Extensive Chromatin Fragmentation Improves Enrichment of Protein Binding Sites in Chromatin Immunoprecipitation Experiments

The Harvard community has made this article openly available. **Please share** how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1093/nar/gkn535</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:8255740">http://nrs.harvard.edu/urn-3:HUL.InstRepos:8255740</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Extensive chromatin fragmentation improves enrichment of protein binding sites in chromatin immunoprecipitation experiments

Xiaochun Fan, Nathan Lamarre-Vincent, Qian Wang and Kevin Struhl*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

Received June 13, 2008; Revised July 31, 2008; Accepted August 4, 2008

ABSTRACT

Extensive sonication of formaldehyde-crosslinked chromatin can generate DNA fragments averaging 200 bp in length (range 75–300 bp). Fragmentation is largely random with respect to genomic region and nucleosome position. ChIP experiments employing such extensively fragmented samples show 2- to 4-fold increased enrichment of protein binding sites over control genomic regions, when compared to samples sonicated to a more conventional size range (300–500 bp). The basis of improved fold enrichments is that immunoprecipitation of protein-bound regions is unaffected by fragment size, whereas immunoprecipitation of control genomic regions decreases progressively along with reduced fragment size due to fewer nonspecific binding sites. The use of extensively sonicated samples improves mapping of protein binding sites, and it extends the dynamic range for quantitative measurements of histone density. We show that many yeast promoter regions are virtually devoid of histones.

INTRODUCTION

Chromatin immunoprecipitation (ChIP) is the standard method to determine the association of proteins with their DNA targets in vivo (1). Typically, living cells are treated with formaldehyde to crosslink proteins to DNA, and the crosslinked chromatin is fragmented by sonication. The resulting ‘input’ sample is immunoprecipitated with an antibody that specifically recognizes the protein of interest, thereby enriching DNA sequences that associate with the protein in vivo. The basic measurement of a ChIP experiment is ‘IP efficiency’ (IE) for a given genomic region, which is defined as the amount of DNA in the IP sample divided by the amount of DNA in the input sample. IP efficiencies for protein-bound regions are higher than for control genomic regions, with the relative level of protein binding being defined as ‘fold enrichment’ over control regions. For standard ChIP analyses of individual genomic regions, amounts of DNA in the input and IP samples are determined by quantitative PCR, often in real time using an appropriate machine. In large-scale ChIP experiments, the samples are amplified by PCR, and the resulting material analyzed on microarrays (2–4) or by high-throughput DNA sequencing (5–8).

The sensitivity of a ChIP experiment ultimately depends on the ability to separate protein-bound DNA fragments from the background of unbound fragments. In this regard, the quality of the antibody and the IP procedure are critical parameters. However, a different type of background arises from nonspecific association of the protein with genomic DNA that is captured by crosslinking, and such background is unaffected by changes in the IP procedure (9). In principle, the level of nonspecific binding should decrease in accordance with decreasing DNA fragment length, because shorter DNA fragments have fewer nonspecific sites. In contrast, binding to specific target sites should depend primarily on the molarity of the DNA fragments and not the length. Here, we show that extensive fragmentation of chromatin significantly increases fold enrichments of protein binding sites in ChIP experiments.

MATERIALS AND METHODS

Chromatin preparation

The yeast experiments utilized a derivative of Saccharomyces cerevisiae strain XF236F8 expressing a Pdr1 derivative tagged at the carboxyl-terminus with three copies of the HA epitope. This strain was generated by gene replacement by integrating an appropriate PCR fragment with the URAS marker followed by looping out this marker through homologous recombination (10). The procedure for micrococcal nuclease (MNase) digestion of noncrosslinked chromatin was described previously (11).
For ChIP experiments in yeast, cells (40 ml) were grown at 30°C in YPD (1% yeast extract; 2% peptone and 2% dextrose) medium to an optical density of 0.6–0.8 at 600 nm, treated with formaldehyde (1% final concentration) for 20 min at room temperature, and then quenched with 10 ml of 2 M glycine for 5 min at room temperature. The fixed cells were disrupted in a Mini Bead Beater (BioSpec Products, Inc. Bartlesville, OK, USA) with six 3-min cycles at maximum speed, and cross-linked chromatin prepared as described previously (1). The resulting samples were transferred to 2 ml microcentrifuge tubes in 0.5 ml aliquots and sonicated by two methods. First, cells were diluted 2-fold in lysis buffer to a volume of 1 ml and sonicated three times (30 continuous pulses at about 20% power) in a Branson Sonifier 450 with a microtip probe at 4°C, with 2 min cooling on ice between pulses. Second, 0.5 ml cells were sonicated in Misonix Sonicator 3000 (Cole-Parmer Instrument Company, Vernon Hills, IL, USA) with a rotating device and a horn for various times (output level of 6; 10 s ON and 10 s OFF). Ice and ice-cold water were filled in the chamber to keep samples cool. For both methods, sonicated samples were then centrifuged for 20 min at 16,000 g to remove insoluble debris. The supernatant was transferred to clean tubes, flash frozen in liquid nitrogen and stored at −80°C.

HeLa cells were grown in MEM supplemented with 10% fetal bovine serum, 100 U/ml penicillin and 100 μg/ml streptomycin. Then 10⁷ cells were trypsinized, fixed in 10 ml MEM supplemented with 1% formaldehyde for 10 min at room temperature, and then quenched by the addition of glycine to a final concentration of 125 mM. After centrifugation, the cells were resuspended in cell lysis buffer (25 mM HEPES pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, 1 mM DTT, 0.1% NP-40) supplemented with EDTA-free protease inhibitor cocktail (Roche, Indianapolis, IN, USA) and 0.5 mM PMSF and incubated on ice for 10 min. After centrifugation, the crude nuclear pellet was resuspended in nuclear lysis buffer (50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS) supplemented with EDTA-free protease inhibitor cocktail and 0.5 mM PMSF. Chromatin was sheared at 4°C, 10 times at 30 second intervals on a Branson Microtip Sonifier 450 (BLD Inc. Garner, NC, USA) set at constant duty and an output level of 4. After centrifugation for 10 min at 16000g, chromatin was sonicated for an additional 0–5 min at 10 s intervals in 0.5 ml aliquots using a cup horn on a Misonix Sonicator 3000 set at level 6.

Chromatin immunoprecipitation

Sonicated chromatin preparations were thawed on ice and a portion of the samples saved as input materials. Input materials were decrosslinked at 65°C overnight in the presence of 50 μl ChIP elution buffer (50 mM Tris–HCl, pH 7.5; 10 mM EDTA; 1% SDS) and purified through QIAGen columns. DNA was eluted with 30 μl TE (10 mM Tris, pH 8.0; 1 mM EDTA), and 10 μl of eluate was diluted to 1 ml as input and 20 μl of eluate was treated with RNase A/T1 and resolved in 1.5% agarose by electrophoresis. Chromatin from about 0.5 to 1 x 10⁸ cells (0.1–0.4 ml chromatin aliquots mentioned above) was incubated with 5 μl 8WG16 [for RNA polymerase II (Pol II)] or 10 μl F-7 (for HA-tagged protein) for 90 min on a rotating wheel at room temperature, after which 10 μl of protein A sepharose was added and incubated for another 90 min. Protein complexes were washed twice with FA lysis buffer (150 mM NaCl), once with FA lysis buffer (500 mM NaCl), once with ChIP wash buffer (10 mM Tris–HCl, pH 8.0; 250 mM LiCl; 1 mM EDTA; 0.5% Nonidet P-40 and 0.5% sodium deoxycholate) and once with TE. Immunoprecipitated complexes were eluted by incubation with 200 μl ChIP elution buffer at 65°C for 1 hour, and samples were decrosslinked at 65°C overnight and purified through QIAgen columns.

Quantitation by real-time PCR

For ChIP experiments, input and immunoprecipitated samples were assayed by quantitative PCR to assess the extent of protein occupancy at different genomic regions. PCR reactions contained 2 μl DNA template, 3 μl of 3.3 μM primer pairs and 5 μl of 2x EvaGreen reaction mix (FluoProbes Interchim Montluçon cedex, France). Quantitative PCR was performed on an Applied Biosystem 7500 Fast unit using a 10 min soak at 95°C, followed by 35 cycles of 5 s at 95°C, 5 s at 53°C and 20 s at 72°C. Threshold cycle (Ct) values were determined at threshold of 0.01. For each amplification product, the IE was determined using the formula IE = a*(10^(-ΔCt(Input)−ΔCt(ChIP)))/Ct(Input), where a is the constant associated with the ratio and dilution of IP and input materials. Fold occupancy of any given region over an open reading frame (ORF)-free region on chromosome V (chrV) control was determined. IP efficiencies and occupancy values presented here represent an average of three independent experiments and have a standard deviation (SD) of less than ±25%. Nucleosome scanning analysis was performed as described previously (11). Sequence information of primer pairs used in this study are available upon request.

RESULTS

Extensive sonication can reduce chromatin fragments to an average size of 200 bp

Essentially all proteins that associate with specific DNA sequences also associate nonspecifically at a low level with virtually any region of the genome. In the context of ChIP experiments, such nonspecific binding can be seen by the slightly higher IP efficiencies at 'control' regions in IPs with the antibody of interest as compared to control IPs lacking an appropriate antibody (9). As mentioned in the Introduction section, such nonspecific binding should increase with fragment size, whereas specific binding should be largely independent of fragment size. Thus, we reasoned that extensive fragmentation of chromatin should preferentially reduce the nonspecific binding, and hence might result in improved fold enrichments in ChIP experiments.

Our heretofore standard procedure for yeast cells utilized a Branson Sonifier 450 with a microtip probe for 90 s,
resulting in DNA fragments in the range of 300–500 bp, which is typical for ChIP experiments described by other laboratories. We were able to reduce DNA fragment size further using a Misonix Sonicator 3000 with a horn (Figure 1A). Sonication for 4 min yielded a size range of DNA fragments similar to that achieved with the Branson Sonifier 450, but sonication for 16 min yielded DNA fragments ranging from 75 to 300 bp in length. In addition, the yield of solubilized chromatin increased about 2- to 3-fold in samples sonicated for longer time (data not shown). Sonication times as long as 60 min did not further reduce DNA fragment size significantly.

**Fragmentation is largely random with respect to genomic region and nucleosome position**

The average size of the extensively sonicated chromatin is roughly similar to that of mono-nucleosomes, and we considered the possibility that fragmentation would preferentially occur within the relatively flexible linker regions between nucleosomes. However, nucleosome-scanning analysis (11) across the *PHO5* promoter and coding region showed that the extensively sonicated sample has constant level of DNA at all positions tested. In contrast, parallel analysis of mono-nucleosomal DNA generated by MNase treatment of noncrosslinked chromatin (MN) were analyzed with the indicated primers pairs spanning the *PHO5* gene. Values are normalized to those obtained with genomic DNA and presented with respect to the average value (defined as 1) in the same sample.

**Increased fold enrichments upon increased fragmentation of chromatin**

We performed ChIP analysis for RNA Pol II on samples generated at various times of sonication. As shown in Figure 2A, for genomic regions tested that show high Pol II occupancy, the various samples showed similar IP efficiencies. In contrast, at control genomic regions, the IP efficiency decreased significantly in samples with longer sonication time. When IP efficiencies were normalized to a negative control, an ORF-free region on chrV, fold enrichments at all Pol II-associated regions increased progressively in accordance with the time and extent of sonication, whereas fold enrichment at the telomeric region (TEL) was unaffected (Figure 2B). As expected from the similar distributions of DNA fragment sizes, the ChIP results obtained from samples prepared with the Branson Sonifier 450 were comparable to those of samples prepared sonication for 4 min with the Misonix Sonicator 3000. Thus, by reducing DNA fragment sites to 75–300 bp (16 min sonication with the horn) from the more conventional 300–500 bp (Branson Sonifier 450), we were able to obtain fold enrichments that are about 2-fold higher than that obtained previously.

**PCR product length does not significantly affect fold enrichments**

One issue with samples containing reduced size of chromatin fragments is that quantitative analyses involving primer pairs producing long PCR products might be adversely affected. In principle, the input and immunoprecipitated samples should be equally affected by the length of the PCR products used in the analysis, thereby having no effect on fold enrichments. To test this, we compared the fold enrichments at *POL1* and *BTN2* using primer pairs that generated small (74 and 64 bp) or large (219 and 349 bp) PCR products (Table 1). In terms of fold enrichment, there was no significant difference between each pair of primer set in chromatin samples prepared with various sonication time and different methods (Figure 2C).
length. Thus, chromatin samples with shorter DNA fragments should yield ChIP profiles with sharper peaks, and hence improve the resolution of the binding site.

To examine the effect of fragment size on mapping protein binding sites, we examined the association of (HA)_3-tagged Pdr1. In accordance with the Pol II results above, IP efficiencies of regions containing Pdr1 binding sites (PDR5, SNQ1 and PDR16 promoters) were unaffected by sonication times, whereas they significantly decreased at control genomic regions (Figure 3A). As a consequence, fold enrichments increased as a function of sonication time, with the most extensively fragmented samples showing 4-fold higher ChIP signals than that obtained with the conventional method using a sonication microtip probe (Figure 3B). We mapped Pdr1 occupancy at the VHR1 promoter region using eight PCR primer pairs spanning a 700 bp region encompassing the VHR1 promoter (Table 1). For chromatin samples prepared by sonication for 16 min with a horn, Pdr1 association appeared as a sharp peak at −752 to −690; while Pdr1 profile was significantly flatter for chromatin samples prepared by sonication for 90 s with a microtip probe (Figure 3C). These results are consistent with the putative Pdr1 binding sites near −736 and −728 at the VHR1 promoter (13). Similar results were observed at the PDR5, PDR16 and TPO1 promoters (data not shown).

Improved mapping of protein binding sites upon extensive fragmentation of chromatin

A protein bound to a specific DNA site in vivo generates a predicted ChIP profile that depends on the size of the fragmented DNA in the input chromatin sample and the length of the PCR products used for the analysis (12). Peak ChIP signals are centered at binding sites and extend the length of the PCR product, whereupon gradually and symmetrically decrease at regions flanking the peak in a manner that depends on DNA fragment length. Thus, chromatin samples with shorter DNA fragments should yield ChIP profiles with sharper peaks, and hence improve the resolution of the binding site.

As expected, however, it took more PCR cycles (higher Ct values) to reach the measurement threshold in samples that were more extensively sonicated (Figure 2D), reflecting the decreasing population of DNA molecules that are long enough to include sequences corresponding to both primers.

Improved measurements of histone density upon extensive fragmentation of chromatin reveals that many yeast promoter regions are virtually devoid of histones

Although most DNA in eukaryotic cells exists in the form of nucleosomes, many promoter (and perhaps other) regions appear to be depleted of histones (11,14–17). However, it has been difficult to quantitate the extent of nucleosome depletion. First, analyses involving mononucleosomal DNA generated by MNase treatment are complicated by the strong DNA sequence preferences of

---

Table 1. Oligos used to compare the size effect of PCR products on fold enrichments in Figure 2D and to map Pdr1 binding sites at VHR1 promoter in Figure 3C, with the coordinates relative to ATG

<table>
<thead>
<tr>
<th>Primer</th>
<th>5’</th>
<th>3’</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>POL1(S)</td>
<td>+713</td>
<td>+786</td>
<td>74</td>
</tr>
<tr>
<td>BTN2(S)</td>
<td>+657</td>
<td>+720</td>
<td>64</td>
</tr>
<tr>
<td>BTN2(I)</td>
<td>+559</td>
<td>+907</td>
<td>349</td>
</tr>
<tr>
<td>A</td>
<td>−1010</td>
<td>−943</td>
<td>68</td>
</tr>
<tr>
<td>B</td>
<td>−962</td>
<td>−891</td>
<td>72</td>
</tr>
<tr>
<td>C</td>
<td>−881</td>
<td>−812</td>
<td>70</td>
</tr>
<tr>
<td>D</td>
<td>−752</td>
<td>−690</td>
<td>63</td>
</tr>
<tr>
<td>E</td>
<td>−657</td>
<td>−597</td>
<td>61</td>
</tr>
<tr>
<td>F</td>
<td>−558</td>
<td>−499</td>
<td>60</td>
</tr>
<tr>
<td>G</td>
<td>−484</td>
<td>−425</td>
<td>60</td>
</tr>
<tr>
<td>H</td>
<td>−345</td>
<td>−286</td>
<td>60</td>
</tr>
</tbody>
</table>

---

Figure 2. RNA Pol II occupancy at some genomic regions. Pol II occupancy at some genomic regions for chromatin preparations with different sonication time and methods was presented as IP efficiency (A) or fold over ORF-free region on chrV (B). (C) Pol II occupancy determined for genes with primer pairs giving different sizes of PCR products. The size of PCR products can be found in Table 1. (D) Differences of Ct values between primer pairs giving different sizes of PCR products. Error bars show SD from three independent experiments.
MNase (18) that preferentially cleave AT-rich promoter regions. Second, measurements of histone density by ChIP using antibodies against histones (or epitope-tagged histones) are problematic because the 300–500 bp DNA fragments from standard ChIP samples will have 2–4 nucleosomes and histone crosslinking to DNA is extremely efficient. In this regard, the 2-fold reduction in histone density at many yeast promoter regions (11) may be a significant underestimate of nucleosome depletion, especially when nucleosome-free regions are shorter than the size of the chromatin fragments.

To address this issue, we performed a ChIP experiment on the extensively sonicated chromatin sample using an antibody against histone H3. As shown in Figure 4, many yeast promoter regions are virtually devoid of histones, with histone densities <10% of those observed at coding regions. As expected, the PHO5 (19) and GAL1,10 (20) promoter regions have much higher histone densities, although these are typically 2-fold below that of coding regions. Also, as expected (16,21), coding regions of highly active genes (RPL3 and ENO2) also show approximately 2-fold lower levels of histone density in comparison to those observed at a poorly transcribed gene (POL1), ORF-free region and TEL. Thus, the use of extensively sonicated samples greatly increases the quantitative accuracy and dynamic range of histone density experiments.

Improved fold enrichment upon extensive fragmentation of crosslinked chromatin from human cells

To extend these results to ChIP experiments in mammalian cells, crosslinked chromatin from HeLa cells was prepared by the standard procedure using the Branson Sonifer 450. This material was then sonicated for an additional 2 or 5 min with the Misonix Sonicator 3000, resulting in a progressive decrease in the size of DNA (Figure 5A). After 5 min, the bulk of the DNA was between 150 and 250 bp, which is slightly larger than the size range obtained for yeast cells. For the experiment shown, generating DNA in the 150–250 bp range required this two-step sonication procedure; treatment with the Misonix Sonicator 3000 alone was insufficient. However, the precise conditions will likely vary among cell types.

We performed ChIP analysis for RNA Pol II on the above samples with the same antibody used in the yeast experiments. In accordance with the results in yeast, longer sonication times resulted in more dramatic reduction in IP efficiencies at control genomic than at the ING1 promoter and EEF1A1 coding regions (Figure 5B). As a consequence, enrichment at the ING1 promoter and EEF1A1 coding region was 3- to 4-fold higher in the most extensively sonicated sample as compared to the sample generated with the Branson Sonifer 450 alone (Figure 5C). Attempts to improve the fold enrichment by more extensive sonication (10 min with the Misonix Sonicator 3000) resulted in reduced IP efficiency and fold enrichment at the ING1 promoter, and we suspect that this is due to damage to Pol II.

DISCUSSION

Extensive sonication of formaldehyde-crosslinked chromatin samples from both yeast and human cells can reduce average fragment size to ~200 bp (range 75–300 bp). Such extensively fragmented chromatin samples yield a 2- to 4-fold increase in fold enrichments in ChIP experiments when compared to samples sonicated to a more conventional size range (300–500 bp). The basis of improved fold enrichments is that IP efficiencies at protein binding sites are unaffected by fragment size,
whereas IP efficiencies at control genomic regions decrease progressively along with reduced fragment size. These experimental observations are in accord with the prediction that the level of nonspecific binding should decrease in accord with decreasing DNA fragment length, because shorter DNA fragments have fewer nonspecific sites, whereas binding to specific target sites should depend primarily on the molarity of the DNA fragments. We do not know whether fragment size affects the background in ChIP experiments due to the population of molecules that are unbound by the protein of interest.

The use of extensively fragmented chromatin samples should increase the sensitivity of all ChIP applications, because sensitivity is essentially defined by fold enrichment. First, extensively fragmented samples will be particularly beneficial for ChIP experiments involving proteins that inefficiently crosslink to chromatin. For example, ChIP experiments that give rise to 3- to 5-fold enrichments by conventional procedures would be expected to yield 6- to 20-fold enrichments, a very significant difference. Second, increased sensitivity of the ChIP assay will also be beneficial when proteins associate weakly with genomic regions, either due to an intrinsically weak binding site or as a consequence of environmental or genetic conditions. Third, highly fragmented chromatin samples improve resolution of where the protein associates with DNA. Fourth, extensive sonication greatly improves quantitative measurements of histone density, and it reveals that many yeast promoter regions are virtually devoid of histones. Lastly, extensively fragmented chromatin samples should significantly improve the ability to distinguish target versus nontarget sites in ChIP-chip and ChIP-sequencing experiments. In both of these large-scale applications, variability due to the amplification procedure, cross-hybridization and counting error due to the limited amount of DNA sequencing decreases the sensitivity of the assay. The higher chromatin yields arising from extensive sonication will help reduce variability due to amplification and counting error. In addition, extensive sonication produces DNA fragments that are of ideal size for making libraries for ChIP sequencing.

Unlike chemical or enzymatic treatments, sonication is an idiosyncratic procedure that depends on many variables and needs to be calibrated empirically. In our laboratory, we were able to fragment yeast chromatin to a size range of 75–300 bp using either the Misonix Sonicator 3000 with a horn for 16 min or the Branson Sonifier 450 with a microtip probe for 8 min. One advantage of the Misonix Sonicator is the ability to process up to 20 samples simultaneously, and the use of a rotating device minimizes variation in fragment size among samples. The degree of sonication might also be affected by crosslinking time, formaldehyde concentration and the concentration of detergent in the broken cells, although we have not investigated these parameters systematically.

In our experiments with the HeLa cell derivative, fragmentation to the 150–250 bp range required a two-step sonication procedure involving both the Branson Sonifier 450 and the Misonix Sonicator 3000. However, as chromatin fragmentation by sonication can differ among mammalian cell types, the procedure needs to be optimized for each cell type. In addition, we found that excessive sonication reduced IP efficiency and fold enrichments, probably due to damage to the protein of interest.

In summary, our results indicate that chromatin can be sonicated to a size range of 75–300 bp, and that such
extensive sonication is highly desirable for ChIP experiments. Thus, we believe that it is well worth the effort to optimize sonication conditions that permit such extensive fragmentation of chromatin.

ACKNOWLEDGEMENTS

We thank Joseph Fleming and Marianne Lindahl Allen for communicating unpublished results, Heather Hirsch and Zarmik Moqtaderi for technical advice and Joseph Geisberg for useful discussions.

FUNDING

Damon Runyon-Walter Winchell Cancer Research Foundation postdoctoral fellowship (Q.W.); National Institutes of Health (GM30186 to K.S.). Funding for open access charge: NIH GM 30186, GM 53720, and Institutes of Health (GM30186 to K.S.). Funding for Foundation postdoctoral fellowship (Q.W.); National Damon Runyon-Walter Winchell Cancer Research Foundation postdoctoral fellowship (Q.W.); National Institutes of Health (GM30186 to K.S.). Funding for open access charge: NIH GM 30186, GM 53720, and Institutes of Health (GM30186 to K.S.).

Conflict of interest statement. None declared.

REFERENCES