Cytometry Part A 58A:157–166 (2004)

Automated Microscopy Identifies Estrogen Receptor Subdomains With Large-Scale Chromatin Structure Unfolding Activity

Anne E. Carpenter, Anousheh Ashouri, and Andrew S. Belmont*

Department of Cell and Structural Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois

Received 3 May 2003; Revision Received 6 August 2003; Accepted 14 October 2003

Background: Recently, several transcription factors were found to possess large-scale chromatin unfolding activity; these include the VP16 acidic activation domain, BRCA1, E2F1, p53, and the glucocorticoid and estrogen steroid receptors. In these studies, proteins were fluorescently labeled and targeted to a multimerized array of DNA sequences in mammalian cultured cells, and changes in the appearance and/or size of the array were observed. This type of experiment is impeded by the low throughput of traditional microscopy.

Methods: We report the application of automated microscopy to provide unattended, rapid, quantitative measurements of fluorescently labeled chromosome regions.

Results: The automated image collection routine produced results comparable to results previously obtained by manual methods and was significantly faster. Using this approach, we identified two subdomains within the E domain of estrogen receptor α capable of inducing large-scale chromatin decondensation.

Conclusions: This work confirms that, like BRCA1, the activation function 2 region of the estrogen receptor has more than one distinct chromatin unfolding domain. In addition, we demonstrate the feasibility of using automated microscopy as a high-throughput screen for identifying modulators of large-scale chromatin folding. *The Supplementary Material referred to in this article can be found at the* CYTO Part A *website* (http://www.interscience.wiley.com/jpages/0196-4763/suppmat/v58A.html) © 2004 Wiley-Liss, Inc.

Key terms: chromatin; automated microscopy; high-throughput screening; estrogen receptor; steroid

Although transcriptional activators are known to recruit components of the transcriptional machinery, it is becoming clear that they are also capable of inducing changes in the chromatin in and around their target genes. At the level of local chromatin structure, such changes include acetylation and other covalent modifications of the histone proteins that package DNA into chromatin (1). In addition, the positioning of histone complexes along the DNA can be altered by chromatin remodeling complexes (2). It has only recently become possible to directly visualize the effects of a transcriptional activator on higher levels of chromatin folding, known as large-scale chromatin structure (3,4).

Several transcription-related proteins have been found to unfold large-scale chromatin structure: VP16 acidic activation domain (5); glucocorticoid receptor (6); E2F1, p53, BRCA1, and COBRA1 (7); and the estrogen receptor (ER) (8,9). These proteins are likely to induce unfolding by recruiting coregulator proteins with large-scale chromatin unfolding activity. The identity of these proteins is unknown and is an area of intense interest. We previously identified chromatin unfolding activity in the ER in the absence of its ligand (8). Based on these initial studies, we wanted to identify protein subdomains of ER with large-scale chromatin unfolding activity. Our long-term goal is to identify those proteins recruited by ER that are responsible for the observed large-scale chromatin unfolding.

However, testing a number of proteins or protein subdomains for large-scale chromatin unfolding activity is tedious. The extent of unfolding induced by a particular protein varies substantially from cell to cell, so a large number of cells must be imaged and the unfolding response averaged. Such tedious data collection and analysis are well suited to automated imaging, also called high-throughput microscopy. Most systems that automate microscope hardware and image analysis are used in pharmaceutical applications, such as high-throughput

Anne E. Carpenter's current address: Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142.

Contract grant sponsor: National Institutes of Health; Contrast grant numbers: R01-GM58460 and R01-GM42516.

*Correspondence to: Andrew S. Belmont, B107 CLSL, 601 S. Goodwin Ave, Urbana, IL, 61801

E-mail: asbel@uiuc.edu

Published online 8 March 2004 in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/cyto.a.10113

Table 1
*Analysis of Variance Table for the Data Presented in Figure 6**

Source	df	SS	MS	F value	P
Protein	13	1,270,107.575	97,700.583	17.45	<0.0001
Hormone	1	162,143.388	162,143.388	267.61	<0.0001
Protein × hormone	13	503,305.537	38,715.811	63.90	<0.0001

*df, degrees of freedom; MS, mean square; SS, sum of square.

screening of drug candidates (10), and in clinical applications (11), such as detecting rare cancer cells (12) or metaphase chromosome rearrangements (13). Also notable are systems for detecting dots visible by fluorescence in situ hybridization (14) and comet tails indicating DNA damage (15). Software development for more sophisticated applications, such as the automated assessment of subcellular localization, has been much more challenging (11,16,17). We describe the development of an automated microscopy data collection and image analysis procedure and its application to the identification of ER protein domains capable of unfolding large-scale chromatin structure.

MATERIALS AND METHODS

Automated Microscopy and Statistical Analysis

The image collection and analysis procedure described in this paper was developed by using an Axiovert 100M motorized inverted fluorescence microscope (Zeiss, Thornwood, NY) with a CoolSnapHQ cooled CCD camera (the pixel width and height at the magnification used is ~0.1 μm; Roper Scientific/Photometrics, Tucson, AZ), a motorized XYZ stage (Ludl Electronic Products, Hawthorne, NY), and motorized filter wheels (Ludl). The hardware was connected via a MAC2002 controller (Ludl) to a PC running Linux, which allowed automation to be controlled by ISee software (ISee Imaging Systems, Raleigh, NC). The light source initially used was an HBO 103W/2 lamp housing powered by an ebq100 isolated power supply (Zeiss) using 100-W mercury lamps (OSRAM, Danvers, MA). We found that the excitation shutter would begin to stick after approximately 1 month of nearly nonstop usage. Using shutters designed to withstand high-temperature conditions did not alleviate the problem, but using a fiber optic light source did (X-Cite microscopy illumination source from EFOS, now EXFO, Richardson, TX). The image collection and analysis routine and a detailed description of its design are available in the Supplemental Data Online. The version of the program used in this paper was called ROInet 127. Measurements of array size were imported into Microsoft Excel and edited manually and with macros.

Each data point used in the statistical analysis shown in Figure 6 was the mean of 80–100 cells. The analysis was a split-plot analysis of variance (Table 1) in which three variables were analyzed for their effects on the area of chromatin arrays: which experiment (because two independent experiments were conducted; see below), which protein was targeted via lac repressor (rep), and whether or not hormone was present. The blocks were the experiments, the "whole" plot was the protein, and the "split" plot was the presence or absence of hormone. Because we multiplied pairwise comparisons, we used a Tukey adjustment. The data were analyzed with SAS proc mixed (18). We noticed (as in previous studies) that for unknown reasons the absolute size of the A03_1 chromatin array is not stable from experiment to experiment. This may be due to a difference in the intrinsic size of the array due to the age or growth conditions of the cells. For these reasons, data from two experiments cannot be combined into one box plot, nor can a mean be obtained by averaging all cell measurements as if they were independent, e.g., by weighting the means from each experiment based on the number of cells within that experiment. The mean for each sample is therefore the unweighted arithmetic mean of the means from the two experiments.

Subcloning

Plasmid construction details have been published for green fluorescent protein (GFP)–lac rep-ER($^{1-595}$) and GFP-lac rep-DEF of ER($^{302-595}$) (8) and for GFP-lac rep (p3′SS EGFP dimer lac repressor) and GFP-lac rep-VP16 (p3′SS EGFP dimer lac repressor-VP16) (5). To make a precursor to reporter plasmid NYE107b, we first cut out eight lac operators plus the TATA box from NYE10 (8 lac op-TATA-CAT reporter) (8) using *Hind*III and *Nco*I and ligated this fragment into the *Hind*III–*Nco*I sites of pGL3 Basic (Promega, Madison, WI), thus creating NYE40. To generate a more sensitive reporter with a more inducible TATA box, the reporter plasmid NYE107b was then constructed by digesting the luciferase reporter pFR-Luc (Stratagene, La Jolla, CA) with *Hind*III, using phosphatase, digesting with *Sma*I, and then simultaneously ligating in two inserts: the eight lac operator sequences from NYE40 excised with *Hind*III and *Xho*I and a synthetic double-stranded oligo containing the E1bTATA box with a *Xho*I sticky end and a blunt end. The oligo sequences were (forward) 5′-TCGAGGGTATATAATGGATCCCC-3′ and (reverse) 5′-GGGGATCCATTATATACCC-3′.

NYE5 (a GFP-lac rep expression plasmid) was made in a manner similar to NYE4 (8) except that an *Asc*I site was added at the end of GFP-lac rep by silent mutations A3167G and A3170C just downstream of a *Pvu*II site. NYE4 and NYE5 are functionally equivalent in coding for the dimer, tight binding form of the lac repressor. NYE5(EYFP) (an YFP-lac rep expression plasmid) was made by cutting the NYE5 plasmid with *Xho*I and *Eco*RV, which removes the coding region for GFP and most of the

lac repressor and replaces it with the corresponding *Xho*I–*Eco*RV fragment from an YFP-lac rep plasmid called pCIneo dimer EYFP tight binding (P. Alvarez-Ortiz and A.S. Belmont, personal communication). Fusions of various subdomains of the ER to YFP-lac repressor were constructed by polymerase chain reaction (PCR) amplifying the human ERα region of interest from the plasmid CMV-ER (19) by using primers that incorporate *Asc*I restriction sites at each end. The PCR products were digested with *Asc*I and ligated into the *Asc*I site of NYE5(EYFP). All PCR-amplified regions were sequenced to ensure fidelity. Further details are provided in the Supplemental Data Online.

Large-Scale Chromatin Unfolding Assay

A03_1 Chinese hamster ovary (CHO) DG44 cells contain a gene-amplified chromosome region containing approximately 400-kb blocks of pSV2-DHFR-8.32 vector repeats separated by an estimated 1,000 kb of flanking, coamplified genomic DNA (20). Each vector copy contains the dihydrofolate reductase (DHFR) cDNA transgene and 256 direct repeats of the lac operator. These cells were cultured at 37°C with 5% CO_2 in F-12 Ham's medium without hypoxanthine or thymidine, with 0.3 μM methotrexate without phenol red, and with 10% dialyzed fetal bovine serum (HyClone Labs, Logan, UT) treated with charcoal/dextran. Phenol-red–free trypsin was used to passage cells. Transfections on coverslips were performed with FuGENE 6 reagent (Roche, Indianapolis, IN) according to the manufacturer's instructions by using 350 ng DNA and 5 μL reagent per 35-mm plate. Fresh medium containing hormone, if applicable, was added 16 h after transfection. Forty-eight hours after transfection, cells were rinsed in calcium, magnesium-free phosphate buffered saline (CMF-PBS), fixed in CMF-PBS with 1.6% formaldehyde (Polysciences, Warrington, PA), and stained with 0.2 μg/ml 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI) in CMF-PBS. Slides were mounted in ProLong Antifade mounting medium (Molecular Probes, Eugene, OR) and stored at 4°C.

Transcription Assays

Wild-type CHO-K1 cells (ATCC CRL no. 9618, American Tissue Culture Collection, Rockville, MD) were cultured at 37°C with 5% CO_2 in phenol-red–free F-12 Ham's medium with 10% charcoal/dextran-treated fetal bovine serum. Transfections for luciferase transcription assays used 0.5 μg NYE107b luciferase reporter, 0.1 μg cytomegalovirus (CMV) β-galactosidase reporter (Clontech, Palo Alto, CA), and 0.25 μg effector plasmid combined with 6 μL of FuGENE 6 reagent per well in 12-well plates. Fresh medium containing hormone treatments, if applicable, was added 16 h after transfection, and cells were harvested and lysed with Passive Lysis Buffer (Promega) 48 h after transfection. Luciferase assays were performed with Luciferase Assay Reagent (Promega) and a Luminoskan luminometer (Thermo LabSystems, Vantaa, Finland). Luciferase readings were normalized for β-galactosidase expression.

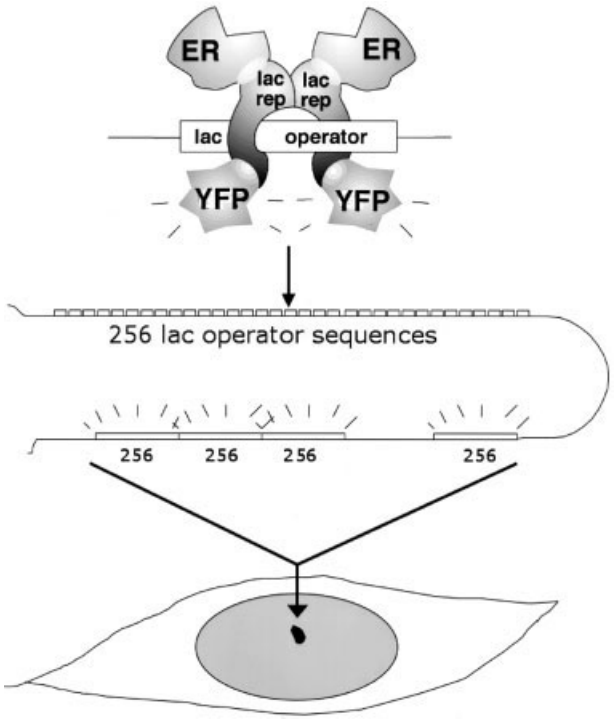

FIG. 1. Schematic of the lac operator/repressor system, whereby a protein of interest (in this case, ER), is targeted to a fluorescently labeled region of a chromosome, consisting of repeated vector copies containing a 256 direct lac operator repeat and a cDNA DHFR transgene interspersed with genomic DNA.

RESULTS

Experimental Design

We used an assay developed for assessing the effects of a protein on large-scale chromatin structure (Fig. 1) (5,21). We used a derivative of a CHO cell line, A03_1, which contains an engineered repetitive array of lac operators and genomic DNA embedded into its genome. This 90 million base pair array replicates middle to late in S phase and forms a condensed chromatin mass roughly 1 μm in diameter through most of interphase (20). These properties are typical of heterochromatin, originally defined as regions of darkly staining, compact chromatin that persist throughout the cell cycle (22).

Because this cell line normally has a compact chromatin array, except during DNA replication, unfolding of the chromatin induced by targeting a protein of interest to the array is easily visible. This targeting is accomplished by fusing the coding region of the gene of interest (in this case, ER) to a gene encoding a yellow fluorescent protein–tagged lac repressor (YFP-lac rep). The resulting mammalian expression plasmid is transiently transfected into A03_1 cells, with the YFP-lac repressor providing simultaneous visualization of the chromatin array and targeting of a protein of interest.

We previously used this system to investigate the effects of ER on large-scale chromatin structure (8). In that work, we investigated the structure of the fibers produced by

targeting various mutations and truncations of the ER to the A03_1 lac operator array. For each fusion protein, we manually collected images of 150 transfected cells and used a macro in NIH Image to identify and measure the fluorescent arrays. Image collection was time consuming, thus preventing rapid analysis of additional ER subdomains and other proteins of interest.

Development of Automated Image Collection and Analysis

We therefore sought to automate data collection, which would require locating rare transfected cells, typically 0.5–5% of the population, by focusing on the YFP-lac repressor-bound chromatin arrays, recording images, and then extracting morphologic measurements. Automated image collection is complicated by the low intensity and small size of the labeled chromosome regions, typically ranging in area from 1 to 5 μm^2. The chromatin arrays must be imaged with a high-power, high numerical aperture lens, and chromatin arrays are not at a predictable focal plane relative to the nuclei.

The program is initiated by placing a slide on the microscope, focusing on the nuclei, specifying the region of the slide to be examined, and naming the location where image and measurement files should be saved. The program (Fig. 2) begins by collecting an image with a short exposure time in the DAPI channel to determine whether any nuclei are in the field of view (Fig. 2, step 1). This decision is made based on whether an image in the DAPI channel with a fixed exposure time surpasses a minimum brightness and whether the bright objects in the field of view are in the size range of nuclei. Neither threshold is very stringent, because the nuclei may initially be out of focus. If no nuclei are present, the program moves the stage to the next field of view.

If nuclei are present, the program collects images at five different positions along the Z axis (4 μm apart and centered over the initial Z position) and chooses the image with the highest maximum pixel intensity as the closest to focus (Fig. 2, step 2). We found this simple focusing method to be sufficiently accurate. Once the nuclei are in focus, an image is collected with a short exposure time in the YFP channel to determine whether any cells in the field of view are transfected with the YFP-lac repressor fusion protein (Fig. 2, step 3). This decision is based on whether the image surpasses a minimum brightness. If not, the program moves to the next field of view and returns to step 1.

If cells are transfected, the program optimally exposes the array(s), beginning by collecting an image with a 0.05-s exposure time (Fig. 2, step 4). This initial image is rarely optimally exposed, which is defined as producing a maximum intensity between 3,000 and 4,095 (i.e. bright, but not saturated, when using a 12-bit camera). If the image is too bright, the exposure time is shortened by 20%. If the image is too dim, a new exposure time is calculated based on the maximum intensity of the initial image. These cycles are repeated until an appropriate image is obtained, usually within three exposures. Next, the location of the YFP-labeled chromatin array(s) is determined by using a feature extraction algorithm based on a simple threshold of an intensity normalized image (Fig. 2, step 5). If the array is within 15 pixels of the edge of the field of view, it is ignored, because partial arrays touching the edge of the field of view would be incorrectly measured.

Each labeled chromosome region is then individually focused within a small, 256 × 256 pixel window using 10 YFP images spaced 2 μm apart (Fig. 2, step 6). Once optimal focus is achieved, the YFP-labeled array and the corresponding DAPI-labeled nucleus are optimally exposed (Fig. 2, step 7) as in step 4. We considered whether to collect all images with a standard exposure time or to optimally expose each image. Because the feature extraction step sets a threshold based on the maximum and minimum pixel intensities of each image rather than on an absolute pixel intensity, and because exposure time and pixel intensity are roughly linearly correlated, either option would be reasonable. We determined that the program would be more robust if arrays were optimally exposed, because the program would adapt to dimmer samples or microscope conditions.

If more than one YFP-labeled object is observed within the 256 × 256 window, the array is ignored, due to uncertainty as whether this represents aneuploidy or a chromosome rearrangement of the amplified chromosome array versus a single decondensed chromosome region that appears split into two separate foci.

Occasionally, fluorescent debris is imaged. These artifacts are minimized by an algorithm that excludes YFP-labeled arrays that do not completely overlap a DAPI-labeled nucleus (Fig. 2, step 8). The 256 × 256 pixel YFP and DAPI images are saved as individual files within a directory (Fig. 2, step 9). The area of the YFP-labeled array is measured by using a feature extraction algorithm and saved in a text file that can be imported into Microsoft Excel. All aspects of the program can be monitored while the program is running via various image windows (Fig. 2). In addition, the user can keep track of the number of transfected cells found ("Counter") and the number of fields of view scanned ("Count").

The program runs until the entire preselected area of the slide has been examined or a user-specified target number of cells has been found and measured. The user then visually scans a montage of stored images to check for and remove measurement artifacts. The most common error (occurring in roughly 4% of images) occurs when the YFP signal is present throughout the nucleus, and this entire area is measured as an array. This occurs because the cell has lost its lac operator chromosome region or the YFP fusion protein is overexpressed and the background level of YFP signal in the nucleus is comparable to that at the labeled chromosome region. Editing this and other less common artifacts typically requires less than 5 min per sample of 150 cells, thus eliminating about 13% of images. Currently, the measurements corresponding to these artifacts are manually edited within the Microsoft Excel spreadsheet storing these values.

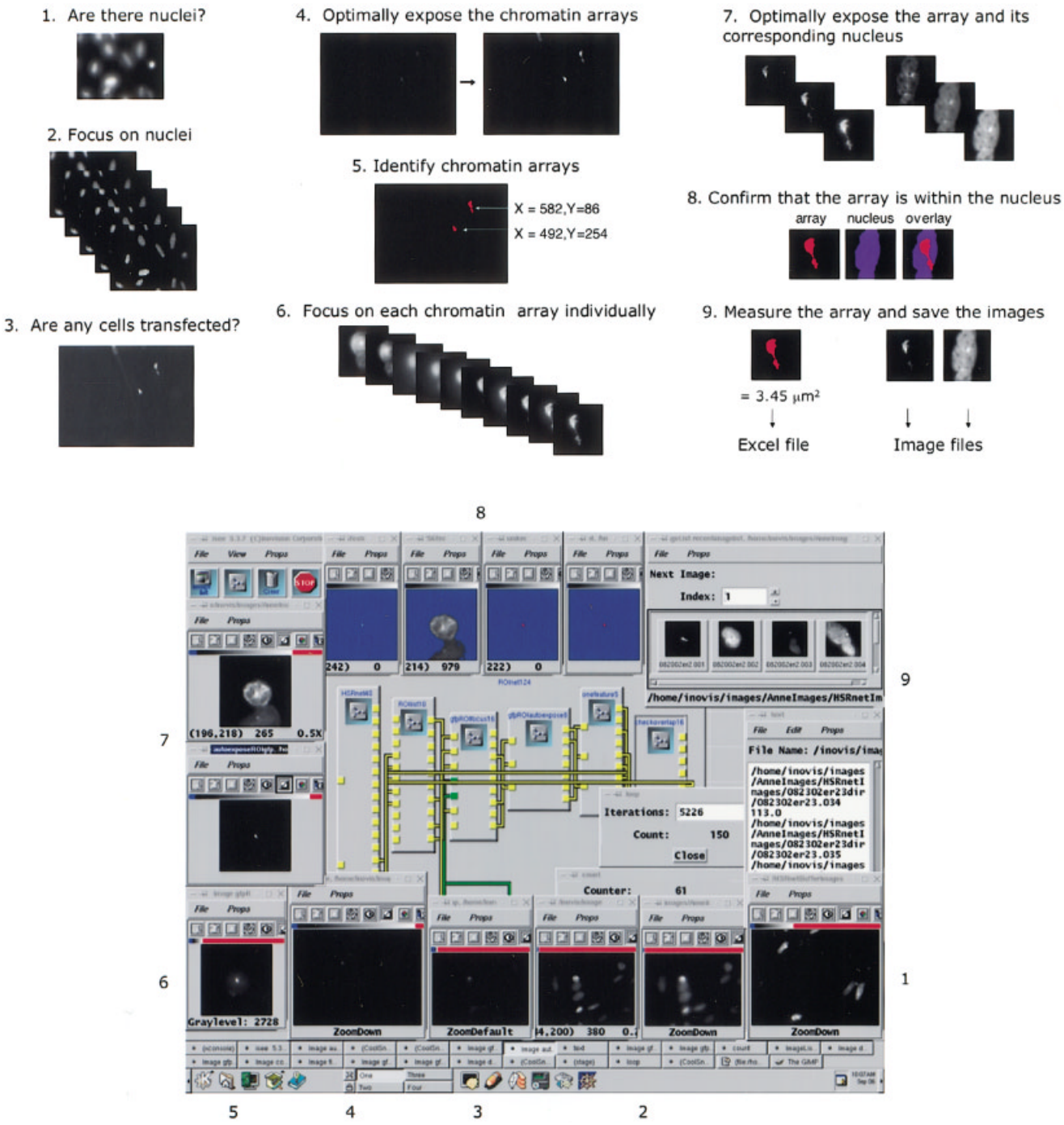

FIG. 2. The automated image collection and analysis procedure (top) and the user interface (bottom). See text for further details.

Validation of the Automated Data Collection Protocol

We first tested this automated data collection program by using well-studied controls. GFP-lac rep does not exhibit chromatin unfolding activity (20), whereas GFP-lac rep-VP16 exhibits strong large-scale chromatin unfolding activity (5). The GFP-lac rep-DEF of ER fusion protein contains the D, E, and F domains of the ER, including the ligand-inducible activation function 2 (AF-2). We previously determined that this region of ER unfolds chromatin dramatically only in the absence of the hormone estradiol (8). The data are presented in two ways: as box plots (Fig. 3A) and plots of mean chromosome region area (Fig. 3B). Results from this automated image collection and analysis

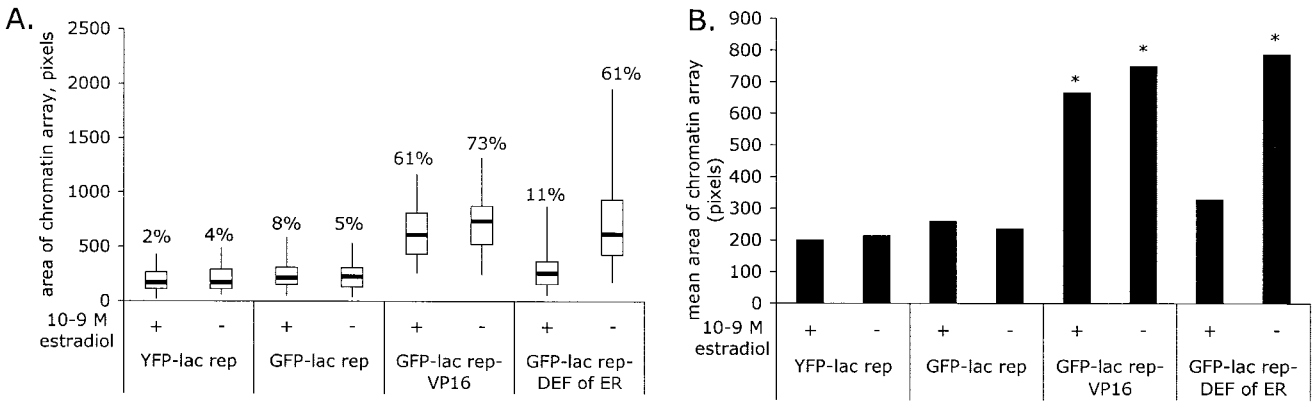

FIG. 3. Control proteins tested with the automated image collection and analysis procedure. Roughly 100 cells were collected for each sample. **A:** Box plots show the spread of sizes observed in each population. The top and bottom of each line show the 95th and 5th percentiles, respectively, and the top and bottom of each box shows the 75th and 25th percentiles, respectively. The horizontal line in the middle of the box shows the 50th percentile (median). The percentage above each box plot indicates the percentage of cells larger than 525 pixels, a threshold between condensed and unfolded structures. **B:** The means for each sample are shown, with asterisks indicating samples significantly different from the GFP-lac rep sample at $P < 0.05$.

program were comparable to previous results. In addition, we confirmed that the YFP-lac rep construct (NYE5[EYFP]) did not significantly unfold chromatin relative to the control GFP-lac rep.

Testing Subdomains of ER for Chromatin Unfolding Activity Using Automated Image Collection

We next sought to further subdivide the ER to identify subdomains that are sufficient to produce large-scale chromatin unfolding when tethered to the A03_1 heterochromatic chromosome region. We therefore constructed a series of truncations and mutations of the ER (Fig. 4) and tested their transcriptional ability in transient transcription assays (Fig. 5). In all cases, transcriptional activity of the YFP-lac repressor-ER fusion proteins was as expected, including the Y537S mutation, which retains significant transcriptional activity even in the absence of estradiol (23), a small portion of domains A and B (amino acids [aa] 35–47) which has negligible transcriptional activity in CHO-K1 cells (24), a region of domain E called AF2a (aa 302–339) which also has negligible transcriptional activity (25), the non-transcriptionally active hinge domain D alone (aa 263–301), the transcriptionally active domains E and F (aa 302–595), and the transcriptionally active domain E alone (aa 302–551). Domain E contains 12 helices based on the crystal structure of the ligand binding domain (26,27). We split domain E helices 1–11 into two halves (aa 302–420, helices 1–6; aa 420–534, helices 7–11) and then removed one to three helices from each end of the 420–534 protein. None of these subdomains of E activated transcription in response to hormone.

We then analyzed these fusion proteins for chromatin unfolding ability by using our automated image collection procedure. Results from two independent transfections were compared to test experimental reproducibility (Fig. 6A and 6B). Similar trends were observed in both experiments. One possible exception is the ER (aa 420–492) protein, which showed greater size variability in the first experiment than in the second, although the medians were similar.

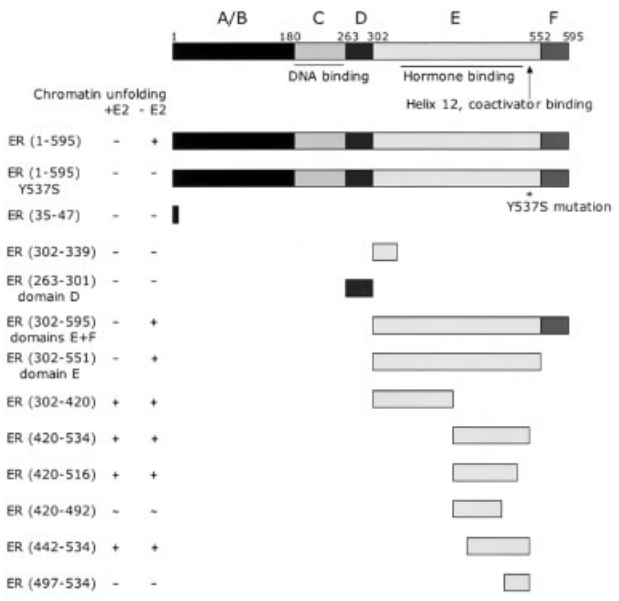

FIG. 4. Schematic of constructs made for this study. The full-length estrogen receptor (ER) is shown at the top, with its six domains, marked A–F, and amino acids marked 1–595. All portions of ER shown were fused to YFP-lac rep. Chromatin unfolding in the presence (+E2) or absence (−E2) of 10^{-9} M estradiol is shown in abbreviated form based on the data shown in Figure 6C, where + indicates significant unfolding, − indicates insignificant activity, and ~ indicates partial activity. These conclusions are the same whether comparing samples statistically ($P < 0.05$) with the negative control (YFP-lac rep) or a positive control (domain E without estradiol). Statistically, ER 420–492 is at the borderline of being significantly different from the negative control ($P = 0.0497$ without hormone and $P = 0.0647$ with hormone) and at the borderline of being significantly different from the positive control ($P = 0.08$ with or without hormone). It is therefore defined has having partial activity.

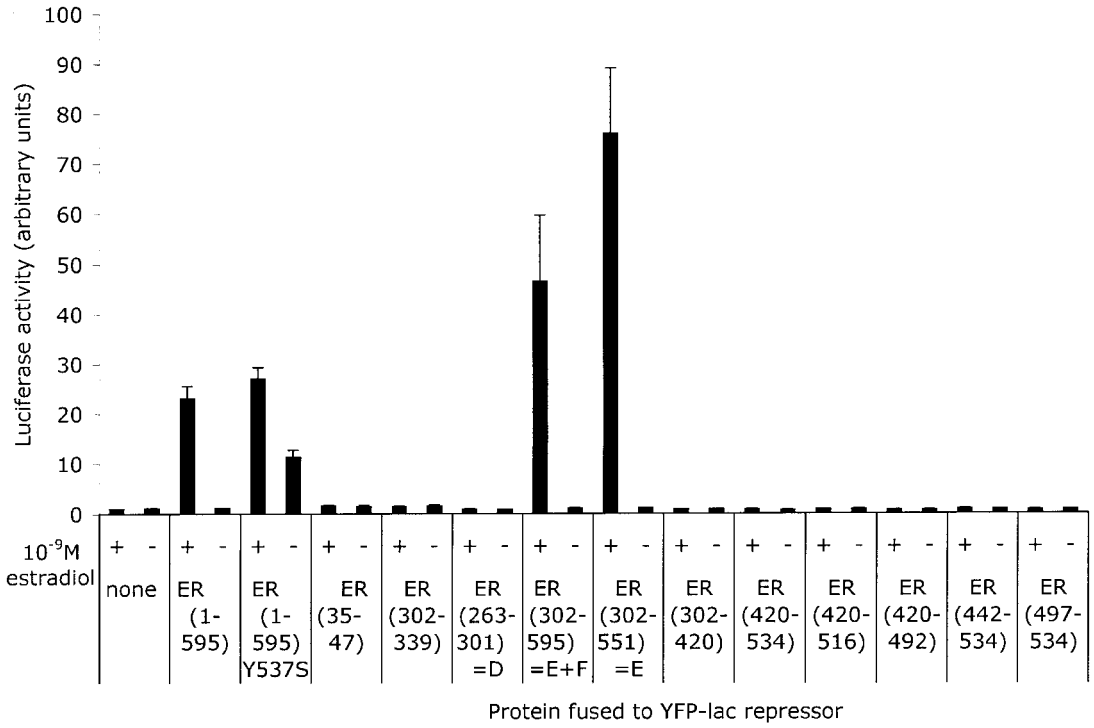

FIG. 5. Transcriptional activity of YFP-lac rep-ER fusion proteins on a transiently transfected 8 lac operator-E1b TATA-luciferase reporter plasmid. Error bars show the standard error from three independent experiments.

We hypothesized that, regardless of whether hormone was added, a constitutively active ER (1–595 Y537S) would resemble the wild-type ER after 48 h of estradiol treatment in its inability to dramatically unfold large-scale chromatin structure. This indeed was the case. The two regions of ER having negligible transcriptional activity (aa 35–47 and aa 302–339) did not exhibit detectable chromatin unfolding activity. Not surprisingly, the hinge region known as domain D also was unable to unfold chromatin significantly. Domains E + F and E alone had unfolding activity, and this unfolding activity was inhibited by the presence of hormone, as is the case for full-length ER. Previously, we showed that helix 12 is required for this inhibition of maximal unfolding activity (8).

Both halves of domain E, excluding helix 12 (ER 302–420 and 420–534), were able to unfold large-scale chromatin structure close to the level produced by the two halves together. However, these subdomains of domain E are unlikely to be able to bind hormone, and their chromatin unfolding ability was not diminished by the presence of hormone. The COOH half (aa 420–534), corresponding to helices 7–11, was further dissected by using truncations from either end. Removing aa 517–534, corresponding to helix 11, had no apparent effect on chromatin unfolding activity. Further removal of aa 493–516, corresponding to helix 10, however, significantly lowered the unfolding activity. This is best seen in Fig. 6B showing box plots with percentile data. Even though the mean array size for construct 420–492 approached that of the other constructs, this was due to a small population of unusually large arrays in the first experiment. We conclude that this construct has diminished activity relative to the intact domain E but still retains partial unfolding activity. Also, removal of aa 420–441, corresponding to helix 7, had no significant effect, but removal of aa 443–496, corresponding to helices 8 and 9, resulted in significant lowering of unfolding activity. Based on these results, it is likely that a major unfolding activity within the COOH half of domain E lies within helices 8–10.

DISCUSSION

Distinct Subdomains of the ER Can Unfold Chromatin

It has become clear that many transcription factors have more than one transcriptional activation domain, including steroid receptors (28), VP16 (29), Sp1 (30), Gcn4 (31), and p65 (32). In this report, we have shown that ER contains multiple large-scale chromatin unfolding domains. In addition to our previous finding that the A–C regions of the ER could unfold chromatin in a small subset of cells (8) is our present finding that each half of the ER's domain E, excluding helix 12, can independently unfold large-scale chromatin structure. This may be a common property of proteins, because BRCA1 was found to have three distinct unfolding domains (7), and VP16 and p65 have more than one unfolding domain (Carpenter AE, Plutz MP, Belmont AS, unpublished observations).

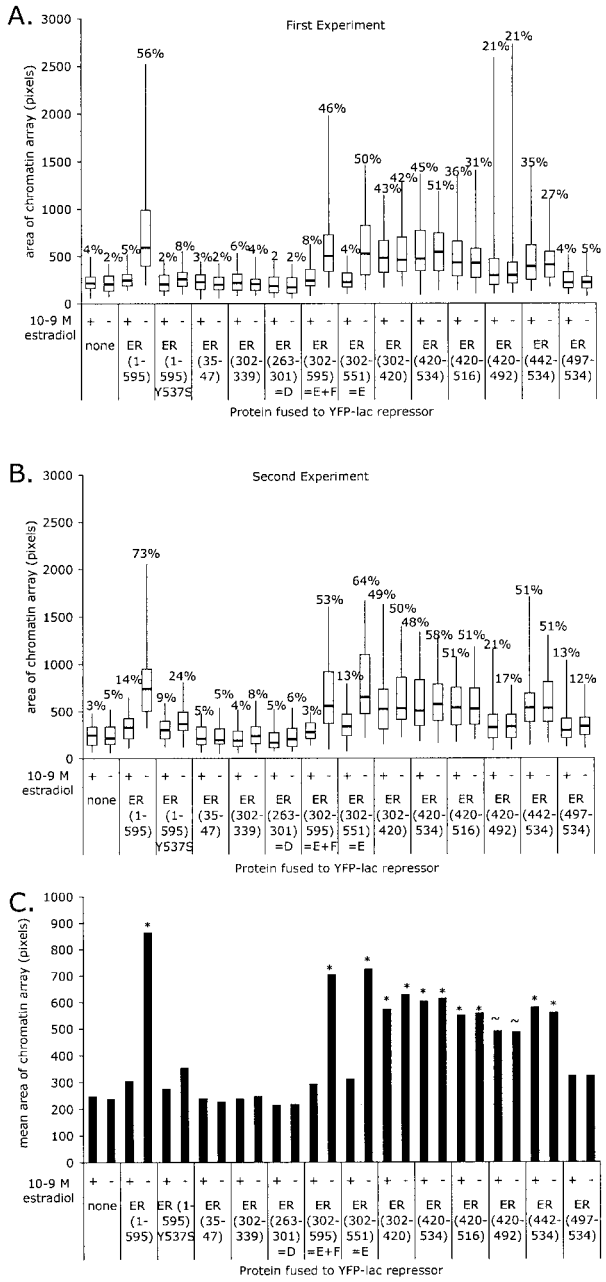

FIG. 6. Large-scale chromatin unfolding assay of YFP-lac rep-ER fusion proteins. Roughly 100 cells were collected for each sample; data are shown as in Figure 3. **A:** Box plots from the first experiment. **B:** Box plots from the second experiment. **C:** Means calculated from both experiments. Asterisks indicate samples with statistically significant unfolding activity, as defined in the caption to Figure 4. +, significant unfolding; −, insignificant activity; ~, partial activity.

Why do transcription factors have multiple domains for transcriptional activation and chromatin unfolding? This apparent redundancy could be present for several reasons: (a) different domains interact with different components of the transcriptional machinery; (b) different domains are regulated differently depending on the cell's environment; and (c) multiple domains synergize, producing stronger effects than expected based on the individual strengths of each domain. In addition, we note the possibility that the apparent redundancy might simply represent a limitation of the experimental assay we used. It is possible that two domains appearing to have comparable independent activity in the assay we used would have distinct functions in other contexts. For example, we assessed only large-scale chromatin unfolding of a single heterochromatic array in a single cell line. If the target gene were in a different chromatin context, perhaps the activities of the two domains would have differed. Alternatively, perhaps in certain cell types, one domain rather than another would be required, based on which coactivator and corepressor proteins were present in those cells.

We emphasize that, although many transcriptional activators appear to have multiple activation domains and multiple unfolding domains, these two types of domains do not necessarily overlap. For the VP16 acidic activation domain, the transcriptional activation and large-scale chromatin unfolding activities map to the same small amino acid motifs (Carpenter AE, Plutz MP, Belmont AS, unpublished observations). However, the present work reveals that several portions of ER unfold chromatin without having transcriptional ability, and that the Y537S constitutive mutation activates transcription in a transient transfection reporter assay but does not significantly unfold chromatin. In addition, portions of BRCA1 unfold chromatin without having transcriptional ability, and several portions that can activate transcription in transient transfection reporter assays do not unfold large-scale chromatin structure (7). The precise relation between large-scale chromatin unfolding and transcription remains to be established.

Proteins Interacting With ER Chromatin Unfolding Domains

Several proteins have been identified that interact in a ligand-dependent manner with helix 12 in AF-2 of ER or with AF-1 at the N terminus of ER (33). However, very few proteins are known to bind to domain E outside of helix 12, the region that we found unfolded large-scale chromatin structure. Among these are $TAF_{II}30$ (with ER aa 283–330) (34), repressor of ER activity (with ER aa 304–530) (35), and possibly CoRNR box-containing proteins (36) and heat shock proteins (37,38). BRCA1 is known to interact with ER aa 282–420 (39) and to unfold large-scale chromatin structure (7). BRCA1 may be responsible, therefore, for the unfolding activity we observed in ER aa 302–420. Now that several small ER subdomains possessing chromatin unfolding activity have been identified, it is feasible to perform a yeast two-hybrid screen to identify novel proteins that may mediate or regulate large-scale chromatin unfolding.

Automated Microscope Outperforms Manual Methods

The primary motivation for our development of an automated program was to allow unattended image col-

lection. However, when using this relatively straightforward algorithm on a commercially available, standard microscope platform, we also significantly accelerated data collection relative to manual microscope operation. An overall rate of 60 transfected cells per hour was obtained in the experiment summarized in Figure 6A as opposed to 36 cells per hour in our previous study (8). Because data collection could proceed unattended for 24 h a day, actual data collection was accelerated even further. With this approach we estimate a rate of data collection of 395 transfected cells per day as opposed to 58 cells per day in our previous work.

Several visual phenotypic screens have been conducted (40,41) that relied on a preliminary nonvisual screen or on human image collection and analysis that is typically tedious and presents the potential for observer bias. For the application described in this paper, results from the automated image collection routine were comparable to previous manually obtained results (5,8,20) and were not subject to bias in the selection of cells for analysis or in the selection of exposure times. In addition, the automated method was much faster. We are now able to test a variety of different proteins for large-scale chromatin unfolding activity in a moderate-throughput manner.

With the current system, the rate-limiting step is the search for cells on the coverslip that have been transfected with the fluorescent protein. These cells are present in 1–25% of the fields of view, depending on transfection efficiency. By using a stable cell line for screening, we estimate that tremendous increases in throughput, on the order of 100-fold, would be obtained. Such rates would be sufficiently high throughput to make feasible chemical biology and siRNA screens to examine specific aspects of nuclear structure and dynamics.

ACKNOWLEDGMENTS

The authors thank Matthew Plutz and Pedro Alvarez-Ortiz for providing plasmid constructs and Ramji Rajendran and Benita Katzenellenbogen for helpful comments. Statistical analysis was provided by Paul Holmes and Dr. Susanne Aref of the Illinois Statistics Office. This work was supported by grants from the National Institutes of Health to Andrew S. Belmont (R01-GM58460 and R01-GM42516). Anne E. Carpenter was a Howard Hughes Medical Institute Predoctoral fellow under the name of Anne C. Nye.

LITERATURE CITED

1. Berger SL. Histone modifications in transcriptional regulation. Curr Opin Genet Dev 2002;12:142–148.
2. Hassan AH, Neely KE, Vignali M, et al. Promoter targeting of chromatin-modifying complexes. Front Biosci 2001;6:D1054–D1064.
3. Belmont AS. Visualizing chromosome dynamics with GFP. Trends Cell Biol 2001;11:250–257.
4. Belmont AS, Li G, Sudlow G, Robinett C. Visualization of large-scale chromatin structure and dynamics using the lac operator/lac repressor reporter system. Methods Cell Biol 1999;58:203–222.
5. Tumbar T, Sudlow G, Belmont AS. Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. J Cell Biol 1999;145:1341–1354.
6. Muller WG, Walker D, Hager GL, McNally JG. Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. J Cell Biol 2001;154:33–48.
7. Ye Q, Hu YF, Zhong H, et al. BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. J Cell Biol 2001;155:911–921.
8. Nye AC, Rajendran RR, Stenoien DL, et al. Alteration of large-scale chromatin structure by estrogen receptor. Mol Cell Biol 2002;22:3437–3449.
9. Stenoien DL, Mancini MG, Patel K, et al. Subnuclear trafficking of estrogen receptor-alpha and steroid receptor coactivator-1. Mol Endocrinol 2000;14:518–534.
10. Taylor DL, Woo ES, Giuliano KA. Real-time molecular and cellular analysis: the new frontier of drug discovery. Curr Opin Biotechnol 2001;12:75–81.
11. Tarnok A, Gerstner AO. Clinical applications of laser scanning cytometry. Cytometry 2002;50:133–143.
12. Kraeft SK, Sutherland R, Gravelin L, et al. Detection and analysis of cancer cells in blood and bone marrow using a rare event imaging system. Clin Cancer Res 2000;6:434–442.
13. Eils R, Uhrig S, Saracoglu K, et al. An optimized, fully automated system for fast and accurate identification of chromosomal rearrangements by multiplex-FISH (M-FISH). Cytogenet Cell Genet 1998;82:160–171.
14. Netten H, Young IT, van Vliet LJ, et al. FISH and chips: automation of fluorescent dot counting in interphase cell nuclei. Cytometry 1997;28:1–10.
15. Petersen AB, Gniadecki R, Wulf HC. Laser scanning cytometry for comet assay analysis. Cytometry 2000;39:10–15.
16. Murphy RF, Boland MV, Velliste M. Towards a systematics for protein subcelluar location: quantitative description of protein localization patterns and automated analysis of fluorescence microscope images. Proc Int Conf Intell Syst Mol Biol 2000;8:251–259.
17. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med 2002;8:1323–1328.
18. Littell RC, Milliken GA, Stroup WW, Wolfinger RD. SAS system for mixed models. Cary, NC: SAS Institute; 1996.
19. Wrenn CK, Katzenellenbogen BS. Structure–function analysis of the hormone binding domain of the human estrogen receptor by region-specific mutagenesis and phenotypic screening in yeast. J Biol Chem 1993;268:24089–24098.
20. Li G, Sudlow G, Belmont AS. Interphase cell cycle dynamics of a late-replicating, heterochromatic homogeneously staining region: precise choreography of condensation/decondensation and nuclear positioning. J Cell Biol 1998;140:975–989.
21. Robinett CC, Straight A, Li G, et al. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J Cell Biol 1996;135(pt 2):1685–1700.
22. Heitz E. Das heterochromatin der moose. Jahresbehr Wiss Botanik 1928;69:762–818.
23. Weis KE, Ekena K, Thomas JA, et al. Constitutively active human estrogen receptors containing amino acid substitutions for tyrosine 537 in the receptor protein. Mol Endocrinol 1996;10:1388–1398.
24. Metivier R, Petit FG, Valotaire Y, Pakdel F. Function of N-terminal transactivation domain of the estrogen receptor requires a potential alpha-helical structure and is negatively regulated by the A domain. Mol Endocrinol 2000;14:1849–1871.
25. Pierrat B, Heery DM, Chambon P, Losson R. A highly conserved region in the hormone-binding domain of the human estrogen receptor functions as an efficient transactivation domain in yeast. Gene 1994;143:193–200.
26. Brzozowski AM, Pike AC, Dauter Z et al. Molecular basis of agonism and antagonism in the oestrogen receptor. Nature 1997;389:753–758.
27. Tanenbaum DM, Wang Y, Williams SP, Sigler PB. Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc Natl Acad Sci USA 1998;95:5998–6003.
28. Kumar R, Thompson EB. The structure of the nuclear hormone receptors. Steroids 1999;64:310–319.
29. Regier JL, Shen F, Triezenberg SJ. Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. Proc Natl Acad Sci USA 1993;90:883–887.
30. Courey AJ, Tjian R. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell 1988;55:887–898.
31. Drysdale CM, Duenas E, Jackson BM et al. The transcriptional activator GCN4 contains multiple activation domains that are critically dependent on hydrophobic amino acids. Mol Cell Biol 1995;15:1220–1233.
32. Moore PA, Ruben SM, Rosen CA. Conservation of transcriptional activation functions of the NF-kappa B p50 and p65 subunits in mammalian cells and *Saccharomyces cerevisiae*. Mol Cell Biol 1993;13:1666–1674.

33. Klinge CM. Estrogen receptor interaction with co-activators and co-repressors. Steroids 2000;65:227–251.
34. Jacq X, Brou C, Lutz Y, et al. Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. Cell 1994;79:107–117.
35. Delage-Mourroux R, Martini PG, Choi I, et al. Analysis of estrogen receptor interaction with a repressor of estrogen receptor activity (REA) and the regulation of estrogen receptor transcriptional activity by REA. J Biol Chem 2000;275:35848–35856.
36. Huang HJ, Norris JD, McDonnell DP. Identification of a negative regulatory surface within estrogen receptor alpha provides evidence in support of a role for corepressors in regulating cellular responses to agonists and antagonists. Mol Endocrinol 2002;16:1778–1792.
37. Schlatter LK, Howard KJ, Parker MG, Distelhorst CW. Comparison of the 90-kilodalton heat shock protein interaction with in vitro translated glucocorticoid and estrogen receptors. Mol Endocrinol 1992;6: 132–140.
38. Chambraud B, Berry M, Redeuilh G, et al. Several regions of human estrogen receptor are involved in the formation of receptor-heat shock protein 90 complexes. J Biol Chem 1990;265:20686–20691.
39. Fan S, Ma YX, Wang C, et al. Role of direct interaction in BRCA1 inhibition of estrogen receptor activity. Oncogene 2001;20:77–87.
40. Peterson RT, Link BA, Dowling JE, Schreiber SL. Small molecule developmental screens reveal the logic and timing of vertebrate development. Proc Natl Acad Sci USA 2000;97:12965–12969.
41. Mayer TU, Kapoor TM, Haggarty SJ, et al. Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. Science 1999;286:971–974.