Cloning of the gene and cDNA for mammalian β-adrenergic receptor and homology with rhodopsin

Richard A. F. Dixon*, Brian K. Kobilka†, David J. Strader‡, Jeffrey L. Benovic†, Henrik G. Dohlman†, Thomas Frielle†, Mark A. Bolanowski†, Carl D. Bennett§, Elaine Rands*, Ronald E. Diehl*, Richard A. Mumford‡, Eve E. Slater‡, Irving S. Sigal*, Marc G. Caron†, Robert J. Lefkowitz† & Catherine D. Strader‡

Departments of *Virus and Cell Biology Research and §Medicinal Chemistry, Merck Sharp and Dohme Research Laboratories, West Point, Pennsylvania 19486, USA
† Howard Hughes Medical Institute, Department of Medicine, Biochemistry and Physiology, Duke University Medical Center, Durham, North Carolina 27710, USA
‡ Department of Biochemistry and Molecular Biology, Merck Sharp and Dohme Research Laboratories, Rahway, New Jersey 07065, USA

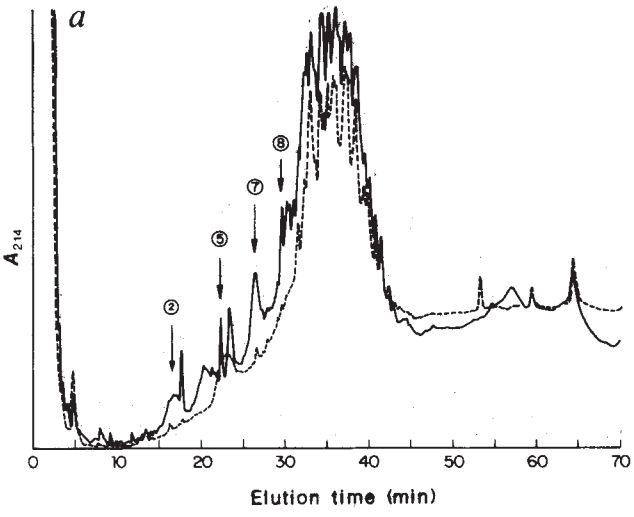
The adenylate cyclase system, which consists of a catalytic moiety and regulatory guanine nucleotide-binding proteins, provides the effector mechanism for the intracellular actions of many hormones and drugs[1]. The tissue specificity of the system is determined by the particular receptors that a cell expresses. Of the many receptors known to modulate adenylate cyclase activity, the best characterized and one of the most pharmacologically important is the β-adrenergic receptor (βAR). The pharmacologically distinguishable subtypes of the β-adrenergic receptor, β_1 and β_2 receptors, stimulate adenylate cyclase on binding specific catecholamines[1]. Recently, the avian erythrocyte β_1, the amphibian erythrocyte β_2 and the mammalian lung β_2 receptors have been purified to homogeneity and demonstrated to retain binding activity in detergent-solubilized form[1–5]. Moreover, the β-adrenergic receptor has been reconstituted with the other components of the adenylate cyclase system *in vitro*[6], thus making this hormone receptor particularly attractive for studies of the mechanism of receptor action. This situation is in contrast to that for the receptors for growth factors and insulin, where the primary biochemical effectors of receptor action are unknown. Here, we report the cloning of the gene and cDNA for the mammalian β_2AR. Analysis of the amino-acid sequence predicted for the βAR indicates significant amino-acid homology with bovine rhodopsin and suggests that, like rhodopsin[7], βAR possesses multiple membrane-spanning regions.

Hamster lung βAR was purified to homogeneity by sequential affinity chromatography and molecular-sieve HPLC as described previously[2,5]. The purified receptor bound ligand with theoretical specific activity and migrated on SDS-polyacrylamide gel electrophoresis as a single broad band at a relative molecular mass (M_r) of 64,000 (64K). Initial attempts to obtain N-terminal sequence data on intact βAR failed, presumably because the N-terminus of this protein was blocked. Therefore, peptide fragments generated by CNBr cleavage of pure βAR were isolated by reverse-phase HPLC.

Figure 1*a* (solid line) shows a peptide map generated from 1 nmol of pure receptor; the broken line shows the HPLC profile resulting from CNBr treatment of the detergent alone. The βAR-derived peptides produced at least nine specific absorbance peaks which were reproducibly observed in five separate βAR preparations. The most prominent of these peptides (marked with arrows in Fig. 1) were subjected to N-terminal sequence analysis, yielding the amino-acid sequences in Fig. 1*b*.

To confirm that the determined amino-acid sequences were those of the βAR polypeptide, we raised anti-peptide antibodies against peptide 7. Peptide 7 was expressed in *Escherichia coli* as a C-terminal peptide fused to the N-terminal domain of the yeast *RAS*^{sc1} protein SC1N (ref. 8). Rabbits injected with the isolated fusion protein produced antibodies which reacted with ^{125}I-labelled immunogen as well as pure ^{125}I-labelled βAR or pure ^{125}I-cyanopindolol-labelled (^{125}I-CYP) βAR. Figure 2*a* shows an immunoprecipitation titration curve of this antibody against ^{125}I-CYP-labelled solubilized βAR from hamster lung, hamster heart, A431 epidermoid carcinoma cells and turkey erythrocytes. No immunoprecipitation of counts above background was observed in control experiments when ^{125}I-CYP was incubated with the antibody in the absence of receptor.

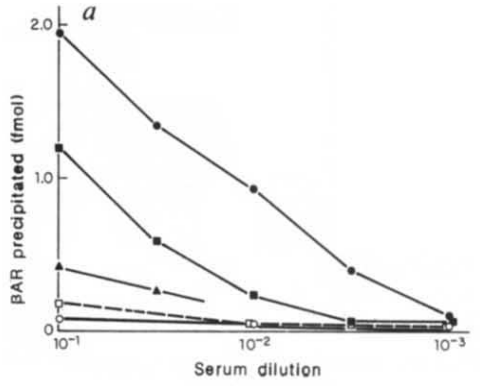
Antibody to the hamster lung βAR-derived peptide was capable of recognizing the human βAR from the A431 line, albeit with a slightly lower sensitivity (50%) (Fig. 2*a*). The antibody also cross-reacted slightly with hamster heart βAR, a tissue containing a β_1 subtype of receptor, but did not immunoprecipitate the β_1AR of turkey erythrocytes. These differences in antibody sensitivity could reflect differences in either primary sequence or the conformation of this region of the protein within the various receptor subtypes and receptors from different species. To confirm that the antibody was recog-

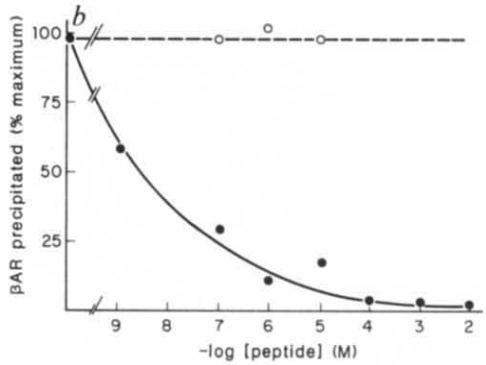


b

Peptide 2	G P P G X X S X F L L T T X G S
Peptide 5	G L A V V P F G A S
Peptide 7	V F V Y S R V F Q V A K R Q L Q K I D K S E G R F H S P N L G Q V E
Peptide 7a	G E A S G S Q L G
Peptide 8	V W I V S G L T S F L P I

Fig. 1 Amino-acid sequence of peptides derived from CNBr-treated β-adrenergic receptor. *a*, Absorbance profiles represent CNBr treatment of pure βAR (solid line) or digitonin (dashed line). The arrows indicate the peptides that were sequenced. *b*, Amino-acid sequences identified by HPLC following each cycle of the sequenator. Two of the four blank cycles (X) in the amino-acid sequence for peptide 2 are presumed to be due to *N*-linked glycosylation. Peptides 7 and 7a were located within the same peak. **Methods.** βAR was purified to homogeneity from hamster lung membranes by the method of Benovic *et al.*[5], using affinity chromatography followed by molecular-sieve HPLC. Binding of ^{125}I-CYP to intact cells or to solubilized βAR was determined according to Caron and Lefkowitz[22]. For peptide preparation, ~1 nmol of pure βAR was treated with CNBr (0.4 mM) in 70% formic acid at 23 °C for 20 h. After lyophilization, the sample was resuspended in 20 mM trifluoroacetic acid (TFA) and the peptides separated by reverse-phase HPLC on a Synchropak C-4 column, eluted with a 10–70% acetonitrile gradient containing 20 mM TFA. The N-terminal sequence analysis was performed by the method of Hewick *et al.*[23], using a gas-phase sequenator (Applied Biosystems). The phenylthiohydantoin (PTH) amino acids produced at each step were separated and quantitated by HPLC[24].





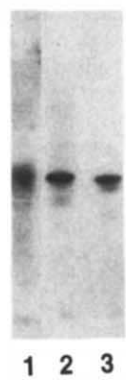


Fig. 2 Immunoreactivity of β-adrenergic receptor. *a*, Immunoprecipitation of ^{125}I-CYP-βAR from hamster lung (●), hamster heart (▲), A431 cells (■) or turkey erythrocytes (○) by serum from rabbits immunized with βAR peptide 7. □, Immunoprecipitation of hamster lung receptor by preimmune serum. *b*, Immunoprecipitation of ^{125}I-CYP-labelled hamster lung βAR with anti-βAR peptide 7 antibody following preincubation with a synthetic peptide containing a portion of the sequence for peptide 7 (●), or with an unrelated peptide, atrial natriuretic factor (○). *c*, Protein immunoblotting of the βAR. Protein samples were separated on a 10% polyacrylamide gel[25], transferred to nitrocellulose and treated sequentially with antibodies and ^{125}I-protein A (10^6 c.p.m. per 25 ml) as described elsewhere[26,27]. Lane 1, pure hamster lung βAR (5 pmol); lane 2, A431 cell lysate from 4×10^5 cells; lane 3, RPMI 1846 lysate from 10^5 cells.

Methods. The immunogen for induction of anti-peptide antibodies was expressed in *E. coli*. Two oligonucleotides encoding the 19 amino acids of peptide 7 (QVAKRQLQKIDKSEGRFHS) were synthesized (see Fig. 1 legend), annealed and ligated[28] into the *Acc*I and *Hin*dIII sites of plasmid pSC1N[8] to give the plasmid pβP1. *E. coli* transformed with pβP1 overexpresses a protein of apparent M_r 23K, while *E coli* containing pSC1N overexpresses a protein of apparent M_r 21K (data not shown). This observed difference in relative molecular mass of the two proteins is consistent with the encoded fusion protein containing 19 additional amino acids. To prepare antigen, plasmid-containing cells were grown in L-broth containing ampicillin and isopropylthiogalactoside at 37 °C for 16 h, harvested by centrifugation, lysed by sonication and the soluble proteins removed by centrifugation at 40,000*g*. The cell pellet was seqentially extracted with 1 M NaCl, 1% Triton X-100 and 1.75 M guanidinium-HCl. The SC1N βAR fusion protein was extracted from the cells with 3.25 M guanidinium-HCl, dialysed against phosphate-buffered saline and used directly as an immunogen. Approximately 100 mg of fusion protein of 90% purity was obtained from 1 litre of starting culture. Antibodies were detected in serum from injected rabbits by incubation of the serum with ^{125}I-CYP-labelled soluble βAR in 10 mM Tris-HCl, 0.1 M NaCl, 0.1% digitonin, 0.5% bovine serum albumin (BSA) *p*H 7.4. After 2 h at 25 °C, the antibody was precipitated by addition of either $(NH_4)_2SO_4$ to 50% or *Staphylococcus aureus* protein A, followed by incubation in ice for 30 min. The precipitated protein was collected by centrifugation, and the radioactivity contained in the antibody pellet measured. For the peptide blocking experiment, the peptide YAKRQLQKIDKSEGR was synthesized[29] using a SAM II peptide synthesizer (Biosearch), and purified on a Whatman C-18 Magnum column in a H_2O/acetonitrile gradient containing 0.2% TFA. The resulting product was judged to be pure by amino-acid sequencing and mass spectral analysis. Increasing concentrations of this peptide were added to a 1:100 dilution of anti-peptide 7 antiserum and incubated for 2 h at 23 °C. The treated antiserum was then mixed with ^{125}I-CYP-labelled βAR and assayed as above.

nizing the amino-acid sequence of peptide 7, a chemically synthesized peptide was used as a specific inhibitor of antibody-receptor interactions. At concentrations ≥100 μM, this synthetic peptide completely prevented the immunoprecipitation of ^{125}I-CYP-labelled βAR by the antibody (Fig. 2*b*). An unrelated peptide, atrial natriuretic factor[9], had no effect on the immunoprecipitation.

The specificity of the antibody for the βAR was demonstrated further by protein immunoblotting. As shown in Fig. 2*c* (lane 1), the antibody reacted specifically with pure hamster lung βAR; a single protein of the same relative molecular mass (64K) was also observed in human A431 and hamster melanoma RPMI 1846 cells (Fig. 2*c*, lanes 2, 3), both of which were found to contain βAR on the basis of ^{125}I-CYP binding[10] (data not shown). This specific immunoreactive band was not observed on prior treatment of the antibody with the synthetic peptide and was not present when normal rabbit serum was substituted for the anti-peptide 7 antibody (data not shown).

To facilitate cloning of the βAR gene, oligonucleotides complementary to the DNA encoding the amino-acid sequence of peptide 7 were synthesized for use as hybridization probes (see Fig. 3 legend). In hybridization experiments performed at high stringency on blots of hamster genomic DNA, a single hybridizing band of 5.2 kilobases (kb) was observed in *Eco*RI digests and a band of 1.3 kb was observed in *Hin*dIII digests (data not shown). When a complete hamster genomic library was screened under the same conditions, five clones were isolated. Restriction analysis of the phage DNA revealed that all these clones contained a 1.3-kb *Hin*dIII and a 5.2-kb *Eco*RI fragment which hybridized to the probes (data not shown). Mapping of the phage DNA indicated that these clones overlap to give a total of 30 kb of contiguous genomic DNA. Figure 3 shows the restriction map of the genomic DNA containing the βAR-related sequences. Sequencing of the 1.3-kb *Hin*dIII fragment revealed a continuous open reading frame encoding 435 amino acids; the sequences of all the CNBr peptides shown in Fig. 1 were contained within this putative polypeptide.

Using the 1.3-kb *Hin*dIII gene fragment as a probe, seven clones were obtained from an unamplified hamster cDNA library (2×10^6 recombinants). Two of these cDNAs hybridized to oligonucleotide probes specific for the N-terminal, middle and C-terminal portions of the βAR gene. The nucleotide sequence of these two cDNA clones (Fig. 4) extends from 210 nucleotides (nt) 5′ to the open reading frame encoding the βAR

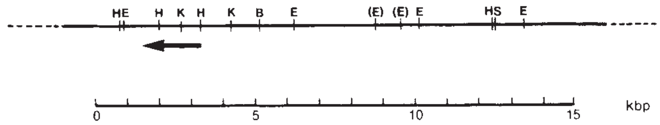


Fig. 3 Restriction map of the hamster βAR gene. A portion of the hamster DNA is shown. The 1.3-kb *Hin*dIII fragment which hybridizes to oligonucleotides specific for peptide 7, is underlined with an arrow indicating the direction of transcription of the βAR gene. The restriction enzyme sites are: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; S, *Sal*I. Those sites shown in parentheses have not been unequivocally ordered.

Methods. All restriction enzymes, *E. coli* DNA polymerase I, T_4 DNA ligase and T_4 polynucleotide kinase were purchased from New England Biolabs. Radiolabelled nucleotides were purchased from Amersham. λ EMBL3A phage arms[30] and λ *in vitro* DNA packaging extracts were purchased from Vector Cloning Systems. Standard recombinant DNA and microbiological procedures were used throughout[28]. Genomic libraries were constructed using high-M_r genomic DNA isolated from hamster lung cells in the vector λ EMBL3 (refs 6, 30). Probes for peptide 7 coding sequences, oligonucleotides ON225(5′ pCTCCACCTGGCCCAGGTTGGG-AGAGTGGAACCTGCCCTCAGACTTGTCGAT) and ON229(5′ pAGGCAGCTGCAGAAGATCGACAAGTCTGAG) and ON168 (5′ pTTCCAGGTGGCCAAGCGGCAGCTGCAGAAGATCGA-CAA) and ON169(5′ pATGGTCTTTGTCTACTCCCGGGTCT-TCCAGGTGGCCAA), were synthesized on an Applied Biosystems Model 3A DNA synthesizer. The oligonucleotides were labelled by either a fill-in reaction (ON225, ON229) using Klenow DNA polymerase and all four [α-^{32}P]dNTPs[31] or by phosphorylation using T_4 polynucleotide kinase and [γ-^{32}P]ATP[28]. Phage libraries were screened by the method of Benton and Davis[32] using the hybridization conditions of Ullrich *et al.*[33]. DNA was isolated from CsCl-banded phage as described elsewhere[28]. For restriction analysis, DNA was digested with the appropriate enzyme and electrophoresed on 0.8% agarose gels. DNA was transferred to nitrocellulose by the procedure of Southern[34] and hybridized as above.

```
                                                                        70
CAGCGTTCAA GCTGCTGTTA GCAGGCACCG CGAGCCCCGG GCACCCCACG AGCTGAGTGT GCAGGACGCG
                                                                       140
CCCCCAGCAC AGCCACCTAC AGCCGCTGAA TGAAGCTTCC AGGAGTCTGC CTCCGGCCGG CTGCGCCCCG
                                                                       210
TCGGAGGTGC ACCCGCTGAG AGCGCCAGGG CACCAGAAAG CCGGTGCGCT CACCTGCTCG TCTGCCAGCG
                                                                        264
ATG GGG CCA CCC GGG AAC GAC AGT GAC TTC TTG CTG ACA ACC AAC GGA AGC CAT
MET Gly Pro Pro Gly Asn Asp Ser Asp Phe Leu Leu Thr Thr Asn Gly Ser His
1                   ***                                 ***
          2
                                                                        318
GTG CCA GAC CAC GAT GTC ACT GAG GAA CGG GAC GAA GCA TGG GTG GTA GGC ATG
Val Pro Asp His Asp Val Thr Glu Glu Arg Asp Glu Ala Trp Val Val Gly MET
19
                                                                        372
GCC ATC CTT ATG TCG GTT ATC GTC CTG GCC ATC GTG TTT GGC AAC GTG CTG GTC
Ala Ile Leu MET Ser Val Ile Val Leu Ala Ile Val Phe Gly Asn Val Leu Val
37
                                                                        426
ATC ACA GCC ATT GCC AAG TTC GAG AGG CTA CAG ACT GTC ACC AAC TAC TTC ATA
Ile Thr Ala Ile Ala Lys Phe Glu Arg Leu Gln Thr Val Thr Asn Tyr Phe Ile
55
                                                                        480
ACC TCC TTG GCG TGT GCT GAT CTA GTC ATG GGC CTA GCG GTG GTG CCG TTT GGG
Thr Ser Leu Ala Cys Ala Asp Leu Val MET Gly Leu Ala Val Val Pro Phe Gly
73
                                             5
                                                                        534
GCC AGT CAC ATC CTT ATG AAA ATG TGG AAT TTT GGC AAC TTC TGG TGC GAG TTC
Ala Ser His Ile Leu MET Lys MET Trp Asn Phe Gly Asn Phe Trp Cys Glu Phe
91
                                                                        588
TGG ACT TCC ATT GAT GTG TTA TGC GTC ACA GCC AGC ATT GAG ACC CTG TGC GTG
Trp Thr Ser Ile Asp Val Leu Cys Val Thr Ala Ser Ile Glu Thr Leu Cys Val
109
                                                                        642
ATA GCA GTG GAT CGC TAC ATT GCT ATC ACA TCG CCA TTC AAG TAC CAG AGC CTG
Ile Ala Val Asp Arg Tyr Ile Ala Ile Thr Ser Pro Phe Lys Tyr Gln Ser Leu
127
                                                                        696
CTG ACC AAG AAT AAG GCC CGA ATG GTC ATC CTA ATG GTG TGG ATT GTA TCC GGC
Leu Thr Lys Asn Lys Ala Arg MET Val Ile Leu MET Val Trp Ile Val Ser Gly
145
                                                         8
                                                                        750
CTT ACC TCC TTC TTG CCC ATT CAG ATG CAC TGG TAC CGT GCC ACC CAC CAG AAA
Leu Thr Ser Phe Leu Pro Ile Gln MET His Trp Tyr Arg Ala Thr His Gln Lys
163
                                                                        804
GCC ATC GAC TGC TAT CAC AAG GAG ACT TGC TGC GAC TTC TTC ACG AAC CAG GCC
Ala Ile Asp Cys Tyr His Lys Glu Thr Cys Cys Asp Phe Phe Thr Asn Gln Ala
181
                                                                        858
TAC GCC ATT GCT TCC TCC ATT GTA TCT TTC TAC GTG CCT CTA GTG GTC ATG GTC
Tyr Ala Ile Ala Ser Ser Ile Val Ser Phe Tyr Val Pro Leu Val Val MET Val
199
                                                                        912
TTT GTC TAT TCC AGG GTC TTC CAG GTG GCC AAA AGG CAG CTC CAG AAG ATA GAC
Phe Val Tyr Ser Arg Val Phe Gln Val Ala Lys Arg Gln Leu Gln Lys Ile Asp
217
        7
                                                                        966
AAA TCT GAG GGA AGA TTC CAC TCC CCA AAC CTC GGC CAG GTG GAG CAG GAT GGG
Lys Ser Glu Gly Arg Phe His Ser Pro Asn Leu Gly Gln Val Glu Gln Asp Gly
235
                                                                       1020
CGG AGT GGG CAC GGA CTC CGA AGG TCC TCC AAG TTC TGC TTG AAG GAG CAC AAA
Arg Ser Gly His Gly Leu Arg Arg [Ser] [Ser] Lys Phe Cys Leu Lys Glu His Lys
253
                                                                       1074
GCC CTC AAG ACT TTA GGC ATC ATC ATG GGC ACA TTC ACC CTC TGC TGG CTG CCC
Ala Leu Lys Thr Leu Gly Ile Ile MET Gly Thr Phe Thr Leu Cys Trp Leu Pro
271
                                                                       1128
TTC TTC ATT GTC AAC ATC GTG CAC GTG ATC CAG GAC AAC CTC ATC CCT AAG GAA
Phe Phe Ile Val Asn Ile Val His Val Ile Gln Asp Asn Leu Ile Pro Lys Glu
289
                                                                       1182
GTT TAC ATC CTC CTT AAC TGG TTG GGC TAT GTC AAT TCT GCT TTC AAT CCC CTC
Val Tyr Ile Leu Leu Asn Trp Leu Gly Tyr Val Asn Ser Ala Phe Asn Pro Leu
307
                                                                       1236
ATC TAC TGT CGG AGT CCA GAT TTC AGG ATT GCC TTC CAG GAG CTT CTA TGC CTC
Ile Tyr Cys Arg Ser Pro Asp Phe Arg Ile Ala Phe Gln Glu Leu Leu Cys Leu
325
                                                                       1290
CGC AGG TCT TCT TCA AAA GCC TAT GGG AAC GGC TAC TCC AGC AAC AGT AAT GGC
Arg Arg [Ser] [Ser] [Ser] Lys Ala Tyr Gly Asn Gly Tyr Ser Ser Asn Ser Asn Gly
343
                                                                       1344
AAA ACA GAC TAC ATG GGG GAG GCG AGT GGA TGT CAG CTG GGG CAG GAA AAA GAA
Lys Thr Asp Tyr MET Gly Glu Ala Ser Gly Cys Gln Leu Gly Gln Glu Lys Glu
361
                              7a
                                                                       1398
AGT GAA CGG CTG TGT GAG GAC CCC CCA GGC ACG GAA AGC TTT GTG AAC TGT CAA
Ser Glu Arg Leu Cys Glu Asp Pro Pro Gly Thr Glu Ser Phe Val Asn Cys Gln
379
                                                                       1452
GGT ACT GTG CCT AGC CTT AGC CTT GAT TCC CAA GGG AGG AAC TGT AGT ACA AAT
Gly Thr Val Pro Ser Leu Ser Leu Asp Ser Gln Gly Arg Asn Cys Ser Thr Asn
397                                                 ***             ***
                                                                       1517
GAC TCA CCG CTG TAA TGCAGGCTTT CTGCTTTTTA AGACCCCTCC CTGACAGGAC ACTAACCAGA
Asp Ser Pro Leu
415
                                                                       1587
CTATTTAACT TGAGTGTAAT AACTTTAGAA TAAAACTGTA TAGAGATTTG CAGAAGGGGA GCATCCTTCT
                                                                       1657
GCCCTTTTTT ATTTTATTTT TTTAAGCCGC AAAAATAGAG AGGGAGAGAA ACTGTACTTG AGTGCTTGTT
                                                                       1727
TGTTTCTTGT GCAATTCAGT TCCTCTTTGC GTGGAACTTA AAAGTTTCTG TCTGAAGTAT GTTGGGTTCT
                                                                       1797
AGAGGACTGT CTGTATGTTT AGATGATTTT CCATGCATCT ACCTCACTCG TCAAGTGTTA GGGGATACGC
                                                                       1867
TGCTAGTAAT TTGTACCTGA AGGAAATTTT CCTTCCTGTA CCCTTACACT TGTCAATCCT GTGTCTTGGA
                                                                       1937
CCTTTCTGCT GTGAATATAT ACTCTCTCCC GCTCCACTTA TTTGCTCAAA TGGAGTGTGT AGACAGGGAT
                                                                       2007
CTTGAGGGAC AGCTTCAGTT GGTTTTTTTT TTTTTTTTGA GCAAAGTCTA AAGTTTACAG T[AAATAAA]TT
                   2026
GTTTGACCAC GAAAAAAAAA
```

Fig. 4 (Right) Nucleotide and deduced amino-acid sequence of the βAR cDNA. The nucleotides are numbered on the right-hand side of each line beginning with the first nucleotide of the most 5′ cDNA clone. The translated amino-acid sequence is shown beneath the corresponding nucleotide sequence and is numbered to the left of each line. Underlined amino acids represent the CNBr peptides whose sequences are given in Fig. 1. All predicted amino acids agree with those determined by peptide sequencing, with the single exception of a cysteine for serine substitution in peptide 7a. All derived peptide sequences are preceded by a methionine, consistent with CNBr cleavage. The underlined nucleotides preceding the βAR sequences denote the first methionine codon and an in-frame termination codon. The boxed nucleotides at the 3′ end of the sequence represent the polyadenylation signal. Postulated glycosylation sites (Asn-X-Ser/Thr) are indicated by asterisks. Putative protein kinase A phosphorylation sites are boxed.

Methods. Reverse transcriptase was purchased from Seikagaku-America. RNase H was from Pharmacia and λ gt10 arms were from Vector Cloning. Total cellular RNA was isolated from growing cultures of DDT1-MF2 cells by the guanidinium isothiocyanate-CsCl method[47]. Poly(A)$^+$ RNA was purified by chromatography on oligo(dT)-cellulose. Double-stranded cDNA was synthesized by oligo(dT)-primed reverse transcription of the poly(A)$^+$ RNA, followed by treatment with *E. coli* DNA polymerase and RNase[48]. The ends of the cDNA were blunted with T_4 DNA polymerase[28]. Following protection of the *Eco*RI sites with *Eco*RI methylase, *Eco*RI linkers were added[28]. *Eco*RI-digested cDNA was size-fractionated by agarose gel electrophoresis to obtain cDNAs between 2 and 7 kb long. The cDNA was ligated to the vector λ gt10 (ref. 49) and packaged *in vitro*. The resulting library was screened unamplified as described in Fig. 3 legend. The 1.3-kb *Hin*dIII fragment was labelled using [α-^{32}P]dCTP by nick-translation for use as a hybridization probe[28]. The cDNA inserts contained in the positive phage were subcloned into pUC13 or M13 mp19 for DNA sequence analysis[35–37]. Both strands of the clones were sequenced with no discrepancies.

peptides to 560 nt 3′ to the termination codon. A hexanucleotide AATAAA occurs near the 3′ end, followed by a poly(A) stretch. The nucleotide sequence of the genomic clones is identical to that of the cDNA clones up to but not including the poly(A) tail. These data demonstrate an absence of introns within the coding and 3′-untranslated regions of the βAR gene but the possibility of introns in the untranslated region 5′ to the sequence isolated cannot be ruled out. While the lack of introns is unusual, it is not unprecedented in that sea urchin histone genes[38] and mammalian α- and β-interferon genes[39,40] are also uninterrupted. Previous studies with simian virus 40 and β-globin indicate that introns are necessary for efficient expression of those genes[41,42], so the lack of introns within the hamster βAR gene may account in part for its low level of expression.

Whereas most eukaryotic genes are translated using the first AUG encountered in the messenger RNA[43], the open reading frame (ORF) encoding the βAR peptides begins at the second AUG. The first AUG is followed by a termination codon after only 19 amino acids. This open reading frame is in-frame with the AUG codon beginning the βAR polypeptide. A similar situation has recently been reported for the oestrogen receptor, where the reading frame of that receptor is preceded by a short 20-amino-acid ORF[44]. Note that the DNA sequence around the AUG codon for the second ORF agrees well with the consensus eukaryotic translation initiation sequence[43] (CAGCGAUGG compared with CCGCCAUGG), whereas the sequence surrounding the first AUG does not.

Translation beginning with the second AUG as the initiation site would produce a polypeptide of 418 amino acids, with a M_r (46K) in close agreement with the apparent M_r of the deglycosylated βAR (49K)[11]. The protein sequence immediately following the initiator methionine residue is identical to the sequence determined for peptide 2. Apparently, like several other integral membrane proteins—for example, bovine opsin[45]—the βAR does not contain a cleavable signal sequence and may use internal signals for the insertion of the protein into the membrane. The receptor has been shown to have two sites of *N*-linked glycosylation (R.J.L. *et al.*, in preparation). Both sites are present in peptide 2, consistent with the results of peptide sequencing. The presence of two consensus protein kinase A and C sites[46] within the coding region (Fig. 4) agrees with *in vitro* phosphorylation results[20].

Hydropathicity profiles of the predicted βAR amino-acid sequence were produced using the analyses of Hopp and Woods[12] and of Kyte and Doolittle[13], with similar results. As shown in Fig. 5*a*, the βAR sequence should encode a largely hydrophobic polypeptide, with the N-terminal region of the receptor being predominantly hydrophobic and the C-terminal region of the molecule being hydrophilic. The βAR hydropathicity profile is remarkably similar to that of the rhodopsins (Fig. 5*b*), of which bacteriorhodopsin[7] is known to contain seven membrane-spanning helices. Not only does a similar pattern of repeating hydrophobic sequences 20–25 residues long occur in the predicted βAR sequence, but also the amino-acid composition of these postulated helices is similar to that of the rhodopsins, having a high proportion of proline and aromatic amino acids. The exact number of transmembrane regions remains to be determined. Amino-acid homology was apparent when the sequences corresponding to the postulated helices V, VI and VII of bovine opsin[14], which comprise the retinal binding site[15], were aligned with those for the analogous regions of βAR (Fig. 5*c*).

The sequence homology between βAR and rhodopsin parallels similarities in their function: both rhodopsin and βAR are involved in signal transduction mechanisms that involve interaction with the guanine nucleotide regulatory proteins transducin[16] and G_s (ref. 17), respectively. Moreover, it seems that phosphorylation has an important role in the regulation of both rhodopsin and βAR[18,19]. Rhodopsin is multiply phosphorylated at its C-terminus, which contains seven serine

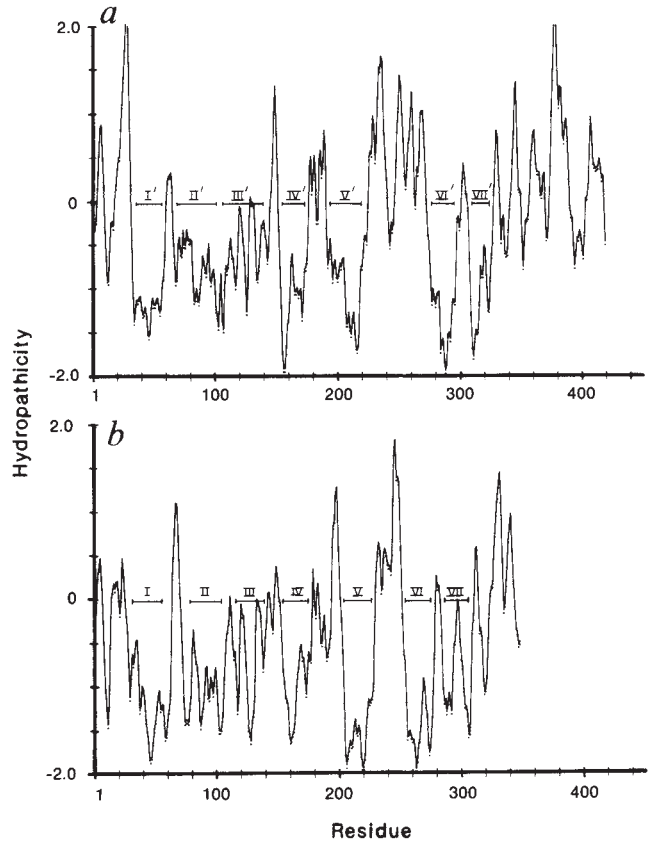


Fig. 5 Structure and sequence homology for hamster βAR and bovine opsin . Hydropathicity profiles are shown for hamster βAR (*a*) and bovine opsin[14] (*b*). Values were calculated by the method of Hopp and Woods[12]. Hydrophobicity increases with decreasing values. Horizontal lines indicate hydrophobic peptide regions 20–25 residues long. The putative transmembrane helices of bovine opsin are numbered as in ref. 14. We have designated the proposed transmembrane helices of βAR based on the analogous hydropathicity patterns of βAR and opsin. *c*, Amino-acid sequence homologies between the hamster βAR and bovine opsin (OPS). Positions of amino-acid identities are designated by asterisks. The underlined lysine residue (296) of bovine rhodopsin is involved in a Schiff base with retinal[15].

and threonine residues[18]. The C-terminus of the postulated sequence for βAR contains several serine and threonine residues that could serve as sites for phosphorylation[19].

The similarity in both the amino-acid sequence and proposed structure for the retinal binding site of rhodopsin and the analogous region in βAR suggests a specific mode of action for the ligands that modulate βAR function. We propose that these compounds interact with the βAR in a manner similar to that by which retinal interacts with opsin[21]; that is, they intercalate among the hydrophobic transmembrane helices and thereby determine whether the receptor is in its active or inactive conformation. According to this model, agonists and antagonists would mimic the action of photoactivated and native retinal. The generality of this hypothesis will be tested once the genes for similar membrane-bound receptors, such as those for leuko-

trienes, prostaglandins, dopamine and histamine, have been isolated. Our proposed model for the structure of βAR and its interaction with pharmacologically important ligands should, together with the biochemical and genetic studies now possible, provide a rational basis for a new approach to the development of more selective drugs.

We thank Lenora Davis, Jill D'Alonzo, Ester Hou, Tong Sun Kobilka, John Rodkey, Claudia Staniszewski, H. Vincent Strout Jr and Barbara Zemcik for technical assistance, Jeanne White and Lynn Tilley for manuscript preparation, and Drs Eugene Cordes, Ronald Ellis, Russell Kaufman, Paul Keller, Pete Kniskern, Richard Lebovitz, Mark Riemen and Michael Rosenblatt for helpful discussions. We particularly thank Dr Edward Scolnick for continued support and direction.

Received 10 March; accepted 11 April 1986.

1. Lefkowitz, R. J., Stadel, J. M. & Caron, M. G. *A. Rev. Biochem.* **52,** 159-186 (1983).
2. Shorr, R. G., Lefkowitz, R. J. & Caron, M. G. *J. biol. Chem.* **256,** 5820-5826 (1981).
3. Shorr, R. G., Strohsacker, M. W., Lavin, T. N., Lefkowitz, R. J. & Caron, M. G. *J. biol. Chem.* **257,** 12341-12350 (1982).
4. Homcy, C. J., Rockson, S. G., Countaway, J. & Egan, D. A. *Biochemistry* **22,** 660-668 (1983).
5. Benovic, J. L., Shorr, R. G. L., Caron, M. G. & Lefkowitz, R. J. *Biochemistry* **23,** 4510-4518 (1984).
6. Cerione, R. A. *et al. J. biol. Chem.* **259,** 9979-9982 (1984).
7. Engelman, D. M., Goldman, A. & Steitz, T. A. *Meth. Enzym.* **88,** 81-88 (1982).
8. Temeles, G. L., Gibbs, J. B., D'Alonzo, J. S., Sigal, I. S. & Scolnick, E. M. *Nature* **313,** 700-703 (1985).
9. Seidah, N. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **81,** 2640-2644 (1984).
10. Delavier-Klutchko, C., Hocheke, J. & Strasberg, A. D. *FEBS Lett.* **169,** 151-155 (1984).
11. Stiles, G. L., Benovic, J. L., Caron, M. G. & Lefkowitz, R. J. *J. biol. Chem.* **259,** 8655-8663 (1984).
12. Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78,** 3824-3828 (1981).
13. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157,** 105-132 (1982).
14. Nathans, J. & Hogness, D. S. *Cell* **34,** 807-814 (1983).
15. Hubbell, W. L. & Bownds, M. D. *A. Rev. Neurosci.* **2,** 17-34 (1979).
16. Stryer, L. *A. Rev. Neurosci.* **9,** 87-119 (1986).
17. Gilman, A. G. *Cell* **36,** 577-579 (1984).
18. Hargrave, P. A. *Prog. Retinal Res.* **1,** 1-51 (1982).
19. Benovic, J. L., Strasser, R. H., Caron, M. G. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
20. Benovic, J. L. *et al. J. biol. Chem.* **260,** 7094-7101 (1985).
21. Thomas, D. D. & Stryer, L. *J. molec. Biol.* **154,** 145-157 (1982).
22. Caron, M. G. & Lefkowitz, R. J. *J. biol. Chem.* **251,** 2374-2384 (1976).
23. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. *J. biol. Chem.* **256,** 7790-7797 (1981).
24. Spiess, J., Rivier, J. E., Rodkey, J. A., Bennett, C. D. & Vale, W. *Proc. natn. Acad. Sci. U.S.A.* **76,** 2974-2978 (1979).
25. Laemmli, U. K. *Nature* **117,** 680-685 (1970).
26. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4350-4356 (1979).
27. Johnson, D. A. & Elder, J. H. *J. exp. Med.* **159,** 1751-1756 (1983).
28. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
29. Merrifield, R. B. *J. Am. chem. Soc.* **85,** 2149-2154 (1963).
30. Frischauf, A. M., Lehrach, H., Paustka, A. & Murray, N. *J. molec. Biol.* **170,** 827-842 (1983).
31. Lauffer, L. *et al. Nature* **318,** 334-338 (1985).
32. Benton, W. D. & Davis, R. W. *Science* **196,** 180-182 (1977).
33. Ullrich, A., Berman, C. H., Dull, T. J., Gray, A. & Lee, J. M. *EMBO J.* **3,** 361-364 (1984).
34. Southern, E. M. *J. molec. Biol.* **98,** 503-517 (1975).
35. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560-565 (1977).
36. Vieira, J. & Messing, J. *Gene* **19,** 259-268 (1982).
37. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463-5467 (1977).
38. Schaffner, W. *et al. Cell* **14,** 655-671 (1978).
39. Nagata, S., Mantei, N. & Weissman, C. *Nature* **287,** 401-408 (1980).
40. Lawn, R. M. *et al. Nucleic Acids Res.* **9,** 1045-1052 (1981).
41. Volckaert, G., Feunteun, J., Crawford, L. V., Berg, P. & Fiers, W. J. *J. Virol.* **30,** 674-682 (1979).
42. Hamer, D. H. & Leder, P. *Cell* **18,** 1299-1302 (1979).
43. Kozak, M. *Nucleic Acids Res.* **12,** 857-872 (1984).
44. Green, S. *et al. Nature* **320,** 134-139 (1986).
45. Friedlander, M. & Blobel, G. *Nature* **318,** 338-343 (1985).
46. Feramisco, J. L., Glass, D. B. & Krebs, E. G. *J. biol. Chem.* **255,** 4240-4245 (1980).
47. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18,** 5294-5299 (1979).
48. Gubler, U. & Hoffman, B. J. *Gene* **25,** 263-269 (1983).
49. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning a Practical Approach* Vol. 1, 49-78 (IRL, Oxford, 1985).

In vitro osteoclast generation from different bone marrow fractions, including a highly enriched haematopoietic stem cell population

Ben A. A. Scheven*, Jan W. M. Visser† & Peter J. Nijweide*

* Laboratory of Cell Biology and Histology, University of Leiden, Rijnsburgerweg 10, 2333 AA Leiden, The Netherlands
† Radiobiological Institute TNO, Lange Kleiweg 151, 2288 GJ Rijswijk, The Netherlands

It is well established that the osteoclast is formed by fusion of post-mitotic, mononuclear precursors[1] derived from circulating progenitor cells[2]. However, the precise haematopoietic origin of the osteoclast is unknown. We have investigated this here by fractionating mouse bone marrow and isolating haematopoietic stem cells using a three-step method combining equilibrium density centrifugation and two fluorescence-activated cell sortings (FACS)[3], and have tested the ability of each bone marrow fraction, including highly purified haematopoietic stem cells, to generate osteoclasts during co-culture with preosteoclast-free embryonic long bones[4,5]. The osteoclast-forming capacity was found to increase with increasing stem cell purity. On the other hand, the culture time needed for osteoclast formation also increased with purification, suggesting the presence of progressively more immature progenitor cells. The pluripotent haematopoietic stem cell fractions with the highest purity needed preincubation with a stem cell-activating factor (interleukin-3) to activate the predominantly quiescent stem cells *in vitro*.

Embryonic long bones stripped of the endogenous osteoclast precursor pool (the periosteum) were used to induce the development of osteoclasts from different bone marrow cell populations. When stripped long bones are co-cultured with cell populations containing osteoclast progenitors, osteoclasts develop which thereby give rise to the formation of a marrow cavity. Culture of stripped bones alone does not lead to appearance of osteoclasts[4,5]. Total bone marrow cell population and both high-density ($1.100 < \rho < 1.078\ g\ cm^{-3}$) and low-density ($\rho < 1.078\ g\ cm^{-3}$) bone marrow fractions, obtained by equilibrium density centrifugation on a discontinuous metrizamide gradient, gave rise to osteoclast formation (Table 1). Autoradiography using continuous ^{3}H-thymidine labelling indicated that the osteoclasts were derived from progenitor cells which had proliferated during co-culture. The osteoclasts exhibited tartrate-resistant acid phosphatase activity, which is specific for these cells. Some mononuclear cells and an occasional multinucleated cell present in the plasma clot surrounding the bones also displayed this osteoclastic enzyme activity. Mature granulocytes accumulated in the developing marrow cavity following the invasion of osteoclasts. With time, other cell types such as megakaryocytes, macrophages and undifferentiated cells were observed in the expanded cavity.

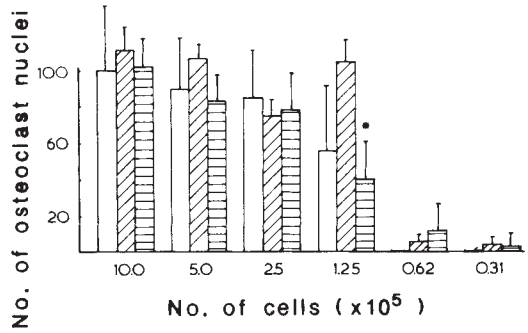


Fig. 1 Osteoclast formation in relation to the number of bone marrow cells (per plasma clot per bone) after 8 days of co-culture. Means ± s.d. of the number of osteoclast nuclei in five axial sections of four to six cultures are shown. □, Unfractionated bone marrow cells; ▨, low-density bone marrow cells; ▤, high-density bone marrow cells. Statistically significant ($P < 0.05$) compared with low-density cells.

Dilution experiments were performed to assess the relationship between the number of cells added to the bones and the