Expression of the Oct-1 Transcription Factor and Characterization of Its Interactions with the Bob1 Coactivator†

Larissa Lee,‡ Elliott Stollar,‡ JuFang Chang,‡,§ J. Günter Grossmann,|| Ronan O'Brien,⊥ John Ladbury,⊥ Brian Carpenter,# Stefan Roberts,# and Ben Luisi*,‡

Department of Biochemistry, Cambridge University, 80 Tennis Court Road, Cambridge CB2 1GA, U.K., CLRC Daresbury Laboratory, Daresbury WA4 4AD, U.K., Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, U.K., and School of Biological Sciences, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, U.K.

Received January 16, 2001; Revised Manuscript Received March 7, 2001

ABSTRACT: The Oct-1 transcription factor regulates a variety of tissue-specific and general housekeeping genes by recruiting specialized coactivators of transcription. It acts synergistically with the B-cell-specific coactivator Bob1 (OCA-B, OBF-1) to stimulate transcription of immunoglobulin genes. To analyze Oct-1's interactions with Bob1 and other regulatory proteins, we have overexpressed and purified different functional domains of the recombinant proteins. A version of Oct-1 that encompasses the amino-terminal activation region and the POU DNA-binding domain was extensively characterized (OctΔC1; comprising residues 1−445). Using an in vitro transcription assay, we demonstrate that this fragment is sufficient and necessary to stimulate transcription from an immunoglobulin promoter with Bob1. It also coactivates from the herpes simplex virus ICPO promoter element in the presence of VP16. Using a range of spectroscopic and biophysical techniques, we demonstrate that the activation domains of Oct-1 and Bob1 have little globular structure and that they do not physically interact. Thus, their functional synergy is likely to arise by the co-recruitment of common factors as part of a larger regulatory assembly. We propose a hypothesis to explain why the activation domains of these and other transcription factors of metazoans have little if any intrinsic structure.

Transcriptional regulation is a highly organized process that directs gene expression in response to developmental and signaling events. The class II genes, which include the protein-encoding genes, require RNA pol II and its associated general transcription factors (GTFs) (TFIIA, -B, -D, -E, -F, and -H) to direct basal transcription from the core promoter (*1*). Transcriptional activation is mediated by activator proteins that recognize cognate DNA sequences and facilitate the formation of a functional preinitiation complex. Activator proteins have been found to interact directly with components of the basal transcription machinery or act through additional factors known as coactivators (*2*−*5*).

A classic example of a transcriptional regulatory protein is Oct-1, which mediates cell-type-specific and temporally restricted gene expression, as well as the control of ubiquitous, cell-general genes. This transcription factor recognizes an octameric sequence in the promoters of various differentially expressed genes, including B-cell-restricted immunoglobulin genes, the cell-cycle-regulated histone H2B gene, and the herpes simplex virus (HSV)[1] immediate-early genes. From the same or a similar octameric site, Oct-1 also mediates the ubiquitous transcription of small nuclear RNA genes by either RNA pol II or pol III (*6*). Oct-1 imparts its specificity in gene regulation through its ability to interact with various gene, tissue, or cell-type-specific coactivators, such as the coactivator VP16 from herpes simplex virus (HSV) (*7*) and SNAP190 in snRNA gene activation (*8*). These cofactors mediate transcriptional activation of DNA-bound Oct-1 through downstream interactions that target the general transcription machinery.

In the case of immunoglobulin gene expression, Oct-1 acts together with its partner, the B-cell-specific coactivator Bob1/OCA-B/OBF-1 (*9*−*11*). The expression of immunoglobulin genes is mediated by Bob1 through direct interaction with DNA-bound Oct-1. Bob1 contacts both subunits of the bipartite DNA-binding domain of Oct-1, known as Oct-1 POU. It also makes specific contacts in the major groove of DNA, by recognizing the fifth base pair of the octamer consensus ATGCAAAT, and is able to discriminate among variants of the octamer site (*12*, *13*).

Bob1 is a proline-rich, intrinsically unstructured protein (*14*) that exhibits unusual transcriptional activation properties. The coactivator has no activation potential of its own when recruited to the promoter by the isolated Oct-1 POU DNA-

† L.L. was supported by the Winston Churchill Foundation, and this work was funded by the Wellcome Trust.

* Address correspondence to this author. E-mail: ben@cryst.bioc.cam.ac.uk. FAX: 44-1223-766002.

‡ Cambridge University.

§ Current address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave., Boston, MA 02115.

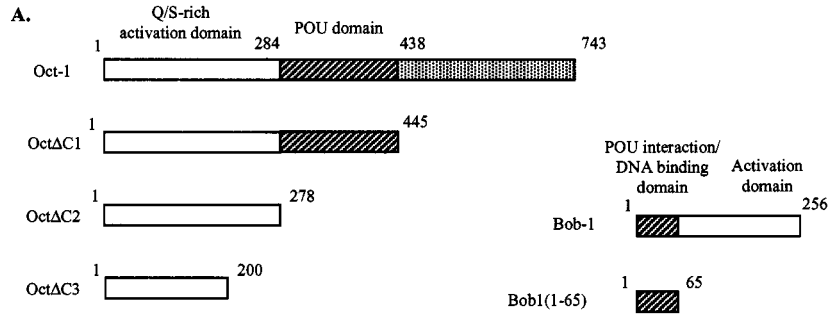
|| CLRC Daresbury Laboratory.

⊥ University College London.

University of Manchester.

[1] Abbreviations: CD, circular dichroism; SAXS, small-angle X-ray scattering; ITC, isothermal titration calorimetry; NMR, nuclear magnetic resonance; HSV, herpes simplex virus.

10.1021/bi010095x CCC: $20.00
Published on Web 05/08/2001



B. OctΔC1

```
  1 MNNPSETSKP SMESGDGNTG TQTNGLDFQK QPVPVGGAIS TAQAQAPLGH
 51 LHQVQLAGTS LQAAAQSLNV QSKSNEESGD SQQPSQPSQQ PSVQAAIPQT
101 QLMLAGGQIT GLTLTPAQQQ LLLQQAQAQA QLLAAAVQQH SASQQHSAAG
151 ATISASAATP MTQIPLSQPI QIAQDLQQLQ QLQQQNLNLQ QFVLVHPTTN
201 LQPAQFIISQ TPQGQQGLLQ AQNLLTQLPQ QSQANLLQSQ PSITLTSQPA
251 TPTRTIAATP IQTLPQSQST PKRIDTPSLE EPSDLEELEQ FAKTFKQRRI
301 KLGFTQGDVG LAMGKLYGND FSQTTISRFE ALNLSFKNMC KLKPLLEKWL
351 NDAENLSSDS SLSSPSALNS PGIEGLSRRR KKRTSIETNI RVALEKSFLE
401 NQKPTSEEIT MIADQLNMEK EVIRVWFCNR RQKEKRINPP SSGGT
```

Bob-1

```
  1 MLWQKPTAPE QAPAPARPYQ GVRVKEPVKE LLRRKRGHAS SGAAPAPTAV
 51 VLPHQPLATY TTVGPSCLDM EGSVSAVTEE AALCAGWLSQ PTPATLQPLA
101 PWTPYTEYVP HEAVSCPYSA DMYVQPVCPS YTVVGPSSVL TYASPPLITN
151 VTTRSSATPA VGPPLEGPEH QAPLTYFPWP QPLSTLPTST LQYQPPAPAL
201 PGPQFVQLPI SIPEPVLQDM EDPRRAASSL TIDKLLLEEE DSDAYALNHT
251 LSVEGF
```

FIGURE 1: Schematic diagrams of functional domains and amino acid sequence of recombinant Bob1 and Oct-1 proteins. (A) Schematic representations of C-terminal truncated Oct-1 proteins, full-length Bob1, and Bob1(1−65). Functional domains of Oct-1 and Bob1 are indicated. (B) Amino acid sequence of Bob1 and OctΔC1. The tryptophan residues are highlighted in boldface type. The Oct-1 POU domain is underlined.

binding domain or when used with other DNA-binding domains. However, Bob1 strongly synergizes with the activation domain of the full-length Oct-1 and with other activation domains, including those of VP16 and GAL4 (*15*). The requirement for an Oct-1 activating region for functional synergy with Bob1 suggests a cooperative effect between the two proteins.

This study addresses a potential mechanism for Bob1-dependent activation that leads to functional synergy with Oct-1. We report here the expression and purification of three C-terminal truncated versions of human Oct-1 in *E. coli* and the structural and functional characterization of the Bob1/Oct-1 activation complex (Figure 1). The first recombinant Oct-1 protein contains residues 1−445, encompassing the N-terminal activation domain and the DNA-binding domain, which we refer to as OctΔC1. We also overexpressed and purified residues 1−278 and 1−200 of Oct-1, both fragments of the activation domain (OctΔC2 and OctΔC3). The smallest fragment was amenable for structural studies by NMR. The overexpression and purification of full-length recombinant Bob1 have already been reported (*14*). With the availability of large amounts of purified truncated versions of Oct-1 and full-length Bob1, it was possible to test for direct interaction between the activation domains and to explore a potential structural transition as part of the assembly process of the transcriptional activation complex.

Biophysical and functional characterization of recombinant Bob1 and Oct-1 revealed that both proteins have unstructured, noninteracting activation domains, but are capable of functional synergy in transcriptional activation. Therefore, it seems that the observed transcriptional synergy results not from mutual association, but rather from the independent recruitment of cofactors in part of a larger assembly.

MATERIALS AND METHODS

PCR Amplification and Cloning. The full-length human Oct-1 clone was provided by Dr. Peter O'Hare, Marie Curie Research Institute, in a pTrcHis vector (Invitrogen). Three truncated versions of Oct-1 were amplified by PCR from the full-length clone, encompassing either the N-terminal activation domain and the Oct-1 POU DNA-binding domain (OctΔC1, residues 1−445), or fragments of the activation domain alone (OctΔC2, residues 1−278; OctΔC3, residues 1−200). The primers were designed such that the amplified DNA would contain the restriction sites for *Bam*HI and *Xho*I at the 5′ and 3′ ends, respectively, as well as the stop codon TTA. Oligonucleotides were synthesized by the PNAC Facility of the Biochemistry Department, University of Cambridge. The forward primer sequence for all three constructs was as follows (restriction sites are underlined): 5′-CCG CGT GGA TCC AAC AAT CCG TCA GGA AC-3′. The backward primers for the three truncated versions

of Oct-1 were as follows: OctΔC1, 5′-GG CCC CTC GAG TTA GGT CCC ACC ACT GCT TG-3′; OctΔC2, 5′-GG CCC CTC GAG TTA GCT GGG AGT ATC AAT TCG C-3′; OctΔC3, 5′-GG CCC CTC GAG TTA ATT GGT GGT TGG ATG CAC-3′. PCR was carried out using Pfu DNA polymerase and standard cycling temperatures: 10 min at 94 °C; 30 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and elongation at 72 °C for 4.5 min; and a final elongation at 72 °C for 10 min. The PCR DNA was digested with the appropriate restriction enzymes, gel-purified, and cloned into a pET41a expression vector. The resulting Oct-1 fusions contained an N-terminal GST fusion, His-tag and S-tag. The plasmids were checked by DNA sequencing by the DNA Sequencing Facility, Department of Biochemistry, Cambridge University.

Expression and Purification. OctΔC1 and OctΔC3 were overexpressed in the pET41a vector (Novagen) in *E. coli* strain BLR(DE3), a recombinase-deficient inducible expression host. The cultures were grown in LB media with 30 μg/mL kanamycin at 37 °C and induced at $OD_{600} = 0.8$ with 1 mM IPTG, before lowering the temperature to 20 °C for 16 h. The cells were harvested by centrifugation at 4000*g* for 20 min and resuspended in lysis buffer [1× PBS, 200 mM NaCl, 0.1% Tween 20, 0.05% (w/v) lysozyme, 1× complete protease inhibitor (Boehringer Mannheim), 2 mM DTT]. The OctΔC3 preparation did not contain DTT. The cells were placed on ice for 30 min, and then a French press was used for more efficient lysis. The whole cell lysate was centrifuged at 4 °C for 30 min at 31000*g*. The soluble fraction was further processed by DNase I treatment; 1:100 (v/v) 1 M $MgCl_2$ and 1:800 (v/v) DNase I (20 units/μL) were added, and the mixture was incubated on ice for 20 min. Then 1:50 (v/v) Polymin P was added with a further incubation of 5 min on ice. The mixture was centrifuged at 4 °C for 20 min at 21000*g*.

The supernatant was decanted from the DNA precipitant and applied to a glutathione-*S*−Sepharose matrix (Pharmacia) and bound in a slurry with stirring for 2 h at 4 °C. The matrix was packed into a column and washed with 20 column volumes of wash buffer (1× PBS, 200 mM NaCl, 2 mM DTT). The buffer was replaced with 1× PBS for a 16 h thrombin cleavage at 4 °C. The cleaved protein was collected in fractions using the wash buffer. The recombinant proteins were further purified by gel filtration using S200 or S75 Superdex high-resolution preparative columns (Pharmacia). OctΔC1 was loaded on an S200 column in 1× PBS, 200 mM NaCl, 2 mM DTT. OctΔC3 was prepared using an S75 column in 1× PBS, 200 mM NaCl.

Purified recombinant Oct-1 POU, full-length Bob1, and Bob1(1−65) were prepared as previously described (*14*). An additional ATP dissociation step was used to remove DnaK that remained bound to the full-length Bob1 protein. Following affinity purification by TALON metal resin (CLONTECH), the suspension was incubated in washing buffer (1× PBS) containing 4 mM ATP/15 mM $MgSO_4$ at 35 °C for 10 min. The Bob1 fusion was cleaved from the beads with thrombin protease. The samples were further purified by anionic exchange on a Hi-Trap Q column (Pharmacia) followed by gel filtration using a HiLoad 16/60 Superdex preparative column as reported previously (*14*).

Histidine-tagged Pin1 was expressed and purified as described in (*16*). Recombinant VP16 was prepared as described by Grossmann et al. (*17*).

Protein concentration was determined by spectrophotometric measurement using a CARY 100 Bio UV−visible spectrophotometer (Varian). The molar extinction coefficients for Oct-1 POU and full-length Bob1 were $\epsilon_{280\ nm} = 24.4$ and 36.8 $mM^{-1}\ cm^{-1}$, respectively, determined by amino acid composition analysis as previously reported (*14*). A predicted extinction coefficient was calculated for OctΔC1 by ProtParam (*18*) as $\epsilon_{280\ nm} = 13.94\ mM^{-1}\ cm^{-1}$.

In Vitro Transcription Assays. In vitro transcription assays were performed using the primer extension method (*19*). Two promoter sequences were cloned sequentially into a modified pGEM-3 vector (Promega) containing an E4T reporter gene. The promoter inserts were generated by annealing 34mer oligonucleotides in a 1:1 molar ratio (10 mM Tris-HCl, pH 8.5) by heating to 90 °C in a water bath for 5 min and allowing to cool slowly. The cloned Igκ promoter sequence contained a *Hin*dIII overhang that did not generate a cut site, a *Bam*HI overhang/restriction site, and a nested *Hin*dIII site. The octamer sequence of the cloned Igκ promoter sequence is shown in boldface as follows: 5′-AGC TGA ATG **ATTTGCAT** GCT CTC AAG CTT GCA CG-3′. The $ICPO_{mut}$ promoter site was 5′-AAG CCT CCG TGC **ATGCAAAT** GAT ATT CTT TG-3′. Three hundred nanograms of recombinant Bob1 and OctΔC1 were titrated into HeLa nuclear extract in various assays. HeLa nuclear extract was obtained from Computer Cell Culture Center, Belgium.

Tryptophan Fluorescence Measurements. Tryptophan fluorescence spectra were collected on an FP-777 spectrofluorometer at room temperature. Samples were diluted into 38 mM Tris-HCl (pH 7.5), 200 mM NaCl, and 2 mM DTT at a concentration of 1 μM for each component. The stoichiometry for complex assembly was 1:1:1. A 17mer duplex containing the Igκ promoter sequence (5′-TCCT **ATGCAAAT** TATTA-3′) was used to form the protein−DNA complexes. The excitation wavelength was 285 nm, to avoid quenching by DNA absorption at lower wavelengths. The emission spectra were recorded from 300 to 400 nm.

Small-Angle X-ray Scattering. X-ray solution scattering data were collected with the low-angle scattering camera at station 2.1 (*20*) at the SRS Daresbury Laboratory, U.K., using a position-sensitive multiwire proportional counter (*21*). At the sample-to-detector distance of 2.5 m and the X-ray wavelength of $\lambda = 1.54$ Å, a momentum transfer interval of $0.013\ Å^{-1} \leq q \leq 0.36\ Å^{-1}$ was covered. The modulus of the momentum transfer is defined as $q = (4\pi \sin \Theta)/\lambda$, where 2Θ is the scattering angle. The q-range was calibrated using an oriented specimen of wet rat tail collagen (based on a diffraction spacing of 670 Å).

SAXS data for Bob1 (the native protein and a mutant with the substitution Ser184 → Glu at concentrations between 0.3 and 3 mg/mL) were collected in buffer containing 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 5 mM DTT at a sample-to-detector distance of 2 m. OctΔC1 in buffer containing 38 mM Tris-HCl (pH 7.5), 200 mM NaCl, 2 mM DTT was examined with a 2.5 m camera at concentrations of 1 and 11.3 mg/mL. All scattering curves were evaluated initially using the Guinier approximation at low q-values ($q < 0.03\ Å^{-1}$); i.e., the radius of gyration, R_g, is extracted from the slope ($=R_g^2/3$; *21*) in the ln $I(q)$ versus q^2 plot. This revealed values for R_g larger than 50 Å. The radius of gyration can be considered as a measure of elongation if the

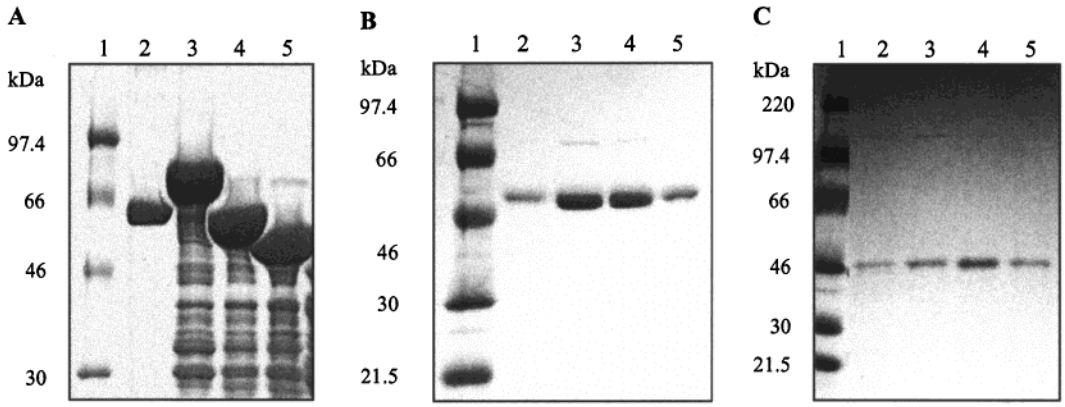


FIGURE 2: SDS−PAGE analysis of affinity purification of GST-OctΔC1−3 and purified OctΔC1 and OctΔC3 following gel filtration. (A) The three truncated Oct-1 constructs (pET41.OctΔC1−3) were expressed in BLR(DE3) at 20 °C (0.5 mM IPTG, 16 h). The lysates were applied to glutathione-*S*−Sepharose (lanes 3−5). The molecular masses of GST-OctΔC1−3 are 80.4, 61.4, and 53.0 kDa. Lane 1: molecular mass markers. Lane 2: BSA standard, 1 mg/mL. (B) OctΔC1 fractions from S200 gel filtration. The sample eluted in two peaks, shown in lane 2 (first peak) or fractions representing the second more predominant peak (lanes 3−5). (C) OctΔC3 also eluted in two peaks from the S75 column (first peak, lane 2; second peak, lanes 3−5).

internal inhomogeneity of scattering densities is negligible. Considering the molecular mass of Bob1 (27.4 kDa) and OctΔC1 (50 kDa), such large values for R_g can only be in accord with a nonglobular macromolecule capable of assuming a large number of conformations. For comparison, a globular protein with a molecular mass between 25 and 50 kDa would be expected to possess a radius of gyration between 15 and 25 Å. It is therefore reasonable to analyze the scattering data in terms of a polypeptide chain with Gaussian statistics, which can be characterized by the Debye function (*23*):

$$I(q) = \frac{12}{q^2R^2}\left[e^{-q^2R^2/6} - \left(1 - \frac{q^2R^2}{6}\right)\right]$$

with the intensity normalized to unity for zero scattering angle. R represents the average distance between the ends of the polypeptide chain and is related to the mean radius of gyration R_g of the chain by the equation:

$$R_g^2 = \frac{R^2}{6}$$

Circular Dichroism. CD spectra were recorded on a Jobin Yvon CD6 spectropolarimeter. Purified recombinant OctΔC1 and OctΔC3 were dialyzed into 10 mM NaH_2PO_4 (pH 7.0) using a MWCO 3500 SpectraPor Membrane. Data were collected at room temperature in a 1.0 mm cuvette with a total protein concentration of 0.2 mg/mL. The ternary complex was prepared in equal molar ratios of each component. A 17mer duplex containing the Igκ promoter (5′-T CCT **ATGCAAAT** TAT TA-3′) was used for ternary complex assembly. CD spectra between 200 and 250 nm were recorded and corrected with a blank buffer reading.

One-Dimensional Nuclear Magnetic Resonance. A sample of OctΔC3 was prepared in 1× PBS and concentrated to 450 μL at 8.3 mg/mL (0.33 mM). Fifty microliters of D_2O was added to the sample with 1 μL of TSP for reference. Spectra were recorded at 22 °C from a 500 MHz NMR spectrometer (Bruker).

Isothermal Titration Calorimetry. Synthetic 25mer oligonucleotides containing the Igκ promoter (5′-C CAG GGT **ATGCAAAT** TAT TAA GGG C-3′) were purified by ion-exchange chromatography and annealed as described previously (*14*). To avoid heat signals from the mixing of nonequivalent buffers, the protein and DNA samples were dialyzed using MWCO 3500 membrane tubing (SpectraPor) into 25 mM Tris-HCl (pH 7.5), 1 mM Na_2-EDTA, 5 mM DTT, 200 mM NaCl, 1 mM sodium azide. The concentration of the Igκ duplex was determined spectrophotometrically with the molar extinction coefficient $\epsilon_{260\ nm} = 409.1\ mM^{-1}\ cm^{-1}$ (*24*).

The ITC experiments were performed at 25 °C using a VP-ITC instrument from MicroCal Inc. (Northampton, MA). Igκ solutions (typically 35−45 μM) were titrated into Oct-1 POU solutions (typically 3−5 μM) using a 250 μL syringe, with each titration consisting of 17 × 15 μL injections. Five minutes was left between each injection, and the filter period for data collection was 2 s. The heat associated with each injection was obtained by integrating the area under the resulting peak using the Origin ITC data analysis software provided with the instrument. Heats of dilution and injection were measured in control experiments in which the Igκ solutions were titrated into dialysis buffer. These heats were found to be similar to those observed at the end of the DNA−protein titrations. The integrated heats from the DNA−protein experiments were corrected by the enthalpy changes observed at the end of the titrations. The data were analyzed to obtain estimates of the observed binding constant K_{obs}, stoichiometry constant n, and enthalpy of binding ΔH_{cal} using the "single set of identical sites" model within the Origin software package.

RESULTS

Expression and Purification of Truncated Oct-1 Proteins. The fusion proteins encompassing residues 1−445, 1−278, and 1−200 of Oct-1 (OctΔC1−3) expressed quite well and were found to be soluble (Figure 2A, lanes 3−5). The Oct-1 proteins were liberated from the glutathione-*S*−Sepharose matrix by thrombin cleavage and eluted from the column. The collected fractions were concentrated and applied to an S200 or S75 size exclusion column (Figure 2B,C). OctΔC1 eluted from the S200 in two peaks, with most of the material collected in the second peak (Figure 2B, lanes 3−5). OctΔC3 also behaved similarly and eluted in two peaks, and material

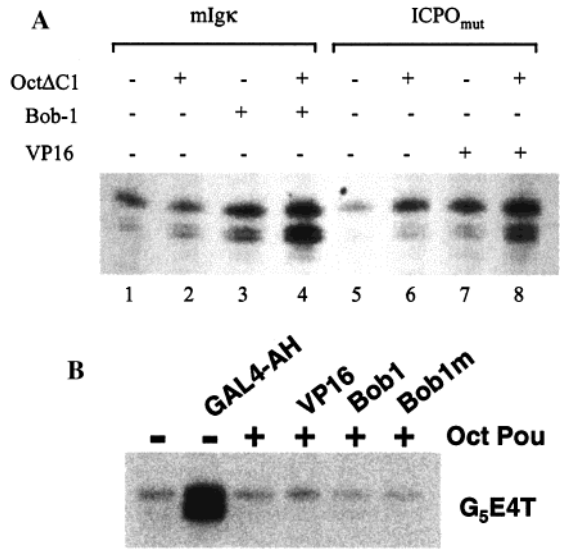


FIGURE 3: OctΔC1 and Bob1 activate transcription from the Igκ promoter. (A) OctΔC1 stimulated weak activation from the Igκ and ICPO$_{mut}$ promoters compared to basal levels (lanes 2, 6 vs lanes 1, 5). When Bob1 or VP16 was added to the HeLa nuclear extract, only weak activation was observed (lanes 3, 7), possibly due to low levels of endogenous Oct-1 in the particular preparation. Addition of both recombinant proteins resulted in strong activation from the respective promoters (lanes 4, 8). (B) Controls showing that Oct-1 and Bob1 do not support activation if the octamer-binding site is replaced with the GAL4 recognition element, while the GAL4−4H fusion activates strongly from this control promoter. 300 ng each of OctΔC1, Bob1, and VP16 was used in each reaction. Bob1m is a mutant S184→E, to mimic phosphoserine.

from the second peak (Figure 2C, lanes 3−5) was used for subsequent structural studies.

OctΔC1 Forms a Functional Complex on Octamer DNA. The purified recombinant OctΔC1 was competent to bind target DNA, as shown by electrophoretic mobility assay (result not shown). It was also capable of recruiting the coactivator Bob1 into a ternary complex, resulting in a supershift.

OctΔC1 was able to stimulate transcription from promoters containing the octamer motif (Figure 3A). The N-terminal activation domain of OctΔC1 was sufficient for weak activation from promoters containing the octamer site (lanes 2 and 6). An approximate 2-fold stimulation was observed from the basal level. OctΔC1 also synergized with recombinant Bob1 and VP16 to strongly stimulate transcription approximately 8-fold from the Igκ and ICPO$_{mut}$ promoters, respectively (lanes 4 and 8). However, strong activation with Bob1 alone was never achieved, possibly due to low levels of endogenous Oct-1 in that particular batch of nuclear extract (lanes 3 and 7). This activation was specific to the octamer site. In controls using the same E4 core promoter, where the octamer site was replaced with five binding sites for the yeast transcription factor GAL4, OctΔC1 and Oct-1 POU failed to activate with either VP16 or Bob1, while the fusion protein GAL4-AH strongly activated from this promoter (Figure 3B).

The isolated DNA-binding domain of Oct-1 could not stimulate transcription either alone or in combination with Bob1. Bob1(1−65) lacking the activation domain also failed to stimulate transcription (result not shown). This is in agreement with earlier findings that show that activation by Bob1 is dependent on the Oct-1 activation function. Our

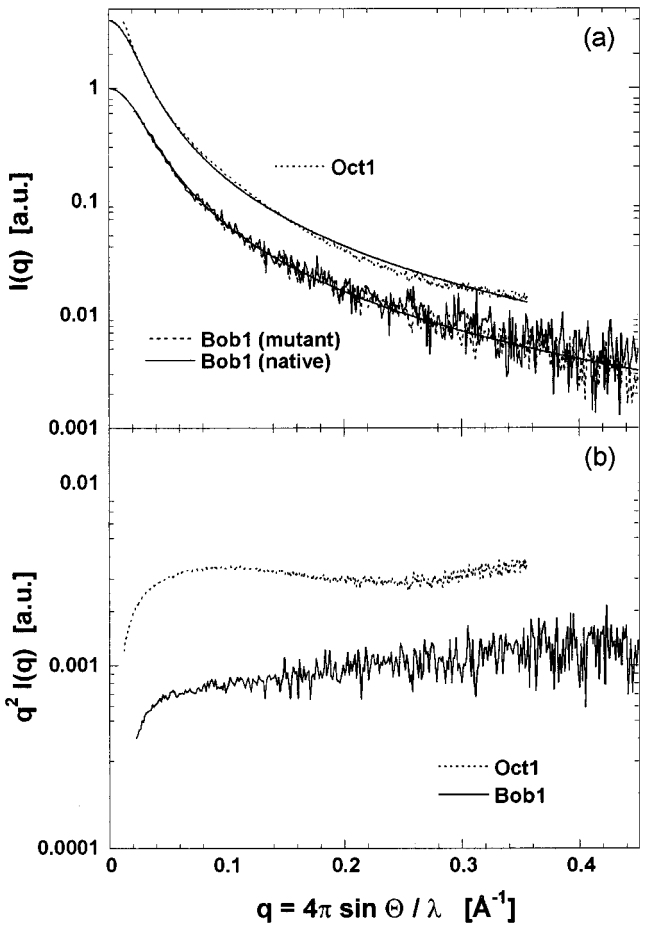


FIGURE 4: OctΔC1 and Bob1 small-angle X-ray scattering profiles. (a) The scattering data indicate that Bob1 exists as a random chain and OctΔC1 is loosely folded (smooth curves represent fits using the description of a Gaussian chain). (b) The Kratky plot emphasizes the nonglobular nature of both proteins from their asymptotic scattering behavior.

results suggest that the N-terminal activation domain of Oct-1 is necessary and sufficient for activation by Bob1. It seems that the C-terminal domain of Oct-1 is not required for Bob1- or VP16-mediated transcriptional activation.

OctΔC1 Is Partially Structured. Small-angle X-ray scattering (SAXS) was used to provide low-resolution structural information for OctΔC1 and Bob1. The SAXS profiles of OctΔC1 and Bob1 indicated that both proteins do not have a folded, globular structure, and a Gaussian chainlike behavior was most appropriate to describe these experimental findings (Figure 4a). The scattering plots were fit with a random-chain model which yields a radius of gyration of 69.4 Å and end-to-end length of 170 Å for OctΔC1, and a radius of gyration of 55.4 Å and length of 136 Å for Bob1. As the polymer dimensions do not reflect the respective protein masses (OctΔC1 is nearly twice as large as Bob1), these results suggest that OctΔC1 retains some structured peptide segments, presumably from the POU domain, while Bob1 exists entirely as a random chain. The latter is supported by the goodness of fit values ($\chi^2_{Bob1} = 0.008$ against $\chi^2_{OctΔC1} = 0.302$). This finding is also confirmed by the asymptotic behavior (for large scattering angles): Whereas Bob1 scattering curves vary as q^{-2} at large q-values, which is characteristic for random chains, OctΔC1 does not fully follow this trend as can be seen in Figure 4b, representing the so-called Kratky-plot [i.e., $q^2I(q)$ versus q].

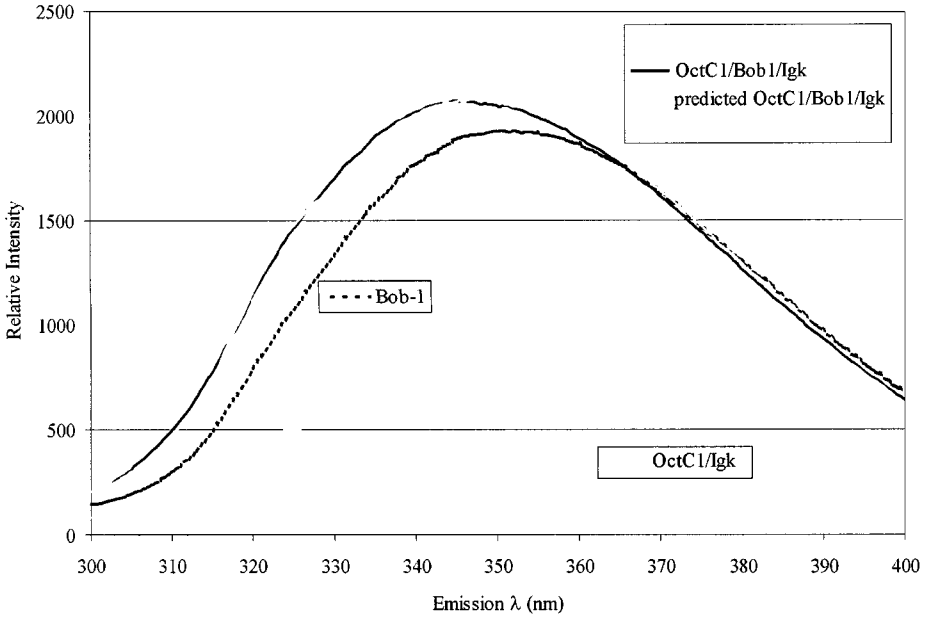


FIGURE 5: Bob1 does not undergo a structural transition in ternary complex formation. The predicted fluorescence emission spectrum was calculated by adding the individual spectra of OctΔC1/Igκ DNA and Bob1. The emission peak of the ternary complex did not differ in intensity or position from the prediction. This result suggested that the environment of the tryptophan residues in Bob1 remained unchanged in complex assembly. A 1 μM sample of each component was used. The excitation wavelength was 285 nm.

We further examined the structural composition of Oct-1 by 1D-NMR of the N-terminal activation domain (residues 1−200). The NMR spectrum of OctΔC3 indicated the absence of a globular fold, as the proton chemical shifts characteristic of tertiary structure were lacking in the regions for α-helix and β-strand (result not shown).

The results of these biophysical studies indicate that the activation domains of Oct-1 and of Bob1 are not structured and exist as elongated polymers.

Activation Domains of Bob1 and Oct-1 Do Not Physically Interact. Molecular genetics studies have shown that the Bob1 activation domain does not function autonomously and requires a partner activator (*10*, *15*). Oct-1 provides an additional activation domain that functions synergistically with the Bob1 activation domain. The DNA-binding domain of Oct-1 (Oct-1 POU) is sufficient to tether Bob1 to the Igκ promoter through protein−protein interactions, but this interaction itself does not result in transcriptional activity. Could the noted synergy arise from a direct interaction between the Oct-1 and Bob1 activation domains? As neither domain is itself structured, any direct association might induce mutual folding.

We first examined the proposed structural transition between the activation domains of Oct-1 and Bob1. Circular dichroism (CD) was used to monitor conformational changes and the associated changes in secondary structure content in the ternary complex assembly of Bob1 and OctΔC1 on Igκ DNA. Individual spectra for OctΔC1/DNA and Bob1 were recorded, as well as a spectrum for the ternary complex. A predicted CD spectrum was calculated from the individual spectra of OctΔC1/Igκ and Bob1 for direct comparison to the experimental result. The predicted and experimental spectra were superimposable, suggesting that there was no appreciable change in the composition of secondary structure upon complex formation (result not shown).

A potential interaction between the activation domains of Oct-1 and Bob1 was further explored by UV fluorescence studies. OctΔC1 contains two tryptophan residues, which are both in the POU domain. Full-length Bob1 has four tryptophan residues scattered throughout the protein (Figure 1B). Because the N-terminal activation domain of Oct-1 is free of tryptophan, changes in fluorescence upon binding of Bob1 to the OctΔC1/DNA complex may be attributed to the burial of tryptophan residues of Bob1. The individual emission spectra of OctΔC1 and Bob1 were recorded from 300 to 400 nm, and correlate well with the anticipated environment of their tryptophan residues (Figure 5). The emission maximum of Bob1 was at 350 nm, near the peak for free tryptophan, typically at ∼360 nm (result not shown). This result is consistent with the observation that Bob1 is unstructured in solution. OctΔC1 has a spectrum typical of globular proteins with the tryptophan residues buried in the hydrophobic core. Indeed, the two tryptophans of OctΔC1 are both in the POU DNA-binding domain (Figure 1B), which is known to be structured (*13*). In ternary complex formation, the emission spectrum does not change in intensity, and the position of the maximum does not shift when compared to a predicted signal calculated from the individual OctΔC1/DNA and Bob1 spectra (Figure 5). Therefore, the environment of the tryptophan residues in Bob1 remained unchanged, further supporting the proposal that there is no interaction between the activation domains.

Isothermal titration calorimetry (ITC) provided the definitive answer to the proposed mechanism for Bob1/Oct-1 activation. ITC was used at 25 °C to monitor heats of reaction upon DNA binding, as well as to detect any heat changes resulting from interactions between the activation domains of Oct-1 and Bob1. Binding constants and enthalpies of binding were determined for Oct-1 POU and OctΔC1 on Igκ DNA. The binding curve of OctΔC1 is shown in Figure 6A. The binding constant for OctΔC1/DNA was 9.3×10^6 M^{-1}, with a binding enthalpy of −6.56 kcal/mol. These parameters are comparable to the Oct-1 POU values, indicating that the N-terminal activation domain of Oct-1

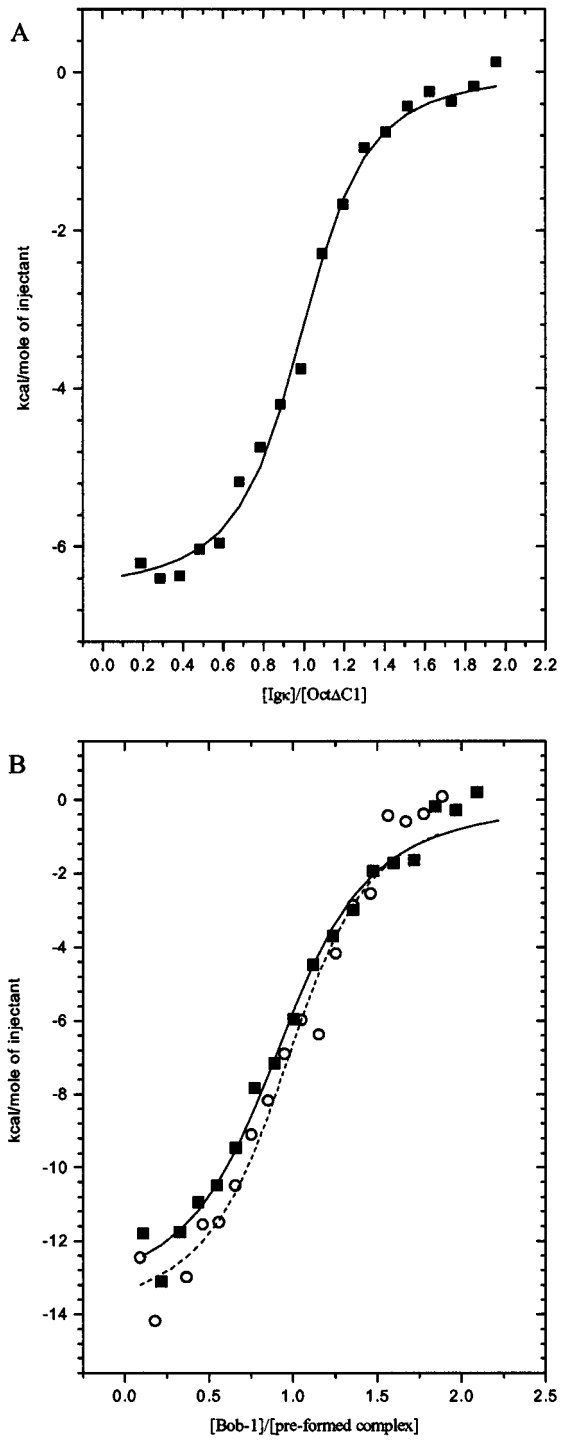


FIGURE 6: Integrated heat of binding for OctΔC1 and ternary complex formation. (A) Titration of the Igκ oligomer (36.6 μM) into a solution containing OctΔC1 (4 μM) at 25 °C in 25 mM Tris-HCl (pH 7.5), 1 mM Na_2-EDTA, 5 mM DTT, 200 mM NaCl, 1 mM sodium azide. (B) Titration of Bob1 (27.4 μM) into a solution of preformed Oct-1-binding complexes: Oct-1 POU/DNA (open circles) (2.9 μM) and OctΔC1/DNA (solid squares) (3.2 μM) at 25 °C in the same buffer.

does not affect DNA binding (Table 1). In the ternary assembly, the enthalpies of binding using either Oct-1 POU or OctΔC1 were quite comparable. The enthalpy change in complex formation of Oct-1 POU, Bob1, and Igκ DNA was −14.1 kcal/mol. A comparable heat of reaction was found when OctΔC1 was used to assemble the ternary complex: −13.6 kcal/mol. The binding constants were 4.7×10^6 and 4.2×10^6 M^{-1} for Bob1 associating with Oct-1 POU on DNA and OctΔC1 on DNA, respectively (Table 1). The binding curves for Bob1 associated with the preformed Oct-1 POU or OctΔC1/DNA complexes are shown in Figure 6B. These enthalpy changes and binding constants are quite similar and support independent, noninteracting activation domains.

DISCUSSION

Three C-terminal truncated versions of human Oct-1 were overexpressed in *E. coli* and purified to homogeneity. OctΔC1 was shown to be a functionally active fragment of Oct-1, and together with recombinant Bob1 was able to stimulate transcription from an immunoglobulin promoter using an in vitro transcription system. Although the activation domain of Oct-1 has been previously mapped to the N-terminus, its activation potential was demonstrated in the context of a β-globulin promoter (*25*). This study demonstrated that the N-terminal activation domain is sufficient for synergistic function with Bob1 in an in vitro assay, although the function of the C-terminus of Oct-1 remains unclear.

The recombinant Oct-1 proteins were also amenable for biophysical characterization. The N-terminal activation domain of Oct-1 was shown to lack a globular fold by NMR and SAXS analysis, a common characteristic shared by other transcription factors with unstructured activation domains, such as p53, AF-1 of the glucocorticoid receptor, and VP16 (*26*−*29*).

In this study, we examined the mechanism of synergistic function between Oct-1 and Bob1 in immunoglobulin gene activation. Because the activation domains of Oct-1 and Bob1 have little intrinsic structure, it seems likely that these elements may become structured upon specific recruitment into cognate complexes. For example, the unstructured activation domain of VP16 becomes ordered upon interaction with $TAF_{II}21$ (*29*). However, the CD, fluorescence, and ITC studies presented here revealed that the functional synergy of Oct-1 and Bob1 does not result from a structural transition or any direct interaction between the activation domains. Although Bob1-mediated activation is not obtained simply by recruitment to the promoter, its mechanism of action does not involve structural cooperativity with Oct-1 either. Rather, the synergistic activation may be mediated by a common coactivator complex. In this model, the activation domains of Oct-1 and Bob1 recruit different subunits of a mediator (Figure 7). Individual recruitment of a subcomplex could be responsible for the observed synergistic function. Such synergy has been noted in the interplay between the two activation domains of the glucocorticoid receptor (AF-1 and AF-2), which interact with two different subunits of the DRIP/TRAP complex (*30*). Therefore, we conclude that the observed transcriptional synergy between Oct-1 and Bob1 does not result from direct contact between the domains, but more likely through the independent recruitment of coactivators.

Bob1 is known to interact directly with the general cofactor PC4/p15. We have observed that recombinant p15 will form a supershift with recombinant Bob1 and OctΔC1 on the immunoglobulin promoter in an electrophoretic mobility retardation assay. Early studies in a reconstituted transcription

Table 1: Thermodynamic Parameters Observed for Binding Reactions Using ITC

A (μM)	B (μM)	stoichiometry	binding constant, K_{obs} ($M^{-1} \times 10^6$)	binding enthalpy, ΔH_{cal} (kcal mol^{-1})
Igκ (36.6)	Oct-1 POU (4)	1.00 ± 0.02	4.4 ± 1.0	−6.06 ± 0.11
Igκ (36.6)	OctΔC1 (4)	0.97 ± 0.01	9.3 ± 1.5	−6.56 ± 0.13
Bob-1 (27.4)	Oct-1 POU/Igκ (2.9)	0.99 ± 0.04	4.7 ± 1.4	−14.10 ± 0.70
Bob-1 (27.4)	OctΔC1/Igκ (3.2)	0.95 ± 0.03	4.2 ± 0.8	−13.60 ± 0.51

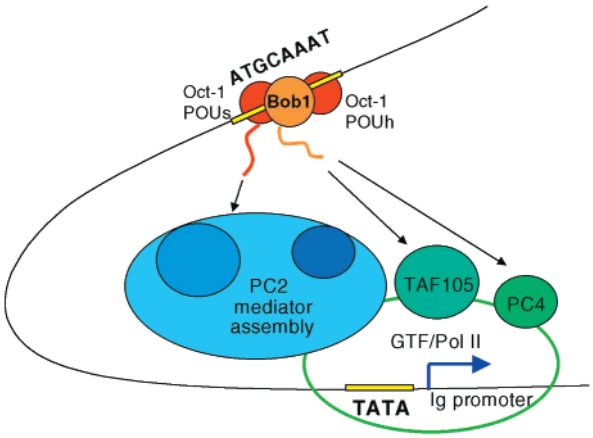


FIGURE 7: Model for transcriptional activation by multiple synergistic contacts. The downstream interactions that mediate Bob1 activation have not yet been elucidated. The involvement of multiple cofactors may lead to synergistic activation or serve functionally redundant requirements. Oct-1 (shown in red) is bound to the octamer sequence. The figure includes the bipartite DNA-binding domain and the unstructured N-terminal activation domain of Oct-1. Bob1 and its activation domain are shown in orange.

system showed that PC4/p15 and PC2 were both required for synergistic activation of Oct-1 and Bob1 (*31*). Further, a recent study has revealed that PC2 is a multisubunit complex of at least 15 polypeptides and a submodule of the DRIP/TRAP complexes (*32*). If the target coactivators are part of a common complex such as the DRIP/TRAP assembly or other subassembly, then the effects of the colocalization will be cooperative. We have recently found that the immobilized ternary complex of OctΔC1/Bob1/octamer-DNA does indeed recruit components of the DRIP complex (Heather Owen, Larissa Lee, et al., unpublished result).

Another component of Bob1-mediated activation has recently been identified. $TAF_{II}105$ is a B-cell-enriched component of TFIID, which includes TBP and TBP-associated factors (TAFs). $TAF_{II}105$ has been found to enhance Bob1 activation in B-cells by binding directly to the C-terminus of Bob1 (*33*). However, the enhancement of Bob1 activation by $TAF_{II}105$ was not observed in HeLa cells, but PC4-depleted HeLa cells remained transcriptionally active, which suggests some functional redundancy in the activity of PC4/p15. The absolute requirements for TAFs and mediators are unclear. In the case of some activators, mediator function can bypass the requirement for TAFs, but in others, the TAFs are essential (*34*). The multiple requirements or sometimes functional redundancy between TAFs and mediators have yet to be resolved, but it is likely they function in concert (Figure 7).

As Bob1 is enriched in proline, we also considered whether the Bob1 activation mechanism might involve a structural transition through proline cis−trans isomerization. The amino acid sequence of Bob1 contains many PXXP motifs, which may be recognized by a prolyl *cis*−*trans*-isomerase. In the stepwise activation process, Bob1 may recruit a proline isomerase to overcome a kinetic bottleneck in a structural transition. To explore this hypothesis, three proline isomerases representing the parvulin and FK-506-binding protein (FKBP) families were tested for their effect on Bob1 activation. The human proline isomerase Pin1 is a member of the parvulin family that recognizes phosphoserine flanked by proline (*16*); intriguingly, Bob1 contains a phosphorylated serine (S184) that is implicated in the inducible activation of Bob1 (*35*). However, neither the human Pin1 nor representatives of the FKBP family had any significant effect on transcription in vitro by Bob1 and Oct-1 (results not shown). It thus seems unlikely that assembly of the cognate complex and its accompanying folding are kinetically regulated by proline conformational transitions.

It is rather surprising that important regulatory proteins such as Oct-1 bear extensive regions with little intrinsic structure, or that the entirety of the 27 kDa Bob1 coactivator protein appears to be unstructured. However, these proteins are not special cases: genomic and structural analyses are revealing that a number of other regulatory proteins share this property of having extensive unstructured regions (*36*). The question comes to mind as to why 'natively' unstructured proteins evolved when they might seem potentially deleterious. On first consideration, they might seem likely to associate nonspecifically to each other or to globular proteins, with intracellular precipitation ensuing. However, these proteins appear not to be extensively hydrophobic, and so may not associate through fortuitous nonpolar interactions. Indeed, the SAXS results presented here show that the activation domains are not only unstructured, but elongated as well, so that they do not resemble a compact entity like the so-called molten globule which has been postulated as an intermediate in the folding of certain globular proteins. Clearly, on recognizing the cognate partner, these unstructured domains undergo a limited type of protein folding that probably involves the docking of a peptide segment into a cleft or surface of a globular partner, or even the mutual folding of two unstructured partners. This recognition by induced-folding is a special case of molecular recognition, and it might be seen to offer an economical means of ensuring specificity, since the required element can be very small. It may also offer a kinetic advantage, in that the tethered recognition segment can sweep out a large volume rapidly to find its cognate partner.

One striking aspect of molecular recognition for biological macromolecules is that the free energy of association is often independent of temperature and other physicochemical changes (*37*−*39*). For instance, we noted that temperature changes over a limited range have little effect on the free energy of association of Oct-1 and DNA or on Bob1's recruitment into the ternary complex with Oct-1 and DNA (JuFang Chang, Ph.D. thesis, University of Cambridge, 1999). This arises because the changes in entropy and

enthalpy, which can both be extremely sensitive to temperature, compensate almost exactly to maintain an invariant free energy. Perhaps this offers the organism an advantage in that the regulation processes are not overly sensitive to temperature or other perturbations. In having specially designed unstructured segments, the protein will be recognized in a folding event that precisely mimics the interaction occurring when a globular DNA-binding factor is docked into its cognate DNA site. Thermodynamically, these might be similar events, with comparable composition of nonpolar/polar residues buried. The result is that the binding of transcription factors to the DNA, either structured or partially structured, and their binding to unstructured coactivators may respond in the same way to variations in the environment. In the multicellular metazoans, which have extensively communicating networks of interacting transcription factors, the variation of activity of a single factor, for instance through condition-sensitive equilibria, could unbalance the regulation with potentially catastrophic effects. Perhaps permitting important regulatory processes to have unstructured components is one means of ensuring stability in a complex and dynamic system.

ACKNOWLEDGMENT

We thank Drs. Peter O'Hare, Sophie Jackson, Andrew Sharff, Kathryn Phillips, and Joe Noel for kindly providing reagents, Dr. Catherine Stott for collecting NMR spectra, and Raymond Low for guidance with the fluorescence experiments. We thank Drs. Kathryn Phillips, Tali Haran, Mercy Devasahayam, and Heather Owen for valuable discussions. B.L., J.L., and S.R. are Wellcome Trust Senior Research Fellows.

REFERENCES

1. Orphanides, G., Lagrange, T., and Reinberg, D. (1996) *Genes Dev. 10*, 2657−2683.
2. Malik, S., and Roeder, R. G. (2000) *Trends Biochem. Sci. 25*, 277−283.
3. Rachez, C., Suldan, Z., Ward, J., Chang, C. P., Burakov, D., Erdjument-Bromage, H., Tempst, P., and Freedman, L. P. (1998) *Genes Dev. 12*, 1787−1800.
4. Fondell, J. D., Ge, H., and Roeder, R. G. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8329−8333.
5. Naar, A. M., Beaurang, P. A., Zhou, S., Abraham, S., Solomon, W., and Tjian, R. (1999) *Nature 398*, 828−832.
6. Lobo, S. M., and Hernandez, N. (1994) *Transcription, mechanisms and regulation*, pp 127−159, Raven Press, New York.
7. Walker, S., Hayes, S., and O'Hare, P. (1994) *Cell 79*, 841−852.
8. Wong, M. W., Henry, R. W., Ma, B. C., Kobayashi, R., Klages, N., Matthias, P., Strubin, M., and Hernandez, N. (1998) *Mol. Cell. Biol. 18*, 368−377.
9. Gstaiger, M., Knoepfel, L., Georgiev, O., Schaffner, W., and Hovens, C. M. (1995) *Nature 373*, 360−362.
10. Luo, Y., and Roeder, R. G. (1995) *Mol. Cell. Biol. 15*, 4115−4124.
11. Strubin, M., Newell, J. W., and Matthias, P. (1995) *Cell 80*, 497−506.
12. Cepek, K. L., Chasman, D. I., and Sharp, P. A. (1996) *Genes Dev. 10*, 2079−2088.
13. Chasman, D., Cepek, K., Sharp, P. A., and Pabo, C. O. (1999) *Genes Dev. 13*, 2650−2657.
14. Chang, J. F., Phillips, K., Lundback, T., Gstaiger, M., Ladbury, J. E., and Luisi, B. (1999) *J. Mol. Biol. 288*, 941−952.
15. Krapp, A., and Strubin, M. (1999) *Mol. Cell. Biol. 19*, 4247−4254.
16. Ranganathan, R., Lu, K. P., Hunter, T., and Noel, J. P. (1997) *Cell 89*, 875−886.
17. Grossmann, J. G., Sharff, A. J., O'Hare, P., and Luisi, B. (2001) *Biochemistry* (in press).
18. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
19. Lin, Y. S., and Green, M. R. (1991) *Cell 64*, 971−981.
20. Towns-Andrews, E., Berry, A., Bordas, J., Mant, P. K., Murray, K., Roberts, K., Sumner, I., Worgan, J. S., and Lewis, R. (1989) *Rev. Sci. Instrum. 60*, 2346−2349.
21. Lewis, R. (1994) *J. Synchrotron. Radiat. 1*, 43−53.
22. Guinier, A. (1939) *Ann. Phys. (Leipzig) 12*, 161−237.
23. Debye, P. (1947) *J. Phys. Colloid. Chem. 51*, 18−32.
24. Lundback, T., Chang, J. F., Phillips, K., Luisi, B., and Ladbury, J. E. (2000) *Biochemistry 39*, 7570−7579.
25. Tanaka, M., and Herr, W. (1990) *Cell 60*, 375−386.
26. Lee, H., Mok, K. H., Muhandiram, R., Park, K. H., Suk, J. E., Kim, D. H., Chang, J., Sung, Y. C., Choi, K. Y., and Han, K. H. (2000) *J. Biol. Chem. 275*, 29426−29432.
27. Dahlman-Wright, K., Baumann, H., McEwan, I. J., Almlof, T., Wright, A. P., Gustafsson, J. A., and Hard, T. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 1699−1703.
28. O'Hare, P., and Williams, G. (1992) *Biochemistry 31*, 4150−4156.
29. Uesugi, M., Nyanguile, O., Lu, H., Levine, A. J., and Verdine, G. L. (1997) *Science 277*, 1310−1313.
30. Hittelman, A. B., Burakov, D., Iniguez-Lluhi, J. A., Freedman, L. P., and Garabedian, M. J. (1999) *EMBO J. 18*, 5380−5388.
31. Luo, Y., Ge, H., Stevens, S., Xiao, H., and Roeder, R. G. (1998) *Mol. Cell. Biol. 18*, 3803−3810.
32. Malik, S., Gu, W., Wu, W., Qin, J., and Roeder, R. G. (2000) *Mol. Cell 5*, 753−760.
33. Wolstein, O., Silkov, A., Revach, M., and Dikstein, R. (2000) *J. Biol. Chem. 275*, 16459−16465.
34. Hahn, S. (1998) *Cell 95*, 579−582.
35. Zwilling, S., Dieckmann, A., Pfisterer, P., Angel, P., and Wirth, T. (1997) *Science 277*, 221−225.
36. Wright, P. E., and Dyson, H. J. (1999) *J. Mol. Biol. 293*, 321−331.
37. Jen-Jacobson, L., Engler, L. E., and Jacobson, L. A. (2000) *Structure 8*, 1015−1023.
38. Sharp, K. (2001) *Protein Sci. 10*, 661−667.
39. Searle, M. S., Westwell, M. S., and Williams, D. H. (1995) *J. Chem. Soc., Perkin Trans. 2*, 141−151.

BI010095X