Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast

Michael M. Wang & Randall R. Reed

Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA

A novel genetic selection in yeast has been used to isolate a complementary DNA for the transcriptional activator, Olf-1, which binds to the regulatory sequences of several olfactory-specific genes. The Olf-1 protein, expressed exclusively in the olfactory receptor neurons and their precursors, contains a new helix–loop–helix motif and functions as an apparent homodimer. Olf-1 may be the first member of a family of related proteins that may direct cellular differentiation in a variety of neuronal tissues.

THE mammalian olfactory system has the remarkable ability to detect odorants with high sensitivity and specificity. The initial events in the olfactory signal transduction pathway occur in the specialized cilia of the sensory neurons[1]. Unlike other neurons, these olfactory sensory cells are continually replaced throughout adult life[2]. This continued neurogenesis makes the olfactory system an excellent model for the genetic and developmental processes that accompany neuronal differentiation.

Cells within the olfactory epithelium follow a highly ordered developmental programme resulting in the high-level expression of gene products essential for odorant signal transduction. Basal cells, which lie near the basal lamina of this pseudostratified epithelium, give rise to progenitors of the mature sensory neurons as well as non-neuronal support cells. The neuronal precursor cells follow a developmental programme that directs their differentiation into mature olfactory neurons. During this process each neuron extends a dendritic process to the surface of the epithelium and a single axon to the olfactory bulb where the mature neuron synapses with second order neurons. At about this time, the mature neuron expresses several olfactory-specific genes, some of which appear to mediate the odorant signal transduction cascade. Evidence supporting the involvement of a G-protein-coupled receptor pathway in odorant signal transduction has come from the isolation of olfactory-specific components that correspond to each step in the pathway. These include $G_{\alpha olf}$ (ref. 3), type III adenylyl cyclase[4] and a cyclic-nucleotide-activated ion channel, OcNC (refs 5,6). Additional olfactory neuron-specific genes have also been identified whose functions remain unknown[7].

The establishment of the mature olfactory neuronal phenotype probably results from the coordinated expression of olfactory-specific genes. Experiments in other systems have suggested that tissue-specific transcription factors are responsible for turning on tissue-specific genes through interactions with specific DNA sequences. The myogenic helix–loop–helix family of transcription factors trigger the conversion of several cell lines into myoblasts through sequence-specific interactions with muscle-specific genes[8–10]. To understand the regulation olfactory neuronal-specific gene expression, we have investigated the structure of six olfactory-specific genes (M.M.W. and R.R.R., manuscript in preparation). Our studies indicate that each of these genes contains at least one binding site for the DNA-binding protein, Olf-1. The binding of an olfactory-specific factor, Olf-1 was first described in the OMP gene[38]. The Olf-1 activity is detected in nuclear extracts from nasal epithelium and is absent from nuclear extracts of a variety of other tissues. The transcriptional activator Olf-1, may be the critical protein factor involved in activating the expression of the signal transduction components and other genes of the olfactory system.

We have identified a cDNA clone encoding the Olf-1 binding activity using a genetic selection in yeast. The deduced amino-acid sequence of Olf-1 is predicted to contain two helical repeats with similarity to the helix–loop–helix class of transcription factors, and may represent the first identified member of a novel subclass containing this protein motif. The Olf-1 protein is restricted to the mature olfactory neurons and their progenitors in olfactory epithelium. Olf-1 is the first transcriptional activator identified likely to direct the coordinate expression of cell-specific genes in a neuronal system.

Cloning of a specific DNA-binding protein

We have characterized the promoter regions for three genes (G_{olf}, type III adenylyl cyclase and the olfactory cyclic nucleotide gated channel) involved in the olfactory second messenger cascade and three additional genes (OMP[7], 50.06, and 50.11) specifically expressed by olfactory neurons (M.M.W. and R.R.R., manuscript in preparation). A *cis*-acting regulatory sequence present in each of these genes was identified and is described in detail elsewhere. Briefly, an activity, Olf-1, detectable only in olfactory nuclear extracts was found to bind to a site 150 nucleotides upstream of the OcNC transcriptional start site. Using this site in gel mobility assays, sites were identified in each of the other olfactory-specific genes. DNA binding sites (U sites) for other factors present in many tissues exist in some of these promoter regions. The sequence of each of the binding sites as well as their position and orientation in the olfactory neuron-specific genes is shown in Fig. 1.

The low abundance of sensory neurons in olfactory tissue preparations suggested that purification of the Olf-1 binding activity by biochemical methods would prove difficult. Although several DNA-binding proteins have been cloned by screening recombinant expression libraries with radiolabelled concatemerized binding sites[11,12], attempts to screen olfactory cDNA expression libraries with the OcNC Olf-1 binding sites were unsuccessful. We developed a novel cloning strategy using a genetic selection of olfactory cDNAs encoding proteins capable of activating a reporter construct in yeast (Fig. 2). Conceptually, we wished to identify plasmids in a cDNA library that could direct the expression of a protein that binds the Olf-1 site. The olfactory cDNA library was expressed in yeast as a translational fusion with a GAL4 transactivating domain. The reporter construct consisted of three copies of the OcNC Olf-1 binding site adjacent to a low activity promoter directing *HIS3*

© 1993 Nature Publishing Group

gene expression. The presence of this reporter is insufficient to rescue the *his3*$^-$ phenotype of the parent yeast strain. Yeast that harbours a GAL4–olfactory fusion protein capable of binding the Olf-1 sites present in the reporter will activate transcription and increase levels of *HIS3*. The use of the selectable marker *HIS3* enabled direct selection of candidate clones. Others have used yeast expression systems to identify subunits involved in protein–protein interactions[13]. The system to identify specific DNA–protein interactions described here is fundamentally similar to the two-hybrid system[14] and is conceptually complementary to previously described binding-site selection schemes in yeast[15].

An olfactory cDNA library of 3.6 million clones was constructed in the yeast expression vector and transformed into

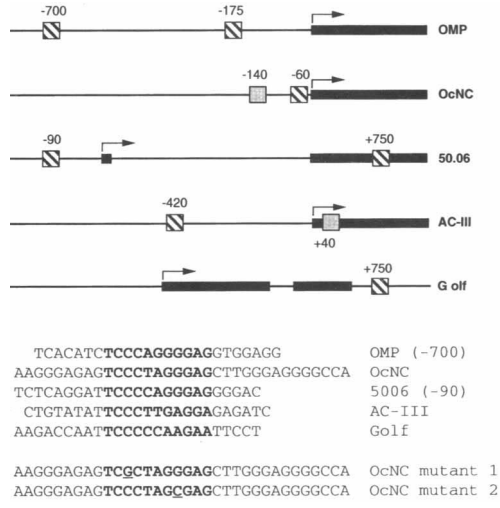


FIG. 1 The genomic structure, location of Olf-1 sites within olfactory specific genes, and sequences of Olf-1 sites. The olfactory cyclic nucleotide channel (OcNC), the olfactory neuron-specific G protein (G_{olf}) and type III adenylyl cyclase (AC-III) are components of the odorant detection pathway whereas olfactory marker protein (OMP) and 50.06 are of unknown function. In each map, the exons of the gene are represented by black bars; the start site for transcription is indicated by an arrow and Olf-1 (striped boxes) and U sites (unpublished data) (bound by a distinct, broadly expressed DNA-binding site, shaded boxes) are indicated. The top strand sequences of the double-stranded oligonucleotides corresponding to Olf-1 sites in the indicated genes are listed with the Olf-1 site in bold. Mutant oligonucleotides used in DNA binding studies are shown at the bottom with the single base changes underlined.

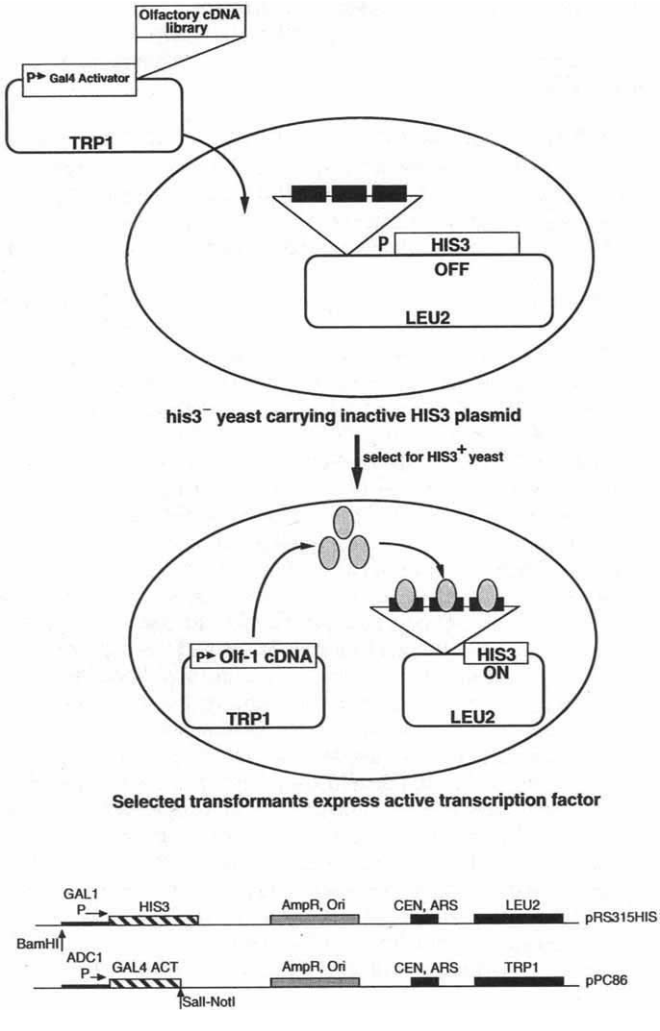


FIG. 2 Scheme and vectors used for cloning of the Olf-1 DNA-binding protein by selection in yeast. The yeast strain yWAM2 was transformed with the *HIS3*-containing plasmid p59.7 to generate the illustrated recipient strain for the expression library. The plasmid in this strain was maintained extrachromosomally by selection for LEU2. This yeast strain was then transformed with an olfactory cDNA library and plated directly on media lacking histidine. Histidine prototrophs that grew in the absence of tryptophan and leucine presumably arose by transactivation mediated by an olfactory cDNA and were further analysed. Bottom, A more detailed schematic of the reporter vector and expression vector used with the unique restriction sites used for insertion of Olf-1 sites and olfactory cDNA, respectively. METHODS. The yeast strain yWAM2 (*MATα Δgal4 Δgal80 URA3::GAL1-lacZ lys2801*^{amber} *his3-Δ200 trp1-Δ63 leu2ade2-101*^{ochre}*CYH2*; provided by P. Hieter and W. Michaud) was derived from PCY2 (ref. 27). The yeast reporter plasmids were constructed by insertion of annealed complementary oligonucleotides containing Olf-1 sites into the *Bam*HI site of the vector pRS315HIS. The plasmid pRS315HIS was generated by inserting a *Bam*HI–*Sal*I fragment containing the GAL1 minimal promoter adjacent to *HIS3* from plasmid HR307a[15] into pRS315[28]. The p59.7 reporter construct used in the selection for Olf-1 binding activity contained three copies of the OcNC Olf-1 binding site. The sites were concatamerized by ligating annealed oligonucleotides to yield *Bam*HI- and *Bgl*II-compatible ends. Yeast transformed with p59.7 were maintained by growth in supplemented synthetic dextrose media (SD) without leucine. Olfactory epithelium cDNA generated from 3-week-old Sprague-Dawley rats was directionally cloned into the *Sal*I–NotI sites of the TRP-1 marked yeast expression vector pPC86[27] using the Superscript cDNA synthesis system (BRL). The ligation products were electroporated into *E.coli* strain DH10B, and the resultant transformants (3.6×10^6 with an average insert size of 1.5 kb) plated onto LB/carbenicillin plates at a density of 1×10^5 colonies per plate. The plasmid library DNA was purified over Qiagen columns after scraping the cells from the plate, and used for transformation into yeast by the PEG/lithium acetate method[29]. All yeast transformations were done with an efficiency of $>10^4$ μg^{-1} vector. Transformations with the olfactory cDNA library were plated directly on SD media without histidine (but including leucine and tryptophan) and incubated for 3 days at 30 °C. Transformation efficiency was determined by plating an aliquot of the transformation mixture onto SD media minus leucine and tryptophan.

© 1993 Nature Publishing Group

a *his3*$^-$ yeast strain carrying the *HIS3* reporter plasmid. Two million transformants were plated on media lacking histidine and 30 HIS$^+$ colonies were isolated. One colony grew on media lacking tryptophan, indicating that it contained the *TRP1*-marked GAL4–olfactory cDNA expression plasmid. This plasmid (Y11) was recovered by transformation into *Escherichia coli*. Y11 is also able to activate *HIS3* transcription in yeast containing a similar reporter plasmid with two copies of an Olf-1 site from the 50.06 gene (Fig. 1). In contrast, the Y11 plasmid fails to rescue the HIS$^+$ phenotype of cells containing the parental *HIS3* vector. These results are consistent with a specific interaction between a protein encoded by Y11 and the Olf-1 sites that leads to transcriptional activation of *HIS3*.

Y11 encodes an unusual helix–loop–helix protein

Nucleotide sequence analysis of the Y11 insert revealed an open reading frame of 570 amino acids which initiated from the first methionine in good context for translation initiation[16] (Fig. 3). An additional ATG codon in poor context for translation initiation precedes the predicted start codon. Analysis of genomic DNA and RACE-PCR products (data not shown) demonstrates that the presumed initiation methionine is preceded in the full-length mRNA by an in-frame stop codon. As expected, the Y11 cDNA insert was in frame with the GAL4 transactivator domain (GAL4TA), enabling Y11 to direct expression of a fusion protein between the GAL4TA and the novel 570-amino-acid polypeptide.

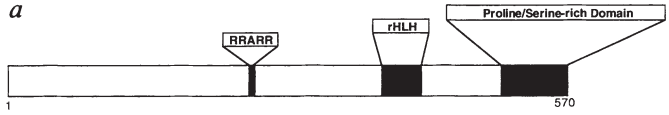
The carboxy-terminal region (residues 441–570) of the Y11 protein is extraordinarily rich in serine (23%) and proline (12%) and may serve as a transactivation domain[17]. In addition, residues 228–232 (Arg-Arg-Ala-Arg-Arg) fit the consensus for a nuclear localization signal. Secondary structure analysis of the Y11-encoded protein predicted that amino acids 354–391 form an α-helix followed by a turn or loop and a second α-helical region. These helices, nearly identical at the amino-acid level, shared modest similarity to the second helix of the helix–loop–helix (HLH) family of transcription factors[18,19]. Specifically, this region of Y11 was identical at several hydrophobic residues that have been conserved among the HLH transcription factors and are believed to participate in inter- and intramolecular coiled-coil interactions[20]. The Y11 protein displayed no homology to the first helix of the HLH family. Though Y11 protein contains several basic residues upstream of the helices, it lacks highly conserved basic residues immediately upstream of the HLH region present in members of the HLH family which activate transcription[21,22]. These observations suggest that Y11 represents the first member of a novel class of proteins containing a repeated helix–loop–helix (rHLH) motif. In light of the properties of the Y11 cDNA, we conclude that Y11 is likely to encode the Olf-1 binding activity. The data presented below strengthens this conclusion.

Previous biochemical characterization of the Olf-1 binding activity suggested that the protein was confined to the olfactory epithelium. To establish that Y11 encoded the expected olfactory-specific binding activity, the level of Y11 messenger RNA expression in six tissues was examined by reverse transcriptase polymerase chain reaction (RT-PCR). A PCR product of the expected size was detected exclusively in template derived from olfactory epithelium (Fig. 4*a*). This expression pattern paralleled that determined for Olf-1 binding [38].

The DNA binding properties of the Y11-encoded protein were assessed by electrophoretic mobility shift assay (EMSA). Double-stranded oligonucleotides containing Olf-1 sites from the five olfactory-specific genes were specifically shifted by protein translated in a reticulocyte lysate (Fig. 4*b*). In contrast, two oligonucleotides containing single base changes in the OcNC site, known to eliminate the ability of native Olf-1 protein to bind DNA, failed to produce the shifted product. Polyclonal antisera raised to amino acids 360–570 of Y11 specifically recognized the protein–DNA complex present in nuclear extracts from olfactory epithelium. When the Y11 antiserum was mixed with preformed Olf-1–DNA complexes, a supershifted band of slower mobility representing a ternary complex of Olf-1 protein/labelled DNA probe/antibody was observed (Fig. 4*c*). The preimmune serum did not affect the mobility of the Olf-1 complex. Moreover, the Y11 antisera did not supershift an unrelated DNA–protein complex (Fig. 4*c*).

The consensus Olf-1 binding site (TCCYYRRGGAG) is nearly symmetric implying that Olf-1 may bind to its site as a dimer. Furthermore, the similarity of the rHLH region of Olf-1 to the basic HLH (bHLH) domains, shown to mediate dimerization in the MyoD family, suggested that Olf-1 participates in stable protein–protein interactions. C-terminal truncated forms of Olf-1 that include the rHLH region retained their ability to bind DNA but resulted in a faster mobility complex than was observed with the full-length protein. The cotranslation of mRNA derived from truncated and full-length Olf-1 cDNA in reticulocyte lysates led to the formation of a

a



b

```
MKEEPLGSGMNAVRTWMQGAGVLDANTAAQSGVGLARAHFEKQPPSNLRK  50
SNFFHFVLALYDRQGQPVEIERTAFVGFVEKEKEANSEKTNNGIHYRLQL 100
LYSNGIRTEQDFYVRLIDSMTKQAIVYEGQDKNPEMCRVLLTHEIMCSRC 150
CDKKSCGNRNETPSDPVIIDRFFLKFFLKCNQNCLKNAGNPRDMRRFQVV 200
VSTTVNVDGHVLAVSDNMFVHNNSKHGRRARRLDPSEAATPCIKAISPSE 250
GWTTGGATVIIIGDNFFDGLQVIFGTMLVWSELITPHAIRVQTPPRHIPG 300
VVEVTLSYKSKQFCKGTPGRFIYTALNEPTIDYGFQRLQKVIPRHPGDPE 350
RLPKEVILKRAADLVEALYGMPHNNQEIILKRAADIAEALYSVPRNHNQL 400
PALANTSVHAGMMGVNSFSGQLAVNVSEASQATNQGFTRNSSSVSPHGYV 450
PSTTPQQTNYNSVTTSMNGYGSAAMSNLGGSPTFLNGSAANSPYAIVPSS 500
PTMASSTSLPSNCSSSSGIFSFSPANMVSAVKQKSAFAPVVRPQTSPPPT 550
CTSTNGNSLQAISGMIVPPM                                 570
```

Helix-Loop-Helix of Olf-1:

```
RLPKEVILKRAADLVEALYGMPHN   helix1 (rat)
  EIILKRAADIAEALYSVPRN     helix 2 (rat)
```

Helix-2 sequence:

```
PQTKLLILHQAVAVILSLEQQ   E12 (rat)
RLPKVEILRNAIRYIEGLGAL   myoD (mouse)
KVSKVEILQHVIDYIRDLQLE   Id (mouse)
KAPKVVILKKATAYILSVQAD   c-myc (rat)
KASRAQILDKATEYIQYMRRK   max (mouse)
KLSKVDTLRIAVEYIRGLQDM   achaete sc. (Dros.)
```

FIG. 3 Sequence and domain structure of Y11 and homology to the helix–loop–helix class of transcription factors. Schematic diagram (*a*) and deduced amino-acid sequence (*b*) of Y11 are shown on top. The locations of the putative nuclear localization signal (RRARR), repeat helix–loop–helix (rHLH) and the serine/proline-rich activation domain are indicated. The sequence of the repeated helix–loop–helix motif predicted by both the Chou–Fasman and Robson–Garnier secondary structure algorithms[30,31] is denoted by arrows. Comparison of the rHLH sequence of Y11 to the sequences of the second helix from members of the HLH family is displayed[18,19]. Residues that are identical in both helices of the predicted rHLH domain are indicated with dots and shaded when a single amino acid appears in at least 5 of the 8 sequences.

METHODS. Total DNA was prepared from yeast Y11 and used to transform *E.coli*. The rescued plasmid Y11 was sequenced on both strands using a combination of exonuclease III deletions and synthetic primers with Sequenase (USB).

© 1993 Nature Publishing Group

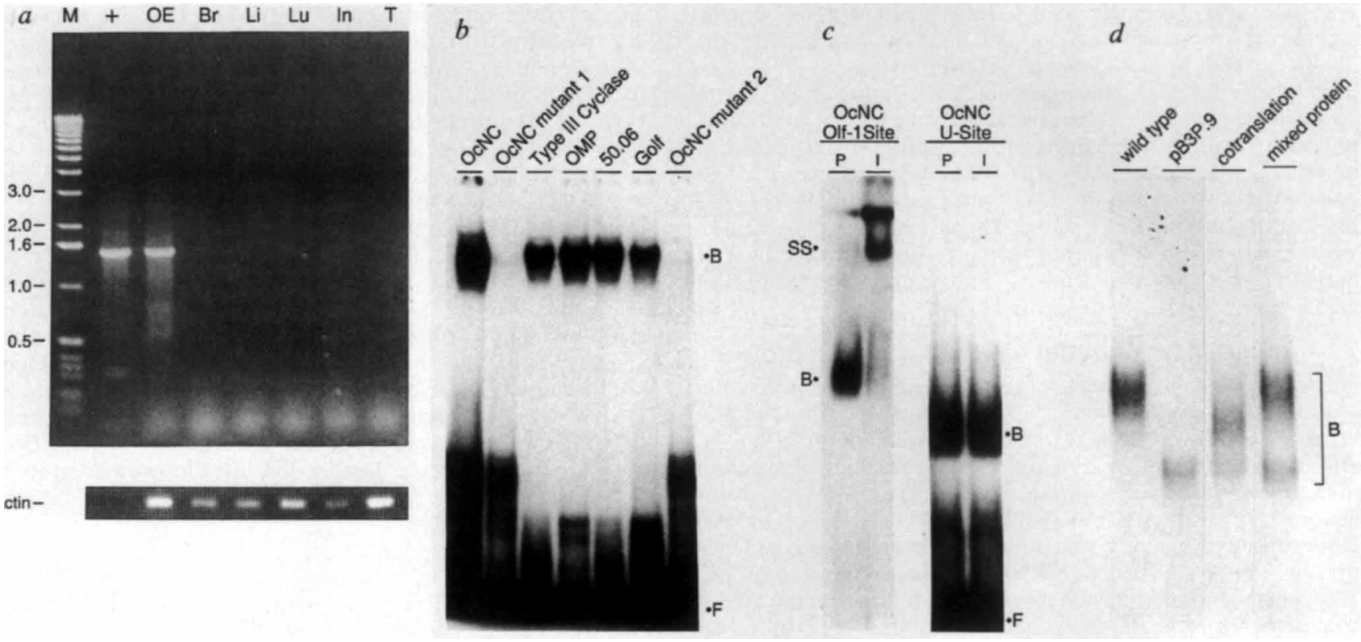
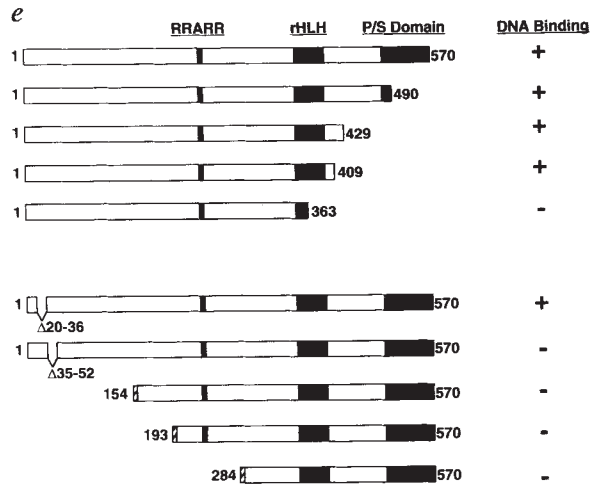


FIG. 4*a*, Expression levels of the Y11 mRNA in various tissues. Y11-specific oligonucleotides were used for RT-PCR on first strand cDNA and the PCR products analysed by gel electrophoresis as previously described[32]. M, molecular mass markers; +, 1 ng purified Y11 plasmid as a positive control; OE, olfactory epithelium; Br, brain; Li, liver; Lu, lung; In, intestine; T, testis. *b*, Gel shifts using *in vitro* expressed protein programmed from Y11. Protein produced by *in vitro* transcription and translation of the Y11 cDNA was incubated with radiolabelled probes encoding Olf-1 sites and analysed by EMSA. The sequence of each of the oligonucleotides sites is shown in Fig. 1. *c*, Antibodies to the Y11-encoded protein supershifts the Olf-1–DNA complex. Binding reactions were done by combining olfactory nuclear extracts with ^{32}P-labelled Olf-1 site from OcNC or a sequence (Fig. 1) present in the OcNC promoter which binds to a broadly expressed nuclear factor (U box - 5′ GATCCTGTCTGGAACTCTGCCAGCATTTAA) (M. M. W. and R. R. R, manuscript in preparation). Pre-immune serum (P) or polyclonal antiserum JH865 (I) was added to each reaction and the mixture analysed by EMSA. Free Olf-1 oligonucleotide probe was run off the gel. *d*, Olf-1 binds to DNA as a multimer. The full-length Olf-1 protein (wild type), truncated Olf-1 (pB3P.9) and cotranslated or post-translationally mixed full-length and truncated Olf-1 were incubated with labelled OcNC Olf-1 oligonucleotide and used for EMSA. The bracket indicates the position of bound complexes in each of the lanes *e*, Deletional analysis of the Olf-1 protein. Olf-1 protein was generated by *in vitro* transcription/translation of appropriate templates resulting in products whose primary structure are illustrated. The ability of truncated and deleted proteins to bind radiolabelled Olf-1 sites was assessed in gel mobility shift assays. The positions of the N- and C-terminal truncations and internal delection endpoints are shown. N-terminal deletions were fused to a synthetic adaptor, depicted by the hatched box, to provide an initiator methionine.

METHODS. RT-PCR was done on total RNA from the indicated tissues using primers designed to amplify Y11 message. The oligonucleotides MW125 (5′ CCACAGTCAACGTGGAT amino acids 203–208) and MW122 (5′ AGTCTTATTTAGAGTGGC 3′ untranslated region) were used in PCR reactions with 5μM primer and 1 ng oligo-dT-primed cDNA template per 20 μl reaction. Reaction conditions were 94 °C for 45 s, 56 °C for 1 min, and 72 °C for 2 min for 35 cycles. The insert of Y11 was subIconed into Bluescript KS$^-$ to create the plasmid pB3P, and protein for gel shifts were prepared by transcribing sense RNA from linearized pB3P and translating 1μg of the RNA in 25 μl of rabbit reticulocyte lysate as described by the manufacturer (Promega). A 2μl aliquot of the translation mix was added to a 20 μl binding reaction (10mM HEPES (pH 7.9 at 4 °C), 10mM $MgCl_2$, 10% glycerol, 50mM KCl, 0.5mM DTT, 50μg ml^{-1} dIdC, 50μg ml^{-1} salmon sperm DNA, and 50pg of ^{32}P-labelled oligonucleotide probe (2,000Ci mmol^{-1}). After a 10 min incubation on ice, the mixture was electrophoresed on a 6% polyacrylamide gel (59:1) acrylamide:bisacrylamide) in 0.25 × TBE (1×TBE is 90mM TrisBorate pH 8.3, 2mM EDTA) at 4 °C. The gel was dried and exposed to film for 6 h. Protein from olfactory nuclear extracts[33] were added to 20-μl binding reactions to a final protein concentration of 0.1mg ml^{-1} and incubated for 10 min on ice. Serum (1 μl) was added to the mixture and the reaction continued for 10 min on ice. Immune serum was obtained from rabbits injected with GST-Y11 fusion protein prepared from *E. coli* BL21 cells carrying a *BglII–HinDIII* fragment of B3P, corresponding to amino acids 363 to 570, ligated into pGEX-1[34]. Protein was eluted from glutathione-agarose beads by incubation in 50mM Tris–HCl pH 8.0,5mM reduced glutathione[35]. Full-length Olf-1 was produced in reticulocyte lysate as described above. Truncated Olf-1 protein was generated by translating RNA from a T3 RNA polymerase transcription from linearized plasmid pB3P.9 which encodes amino acids 1–430 of Y11. pB3P.9 was derived by subcloning into Bluescript the *SalI–SacII* fragment of an exonuclease III-generated 3′ deletion of Y11. DNA templates for generating C-terminal truncations were prepared by linearizing pB3P with a restriction enzyme which cleaved within the coding region of the Olf-1 protein or by cloning partial exonuclease III deletions of the 3′ end of the cDNA. Plasmids used to synthesize N-terminal-deleted proteins were derived from partial clones isolated from an olfactory cDNA library. The *Eco*RI inserts of these clones were ligated to a *SalI–Eco*RI adaptor encoding an in-frame initiatior methionine followed by nucleotides encoding the sequence VPPKKKRKVLR (hatched leader sequence), digested at a unique *Hin*dIII site in the 3′ untranslated region, and the *SalI–Hin*dIII fragment subcloned into a plasmid capable of directing *in vitro* transcription. DNA that coded for the internal deletion of residues 20 to 36 was created by digestion of pB3P with *NarI* and *SmaI*, blunting the ends with the large fragment of DNA polymerasel I, and recircularizing with DNA ligase. This procedure resulted in the substitution of a glycine at the deletion junction. An internal deletion of residues 35 to 52 was generated by partial exonuclease III digestion of *SmaI*-cleaved pB3P. All deletion junctions were sequenced. The yield and size of the protein products generated from these constructs were verified by SDS–PAGE analysis of ^{35}S-labelled translation products.



© 1993 Nature Publishing Group

protein–DNA complex of intermediate mobility by EMSA (Fig. 4*d*). The intermediate position of this complex is consistent with the conclusion that Olf-1 binds DNA as a homooligomer. When the two proteins were mixed after translation, the novel complex was not produced.

The divergence of the Olf-1 helix–loop–helix region from the consensus bHLH and the absence of a conserved basic region in Olf-1 implied that Olf-1 interacted with DNA in a fundamentally different manner. Deletional analysis of the Olf-1 protein provided evidence that the rHLH region was necessary but not sufficient for DNA binding (Fig. 4*e*). Specifically, sequence contained within an 18 amino-acid deletion, situated more than 350 amino acids away from the rHLH, were necessary for binding DNA. This is in contrast to myoD, where a 68-amino-acid region containing the bHLH region of the protein is sufficient for high-affinity DNA binding and transcriptional activation[10,23]. Given the modular nature of DNA-binding proteins, we hypothesize that the rHLH region of Olf-1 might mediate oligomerization, whereas the N-terminal domain might be responsible for sequence-specific contacts with DNA.

Olf-1 localized to nuclei of olfactory neurons

To determine which cells in rat olfactory epithelium express the Olf-1 protein, we used immunohistochemistry with the Olf-1 antisera. Immunoreactive material was detected in the nuclei of mature olfactory neurons as well as in the region of the olfactory epithelium closer to the basal lamina which contains a high fraction of immature neuronal precursors[2,24,25] (Fig. 5). In contrast, the nuclei of sustentacular cells and the most basal cell population did not stain with the Olf-1 antibody. This was seen most clearly when the nuclei were revealed with the DNA-binding dye, DAPI. The presence of Olf-1 immunoreactivity in the mature olfactory neurons was consistent with the high level of expression of putative target genes in these cells. The possibility that Olf-1 immunoreactivity extends to the neuronal precursors suggests that Olf-1 may play an early role in directing

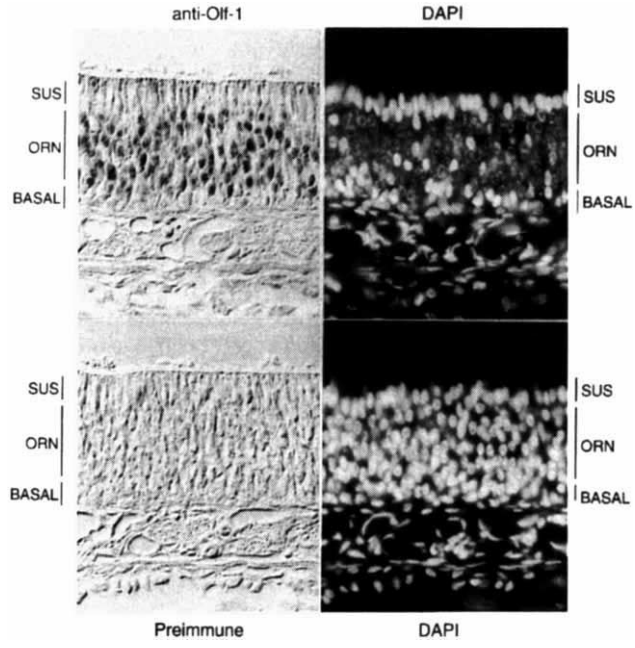


FIG. 5 Pattern of Olf-1 immunoreactivity in adult olfactory epithelium. Sections were incubated with the Olf-1-specific antiserum (top) or preimmune serum (bottom) and visualized with the Vectastain ABC Elite kit. SUS, Sustentacular cell nuclei; ORN, olfactory receptor neurons; BASAL, basal cells. To reveal the location of all the cell nuclei in the epithelium, sections were stained with the fluorescent DNA-binding dye DAPI (1 μg ml^{-1}) before mounting. Note that the diaminobenzidine immunoreaction product quenches the DAPI fluorescence and serves to identify the immunoreactive cell nuclei under bright field (left) and ultraviolet illumination (right).
METHODS. Adult male SD rats were anaesthetized and perfused. Paraformaldehyde-fixed tissue from olfactory areas were embedded in paraffin and sectioned at 8μm. Immunohistochemistry was as described with the ABC kit (Vector Labs) except that sections were blocked with 10% goat serum. The immune serum JH865 was preabsorbed with an excess of GST-bound glutathione beads.

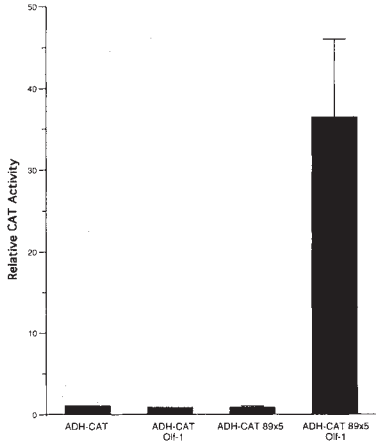


FIG. 6 Functional expression of Olf-1 in HEK 293 cells. Cells were cotransfected with a basal promoter CAT reporter construct (ADH-CAT) or a CAT reporter downstream of five copies of the distal 50.06 Olf-1 site (ADH-CAT89X5) in the presence or absence of a plasmid expressing Olf-1. CAT activity, normalized for protein concentration, was determined from four independent transfection experiments.
METHODS. Cotransfections were essentially as described [36]. Each transfection included 1μg CAT vector, 5μg expression vector, 5μg salmon sperm DNA and, when specified, 50ng Olf-1 expression plasmid per 6-cm dish. The reporter construct pADHCAT was constructed by ligating into Bluescript the *Hin*dIII fragment of pD-33CAT[37], containing the minimal promoter of the *Drosophila* alcohol dehydrogenase gene fused to CAT. Concatemerized oligonucleotides (Fig. 1; 5006-710) were cloned into the *Bam*HI site adjacent to the ADH sequences of pADHCAT to create pADHCAT89X5. The expression plasmid used has been described elsewhere[36]. CAT activity was determined by the phase extraction method[35]

© 1993 Nature Publishing Group

the expression of these genes and that additional levels of transcriptional or translational control contribute to the high levels of target gone expression seen in the most mature neuronal cells.

Activation

The initial yeast screen which identified Olf-1 was based on the ability of the cloned product, potentially fused to the GAL4 activator sequence, to stimulate transcription of *HIS3*. To assess the potential of Olf-1 to transactivate in a mammalian cell line, HEK 293 cells were cotransfected with Olf-1 cDNA and a reporter construct containing five Olf-1 sites upstream of a minimal promoter and the bacterial chloramphenicol acetyl transferase (CAT) gene. Cotransfection with the Olf-1 cDNA increased CAT enzyme activity more than 25-fold in an Olf-1 site-dependent fashion (Fig. 6). These data confirm that Olf-1 protein was sufficient to direct transcriptional activation in a mammalian cell line.

When extracts prepared from Olf-1 transfected cells are used in mobility shift assays, the complex detected is significantly slower in mobility than that detected in nuclear extracts prepared from olfactory tissue (data not shown). The mobility of the shifted product detected in the cell line is identical to that seen with *in vitro* translated Olf-1 protein. Proteolysis of the Olf-1 activity derived from tissue extracts may produce this difference in mobility. Although the Olf-1 protein that we have cloned demonstrates appropriate binding specificity and transcriptional activation as a homomultimer, we are currently examining whether heteromultimeric forms exist *in vivo*.

Discussion

We have successfully exploited a yeast genetic system to clone a specific transcription factor using its known target genes. In contrast to biochemical methods, selection in yeast does not require large amounts of starting material. Moreover, the expression of libraries in a eukaryotic system may allow for more efficient identification of candidate clones in comparison to bacterial expression libraries. Yeast represents a powerful system for the genetic selection of candidate clones that may encode specific DNA-binding proteins and transcriptional activators.

The myogenic and achaete schute bHLH genes encode transcriptional activators which specify cell fate and require the helix–loop–helix for binding. We have demonstrated that the Olf-1 gene encodes an olfactory tissue-specific transcriptional activator containing a novel helix–loop–helix domain essential for DNA binding. Although the Olf-1 protein has the capacity to activate synthetic promoters in transformed cell lines, it remains to be shown whether Olf-1 expression is sufficient to induce expression of endogenous olfactory-specific genes. The data presented are consistent with Olf-1 functioning as a homodimer *in vivo* that is present as a proteolysed dimer in nuclear extracts. Other transcription factors, most notably NFκB (ref. 26), undergo regulated proteolysis, which activates function.

The organization of sequences required for oligomerization and DNA binding distinguish Olf-1 from the previously known HLH transcription regulators and suggests that the bHLH family of proteins may also have unrecognized binding activities mediated by residues outside of the basic domain. The rHLH of Olf-1 may be a conserved structural motif present in other transcription factors; low-stringency hybridization of genomic Southern blots with a small fragment containing this region of the Olf-1 cDNA reveals several bands. Related rHLH proteins could function in olfactory differentiation in a combinatorial fashion as seen with the myogenic HLH proteins and their partners. Alternatively, other members of the rHLH family may function during the differentiation of other neuronal tissues.

Note added in proof: Hagman *et al.* have described the purification and cloning of the transcription factor EBF[39] from an early B-cell line. This gene appears to be an alternatively spliced form of the rat Olf-1 factor described here. The relationship between Olf-1 and EBF, the role of possible heterodimerization, post-translational modification and splice variants, and the nature of the respective binding sites in target genes will be subjects of future interest. □

Received 26 February; accepted 2 June 1993.

1. Reed, R. R. *Neuron* **8**, 205–209 (1992).
2. Graziadei, P. P. C. & Monti Graziadei, G. A. *The Olfactory System: A Model for the Study of Neurogenesis and Axon Regeneration in Mammals* 131–153 (Raven, New York, 1978).
3. Jones, D. T. & Reed, R. R. *Science* **244**, 790–795 (1989).
4. Bakalyar, H. A. & Reed, R. R. *Science* **250**, 1403–1406 (1990).
5. Dhallan, R. S., Yau, K. W., Schrader, K. A. & Reed, R. R. **347**, 184–187 (1990).
6. Ludwig, J., Margalit, T., Eismann, E., Lancet, D. & Kaupp, B. *FEBS Lett.* **270**, 24–29 (1990).
7. Rodgers, K. E. *et al Proc. Natn. Acad. Sci. U.S.A.* **84**, 1704–1708 (1987).
8. Davis, R. L., Weintraub, H. & Lassar, A. B. *Cell* **51**, 987–1000 (1987).
9. Lassar, A. B. *et al. Cell* **58**, 823–831 (1989).
10. Weintraub, H. *et al. Science* **251**, 761–766 (1991).
11. Vinson, C. R., LaMarco, K. L., Johnson, P. F., Landschultz, W. H. & McKnight, S. L. *Genes Dev.* **2**, 801–806 (1988).
12. Singh, H., LeBowitz, J. H., Baldwin, J. A. S. & Sharp, P. A. *Cell* **52**, 415–423 (1988).
13. Grueneberg, D., Natesan, S., Alexandre, C. & Gilman, M. *Science* **257**, 1089–1095 (1992).
14. Fields, S. & Song. O.-K. *Nature* **340**, 245–246 (1989).
15. Wilson, T. E., Fahrner, T. J., Johnston, M. & Milbrandt, J. *Science* **252**, 1296–1300 (1991).
16. Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
17. Henthorn, P., Kiledjian, M. & Kadesch, T. *Science* **247**, 467–470 (1990).
18. Blackwood, E. M. & Eisenman, R. N. *Science* **251**, 1211—1217 (1991).
19. Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. & Weintraub, H. *Cell* **61**, 49–59 (1990).
20. Anthony-Cahill, S. J. *et al. Science* **255**, 979–983 (1992).
21. Voronova, A. & Baltimore, D. *Proc. Natn. Acad. Sci. U.S.A.* **87**, 4722–4726 (1990).
22. Davis, R. L., Cheng, P.-F., Lassar, A. B. & Weintraub, H. *Cell* **60**, 733–746 (1990).
23. Tapscott, S. J. *et al. Science* **242**, 405–411 (1988).
24. Monti-Graziadei, P. P. C. in *Neural Growth and Differentiation* (eds Meisami, E. & Brazier, M. A. B.) 373–396 (Raven, New York, 1979).
25. Verhaagen, J., Oestreicher, W. H., Gispen, W. H. & Margolis, F. L. *J. Neurosci.* **9**, 683–691 (1989).
26. Fan, C. M. & Maniatis, T. *Nature* **354**, 395–398 (1991).
27. Chevray, P. M. & Nathans, D. *Proc. Natn. Acad. Sci. U.S.A.* **89**, 5789–5793 (1992).
28. Sikorski, R. S. & Heiter, P. *Genetics* **122**, 19–27 (1989).
29. Gietz, D., St. Jean, A., Woods, R. A. & Schiestl, R. H. *Nucleic Acids Res.* **20**, 1425–1425 (1992).
30. Chou, P. Y. & Fasman, G. D. *Rev. Biochem.* **47**, 251–276 (1978).
31. Garnier, J., Osguthorpe, D. J. & Robson, B. *J. molec. Biol.* **120**, 97–120 (1978).
32. Levy, N. S., Bakalyar, H. A. & Reed, R. R. *J. Ster. Biochem. molec. Biol* **39**, 633–637 (1991).
33. Dignam, J. D., Lebowitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475–1489 (1983).
34. Smith, D. B. & Johnson, K. S. *Gene* **67**, 31–40 (1988).
35. Ausubel, F. M. *et al. Current Protocols in Molecular Biology* (Wiley, New York, 1990).
36. Gorman, C. M. Gies, D. R. & McCray, G. *DNA and Protein Engineering Techniques* **2**, 3–10 (1990)
37. England, B. P., Ulrike, H & Tjian, R. *J. biol. Chem.* **265**, 5086–5094 (1990).
38. Kudrycki, K. *et al. Molec. cell. Biol.* **13**, 3002-3014 (1993).
39. Hagman, J. *et al. Genes Dev.* **7**, 760-763 (1993).

ACKNOWLEDGEMENTS. We thank P. Chevray, W. Michaud, P. Hieter, H. Shin, A. Cunningham and S. Desiderio for advice and for plasmids and yeast strains, and L. Levin for comments on the text. M.W. is supported by an MSTP award (NIH).

© 1993 Nature Publishing Group