Volume 14 Number 24 1986 Nucleic Acids Research

Selective enrichment of specific DNA, cDNA and RNA sequences using biotinylated probes, avidin and copper-chelate agarose

Andrew A.Welcher[1], Anthony R.Torres[2] and David C.Ward[1,3]*

Departments of [1]Molecular Biophysics and Biochemistry, [2]Laboratory Medicine and [3]Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA

Received 5 August 1986; Accepted 17 November 1986

ABSTRACT

We have developed a general procedure for the rapid and efficient enrichment of specific DNA, RNA or cDNA sequences. Biotinylated DNA or RNA is used as a hybridization probe in solution, avidin is then added to label both the probe and hybrid molecules, and the hybridization mixture chromatographed over cupric-iminodiacetic acid agarose beads. Avidin-probe and avidin-hybrid molecules are selectively retained on the column; non-hybridized sequences are contained in the flow-through fraction. Sequences retained on the column are recovered in high yield by the addition of ethylenediamine tetracetic acid in the buffer. The method can be used in both subtractive enrichment and positive selection protocols. Here we report its application to the isolation of Neisseria gonorrhoeae specific genomic DNA clones and the purification of a cDNA subpopulation representing mRNA sequences that are over-expressed in murine Friend cells after dimethylsulfoxide induction.

INTRODUCTION

The preparation of recombinant DNA clones containing specific DNA or cDNA sequences is often time consuming, requiring either a major effort to isolate the appropriate mRNA or extensive screening of large plasmid, phage or cosmid libraries with nucleic acid or antibody probes. The availability of rapid and efficient methods for preenriching target nucleic acids prior to cloning, particularly if they are of unknown sequence or molecular complexity, can significantly reduce the time and effort of the overall process. Several positive selection procedures have been described for the enrichment of target sequences for which specific hybridization probes exist; most of these involve hybridization to, and elution from, probe sequences immobilized on a solid support, such as agarose, dextran, cellulose or nitrocellulose (reviewed in 1). These procedures generally work well, although the hybridization efficiency is often low due to polynucleotide immobilization on the solid matrix.

Preferential enrichment of specific mRNAs has been achieved, in the absence of nucleotide or amino acid sequence information, when an antibody to

the translation product of the mRNA is available. Positive selection by immunoprecipitation of nacent RNA on polysomes (2) has resulted in the extensive purification of several mRNAs that are relatively rare (3,4), however, size selection of mRNA, based on the immunoprecipitation of in vitro translation products (5), gives a more modest enrichment. In the absence of either specific sequence data or antibody reagents few preenrichment strategies exist. However, when cell populations which differentially express a mRNA of interest are available, cDNA can be made from mRNA of the "producer" cell, hybridized to an excess of mRNA from the "non-producer" cell, and the double-stranded hybrids, reflecting shared sequences, separated from free cDNAs by chromatography on hydroxylapatite [HAP] (6). The efficiency of this subtractive enrichment can be improved further by hybridizing the enriched cDNA in the flow-through fraction to an excess of "producer" mRNA and recovering the resultant hybrids from HAP (7). cDNA clones specific for interferon induced genes (8) murine histocompatibility complex genes (9) Aspergillus differention genes (10) and α_{2u} globulin (11) have been obtained by screening libraries with HAP-selected cDNAs. Nevertheless, since fractionation on HAP is based on the extent of doublestrandedness, RNAs with extensive secondary structure can heavily contaminate the hybrid fraction and reduce the selectivity.

In this paper we describe a method that can be used for both positive and subtractive sequence enrichment which exploits the high specificity and avidity of the interaction between biotin and avidin (12) or Streptavidin (13) and a metal chelate adsorbent (14) that reversibly binds these protein-biotinylated nucleic acid complexes. We also report the isolation of a set of Neisseria gonorrhoeae specific DNA clones by screening a genomic library with DNA obtained after subtractive hybridization with biotinylated Neisseria meningitidis DNA, which shares $\sim$ 80% sequence homology with N. gonorrhoeae (15). cDNAs complementary to mRNA sequences overexpressed in dimethylsulfoxide treated Friend leukemia cells were also prepared by subtractive hybridization with photobiotinylated RNA from uninduced cells.

MATERIALS AND METHODS

Bacterial Isolates

Neisseria meningitidis (ATCC 13090, group B), Neisseria gonorrhoeae (ATCC E 27630), Escherichia coli RC109, and clinical isolates from the Clinical Microbiology Laboratory at Yale-New Haven Hospital, were used in the experiments reported here. The N. gonorrhoeae strain was shown (M. Li,

personal communications) to be free of the high copy number plasmid that is present in 95% of N. gonorrhoeae isolates (16). Neisseria cultures were propagated on chocolate agar plates.

DNA Isolation

The Marmur method for DNA isolation (17) was used with minor modifications. Neisseria organisms were scraped with a sterile cotton swab from 8 chocolate agar plates (10 cm in diameter) after overnight growth, and suspended in 20 mls of lysis buffer (0.15 M NaCl-10 mM EDTA-1% SDS). E. coli was grown in YT broth, concentrated by centrifugation and suspended in 20 ml of lysis buffer. After lysis at 65°C for 30 minutes, 3.5 mls of 6 M sodium perchlorate was added. The solution was extracted with 12.5 mls chloroform-isoamyl alcohol (24:1) in a 50 ml polypropylene centrifuge tube placed on a shaker platform for 30 minutes. The aqueous and organic phases were separated by centrifugation at 12,000 RPM for 10 minutes, using a SS34 rotor in a Sorvall centrifuge. The aqueous layer was removed and the extraction procedure repeated. The DNA was precipitated overnight with 2.5 volumes of absolute ethanol at -20°C, pelleted by centrifugation at 10,000 RPM for 10 minutes, and then resuspended in 0.5 mls of 10 mM Tris-HCl, pH 7.5 - 1mM EDTA (T.E.). Ribonuclease A (Cooper Biomedical) was added to a final concentrations of 100 µg/ml and the solution incubated at 37°C for 1 hr. Proteinase K (EM Biochemicals) and sodium dodecylsulfate (BDH Chemical) were then added at concentrations of 200 µg/ml and 0.2% w/v, respectively, and the solution incubated a further 2 hrs at 37°C. The sample was extracted twice with phenol and once with ethyl ether. Sodium chloride was added to a final concentration of 0.2 M, and the DNA ethanol precipitated as before. The DNA was resuspended in 100 µl of T.E. and its concentration measured spectrophotometrically. If the 260/280 ratio was below 1.6, the Proteinase K digestion and phenol extraction steps were repeated a second time.

DNA Digestion

Purified bacterial DNAs were digested with restriction endonucleases (New England Biolabs) using buffers and incubation conditions recommended by the manufacturer. Routinely, 100 µg of N. meningitidis DNA was digested with 50 U of AluI in a volume of 200 µl at 37°C for 4 hours, while 10 µg of N. gonorrhoeae DNA was digested with 10 U Sau3A I in a volume of 100 µl at 37°C for 4 hours. The DNAs were extracted once with phenol, once with ether, ethanol precipitated and pelleted as before, and then resuspended in 50 µl of T.E.

DNA Labeling

N. meningitidis DNA, digested with AluI, was labeled with ^{3}H dATP (NEN)

and Bio-11-dUTP (18) by nick-translation (19). Four reactions each containing 10 µg of N. meningitidis were nick-translated with 30 U DNA Polymerase I (NEB), DNase I (Sigma) at a final concentration of 20 pg/ml, in a volume of 200 µl, at 14°C for one hour. The four reactions were pooled, EDTA was added to a concentration of 10 mM, and the DNA was concentrated by centrifugal lyophilization (Savant speed vac) to a volume of approximately 200 µl. Unincorporated nucleotides were removed from the solution by chromatography on a 3 ml Sephadex G-50 (Pharmacia) column equilibrated in T.E.

The N. gonorrhoeae DNA was labeled with ^{32}P dCTP (Amersham) by filling in the Sau3A I restriction sites with DNA Polymerase I Klenow fragment (NEB). The reaction mixture (80 µl), containing 1 µg N. gonorrhoeae (digested with Sau3A I), 250 µCi ^{32}P dCTP (specific activity ∿ 3000 Ci/mmole), 20 U Klenow enzyme, dGTP, dATP, and dTTP, each at a concentration of 100 µM, was incubated at room temperature for 15 minutes. The DNA was phenol extracted two times, extracted once with ether, and then chromatographed through a 1 ml G-50 spin column (21) equilibrated in T.E. The specific activity of the N. gonorrhoeae DNA was $5x10^7$ cpm/µg.

Growth and Induction of Erythroleukemic Cells

The murine Friend cell line MEL 745 was grown in Dulbecco's Modification of Eagle's Medium (Flow) supplemented with 13% (V/V) heat inactivated fetal calf serum (Gibco), glutamine 2 mM, penicillin 102 U/ml, and streptomycin 75 U/ml. Cells were propagated at 37°C in an atmosphere containing 5% CO_2, with the cell density being maintained between $1x10^5$ and $1x10^6$ cells per ml. Erythroid differentiation was induced by the addition of dimethylsulfoxide (Aldrich) to a final concentration of 1.8% for a period of 72 hrs. The extent of β-globin induction was monitored by the benzidine color assay (20).

Preparation of poly(A)$^+$ RNA and cDNA Synthesis

Cytoplasmic RNA was prepared from induced and uninduced MEL 745 cells and poly(A)$^+$ RNA was purified by chromatography on oligo(dT)-cellulose (Collaborative Research) as described by Maniatis et al. (21) . First strand cDNA synthesis was carried out in a reaction volume of 50 µl, containing 50 mM Tris-HCl, pH 8.1, 5 mM $MgCl_2$, 2 mM DTT, 40 mM KCl, 100 µM dGTP, 100 µM dTTP, 100 µM dATP, 50 µM dCTP, 150 µCi of α ^{32}P dCTP (∿ 3000 Ci/mmole), 50 µg of oligo (dT_8) (Pharmacia), 0.5 µg poly(A)$^+$ RNA from DMSO induced cells, 2.5 µg Actinomycin D, and 100 units of AMV reverse transcriptase (Life Sciences), for 1 hour at 41°C. RNA was removed by treating the reaction with NaOH at a final concentration of 0.2 M for 1 hr at 41°C. EDTA was then added to 20 mM, and unincorporated nucleotides and excess primer removed by chromatography over a

1.5 ml G-50 Sephadex column. The cDNA in the void volume was then extracted once with phenol and concentrated by butanol extraction to a volume of 100 µl. The cDNA was then extracted with ether one time, ethanol precipitated and the cDNA resuspended in 5 µl of distilled water. The specific activity of the cDNA pool was $9x10^7$ cpm/µg.

Labeling of RNA with Photobiotin

Non-induced poly(A)$^+$ RNA (40 µg in 40 µl of T.E.) was mixed with 40 µg of photobiotin (Vector Labs) dissolved in 40 µl dH_2O for a total volume of 80 µl. The solution, in a 1.5 ml Eppendorf tube, was placed on ice and irradiated for 20 minutes by a sunlamp (GE model RSM, 275W) at a height of 10 cm. After adjusting the solution to 0.1 M Tris-HCl - 1 mM EDTA, pH 8.0, it was extracted twice with an equal volume of 1-butanol. Sodium acetate, pH 5.2, was added to the aqueous phase at a final concentration of 0.25 M and the RNA precipitated with 2.5 volumes of 100% ethanol. After 30 min at -20°C, the RNA was pelleted by centrifugation at 4°C for 10 minutes in a Microfuge (Beckman) at 13,000 RPM. The RNA was washed with 70% ethanol, repelleted as above, and then resuspended in 40 µl of T.E.

Solution Hybridization with Biotinylated DNA

Ten µg of N. meningitidis DNA (labeled with bio-11-dUTP and ^{3}H-dCTP) was added to 1 µg N. gonorrhoeae DNA (labeled with ^{32}P-dCTP) in a 1.5 ml Eppendorf tube. The mixture was concentrated by centrifugal lyophilization to a volume of approximately 100 µl, and desalted on a 1 ml G-50 spin column equilibrated in T.E., pH 7.5. The solution was concentrated as above, to a small volume (10-20 µl). Sodium chloride and sodium phosphate (pH 6.8) were added to a final concentration of 0.75 M and 0.05 M, respectively. The solution was heat denatured at 100°C for 4 minutes, quick chilled on ice, overlaid with paraffin oil (J.T. Baker Chemical Co.), and placed in a heating block set at 55°C. The hybridization was allowed to proceed for 15 hrs for a Cot of 40.

At the end of the hybridization, the paraffin oil was removed with a Pasteur pipet and the volume adjusted up to 90 µl with freshly prepared 20 mM $NaHCO_3$-1 M NaCl (pH 7.7), 60 µl of 5 M NaCl was added, followed by 150 µl of a 1 mg/ml solution of Avidin-DN (Vector labs) in 0.05 M $NaHCO_3$ pH 8.2, - 0.1% thimerosal. The tube was gently mixed then placed in a heating block set at 55°C until the sample was applied to the chelate column.

Chromatography on Cupric-iminodiacetic Acid-Agarose

A silanized 1 ml disposable syringe (Becton-Dickinson) plugged with silanized glass wool was packed with 0.4 mls of iminodiacetic acid-agarose (Pierce Chemical). The column was washed with ten column volumes of dH_2O,

then 0.2 mls of a $CuSo_4$ solution (5 mg/ml in dH_2O) was passed through the column; the copper charged column becomes pale blue. The column was washed with ten column volumes dH_2O, then ten column volumes 20 mM $NaHCO_3$-1 M NaCl (pH 7.7). At this point, 300 µl of the bicarbonate solution was added to the DNA-avidin solution (see above), the sample gently mixed, and the solution was slowly passed over the column. Fractions were collected and the column was washed with the bicarbonate solution until the radioactivity of the eluate was reduced to background levels. Bound DNA was eluted by washing the column with 50 mM EDTA-1 M NaCl-0.1% SDS. As the copper ion was released, the column turned from blue to white; the DNA was routinely found in the first drop or two of blue elution buffer. The distribution of N. gonorrhoeae DNA in the unbound and bound fractions was quantitated by Cerenkov counting or with a Beckman Model 170 flow-through radioisotope detector.

Preparing DNA for Subsequent Hybridization Cycle

Unbound DNA fractions were pooled and concentrated by butanol to a volume of approximately 0.5 mls. The solution was dialyzed overnight at 4°C against one liter of T.E. pH 7.5-10 mM NaCl, concentrated by butanol to approximately 100 µl, and then extracted once with ether. The selection procedure was then continued by adding 10 µg of labeled N. meningitidis DNA and repeating the hybridization and chromatographic steps described before. Four or five hybridization cycles were done in the experiments reported here.

Dot Blots

N. gonorrhoeae, N. meningitidis, and E. coli DNAs (300 ng of each) were immobilized on nitrocellulose (22) and dot blot-hybridization used to monitor the enrichment of N. gonorrhoeae DNA fragments. The nitrocellulose membranes were prehybridized for two hours at 42°C in 50% formamide, 5XSSC, 25 mM $NaPO_4$ pH 6.8, 5 x Denhardt's, 250 µg/ml sonicated salmon sperm DNA. 2 ng/ml of selected ^{32}P-labeled N. gonorrhoeae DNA and 10% dextran sulfate were added to the prehybridization cocktail and the solution hybridized for 16 hours at 42°C. The membranes were washed twice for 3 minutes in 2XSSC-0.1% SDS, once for 3 minutes in 0.2XSSC-0.1% SDS, twice for 15 minutes each in 0.1XSSC-0.1% SDS at 55°C, then 1 time for 3 minutes in 2XSSC-0.1% SDS. The membranes were air dried, and placed under X-ray film (Kodak XAR) with a DuPont Cronex Lightning intensifying screen for 4 to 15 hours at -70°C.

Colony Hybridization

N. gonorrhoeae DNA digested with Sau3AI, was shotgun cloned into the Bam HI site of pUC-13 (P.L.-Pharmacia) and the DNA used to transform the bacterial strain JM103. Transformants were plated out on YT-Amp-XGal plates, and white

colonies (containing inserts) were streaked with a toothpick onto YT grid plates, with 100 colonies per plate. Replica grids were prepared on nitrocellulose filters as previously described (21).

The selected DNA from the fifth hybridization cycle, extensively enriched in N. gonorrhoeae sequences, was used as a probe to screen the colony bank. Hybridization conditions were the same as for the dot blots; the washing procedures were also the same with two modifications. The washes with 0.1XSSC-0.1% SDS were done at 60°C instead of at 55°C, and the washes were done for a total of 45 minutes instead of 30 minutes to reduce the background signal. Washed filters were air dried, and placed under XAR-5 film with an intensifying screen for 40 hrs at -70°C.

Isolation of N. gonorrhoeae Specific Colonies

Colonies which hybridized to the affinity selected N. gonorrhoeae DNA were picked by toothpick and grown in 2 mls of YT-broth. Mini-plasmid preparations were done as described (21). Plasmid DNA was digested with EcoRI and Xbal, subjected to electrophoresis on a 1.4% agarose gel, run in a buffer of 40 mM Tris-HCl, pH 8.5 - 5 mM Na acetate-1 mM EDTA, at 100 volts for 6 hrs. The DNA was visualized by EtBr staining, inserts were cut out of the gel, and the DNA was electroeluted from the agarose slices (21). The DNA was concentrated with butanol, ether extracted, then desalted on a 1 ml G-50 spin column equilibrated in T.E. DNA was then nick-translated (19). ^{32}P-labeled DNA was used as a hybridization probe for the dot-blots, as described before, to test if the inserts were specific for N. gonorrhoeae. Random inserts from the same library were tested in the same way to determine the percentage of N. gonorrhoeae specific clones.

Solution Hybridization with Biotinylated RNA

Ten µg of non-induced RNA (labeled with biotin) was added to 50 ng of ^{32}P-labeled cDNA made from induced RNA. The final solution hybridization mixture contained 0.65 M NaCl, 0.04 M $NaPO_4$ pH 6.8, 1 mM EDTA, and 0.05% SDS in a total volume of 20 µl. The RNA was denatured by heating at 100°C for 3 minutes, quick cooled on ice, then covered with mineral oil. Hybridization was done at 60°C for 24 hrs. From the nucleotide complexity of Erythroleukemia cell mRNAs (6), these conditions should give a Rot of 20. After hybridization, the mineral oil was removed, 10 µl of 5 M NaCl and 75 µl of a 1 mg/ml solution of Avidin-DN (Vector Labs) added, and the mixture was incubated for 30 minutes at 60°C. The solution was then passed through a copper chelate column (0.3 ml) with the initial column eluate being recycled through the column twice before initiating the washing steps. Subsequent

chromatographic steps and dot blot analysis of cDNA enrichment were done as described for the N. gonorrhoeae DNA selection experiments.

RESULTS

Separation of Biotinylated and Non-biotinylated Nucleic Acids on Metal Chelate Columns

Metal chelate chromatography (14) has been used previously to fractionate cationic proteins which exhibit differential affinities for heavy metal ions (23,24). Exposed histidine, tryptophan and cysteine residues form relatively stable complexes with immobilized copper or zinc ions in neutral aqueous solutions, however, these complexes are readily reversed by the addition of EDTA or by lowering the pH. We reasoned that metal chelate columns could provide an effective, and inexpensive, means of separating biotinylated nucleic acid species from their non-biotinylated counterparts if avidin, streptavidin or antibiotin antibodies could be used as bridging proteins between the column and the biotinylated molecules. The data shown in Figure 1

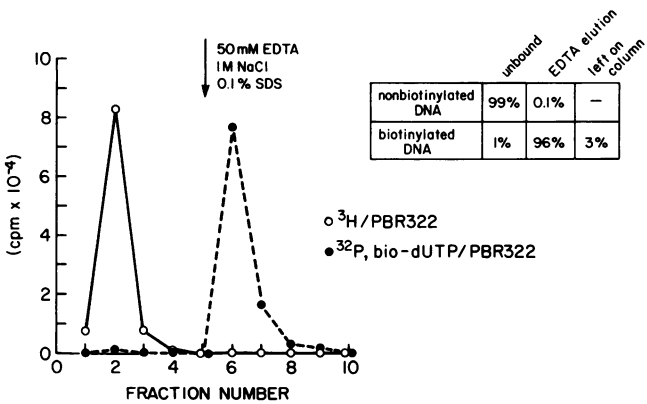

Figure 1. Separation of biotinylated DNA from non-biotinylated DNA by chromatography on copper-chelate agarose. Two samples of pBR322 DNA were nick-translated, one being labeled with ^{3}H-dCMP (o), the other with ^{32}P-dCMP and Bio-11-dUMP (•), each at a specific activity of 10^{7} cpm/µg. To 10 ng of each DNA in 50 µl of T.E. was added 50 µl of 2 M NaCl and 100 µl of Avidin-DN (1 mg/ml). The solutions were kept on ice for 5 min and then chromatographed over 0.2 ml columns of iminodiacetic acid agarose precharged with 0.11 ml of $CuSo_4$ (5 mg/ml). 0.2 ml fractions were collected and counted in hydrofluor. Columns were washed with 1 ml of 20 mM $NaHCO_3$-1 M NaCl and then the column bound DNA was eluted with 50 mM EDTA-1 M NaCl-0.1% SDS. The inset shows the distribution of the two DNA samples during the chromatography.

demonstrates that this is indeed the case. Two samples of pBR322 DNA were nick-translated, one labeled with Bio-11-dUTP and ^{32}P-dCTP, the other labeled with ^{3}H-dCTP in the absence of Bio-11-dUTP. Both samples were incubated with an excess of avidin-DN for 5 minutes and then chromatographed over iminodiacetic acid-agarose charged with copper sulfate (see Methods). Greater than 99% of the ^{3}H-pBR322 DNA was found in the flow-through fraction while ∿ 99% of the ^{32}P-biotinylated pBR322 DNA was retained on the column under identical conditions. Retention of the biotinylated DNA was absolutely dependent on the prior addition of avidin; the chromatographic profile of the biotinylated DNA (Bio-DNA) was identical to that of the control DNA without added avidin. The column-bound Bio-DNA was recovered with high efficiency (∿ 96%) by washing the adsorbent with a solution of 50 mM EDTA-1 M NaCl-0.1% SDS. In contrast, the EDTA-eluate from the control DNA column contained 0.1% or less of the input counts.

In some experiments a small fraction, here 3%, of the Bio-DNA remained on the column even after extensive washing with elution buffer. This residual material appears to be high molecular weight DNA-avidin intermolecular aggregates that are physically trapped on the gel matrix. Since avidin is a tetrameric protein capable of binding 4 molecules of biotin simultaneously, it can, particularly at low avidin-DNA ratios, generate Bio-DNA - Bio-DNA crosslinks. By using a judicious avidin excess (10-fold or more relative to Bio-dNMP incorporated) when forming the DNA-protein complexes in solution, this non-specific entrappment can be maintained at low levels (0.5-3%). However, if necessary, this DNA subset can be recovered by incubating the column in proteinase K and pronase or by incubating at elevated temperatures (50-60°C) in the presence of 1-2% SDS.

The copper-chelate agarose has a high binding capacity for avidin and associated biotinylated molecules. Titration studies demonstrated that several different commercial lots of iminodiacetic acid-agarose bound 15-20 mg of avidin per ml of resin when charged with copper ions; this binding occurs rapidly, virtually instantaneously, between pH 6 and 8 and is not affected by NaCl concentrations as high as 2 molar. Efficient fractionation of Bio-DNA from non-biotinylated DNA also was achieved with streptavidin as the bridging protein or when the chelate was charged with zinc ions (data not shown). Bio-DNA, however, does not bind as tightly to the Zn-chelate, as suggested by prior experiments on protein chromatography (14). Metal chelate agarose also yielded excellent results on both micro (ng) and bulk (multi mg) scales.

We next tested the specificity and efficiency of the copper-chelate

agarose in selecting hybrid molecules formed during solution hybridization reactions with biotinylated probes. ^{32}P-labeled pBR322 DNA was prepared as before and hybridized in solution with a five-fold molar excess of pBR322 DNA, linearized by EcoRI digestion and labeled with ^{3}H-dAMP and Bio-11-dUMP using a DNA polymerase "fill-in" reaction. Following hybridizaion, the mixture was divided into three parts, one part was chromatographed over avidin-agarose (Sigma), a second chromatographed over avidin-DN-agarose (Vector Labs) while the third was incubated with avidin-DN in solution for 5 min and then chromatographed over copper-chelate-agarose. All three chromatographic methods resulted in the retention of > 90% of the biotinylated probe (containing only 4 biotin residues per pBR322 DNA molecule) in a single column passage. However, only 18% and 21% of the ^{32}P-labeled target molecules were retained on the two avidin-agarose resins while 48% was recovered on the copper-chelate column. Similar results were obtained in a comparison of the ability of the copper-chelate agarose and streptavidin-agarose (Bethesda Research Labs) to bind RNA-cDNA hybrids formed with an RNA probe labeled with photobiotin (25). Although both columns retained > 95% of the input probe, 54% of the target cDNA was retained on the copper-chelate column while only 15% was bound to the streptavidin-agarose. Efficient recovery of hybrid molecules on avidin-agarose or streptavidin-agarose required more extended column incubation periods or multiple cycles through the affinity matrix. These observations suggest the formation of avidin-biotin or streptavidin-biotin nucleic acid complexes is appreciably more efficient in solution than when either protein is immobilized on a solid matrix. The results further demonstrate that the capture efficiency of such complexes on copperchelates is quite high.

A similar comparison of the binding of ^{32}P-labeled, biotin-free, DNA to these various affinity matrices indicated that the copper-chelate column exhibited the lowest level of interaction with non-biotinylated DNA; by washing extensively with 20 mM $NaHCO_3$-1 M NaCl the non-specific binding can easily be reduced to less than 1 part in 10^4. In contrast, the non-specific sticking of "control" DNA to avidin-agarose or streptavidin-agarose was generally in the range of 0.1-0.5%.

Enrichment of N. gonorrhoeae DNA by Subtractive Hybridization

As a first test of the utility of this fractionation procedure tor selective sequence enrichment, we set out to isolate a family of Neisseria gonorrhoeae genomic DNA clones containing sequences which were absent from the closely related organism Neisseria meningitidis. These two human pathogens,

causitive agents of gonorrhoeae and acute bacterial meningitis, respectively, have approximately 80% sequence homology (15). The basic subtractive hybridization strategy we used is outlined below.

The DNA of interest (N. gonorrhoeae) first was digested with the restriction enzyme Sau3A. This enzyme gives DNA fragments between 0.5 and 5 kb with an average size of ∿ 2 kb, thus maximizing the amount of potential N. gonorrhoeae specific sequences to be obtained during the selection procedure, and it yields fragments which, after reannealing, could be cloned directly into a BamHI site of an appropriate plasmid vector. Alternatively, the DNA was ^{32}P-labeled by "filling-in" the single-stranded ends of the Sau3A fragments with DNA polymerase and ^{32}P-dCTP. The N. meningitidis (don't-want) DNA was cut with a second restriction enzyme, AluI, and labeled with Bio-11-dUTP by nick-translation. Both steps minimize the risk of cloning any meningitidis DNA which might escape subsequent fractionation procedures. A 10-fold excess of the biotinylated meningitidis DNA was hybridized with N. gonorrhoeae DNA,

Chromatographic Cycle	^{32}P N.Gon unbound	^{32}P N.Gon EDTA eluted
1	66%	34%
2	74%	26%
3	87%	13%
4	92%	8%
5	99%	1%

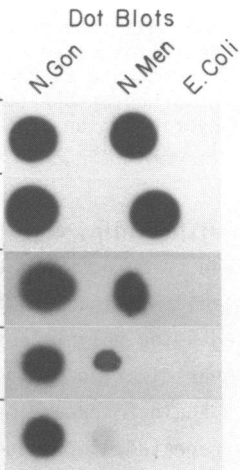

Figure 2. Purification of ^{32}P-labeled N. gonorrhoeae DNA by subtractive hybridization with biotinylated N. meningitidis DNA. After each of the five rounds of solution hybridization the samples were adjusted to 1 M NaCl, Avidin-DN added to a final concentration of 0.5 mg/ml, the samples incubated at 55°C for 10 minutes and then chromatographed over a 0.4 ml bed of cupric-iminodiacetic acid agarose. The distribution of the ^{32}P-labeled N. gonorrhoeae DNA in the non-hybridized (unbound) and hybridized (EDTA-eluted) fractions after each cycle is tabulated. The specificity of the unbound DNA fraction at each cycle against N. gonorrhoeae (N. gon), N. Meningitidis (N. men) and E. coli DNA is illustrated by the dot-blot hybridization reactions on the right.

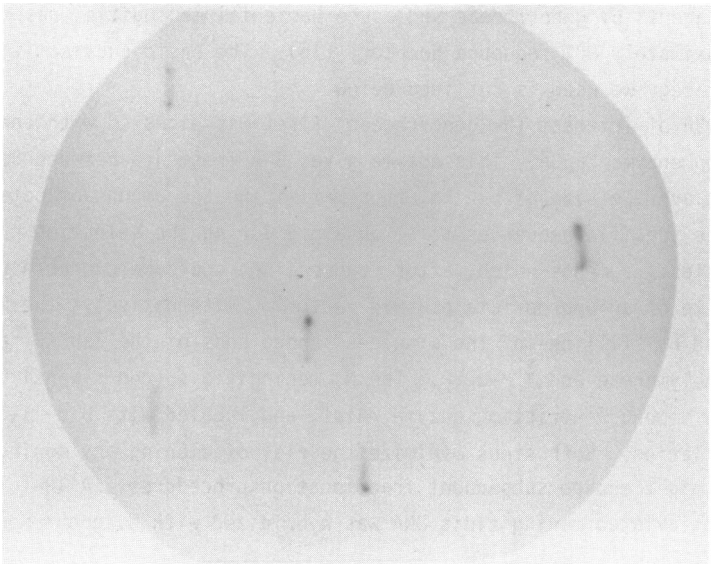

Figure 3. Hybridization to a grid of 100 colonies from a shot-gun library of Sau3AI digested N. gonorrhoeae DNA, cloned into pUC-13 and transformed into E. coli strain JM103. The ^{32}P-labeled probe was the N. gonorrhoeae DNA subset obtained after five cycles of subtractive hybridization and copper-chelate chromatography. Only five of the 100 colonies show positive hybridization.

incubated with avidin, and then chromatographed over a copper-chelate column as described in Methods and Materials. In cases where the target DNA was not radiolabeled, an aliquot of the DNA in the flow-through fraction was ^{32}P-labeled using polynucleotide kinase and then used to monitor enrichment by dot-blot hybridization to the two bacterial DNAs. The remaining DNA was hybridized again to excess meningitidis DNA. The hybridization-fractionation cycle was repeated 3 or 4 additional times or until a strong dot-blot hybridization signal was obtained only from the N. gonorrhoeae DNA.

The results of a subtractive enrichment experiment, using N. gonorrhoeae DNA prelabeled with ^{32}P, is shown in Figure 2. With each additional hybridization-chromatography cycle the percentage of the ^{32}P-N. gonorrhoeae DNA that was retained on the copper-chelate resin decreased; after the fifth round of hybridization only 1% of the input DNA was bound. The remaining 99% of the DNA in the flow-through fraction, representing 5% of the original N. gonorrhoeae DNA, also give virtually no signal when hybridized to N. meningitidis DNA. This preenriched, ^{32}P-labeled, DNA was then used to screen

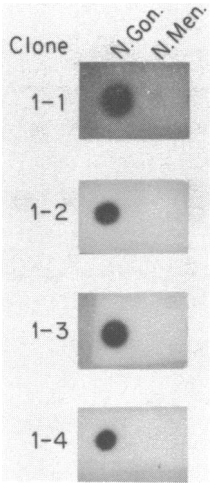

Figure 4. Plasmid clones identified by the ^{32}P-labeled DNA subset obtained after five cycles of subtractive hybridization and copper-chelate chromatography are specific for N. gonorrhoeae by dot-blot hybridization. Clones 1-1, 1-2, 1-3 and 1-4 had DNA inserts of 1.2, 0.5, 0.3 and 1.0 kb, respectively.

a library of Sau3A-digested N. gonorrhoeae DNA that had been cloned directly into the Bam HI site of pUC-13. Nitrocellulose filters, each containing 100 transformants, were prepared and subjected to colony hybridization as described in Methods. On average only 5% of the colonies in any grid gave a positive hybridization signal (Figure 3). Screening the shot-gun library with preenriched DNA as a probe was the most efficient way to identify multiple N. gonorrhoeae specific clones. Direct cloning of the enriched DNA subpopulation after reannealing gave less reproducible results, mainly due to variations in the ligation and transformation efficiences.

Plasmid DNA preparations were prepared from a total of 8 positive clones and the DNA insert sizes determined by gel electrophoresis; the size distribution, from 0.2 to 1.5 kb with an average of 1.1 kb, was only slightly smaller than that seen for the total genomic DNA after Sau3A digestion. These plasmid DNAs were ^{32}P-labeled by nick-translation and hybridized to dots of N. gonorrhoeae and N. meningitidis DNA. Seven of the 8 clones tested were found to be specific for N. gonorrhoeae. Typical hybridization results are exemplified by the four clones illustrated in Figure 4. In contrast, when 22 clones were picked randomly from the library and screened in a similar fashion only 1 of the plasmids hybridized uniquely to N. gonorrhoeae DNA.

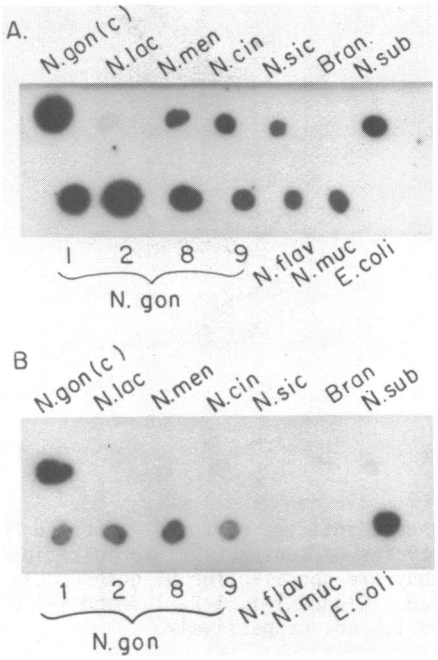

Figure 5. N. gonorrhoeae clones, detected by preenriched DNA probes, specifically hybridize to N. gonorrhoeae DNA. Panel A. N. gonorrhoeae genomic DNA, ^{32}P-labeled by nick-translation, hybridizes both to N. gonorrhoeae strains (the origin strain [N. gon (c)] and clinical isolates 1,2,8 and 9) as well to DNA of other Neisseria strains; N. lacamica (N. lac), N. meningitidis (N. men), N. cinerea (N. cin), N. sicca (N. Sic), N. subflava (N. sub), N. flavescens (N. flav) and N. mucosa (N. muc). No hybridization is observed with Escherichia coli (E. coli) or Branhamella catarrhlis (Bran) DNAs. Panel B. ^{32}P-labeled Clone 1-3 DNA, containing a 0.3 kb fragment of N. gonorrhoeae DNA, hybridized to all N. gonorrhoeae isolates but not to the DNA of other Neisseria strains.

Several of the clones identified with the preenriched DNA probe set were also tested for cross-hybridization with 7 other Neisseria species and for their ability to identify clinical isolates of N. gonorrhoeae. As shown in Figure 5A, total genomic DNA from N. gonorrhoeae hybridized strongly to all Neisseria species tested except N. lactamica. No hybridization was observed with E. coli or Branhamella catarrhlis, two non-Neisseria DNA controls. In contrast, clone 1-3 DNA hybridized significantly only to the 5 N. gonorrhoeae isolates (Figure 5B). Similar results were obtained with other clones. The signal seen with E. coli DNA in Figure 5B is an artifact, and reflects traces of contaminating E. coli sequences in the plasmid preparation which become

Chromatographic Cycle	^{32}P cDNA Unbound	^{32}P cDNA EDTA eluted	Dot Blots: Induced RNA	Dot Blots: Uninduced RNA
0	—	—		
1	45 %	55 %		
2	52 %	48 %		
3	72 %	28 %		
4	93 %	7 %		

Figure 6. Purification of cDNAs specific to DMSO-induced MEL-745 cells. ^{32}P-labeled cDNAs prepared from poly(A)$^+$ RNA expressed 72 hours after treatment with DMSO, were hybridized with an excess of poly(A)$^+$ RNA from induced cells a total of 4 times. After each hybridization reaction the sample was incubated with Avidin-DN and chromatographed over a copper-chelate resin (see Methods). The distribution of ^{32}P-labeled cDNA in the non-hybridized (unbound) and hybridized (EDTA-eluted) fractions after each cycle is tabulated and the hybridization signal of each unbound cDNA fraction against induced and uninduced poly(A)$^+$ RNA is shown.

radiolabeled during nick-translation (data not shown). Such artifactual signals can be eliminated by subcloning the insert into SP-6 or T-phage transcription vectors and using the insert-specific transcripts as hybridization probes.

Purification of cDNAs Complementary to mRNAs Overexpressed in DMSO-induced Friend Cells

Murine erythropoietic precursor cells transformed by Friend Virus, termed Friend cells, morphologically resemble normal proerythroblasts. When treated in culture with dimethylsulfoxide (DMSO) they differentiate along the erythroid pathway. During these changes there is an accumulation of α and β-globin mRNA and hemoglobin as well as other products characteristic of terminal erythropoiesis (26). To test the fractionation method further we have used poly(A)$^+$ RNA from an uninduced Friend cell (MEL-745), chemically labeled with photobiotin, to purify a cDNA subpopulation which hybridizes only to poly(A)$^+$ RNA prepared from MEL 745 cells 72 hours after DMSO-induction. Biotinylated RNA and ^{32}P-labeled first-strand cDNA were prepared, hybridized,

and chromatographed over the copper-chelate resin as described in Methods. As shown in Figure 6, the cDNA hybridizes extensively to both induced and uninduced RNA prior to subtractive enrichment. In contrast, after four cycles of hybridization and chromatography the ^{32}P-cDNA hybridized exclusively with RNA from induced cells. This cDNA subpopulation represents 7 percent of the original cDNA pool. Affara and Daubas (6) had previously shown that 70-75% of the induction specific cDNA, separated by HAP chromatography, is complementary to α and β-globin mRNA. The cDNA subpopulation selected via the copper-chelate agarose also hybridized extensively with α and β-globin genomic DNA clones (data not shown). The cloning and further characterization of this cDNA subpopulation will be reported elsewhere. It is clear, however, that the biotin-avidin-copper-chelate separation method described here can be used as an efficient alternative to HAP-chromatography for selective cDNA purification.

DISCUSSION

We have described a versatile method for the purification of selected nucleic acid sequences and demonstrated its utility in preparing N. gonorrhoeae specific genomic DNA clones and a differentiation specific cDNA subpopulation. Although the experiments reported here required four or five cycles of subtractive hybridization and chromatography, two cycles are usually sufficient for enrichment when higher concentrations of driver probes are used. Furthermore, when selecting target sequences of relatively low abundance, it is recommended that the biotinylated probe pool be incubated with an excess of avidin, then bound to and eluted from the copper-chelate agarose prior to use. This ensues that all probe molecules added to hybridization reactions can be quantitatively recovered during subsequent chelate chromatography, thus minimizing the potential of contaminating target sequences with non-biotinylated probes. Since biotinylated probes eluted from the copper-chelate column with EDTA will be complexed with avidin, the protein must be removed, by proteinase K-pronase digestion and phenol extraction, in order to maintain optimum hybridization kinetics.

This technique can be applied to a wide variety of separation problems where no information concerning specific target DNA sequences, their transcripts or their gene products is available. For example, in species such as bacteria, where the genome contains few, if any, highly repetitive sequence elements that can lower selective enrichment by generating interstrand networks, one could readily isolate large sets of organism specific sequences.

Enrichment by subtractive hybridization potentially provides a rapid means of preparing extensive probe sets which could be useful in the clinical diagnosis of specific pathogens. Conversely, one could analyze the sequence subsets retained on copper-chelate agarose to identify conserved sequences that are common to an entire family or genera of organisms. The subtraction hybridization method also could be used to identify and clone DNA segments of unknown sequence which have been deleted during strain evolution, mutagenesis, or as a direct or indirect consequence of pathologic disease (27). In screening the genomes of higher organisms that possess numerous repetitive elements, unique sequence subsets could be prepared from somatic cell hybrids, clone banks, or individual chromosomes by subtractive hybridization with appropriate repetitive sequence probes, e.g., the Alu, Kpn or centromer family repeats. These and other applications of this fractionation method are currently being evaluated.

ACKNOWLEDGEMENTS: This work was supported by Public Health Service Grants AI-19973, CA-16038 and GM-32156 from the National Institutes of Health. We thank Vincent Piscitelli for supplying the various Neisseria strains and D. Greenberg for word processing assistance.

*To whom correspondence should be addressed

REFERENCES

1. Potuzak, H. and Dean, P.D.G. (1978) FEBS Letters 88, 161-166.
2. Shapiro, S.Z. and Young, J.R. (1981) J. Biol. Chem. 256, 1495-1498.
3. Kraus, J.P. and Rosenberg, L.E. (1982) Proc. Natl. Acad. Sci. USA 79, 4015-4019.
4. Korman, A.J., Knudsen, P.J., Kaufman, J.F. and Strominger, J.L. (1982) Proc. Natl. Acad. Sci. USA 79, 1844-1848.
5. Parnes, J.R., Velan, B., Felsenfeld, A., Ramanthan, L., Ferrini, U., Appella, E. and Seidman, J.G. (1981) Proc. Natl. Acad. Sci. USA 78, 2253-2257.
6. Affara, N. and Daubas, P. (1979) Dev. Biol. 72, 110-125.
7. Timberlake, W.E. (1980) Dev. Biol. 78, 497-510.
8. Samanta, H., Dougherty, J.P., Brawner, M.E., Schmidt, H. and Lengyel, P. (1982) in UCLA Symp. on Mol. and Cell. Biol. 25, 59-72.
9. Davis, M.M., Cohen, D.I., Nielsen, E.A., Steinmetz, M., Paul, W.E. and Hood, L. (1984) Proc. Natl. Acad. Sci. USA 81, 2194-2198.
10. Zimmerman, C.R., Orr, W.C., Leclerc, R.F., Barnard, E.C. and Timberlake, W.E. (1980) Cell 21, 709-715.
11. Kurtz, D. and Feigelson, P. (1977) Proc. Natl. Acad. Sci. USA 74, 4791-4795.
12. Green, N.M. (1975) Adv. Prot. Chem. 29, 85-133.
13. Argarana, C.E., Kuntz, I.D., Birken, S., Axel, R. and Cantor, C.C. (1986) Nucl. Acids Res. 14, 1871-1882.

14. Porath, J., Carlsson, J., Olsson, I. and Belfrage, G. (1975) Nature 228, 598-599.
15. Kingsbury, D.T. (1967) J. Bact. 94, 870-874.
16. Totten, P.A., Holmes, K.K., Handsfield, H.H., Knapp, J.S., Perine, P.L. and Falkow, S. (1983) J. Inf. Dis. 148, 462-471.
17. Marmur, J. (1961) J. Mol. Biol. 3, 208-218.
18. Brigati, D.J., Myerson, D., Leary, J.J., Spalholz, B. Travis, S.Z., Fong, C.K.Y., Hsiung, G.D. and Ward, D.C. (1983) Virology 126, 32-50.
19. Rigby, P.W.J., Dieckmann, J., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
20. Gusella, J. and Houseman, D. (1976) Cell 8, 263-269.
21. Maniatis, T., Fritch, E.F. and Sambrook, J. (1982) Molecular Cloning (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
22. Thomas, P. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
23. Lönnerdal, B., Carlsson, J. and Porath, J. (1977) FEBS Letters 75, 89-92.
24. Torres, A.R., Peterson, E.A., Evans, W.H., Mage, M.G. and Wilson, S.M. (1979) Biochim. Biophys. Acta. 576, 385-392.
25. Forster, A.C. McInnes, J.L., Skingle, D.C. and Symons, R.H. (1985) Nucl. Acids Res. 13, 745-761.
26. Harrison, P.R. (1976) Nature 262, 353-356.
27. Kunkel, L.M., Monaco, A.P., Middlesworth, W., Ochs, H.D. and Latt, S.A. (1985) Proc. Natl. Acad. Sci. 82, 4778-4782.