Proc. Natl. Acad. Sci. USA
Vol. 78, No. 11, pp. 6613–6617, November 1981
Biochemistry

Use of synthetic oligonucleotides as hybridization probes: Isolation of cloned cDNA sequences for human β_2-microglobulin*

(mixed sequence oligonucleotides/colony screening)

SIDNEY V. SUGGS, R. BRUCE WALLACE, TADAAKI HIROSE†, ERIC H. KAWASHIMA‡, AND KEIICHI ITAKURA§

Division of Biology, Molecular Genetics Section, City of Hope Research Institute, Duarte, California 91010

Communicated by James F. Bonner, June 22, 1981

ABSTRACT **We have synthesized two sets of 15-base-long oligodeoxyribonucleotides corresponding to all possible coding sequences for a small portion of human β_2-microglobulin. Labeled oligonucleotides were used as hybridization probes to screen bacterial clones containing cDNA sequences primed with oligo(dT) and inserted into the plasmid vector pBR322. One β_2-microglobulin cDNA clone was detected in the 535 bacterial plasmid clones that were screened. The clone has been characterized by blotting and nucleotide sequence analysis. The cloned β_2-microglobulin sequence contains 217 base pairs of the 3′ untranslated region of the mRNA and 328 base pairs (97%) of the coding region.**

The hybridization properties of oligodeoxyribonucleotides have been characterized by a number of techniques (1–4). Under appropriate conditions, oligonucleotides hybridize to specific sites in DNA (4, 5). Furthermore, perfectly base paired oligonucleotide duplexes can be discriminated from duplexes containing a single mismatched base pair (4–6). We have taken advantage of the hybridization properties of oligonucleotides in developing a method for the isolation of specific cloned DNA sequences (5). Our general approach is to chemically synthesize a mixture of oligonucleotides that represent all possible codon combinations for a small portion of the amino acid sequence of a given protein. Within this mixture must be one sequence complementary to the DNA coding for that part of the protein. This complementary oligonucleotide will form a perfectly base paired duplex with the DNA from the coding region for the protein, whereas the other oligonucleotides in the mixture will form mismatched duplexes. Under stringent hybridization conditions only the perfectly matched duplex will form, allowing the use of the mixture of oligonucleotides as a specific hybridization probe. Mixed sequence oligonucleotide probes should allow isolation of cloned DNA sequences for any protein for which the amino acid sequence is known.

We have applied this method to the isolation of cloned cDNA sequences for human β_2-microglobulin (β_2m). β_2m is a small protein (molecular weight 11,800) that was isolated from urine (7). Subsequently, β_2m was found associated with cell surface antigens of the major histocompatibility locus (8, 9). The exact function of β_2m is unclear, although recent evidence suggests that the molecule may stabilize the tertiary structure of associated proteins (10). The amino acid sequence has been determined for β_2m from four species, including human (11). We have used the amino acid sequence to design probes for the isolation of a cloned cDNA for human β_2m.

MATERIALS AND METHODS

General Methods. Plasmid DNA was isolated by a cleared lysate procedure (12) and purified by chromatography on Bio-Gel A-50m (Bio-Rad) followed by ethidium bromide/CsCl equilibrium sedimentation. Restriction enzymes were purchased from Bethesda Research Laboratories (Rockville, MD). Restriction enzyme reactions were performed at 37°C in 60 mM NaCl/10 mM Tris·HCl (pH 7.6)/7 mM $MgCl_2$/6 mM 2-mercaptoethanol. Agarose and polyacrylamide gels (acrylamide to methylenebisacrylamide ratio, 20:1) were run in Tris/borate electrophoresis buffer (13). Electroelution of DNA fragments from polyacrylamide gels has been described (14). Transfer of DNA fragments from agarose (SeaKem; Marine Colloids, Rockland, ME) gels to nitrocellulose was by the standard Southern procedure (15). Prehybridization and hybridization conditions have been described (5). Recombinant DNA was handled in accordance with National Institutes of Health Guidelines.

Table 1. Oligonucleotide probes for the isolation of β_2m

Amino acid sequence		95 Trp	96 Asp	97 Arg	98 Asp	99 Met	
Possible codons	5′	UGG	GA^U_C	AG^A_G CGN	GA^U_C	AUG	3′
Probe β_2mI	3′	ACC	CT^A_G	TC^T_C	CT^A_G	TAC	5′
Probe β2mII	3′	ACC	CT^A_G	GCN	CT^A_G	TAC	5′
		75	76	77	78		
Amino acid sequence		Lys	Asp	Glu	Tyr		
Possible codons	5′	AA^A_G	GA^U_C	GA^A_G	UA^U_C	3′	
Probe β_2mIII	3′	TT^T_C	CT^A_G	CT^T_C	AT	5′	

N = A, C, G, or T (U).

Simultaneous Synthesis of Oligodeoxyribonucleotides of Mixed Sequence. The oligonucleotide mixtures shown in Table 1 were synthesized by the strategy described previously (5). Probes β2mI and β2mII were synthesized by the triester method in solution (16). β2mIII was synthesized by a solid-phase method (17). All oligonucleotides were purified by high-performance liquid chromatography on an ion-exchange (Du Pont Permaphase AAX) column (5).

Labeling of Oligonucleotides. Oligonucleotides were labeled at the 5′ end by transfer of ^{32}P from [γ-^{32}P]ATP, using bacteriophage T4 polynucleotide kinase (New England Nuclear) as described (4). [γ-^{32}P]ATP was synthesized by the method of Walseth and Johnson (18).

Synthesis of Double-Stranded cDNA. The source of RNA used to prepare cDNA was a human lymphoblastoid cell line, Raji [obtained from S. Ohno (19)]. Cytoplasmic RNA was isolated and fractionated on an oligo(dT)-cellulose (Collaborative Research, Waltham, MA) column as described (20). Double-

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.

Abbreviation: β_2m, β_2-microglobulin.

* This paper is no. 3 in a series. Paper no. 2 is ref. 5.

† Present address: Keio University, Tokyo, Japan.

‡ Present address: Biogen S.A., Geneva, Switzerland.

§ To whom reprint requests should be addressed.

stranded cDNA was synthesized by using successive reactions with oligo(dT) (P-L Biochemicals) plus reverse transcriptase (RNA-dependent DNA polymerase; a gift from J. Beard), *Escherichia coli* DNA polymerase I Klenow fragment A (Boehringer Mannheim), and S1 nuclease (Sigma) as described by Goeddel *et al.* (21). The 3′ termini of the double-stranded cDNA were extended with dC residues by using terminal deoxynucleotidyltransferase (Bethesda Research Laboratories) as described (22).

Isolation of cDNA Clones. The C-tailed cDNA was subjected to electrophoresis on a 5% polyacrylamide gel. The fragments 500–800 base pairs in length were recovered by electroelution. Fifteen nanograms of size-selected cDNA was annealed with 100 ng of pBR322 that had been cleaved with restriction endonuclease *Pst* I and tailed with dG residues (23). The mixture was heated to 65°C for 3 min and then allowed to gradually cool from 43°C to 22°C over the course of 20 hr. The annealed mixture was used to transform *E. coli* strain MC1061 (24) by the Kushner procedure (25).

Screening of cDNA Clones. After transformation, recombinant clones were selected on L agar plates containing 25 μg of tetracycline per ml (23). Individual clones were picked and put on fresh tetracycline plates in an ordered array in duplicate. The bacterial clones were allowed to grow overnight, transferred to Whatman 541 filter paper, amplified with chloramphenicol, and prepared for hybridization as described by Gergen *et al.* (26). The filters were prehybridized, hybridized, and washed as described in our previous paper (5).

Nucleotide Sequence Determination. Restriction endonuclease fragments of cloned cDNA were labeled at their 5′ ends by the sequential action of bacterial alkaline phosphatase (P-L Biochemicals) and polynucleotide kinase plus [γ-^{32}P]ATP (27). *Eco*RI-digested and *Sau*3A-digested DNA was labeled at the 3′ end by repair labeling of the termini: after restriction enzyme digestion, 2 μM [α-^{32}P]dATP (New England Nuclear), polymerase I Klenow fragment A at 2 units/ml, and, in the case of *Sau*3A-digested DNA, 400 μM dGTP (P-L Biochemicals), were added and the reaction mixture was incubated 5 min at 20°C. *Pst* I-digested DNA was labeled at the 3′ termini with ^{32}P-labeled 3′-dATP and terminal transferase as described in the kit supplied by New England Nuclear.

RESULTS

Strategy for Designing β_2m-Specific Oligodeoxyribonucleotide Probes. Our general approach for the isolation of cloned DNA sequences specific for β_2m involved synthesis of a set of oligodeoxyribonucleotides complementary to all the possible coding sequences for a small portion of the protein. In designing the probe sequences, we chose two regions of the protein for which there are relatively few potential coding sequences. The two regions of amino acid sequence we used in designing the β_2m-specific probes are shown in Table 1. Amino acid residues 95–99 of β_2m can be coded for by 24 possible sequences in the mRNA. We synthesized two sets of pentadecanucleotides corresponding to this region; β2mI is a mixture of 8 sequences and β2mII is a mixture of 16 sequences. Amino acid residues 75–78 of β_2m were used to design another probe, β2mIII, which is a set of 8 undecanucleotides. Each of the three sets of probes was synthesized as a mixture of sequences (5).

Isolation of a Bacterial Clone Containing β_2m cDNA Sequences. Using poly(A)-containing cytoplasmic RNA from a human lymphoblastoid cell line, we prepared double-stranded

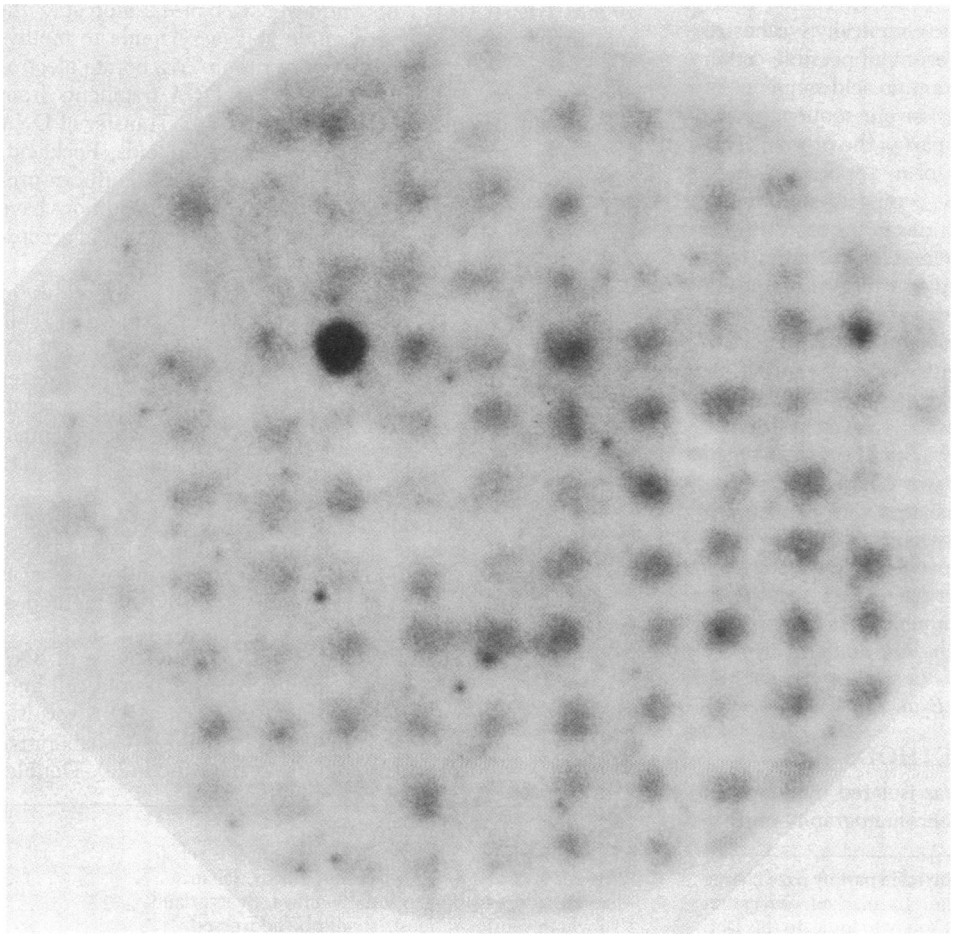

FIG. 1. Hybridization of oligonucleotide probe β2mII to colonies transformed with G-tailed plasmid plus C-tailed cDNA. Bacterial cDNA clones were isolated and screened. ^{32}P-Labeled β2mII probe was hybridized to the filters at 41°C overnight and washed with 0.9 M NaCl/0.09 M sodium citrate at 41°C.

cDNA 500–800 base pairs in length (20) and inserted it into the *Pst* I site of the plasmid vector pBR322 by the standard G·C tailing method (22). The recombinant DNA was used to transform *E. coli* strain MC1061. From the transformants, 535 tetracycline-resistant clones were obtained and placed on fresh plates in an ordered array. The bacterial clones were transferred to Whatman 541 filter paper, amplified with chloramphenicol, and prepared for hybridization as described by Gergen *et al.* (26). The filters were hybridized with ^{32}P-labeled oligonucleotide probes and washed as described in *Materials and Methods*. An autoradiogram of one of the five filters hybridized with the β2mII probe is shown in Fig. 1. The amount of labeled probe hybridized to one of the clones is clearly greater than that hybridized to any of the other clones. Using the β2mI and β2mIII probes to screen the same collection of clones, we could not distinguish specific hybridization to any clone (data not shown).

Characterization of the Cloned cDNA for β_2m by Southern Blot Analysis. Plasmid DNA was prepared from the presumptive bacterial clone for β_2m cDNA observed in Fig. 1. The β_2m plasmid DNA was cleaved with restriction endonucleases *Hin*dII + *Hin*dIII or *Sau*3A. Fig. 2*A* shows the results of a Southern blot analysis of *Hin*dII + *Hin*dIII-cut DNA. Both the β2mII and β2mIII probes hybridize specifically with the smallest *Hin*dII + *Hin*dIII fragment (lanes *A*2 and *A*3). The hybridization of β2mII to the largest *Hin*dII + *Hin*dIII fragment observed in lane *A*2 represents background hybridization and is not detectable with shorter exposure times. Fig. 2*B* shows the results of blot analysis of *Sau*3A-digested β_2m plasmid DNA. The β2mIII probe hybridizes to a restriction fragment 620 base pairs in length (lane *B*3). There is no detectable hybridization of the β2mII probe to any of the major DNA bands (lane *B*2) even though this autoradiogram was exposed for the same amount of time as that in lane *A*2 showing hybridization of β2mII to *Hin*dII + *Hin*dIII-digested DNA. The absence of hybridization of β2mII to *Sau*3A-digested plasmid DNA is due to the presence of a *Sau*3A recognition site within the probe hybridization site. By cleaving the insert DNA with *Sau*3A at a position in the center of the region to which β2mII hybridizes, hybridization of the probe is eliminated. There is a faint band of hybridization observed in lane *B*2. This represents hybridization of the probe to a partial digestion product in which the *Sau*3A site within the probe hybridization site is not cleaved. This partial digestion product is not detectable in the photography of the stained gel. Fig 2*C* is a map of restriction endonuclease sites showing the sites of hybridization of β2mII and β2mIII probes.

Nucleotide Sequence Analysis of the Cloned cDNA for β_2m. The nucleotide sequence of the β_2m plasmid DNA was determined by using the base-specific cleavage reactions of Maxam and Gilbert (27). The nucleotide sequence of the cloned β_2m cDNA is shown in Fig. 3. The cloned cDNA contains most of the coding region for the protein, including part of the coding region for the leader peptide, and a large portion of the 3′ untranslated region.

The amino acid sequence for β_2m predicted from the nucleotide sequence differs from the published amino acid sequence (11) at two positions: (*i*) position 42 is asparagine and not aspartic acid, and (*ii*) the serine at position 67 of the published amino acid sequence is not present in the sequence deduced from the cloned cDNA. While these differences may reflect polymorphism in human β_2m, the latter difference in the amino acid sequence would require the insertion of three nucleotides (serine codon) in the gene. Accordingly, the amino acid sequence for human β_2m determined in the present study establishes the length of the mature protein as 99 residues, the same length as β_2m from the three other species for which the sequence has been determined (28, 29).

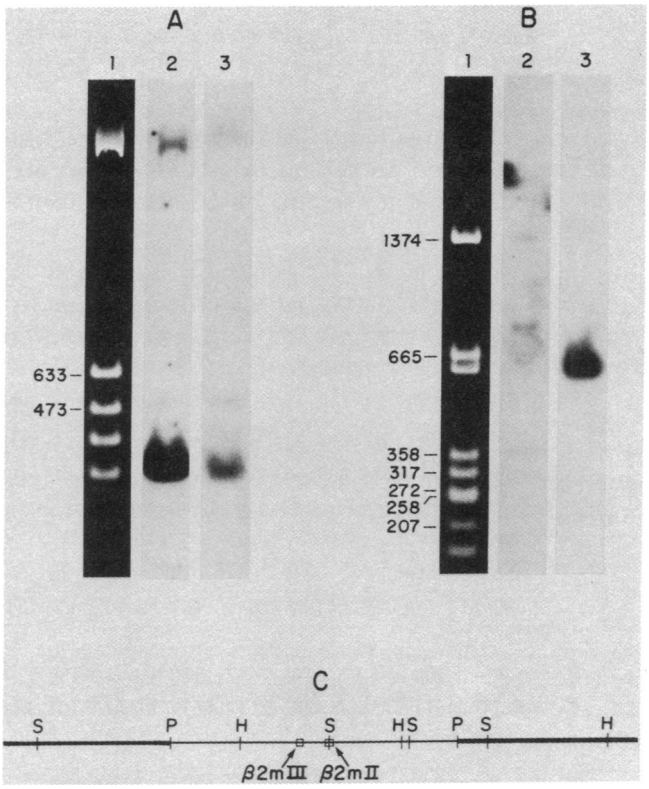

FIG. 2. Southern blot analysis of the cloned cDNA for β_2m. The plasmid DNA containing the β_2m cDNA sequences was digested with restriction enzymes *Hin*dII + *Hin*dIII or *Sau*3A in *A* and *B*, respectively. The DNA fragments were subjected to electrophoresis on a 2% agarose gel, visualized by staining with ethidium bromide, and transferred to nitrocellulose filter paper by the standard Southern blotting technique (18). Duplicate blots were hybridized with ^{32}P-labeled β2mII at 41°C or ^{32}P-labeled β2mIII at 14°C. The blots were washed at 37°C and 20°C, respectively. In both *A* and *B*, lane 1 is a photograph of a lane from the stained gel, lane 2 is an autoradiogram of the blot hybridization with ^{32}P-labeled β2mII, and lane 3 is an autoradiogram of the blot hybridization with ^{32}P-labeled β2mIII. The sizes (in base pairs) of the restriction fragments derived solely from the vector pBR322 are indicated. (*C*) Restriction maps of the region of the plasmid containing the inserted β_2m sequences. The thick line indicates sequences from the plasmid vector pBR322 and the thin line indicates β_2m sequences. The sites of hybridization of the two probes are indicated. H, *Hin*dII site; P, *Pst* I site; and S, *Sau*3A site.

The sequence of the leader peptide for human β_2m determined from the nucleotide sequence differs from the previously published sequence for mouse β_2m (30) at several positions: (*i*) position −2 is glutamic acid in the human sequence and not tyrosine as in the mouse sequence; (*ii*) position −8 is leucine, not valine; and (*iii*) position −10 is alanine, not valine. Only the first of these changes represents a nonconservative amino acid substitution.

The sequence of the 3′ untranslated region of β_2m cDNA is, in general, A+T-rich (65% A·T base pairs). The sequence A-A-T-A-A-A, which is common to many polyadenylylated mRNAs (31), is not observed in the sequence. The absence of this sequence may indicate that a portion of the 3′ untranslated region is not present in the clone.

DISCUSSION

Our results demonstrate the usefulness of mixtures of chemically synthesized oligodeoxyribonucleotides as hybridization

```
 -11 -10  -9  -8  -7  -6  -5  -4  -3  -2  -1   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16
 Leu Ala Leu Leu Ser Leu Ser Gly Leu Glu Ala Ile Gln Arg Thr Pro Lys Ile Gln Val Tyr Ser Arg His Pro Ala Glu
5'-CTC GCG CTA CTC TCT CTT TCT GGC CTT GAG GCT ATC CAG CGT ACT CCA AAG ATT CAG GTT TAC TCA CGT CAT CCA GCA GAG

  17  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43
 Asn Gly Lys Ser Asn Phe Leu Asn Cys Tyr Val Ser Gly Phe His Pro Ser Asp Ile Glu Val Asp Leu Leu Lys Asn Gly
 AAT GGA AAG TCA AAT TTC CTG AAT TGC TAT GTG TCT GGG TTT CAT CCA TCC GAC ATT GAA GTT GAC TTA CTG AAG AAT GGA

  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70
 Glu Arg Ile Glu Lys Val Glu His Ser Asp Leu Ser Phe Ser Lys Asp Trp Ser Phe Tyr Leu Leu Tyr Tyr Thr Glu Phe
 GAG AGA ATT GAA AAA GTG GAG CAT TCA GAC TTG TCT TTC AGC AAG GAC TGG TCT TTC TAT CTC TTG TAT TAT ACT GAA TTC

  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97
 Thr Pro Thr Glu Lys Asp Glu Tyr Ala Cys Arg Val Asn His Val Thr Leu Ser Gln Pro Lys Ile Val Lys Trp Asp Arg
 ACC CCC ACT GAA AAA GAT GAG TAT GCC TGC CGT GTG AAC CAC GTG ACT TTG TCA CAG CCC AAG ATA GTT AAG TGG GAT CGA

  98  99
 Asp Met Stop        10          20          30          40          50          60          70          80          90
 GAC ATG TAA GCAGCATCAT GGAGGTTGA AGATGCCGCA TTTGGATTGG ATGAATTCAA AATTCTGCTT GCTTGCTTTT TAATATTGAT ATGCTTATAC

        100         110         120         130         140         150         160         170         180         190
 ACTTACACTT TATGCACAAA ATGTAGGGTT ATAATAATGT TAACATGGAC ATGATCTTCT TTATAATTCT ACTTTGAGTG CTGTCTCCAT GTTTGATGTA

        200         210
 TCTGAGCAGG TTGCTCCACA GGTAGCT-3'
```

FIG. 3. Nucleotide sequence of the cloned cDNA for human β_2m. The nucleotide sequence of the noncoding strand of the cloned cDNA is shown. Above the nucleotide sequence is the amino acid sequence of the protein, with the amino acids of the leader peptide indicated by negative numbers.

probes for the isolation of specific cloned DNA sequences. Our previous results with model systems (4, 5) indicated that the screening procedure should be highly specific. The results with β_2m show that this is indeed the case: (*i*) The only clone detected with the β_2m probes was the correct one. (*ii*) the β2mI probe, a mixture of eight 15-base-long oligonucleotides, did not hybridize specifically with the β_2m clone in the colony screening experiment. One of the sequences within the β2mI mixture differs from the corresponding sequence in the cloned β_2m cDNA at only a single nucleotide (G in place of T). This result indicated that under stringent hybridization criteria the oligonucleotide probes do not hybridize with closely related sequences containing a single base substitution.

The screening method described in this study is very rapid. The most tedious step in screening for the β_2m clone was the picking of individual bacterial clones and placing them in an ordered array. This step has been eliminated in subsequent experiments in which large numbers of unordered colonies were successfully screened for mouse H-2K^b antigen cDNA clones (unpublished data).

Apparently contradictory results were obtained with the β2mIII probe, a mixture of eight 11-base-long oligonucleotides. One of the sequences in this mixture is perfectly complementary to the cloned cDNA. In the Southern blot analysis, specific hybridization of the β2mIII probe to the cloned β_2m cDNA was observed (Fig. 2); however, in the colony screening experiment, no specific hybridization of β2mIII to the β_2m clone was observed. This discrepancy in the results of the blot and colony screening experiments may reflect differences in the hybridization properties of oligonucleotides to nitrocellulose filters (used in the blot) versus cellulose filters (used in the colony screening) or may reflect a lower background hybridization of oligonucleotides to purified DNA (in the blot) versus DNA in lysed colonies (in the colony screening). In our subsequent experiments isolating cloned sequences for other proteins, oligonucleotide probes longer than 11 bases have been used for colony screening.

In previous studies, chemically synthesized oligodeoxyribonucleotides have been employed in the isolation of specific cloned sequences in basically two ways: oligonucleotides have been used directly as hybridization probes or used as primers for the synthesis of radiolabeled cDNA that is used as the probe. Montgomery *et al.* (32) demonstrated the usefulness of synthetic oligonucleotides as hybridization probes in the isolation of the cloned yeast iso-1-cytochrome *c* gene. In this special case, an oligonucleotide probe of unique sequence could be predicted by comparison of the amino acid sequences of cytochrome *c* from several species. Subsequently, Goeddel *et al.* (33) have screened purified plasmid DNAs enriched for human leukocyte interferon sequences by using labeled oligonucleotides as hybridization probes. Several groups have isolated specific cloned sequences by using synthetic oligonucleotides as primers for the synthesis of cDNA probes (21, 34, 35).

We believe the method using oligonucleotides as hybridization probes is superior to the method employing oligonucleotides as primers in the synthesis of cDNA probes for two main reasons. First, greater specificity can be obtained by using the hybridization approach than by using the priming approach. Under appropriate conditions, a mismatched base pair does not allow formation of oligonucleotide-polynucleotides duplexes (4, 5), whereas base mismatches can be tolerated in the priming method (21, 35, 36). Second, with the priming approach, the amount of probe obtained is dependent on the amount of mRNA available for template. In isolating cloned sequences for very low abundance mRNAs, use of the primer approach would entail the isolation of large amounts of mRNA to produce sufficient

cDNA probe for screening. In addition, it may be increasingly more difficult to produce a specific cDNA probe as the abundance of a mRNA decreases.

In summary, we believe the screening technique described in this report may be applied to the isolation of cloned DNAs for any protein for which the amino acid sequence is known.

Note Added in Proof. Since the preparation of this manuscript the isolation and partial DNA sequence analysis of three mouse β_2m cDNA clones has been reported (37).

We thank Ting H. Huang for purification and sequence analysis of synthetic oligonucleotides. This work was supported by National Institutes of Health Postdoctoral Fellowship GM07591 (to S.V.S.) and National Institutes of Health Grants GM26391 (to R.B.W.) and GM25658 (to K.I.). K.I. and R.B.W. are members of the Cancer Research Center (CA16434) at the City of Hope Research Institute.

1. Astell, C. R. & Smith, M. (1972) *Biochemistry* **11**, 4114–4120.
2. Astell, C. R., Doel, M. T., Jahnke, P. A. & Smith, M. (1973) *Biochemistry* **12**, 5068–5074.
3. Dodgson, J. B. & Wells, R. D. (1977) *Biochemistry* **16**, 2367–2372.
4. Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T. & Itakura, K. (1979) *Nucleic Acids Res.* **6**, 3543–3557.
5. Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H. & Itakura, K. (1981) *Nucleic Acids Res.* **9**, 879–894.
6. Gillam, S., Waterman, K. & Smith, M. (1975) *Nucleic Acids Res.* **2**, 625–633.
7. Berggård, I. & Bearn, A. G. (1968) *J. Biol. Chem.* **243**, 4095–4103.
8. Nakamuro, K., Tanigaki, N. & Pressman, D. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2863–2865.
9. Peterson, P. A., Rask, L. & Lindblom, J. B. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 35–39.
10. Lancet, D., Parham, P. & Strominger, J. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3844–3848.
11. Cunningham, B. A., Wang, J. L., Berggård, I. & Peterson, P. A. (1973) *Biochemistry* **12**, 4811–4821.
12. Kahn, M., Kolter, R., Thomas, C., Figurski, D., Meyer, R., Remant, E. & Helinski, D. R. (1979) *Methods Enzymol.* **68**, 268–280.
13. Peacock, A. C. & Dingman, C. W. (1968) *Biochemistry* **7**, 668–674.
14. Rossi, J. J., Ross, W., Egan, J., Lipman, D. & Landy, A. (1979) *J. Mol. Biol.* **128**, 21–47.
15. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
16. Hirose, T., Crea, R. & Itakura, K. (1978) *Tetrahedron Lett.* **28**, 2449–2452.
17. Broka, C., Hozumi, T., Arentzen, R. & Itakura, K. (1980) *Nucleic Acids Res.* **8**, 5461–5471.
18. Walseth, T. F. & Johnson, R. A. (1979) *Biochim. Biophys. Acta* **526**, 11–31.
19. Epstein, M. A., Achong, B. G., Barr, Y. M., Zajac, B., Henle, G. & Henle, W. (1966) *J. Natl. Cancer Inst.* **37**, 547–555.
20. Singer, R. H. & Penman, S. (1973) *J. Mol. Biol.* **78**, 321–334.
21. Goeddel, D. V., Shepard, H. M., Yelverton, E., Leung, D. & Crea, R. (1980) *Nucleic Acids Res.* **8**, 4057–4074.
22. Roychoudhury, R. & Wu, R. (1980) *Methods Enzymol.* **65**, 43–62.
23. Chang, A. C. Y., Nunberg, J. H., Kaufman, R. J., Ehrlich H. A., Schimke, R. T. & Cohen, S. N. (1978) *Nature (London)* **275**, 617–624.
24. Casadaban, M. & Cohen, S. N. (1980) *J. Mol. Biol.* **138**, 179–207.
25. Kushner, S. R. (1978) in *Genetic Engineering*, eds. Boyer, H. W. & Nicosia, S. (Elsevier/North-Holland, Amsterdam), pp. 17–23.
26. Gergen, J. P., Stern, R. H. & Wensink, P. C. (1979) *Nucleic Acids Res.* **7**, 2115–2136.
27. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
28. Gates, F. T., III, Coligan, J. E. & Kindt, T. J. (1979) *Biochemistry* **18**, 2265–2272.
29. Gates, F. T., III, Coligan, J. E. & Kindt, T. J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 554–558.
30. Lingappa, V. R., Cunningham, B. A., Jazwinski, S. M., Hopp, T. P., Blobel, G. & Edelman, G. M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3651–3655.
31. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
32. Montgomery, D. L., Hall, B. D., Gillam, S. & Smith, M. (1978) *Cell* **14**, 673–680.
33. Goeddel, D. V., Yelverton, E., Ullrich, A., Heyneker, H. L., Miozzari, G., Holmes, W., Seeburg, P. H., Dull, T., May, L., Stebbing, N., Crea, R., Maeda, S., McCandliss, R., Sloma, A., Tabor, J. M., Gross, M., Familletti, P. C. & Pestka, S. (1980) *Nature (London)* **287**, 411–416.
34. Noyes, B. E., Mevarechi, M., Stein, R. & Agarwal, K. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1770–1774.
35. Chan, S. J., Noyes, B. E., Agarwal, K. L. & Steiner, D. F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5036–5040.
36. Houghton, M., Stewart, A. G., Doel, S. M., Emtage, J. S., Eaton, M. A. W., Smith, J. C., Patel, T. P., Lewis, H. M., Porter, A. G., Birch, J. R., Cartwright, T. & Carey, N. H. (1980) *Nucleic Acids Res.* **8**, 1913–1931.
37. Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E. & Seidman, J. F. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2253–2257.