressed in *E. coli* and to com-

plement *thyA* mutants (12). Further, these studies should provide a basis for understanding the different properties between *E. coli* thymidylate synthetase and the corresponding enzyme produced on infection with T2, T4, or T6 phage (14).

Our studies are based on the *thyA* gene, which was first isolated on a specialized λ thyA⁺ transducing phage in the laboratory of Murray (10). There have been subsequent reports on the cloning of the *E. coli thyA* gene, by Rubin *et al.* (16) and recently by Hickson *et al.* (11). Major discrepancies in the genetic and enzymic properties of the two former *thyA* isolates (10, 16) have required us to rigorously confirm in the ensuing experiments that the *thyA* fragment in λ NM589*thyA* is directly derived from the *E. coli* genome. The major objectives of this study are embodied in the subsequent subcloning and amplification experiments, which facilitate investigation of the *E. coli thyA* gene and its thymidylate synthetase product.

MATERIALS AND METHODS

Bacterial and Bacteriophage Strains. Wild-type bacterial strains used in this study are the *E. coli* K-12 strain W3110 and *Bacillus subtilis* 168. Other *E. coli* K-12 strains used include Rue10, which is HB101*thyA* (received from G. Wilson), and the lysogens Rue10(λ cI⁺) and Rue10(λ cI857S7), which carry prophages with either a wild-type (cI⁺) or a temperature-sensitive (cI857) phage λ repressor. These strains were used as our standard highly transformable *thyA*[−] host cells. UC5826, which has the genotype N99*hfl*[−] *recAsu*[−] (λ int6cI857*cro*27Pam3) (obtained from D. Wulff), was used for enzyme overproduction by resident plasmids constructed during the course of this work. The defective prophage in this strain carries a *cro*[−] mutation, which allows for enhanced enzyme production (17). λ NM589*thyA* is a λ thyA⁺ transducing phage generously provided to us by N. Murray. The phage consists of a 7.8-kilobase (kb) *thyA*-containing *Hind*III fragment of *E. coli* cloned in the *Hind*III insertion vector λ NM540 (10).

DNA Procedures. Plasmid DNA was prepared by the method of Clewell and Helinski (18), and filter hybridization was carried out as described by Thomas (19). DNA restriction and modifying enzymes and *Hind*III linkers were obtained from Bethesda Research Laboratories or New England BioLabs and were used according to the manufacturers' instructions.

The strategy for resecting flanking sequences from the *thyA* gene involved partial digestion of pBTA with the 4-base-pair (bp) recognition enzymes *Alu* I and *Hae* III (20) (Fig. 1). Three reactions were set up, each containing 10 μ g of DNA and both restriction endonucleases at enzyme/DNA ratios of 0.1, 0.3, and 1.0 unit/ μ g. After a 1-hr incubation at 37°C, digestion products ranged in size from 0.3 to 10 kb. Fragments between 1.0 and 1.6 kb were coeluted from a 0.8% agarose gel onto DEAE-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: 5,10-CH₂-H₄PteGlu, 5,10-methylenetetrahydrofolate; kb, kilobase(s); FdUMP, fluorodeoxyuridylate; bp, base pairs.

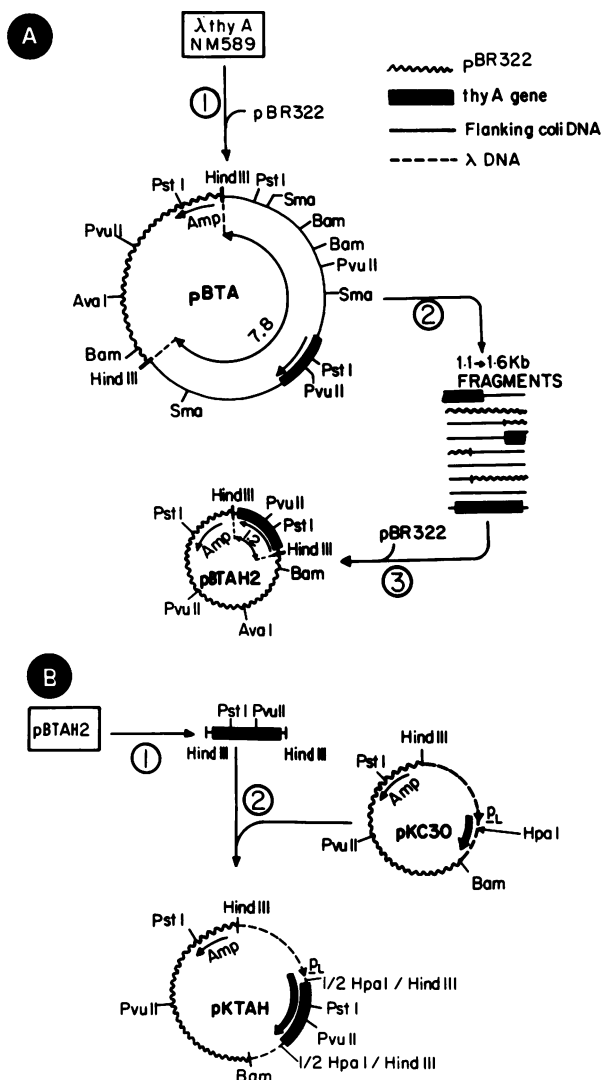


FIG. 1. (A) Subcloning the *thyA* gene into pBR322 and removal of flanking sequences. The 7.8-kb *E. coli thyA HindIII* fragment was excised from λ thyA and ligated into the *HindIII* site of pBR322 to yield pBTA recombinants that impart *Thy*⁺ character to *thyA*⁻ *E. coli* (①). A quasi-random population of fragments in the 1.0- to 1.6-kb range was generated as described (②). The fragments were ligated to *HindIII* linkers and cloned into pBR322. Clones containing an intact *thyA* gene were selected by complementation of *thyA*⁻ cells. pBTAH2, which contains a 1.2-kb insert, was chosen for further study (③). (B) Construction of expression plasmid pKTAH. The 1.2-kb *thyA* fragment, purified from pBTAH2, was cloned into the *Hpa*I site of pKC30 adjacent to phage λ *p_L*. pKTAH denotes the recombinant shown, whereas pKHAT hybrids have the insert in reverse orientation. Single-headed arrows in A and B point to the direction of transcription.

cellulose membrane (Schleicher & Schuell, NA45) (21). *HindIII* linkers were ligated onto the flush *Alu*I or *Hae*III ends of the purified fragments (linker-to-fragment ratio, \approx 100:1), which were then cloned into the *HindIII* site of pBR322, and the mixture was used to transform competent Rue10 cells. The 1.1- to 1.2-kb fragment residing in one *Thy*⁺ recombinant (pBTAH2) was purified. The 5' extensions created by *HindIII* were filled in with DNA polymerase I (22), and the fragment was ligated into the *Hpa*I site of expression plasmid pKC30, adjacent to the *p_L* promoter of phage λ , as first described by Shimatake and Rosenberg (23).

Enzyme Procedures. The conditions for cell growth, temperature shift, extract preparation, and ternary complex formation as outlined in the figure legends have been detailed (13).

Thymidylate synthetase was assayed spectrophotometrically (24) and purified by the quinazoline affinity-column procedure (25). One unit of activity is defined as the amount of enzyme required to convert 1 μ mol of dUMP to dTMP per min at 30°C. Conditions for NaDodSO₄ gel electrophoresis are those as described (26). Immunodiffusion of antibody to the purified synthetase was performed on a microscope slide (27).

RESULTS AND DISCUSSION

Identification of the Origin of the Cloned *thyA* Fragment. Verification that the 7.8-kb *HindIII* insert, which confers *Thy*⁺ character to transducing phage λ NM589*thyA*, is a direct derivative of the *E. coli* genome was necessitated by the following findings: (i) structural, kinetic, and immunologic properties of the enzymes encoded by the two putative *thyA* clones available to us (10, 16) were inconsistent (unpublished data); (ii) no homology was detectable between the two *Thy*⁺ fragments in hybridization experiments (unpublished data); and (iii) contradictions in restriction enzyme sites surrounding these two genes are apparent (16, 28). The lack of correspondence of restriction enzyme sites between the clone of Rubin *et al.* (16) and the *thyA* clone of Hickson *et al.* has been reported also (11).

To confirm the identity of the *Thy*⁺ insert in λ NM589*thyA*, a Southern blot analysis was performed (Fig. 2). pBTA, which was constructed by subcloning the 7.8-kb *Hind* fragment into pBR322 (Fig. 1), was used as a hybridization probe in this experiment. It is clear from the blot that the 7.8-kb insert in pBTA (Fig. 2, lane 2) corresponds in size to the hybridizing band in the *HindIII* digest of *E. coli* chromosomal DNA. Double digestions of pBTA and *E. coli* DNA with *HindIII*/*Ava*I, which cleaves the insert fragment at least four times, further support the or-

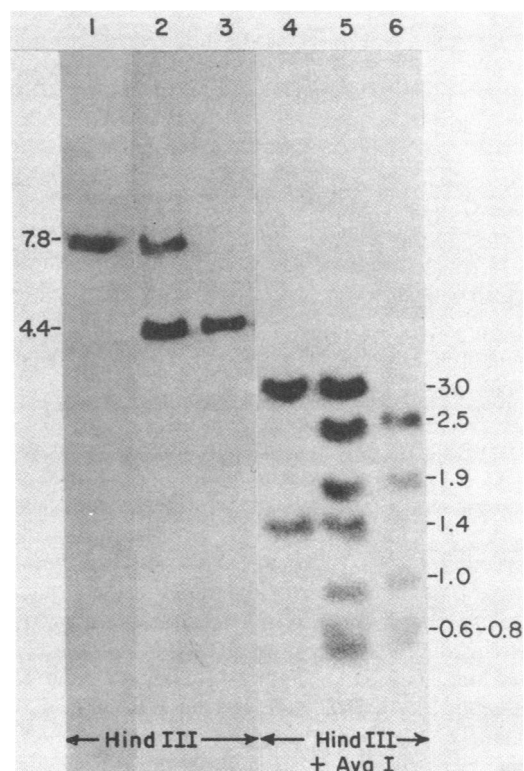


FIG. 2. Hybridization analysis of the *thyA* insert in pBTA. DNA fragments resulting from digestion with *HindIII* (Left) and *HindIII*/*Ava*I (Right) and separated on an 0.8% agarose gel were probed with ³²P-labeled pBTA after transfer onto nitrocellulose membrane (19). Lanes: 1 and 6, *E. coli* DNA; 2 and 5, pBTA DNA; 3 and 4, pBR322 DNA. Sizes are shown in kb.

igin of the 7.8-kb fragment. In addition to the two vector bands (Fig. 2, lane 4) pBTA has five clearly visible insert bands in the 0.6- to 2.5-kb range (Fig. 2, lane 5), and five corresponding bands hybridize to the probe in the lane containing total *E. coli* chromosomal DNA (lane 6).*

The ability to demonstrate the ternary complex formed by thymidylate synthetase, the substrate analog FdUMP, and 5,10-CH₂-H₄PteGlu in crude cell-free extracts (13) has enabled us to compare the size of the native *E. coli* K-12 enzyme with that encoded by the cloned gene. The fluorogram depicted in Fig. 3 shows that the ternary complex produced in extracts of wild-type *E. coli* K-12 (Fig. 3, lane 2) appears identical in *M_r* to that which is present after infection or transformation of *thyA*⁻ *E. coli* with λNM589*thyA* (Fig. 3, lane 3) or pBTA (Fig. 3, lane 4), respectively. Based on these corroborative data, we have used the clone of Borck *et al.* (10) for all subsequent experiments.

Resection of Flanking DNA from the *thyA* Gene. Restriction sites for the enzymes *Bam*HI, *Pst* I, *Pvu* II, and *Sma* I occurring with the 7.8-kb *Hind*III fragment are shown in Fig. 1A. It was possible to show by subcloning experiments that the entire *thyA* gene resides within the 4-kb *Sma* I fragment. However, this fragment proved unsuitable for amplifying expression of the *thyA* gene. Because of inconvenient placement of cleavage sites for all of 13 other 6-bp recognition restriction enzymes tested, we resorted to excising the *thyA* gene by partial digestion with the two 4-bp recognition endonucleases *Alu* I and *Hae* III (20). Based on our estimate of the size of the *thyA* gene as being ≈1.3 kb (*ca.* 900 bp, corresponding to *ca.* 300 amino acids plus regulatory sequences), fragments in the 1.0- to 1.6-kb range were purified from the above partial digest of pBTA (Fig. 1A). After this mixed population of DNA fragments was cloned into pBR322 by using *Hind*III linkers, recombinants containing the intact *thyA* gene were selected on the basis of their ability to complement *thyA*⁻ host cells. *Thy*⁺ clones, representing *ca.* 0.4% of the total recombinant population, contained *Hind*III inserts in the 1.1- to 1.6-kb range. Six of these were screened on gel fluorograms to obtain the size of the gene product by ternary complex formation. All produced an enzyme indistinguishable in size from the wild-type *E. coli* synthetase. The data are shown only for representative recombinant pBTAH2 (Fig. 3, lane 6), which carries the intact *thyA* gene on a *Hind*III fragment of ≈1.2 kb. This approach to subcloning a gene that has its major functional components intact and yet is flanked by a minimum of extraneous DNA is readily applicable to any gene for which there is a positive selection.

Amplification of *thyA* Gene Expression. Amplification of *thyA* expression was achieved by fusing the 1.2-kb fragment to the phage λ *p_L* promoter in multicopy plasmid pKC30 (Fig. 1B). These pKC30-*thyA* hybrids were grown in phage λ lysogens carrying a thermolabile *cI857* repressor. Upon elevation of the temperature to 42°C, an additional band became apparent in extracts of cells carrying pKTAH recombinants, which have the *thyA* insert in transcriptional alignment to *p_L* (Fig. 4A, lane 3). The identity of the protein band, which is indicated by a pointer, was confirmed by its ability to form a ternary complex on incubating the extract with [¹⁴C]FdUMP and 5,10-CH₂-H₄PteGlu. This was evidenced by a small but reproducible mobility shift of the band relative to uncomplexed protein and by its specific retention of label (Fig. 4A, lanes 4 and 4', respec-

* Contrary to the published report, no homology between the putative *thyA* clone of Rubin *et al.* (16) and the *E. coli* chromosome was detectable. Therefore, all of the spurious properties of this clone seem attributable to its having been inadvertently derived from an organism other than *E. coli*.

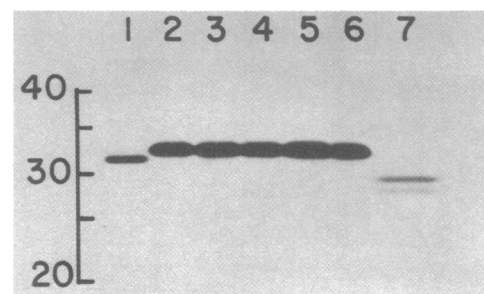


FIG. 3. Thymidylate synthetase ternary complexes in crude extracts. Cell-free sonic extracts containing 0.02–0.1 milliunits of enzyme were treated with [³H]FdUMP and 5,10-CH₂-H₄PteGlu as described (13) before separation on a 12.5% NaDodSO₄/polyacrylamide gel. This fluorogram shows the ternary complex made with extracts prepared from W3110 (*thy*⁺) (lane 2), Rue10 (*thyA*⁻) infected with λNM589*thyA* (lane 3), Rue10/pBTA (lane 4), Rue10/pBTAH2 (lane 5), and Rue10/pKTAH (lane 6). The complex is undetectable in the Rue10 host itself under these conditions (13). For comparative purposes, lane 1 shows the subunit ternary complex of phage T4 synthetase, while lane 7 shows the 2 synthetases present in extracts of *B. subtilis* 168. The *M_r* scale (*M_r* shown × 10⁻³) was derived from a series of protein standards separated on the same gel.

tively). It also should be noted that this band is not inducible in pKHAT recombinants, where the insert is in reverse orientation (Fig. 4B, compare lanes 3 and 4), and this allowed us to infer the direction of transcription of the *thyA* gene as shown in Fig. 1.

Specific activities of synthetase in the range of 80–100 milliunits/mg of protein were initially obtained in temperature-sensitive lysogens that carry a *P*⁻ prophage replication defect

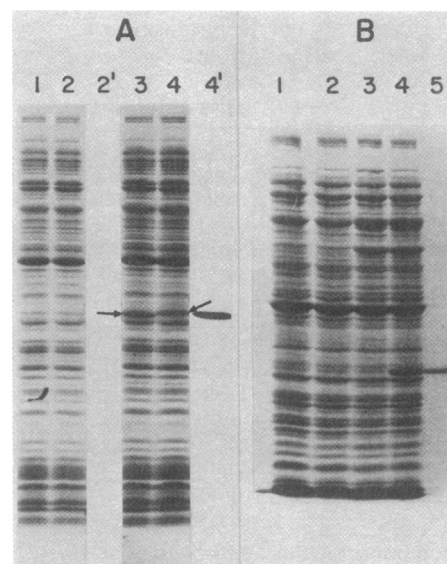


FIG. 4. Identification and purification of the amplified *thyA* gene product. (A) Cell-free sonic extracts were prepared from a logarithmic phase culture of UC5826/pKTAH grown at 32°C (lanes 1 and 2) and after incubation at 42°C for 1 hr (lanes 3 and 4). The samples in lanes 2 and 4 were pretreated with [¹⁴C]FdUMP and 5,10-CH₂-H₄PteGlu before electrophoresis on a 12.5% NaDodSO₄/polyacrylamide gel. Lanes 2' and 4' are fluorograms of the corresponding portion of the gel. (B) Crude extracts prepared from cultures of UC5826/pKHAT (lanes 1 and 3) or UC5826/pKTAH (lanes 2 and 4) before (lanes 1 and 2) and after (lanes 3 and 4) shifting the cultures to the elevated temperature were separated on a 10% gel. The amplified synthetase band, which appears only when the *thyA* gene is correctly oriented and only after incubation at 42°C (lane 4), corresponds to the enzyme purified from such an extract on an affinity column (25) (lane 5).

(17). These yields represent an amplification factor of about 200, several-fold lower than was achieved for other synthetase genes (ref. 13; unpublished data). This is possibly due to the presence of a control element that we have observed in the 5' region of the gene, which acts in the opposite direction to the *thyA* promoter and which may inhibit p_L by initiating convergent transcription (29).

After we switched to a lysogen that is not only P^- but also lacks the Cro function, which ordinarily shuts off the p_L promoter late in a phage λ infection (17), yields were improved 3- to 4-fold to 300–320 milliunits/mg of protein, which corresponds to 6–7% of total soluble protein. This represents a 700-fold increase over basal enzyme levels found in wild-type *E. coli* K-12 extracts (0.4 milliunit/mg) and a 28-fold increase over the reported amplification due to high gene copy number alone (11).

Purification and Properties of *E. coli* K-12 Thymidylate Synthetase. The enzyme was purified 20-fold to apparent homogeneity from a large-scale culture of UC5826/pKTAH after heat induction by using the quinazoline affinity column of Rode *et al.* (25). The purified synthetase at a specific activity of 5 units/mg of protein is shown as a single band in Fig. 4B (lane 5). Although the M_r estimate for the subunit is about 33,000 (based on the gels shown in Figs. 3 and 4) and is in close agreement with published values (11, 30), this may well be an overestimate according to sedimentation equilibrium studies, which suggest a M_r of 59,000 for the dimer. Other gel systems as well as our DNA sequence analyses also support the lower estimates (unpublished data).

The immunoidentity between the *E. coli* B enzyme and that from *E. coli* K-12 is shown in Fig. 5. This result is not surprising in view of the relatedness of these genomes in the *thyA* region as revealed by hybridization analysis with λ *thyA*⁺ DNA as a probe (28). Additional evidence for their relatedness is revealed by amino acid composition of the two enzymes, which appear to be identical (15), and the finding that the DNA sequence of the *E. coli* K-12 *thyA* gene agrees exactly with the amino acid sequence that we have determined for portions of both the *E. coli* B and K-12 enzymes (data to be published elsewhere). Included in this determination are the 20 amino-terminal amino acids, 10 carboxyl-terminal amino acids, and the sequence of the "active site" peptide that binds FdUMP.

The synthetases of T2, T4, and T6 phages, on the other hand, have strikingly dissimilar properties from that of their *E. coli*

host at the level of both protein and gene (Fig. 5) (13–15). This stands in dramatic contrast to the conservation of *thyA* homologous sequences in genomes of many enteric bacteria (28, 31) and to sequence homologies between the *E. coli* enzyme (ref. 15; unpublished data) and that of the relatively distantly related *L. casei* (6). These similarities and differences should ultimately relate to the various functional properties of this metabolically important enzyme.

We thank Norreen Murray, Marty Rosenberg, Dan Wulff, and Gary Wilson for providing us with strains and for useful discussions and Dennis Panicali for helpful advice. We appreciate the contributions of Glenn Eiger, Allan Moelleken, Joan Pedersen-Lane, Judy Reidl, Don Guarino, and Judy Valentino to various aspects of this work and that of Robert MacColl for the sedimentation equilibrium analyses. This investigation was supported in part by grants from the National Science Foundation (PCM8118368) and the National Institutes of Health (GM26387 and GM26645).

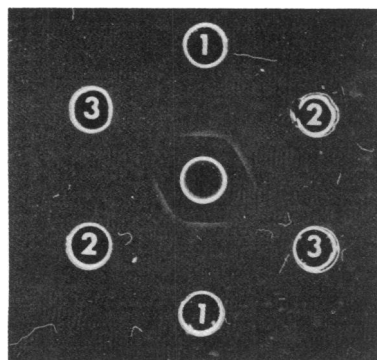


FIG. 5. Immunochemical crossreactivity. Antibody to *E. coli* B thymidylate synthetase (center well) is shown to react with purified synthetase from *E. coli* B (well 1) and from a heat-induced extract of pKTAH (well 2) but not with purified phage T4 thymidylate synthetase (well 3).

1. Friedkin, M. & Kornberg, A. (1957) in *Chemical Basis of Heredity*, eds. McElroy, W. D. & Glass, B. (Johns Hopkins Univ. Press, Baltimore, MD), pp. 609–614.
2. Danenberg, P. V. (1977) *Biochim. Biophys. Acta* **473**, 73–92.
3. Pogolotti, A. L. & Santi, D. V. (1977) *Bioorg. Chem.* **1**, 277–311.
4. Dunlap, R. B., Harding, N. G. L. & Hueneekens, F. M. (1971) *Biochemistry* **10**, 88–97.
5. Leary, R. P. & Kisliuk, R. L. (1971) *Prep. Biochem.* **1**, 47–54.
6. Maley, G. F., Bellisario, R. L., Guarino, D. U. & Maley, F. (1979) *J. Biol. Chem.* **254**, 1301–1304.
7. Duncan, C. H., Wilson, G. A. & Young, F. E. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3664–3668.
8. Mileham, A. J., Revel, H. R. & Murray, N. E. (1980) *Mol. Gen. Genet.* **179**, 227–239.
9. Gryczan, T. J. & Dubnau, D. (1982) *Gene* **20**, 459–469.
10. Borck, K., Beggs, J. D., Brammar, W. J., Hopkins, A. S. & Murray, N. E. (1976) *Mol. Gen. Genet.* **146**, 199–207.
11. Hickson, I. D., Atkinson, K. E. & Emmerson, P. T. (1982) *Gene* **18**, 257–260.
12. Taylor, G. R., Barclay, B. J., Storms, R. K., Friesen, J. D. & Haynes, R. H. (1982) *Mol. Cell. Biol.* **2**, 437–442.
13. Belfort, M., Moelleken, A., Maley, G. F. & Maley, F. (1983) *J. Biol. Chem.* **258**, 2045–2051.
14. Maley, G. F., Maley, F. & Baugh, C. M. (1979) *J. Biol. Chem.* **254**, 7485–7487.
15. Maley, G. F., Belfort, M. & Maley, F. (1982) in *Chemistry & Biology of Pteridines*, ed. Blair, J. A. (Gruyter, New York), in press.
16. Rubin, E. M., Wilson, G. A. & Young, F. E. (1980) *Gene* **10**, 227–235.
17. Herskowitz, I. (1973) *Annu. Rev. Genet.* **7**, 289–324.
18. Clewell, D. D. & Heliniski, D. R. (1969) *Proc. Natl. Acad. Sci. USA* **62**, 1159–1166.
19. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
20. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15**, 687–701.
21. Dretzen, G., Bellard, M., Sassone-Corsi, P. & Chambon, P. (1981) *Anal. Biochem.* **112**, 295–298.
22. Heffron, F., So, M. & McCarthy, B. J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 6012–6016.
23. Shimatake, H. & Rosenberg, M. (1981) *Nature (London)* **292**, 128–132.
24. Wahba, A. J. & Friedkin, M. (1962) *J. Biol. Chem.* **237**, 3794–3801.
25. Rode, W., Scanlon, K. J., Hynes, J. & Bertino, J. R. (1979) *J. Biol. Chem.* **254**, 11538–11543.
26. Belfort, M. (1978) *J. Virol.* **28**, 270–278.
27. Garvey, J. S., Gemer, N. E. & Sussdorf, D. H. (1977) *Methods in Immunology* (W. A. Benjamin, Reading, MA), 3rd Ed., pp. 313–321.
28. Anilionis, A. & Riley, M. (1980) *J. Bacteriol.* **143**, 355–365.
29. Ward, D. & Murray, N. (1979) *J. Mol. Biol.* **133**, 249–266.
30. Haertle, T., Wohlrab, R. & Guschlbauer, W. (1979) *Eur. J. Biochem.* **102**, 223–230.
31. Riley, M. & Anilionis, A. (1980) *J. Bacteriol.* **143**, 366–376.