The use of biotin tagging in *Saccharomyces cerevisiae* improves the sensitivity of chromatin immunoprecipitation

Folkert J. van Werven and H. Th. Marc Timmers*

Department of Physiological Chemistry, Division of Biomedical Genetics, University Medical Centre Utrecht, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands

Received January 20, 2006; Accepted February 9, 2006

ABSTRACT

Affinity tagging has been used in many global studies towards protein function. We describe a highly efficient system for *in vivo* biotinylation of transcription factors in the yeast *Saccharomyces cerevisiae*, which is based on the bacterial BirA biotin ligase. The strength of the biotin–streptavidin interaction was exploited to improve detection of *in vivo* protein–DNA complexes in chromatin immunoprecipitation (ChIP) experiments. In a test system using the biotin-tagged LexA DNA-binding protein, we found that stringent washing conditions resulted in a strong improvement of the signal-to-noise ratios. Yeast strains with chromosomally integrated versions of tagged transcription factor genes were generated using N- or C-terminal biotin-tagging cassettes. ChIP experiments with biotinylated Rpb3p, a RNA polymerase II subunit, showed that Rpb3p-binding could even be detected at weakly expressed genes. Other methods failed to detect RNA polymerase II binding at such genes. Our results show that biotinylation of yeast transcription factors improves the detection of *in vivo* protein–DNA complexes.

INTRODUCTION

Transcription regulation in eukaryotic cells requires the coordinated and sequential action of a plethora of gene-regulatory proteins, cofactor complexes and basal transcription factors (1–3). These proteins interact directly with the regulatory DNA sequences encompassing enhancers and promoter elements or indirectly via recruitment through DNA-specific gene regulators. Approaches depending on *in vitro* DNA-binding and transcription assays or on yeast genetic screens were instrumental in identification of transcription regulatory proteins. In many cases it remained uncertain, however, whether the interaction between transcription regulators and promoter DNA occurred in living cells or whether the observed effects were indirect (4,5).

This has motivated development of *in vivo* protein-DNA crosslinking methods (6,7). The so-called chromatin immunoprecipitation (ChIP) assay has found widespread use to study transcription and chromatin regulation. ChIP relies on fixation of protein–DNA interactions by crosslinking agents such as formaldehyde, immunopurification of the covalently cross-linked complexes and detection of the purified DNA (6,7). As formaldehyde instantly penetrates living cells, *in vivo* binding of transcription factors to DNA can be determined in a time-resolved manner. This has been illustrated by several studies in yeast and human cells (8–11). In addition, sequential ChIP assays allow determination of simultaneous binding of transcription factors to the same DNA fragment (12). Combining ChIP experiments with DNA microarray detection allows determination of the genome-wide localization of transcription factors. Recent studies using tiling microarrays to determine histone modifications on yeast chromosomes showed that nearly all fragments can be mapped provided that the signal-to-noise ratios of the ChIP are acceptable (13).

The immunoprecipitation step requires highly specific antibodies with high affinity for the antigen. In yeast, this can be often circumvented by fusing the gene of interest with one or more copies of generic epitopes like hemagglutinin (HA), Myc and Flag for which high quality antibodies are available. Alternative strategies involve modification with the tandem affinity purification (TAP) tag. However, all approaches suffer from low signal-to-noise ratios compromising the analysis of low-abundant and transient protein–DNA interactions.

To increase signal-to-noise ratios we explored the use of *in vivo* biotinylation in the yeast *Saccharomyces cerevisiae*. The binding between biotin and streptavidin represents one of the strongest non-covalent interactions known ($K_d = 10^{-15}$ M). Several *in vivo* protein biotinylation systems for bacterial and mammalian cells relying on the BirA biotin...
ligase from *Escherichia coli* have been described. These have been applied to protein (complex) purification, and ChIP analysis in mammalian cells (14–16). Here we describe an *in vivo* biotinylation system in yeast for highly sensitive detection of protein–DNA complexes using ChIP approaches.

**MATERIALS AND METHODS**

**Yeast strains**

For the LexA experiments strain EGY48 was used (17). In other experiments we used W303-1B. The Rpb3-TAP strain was generated by PCR-mediated chromosomal integration from genomic DNA of SC1126 strain purchased from EUROSCARF, using primers representing the upstream and downstream sequence of the integrated allele (Table 1).

The strategy for Rpb3-Avitag PCR-mediated chromosomal tagging has been described previously (18,19). In short, two 100 μl PCRs (5 U of Taq DNA polymerase, 0.1 mM of each dNTP and 0.2 μM of each primer) were performed according to the manufacturer’s instruction (NEB) for 30 cycles and transformed into W303-1B yeast cells using a high efficient transformation protocol (20). Cells were grown for 2 h in 1 ml YPD, after which they were plated out on YPD plates supplemented with G418 (200 μg/ml). Rpb3-Avitag and Rpb3-TAP integrations were verified by immunoblot analysis and PCR using primers corresponding to the tagged allele.

**Plasmid construction**

For the LexA experiments sequences for BirA and BirA-NLS were cloned from CMV3xHAbirA (gift from J. Strouboulis, Erasmus University, Rotterdam) and CMV3xHAbirA-NLS (gift from W. W. M. Pijnappel), respectively, using SpeI (blunted) and EcoRI restriction sites into pGEN vector (2 μm and TRP1 marker) (21) in front of the *PGK1* promoter, BirA, and *CYC1* terminator sequences from pGEN-BirA-NLS (22). For the LexA expression plasmids, we used a LexA derivative with the SV40 NLS fused to its C-terminal end. Using the BamHI restriction site of pLexA-NLS we inserted double-stranded oligonucleotides representing the Avitag or Biotag sequences which were flanked by BamHI restriction sites. This resulted in the pLexA-NLS-Avitag and pLexA-NLS-Biotag plasmids, respectively. For the pLexA-NLS-TAP plasmid, we inserted the PCR product of TAP sequences flanked by BamHI restriction sites into the corresponding restriction site of pLexA-NLS.

**Figure 1.** Comparison of biotinylation efficiencies of LexA proteins by BirA or BirA-NLS. (A) LexA-Avitag or LexA-Biotag proteins were coexpressed with either BirA or BirA-NLS proteins as indicated. Biotinylated proteins were detected with SA-HRP as the primary detection agent. Endogenously biotinylated proteins are indicated by the asterisk. (B) Yeast cells were grown in the presence of increasing concentrations of biotin. Protein extracts were incubated with streptavidin (SA) as indicated, and immunoblots were developed with antibodies recognizing LexA protein.

PUG6-Myc-C-Avitag was constructed by insertion of a fragment covering the Avitag and ADH1 terminator sequence of pLexA-NLS-Avitag. The fragment was isolated using SmaI and EcoRV restriction sites, gel purified and inserted into the EcoRV site of pUG6. Next, a double-stranded oligonucleotide representing the myc-tag sequence and flanked by BamHI restriction sites was inserted into the corresponding site of pUG6-C-Avitag. For pUG-Myc-N-Avitag construction, a double-stranded oligonucleotide representing the N-Avitag sequence flanked by a SalI restriction site was inserted into

<table>
<thead>
<tr>
<th>Yeast strains used</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>W303-1B MATa ade2-1 ura3-1 his3-11,15 leu2-3,112 trp1-1 can1-100</td>
<td>(35)</td>
</tr>
<tr>
<td>FWY002 Isogenic to W303-1B except RPB3::Myc-C-Avitag::KANA</td>
<td>This paper</td>
</tr>
<tr>
<td>FWY003 Isogenic to W303-1B except RPB3::TAP::URA3</td>
<td>This paper</td>
</tr>
<tr>
<td>EGY48 MATa his3 trp1 ura2::LexA6Op::LEU2</td>
<td>(17)</td>
</tr>
</tbody>
</table>
the SalI and PvuII site of pUG6 to obtain pUG-N-Avitag. Next, a double-stranded oligonucleotide representing the myc epitope sequence and flanked by BamHI restriction sites was inserted in the corresponding site of pUG6-N-Avitag to obtain the pUG6-Myc-N-Avitag plasmid. All oligonucleotide sequences are available upon request.

**Streptavidin gel-shift assay**

Proteins were extracted as described previously (23). After boiling in sample buffer, the extract was briefly incubated with or without 15 μg streptavidin before loading onto SDS–PAGE gel, which was followed by immunoblot detection using anti-LexA antibody (Santa Cruz) for LexA-Avitag, and anti-myc antibody (clone 4A6; Upstate) for Rpb3-Avitag.

**Chromatin immunoprecipitation**

Chromatin was isolated essentially as described previously (24). Yeast cells were grown in synthetic complete (SC) medium containing 2% glucose lacking histidine and tryptophan for the LexA experiments or lacking histidine for the Rpb3 experiments. To a 100 ml culture of OD 0.5–0.6, 1% formaldehde was added and incubated at room temperature for 20 min before chromatin was extracted. Cells were disrupted by vortexing five times for 2 min (Vortex Genie) in 1 ml of FA lysis buffer (50 mM HEPES–KOH, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate and 0.1% SDS), after which they were centrifuged to remove debris and unbound proteins (100 000 g for 20 min at 4°C). The supernatant was removed and the pellet was resuspended in FA lysis buffer. After 20 min of washing, the pellet was centrifuged again, resuspended in FA lysis buffer and sonicated for 7 min on 30 s on/off cycles at medium setting (Bioruptor, Diagenode) to produce an average fragment size of 400 bp.

Protein–DNA complexes were isolated by immunoprecipitation for 1 h and 45 min at room temperature from 200 μl of chromatin extract. Complexes were washed twice in FA lysis buffer, after which the sample was split in two. One part was washed twice with FA lysis buffer containing 0.5 M NaCl, twice in 10 mM Tris–HCl, pH 8.0, 0.25 M LiCl, 1 mM EDTA, 0.5% Nonidet P-40, 0.5% Na-deoxycholate, and once in 10 mM Tris–HCl, pH 8.0 and 1 mM EDTA. The other part was washed twice with 0.5 M LiCl, 1 mM EDTA, 1% Nonidet P-40 and 1% Na-deoxycholate, and three times with 10 mM Tris–HCl, pH 8.0, 1 mM EDTA and 3% SDS. Rpb1 was immunoprecipitated with 10 μg anti-RNA pol II (8WG16) mouse monoclonal antibodies coupled to 50 μl protein

---

Figure 2. ChIP analysis of LexA binding to chromosomal LexOp sequences. Chromatin was extracted from crosslinked EGY48 cells expressing the indicated LexA proteins. Captured protein–DNA complexes were washed under standard or stringent conditions (containing 3% SDS). (A and B) ChIP analysis of LexA-Avitag and LexA-TAP binding. The bound material was analyzed by LexAOp and POL1 primer sets, and the signals are expressed relative to the input. Please note the difference in scale of the specific signal to the left and the non-binding control to the right of the figure. (C) Similar to (A and B) except that enrichment over the POL1 control is indicated. (D) Control ChIP analysis of LexA-Avitag proteins in the absence of BirA expression (left columns) and of non-tagged LexA protein in the presence of BirA expression (right columns). ChIP analysis was performed as shown in (A) and the signals are expressed relative to the POL1 control sequences.
G–agarose beads (Roche). TAP-tagged and Avitag-tagged proteins were immunoprecipitated with 50 ml immunoglobulin G Sepharose 6 fast flow beads (Amersham) or 60 ml Dynabeads M-280 Streptavidin (Dynal). Immunoprecipitated chromatin was eluted by incubating two times for 10 min at 65°C in 10 mM Tris–HCl, pH 7.5, 1 mM EDTA and 1% SDS. The formaldehyde crosslinks were reversed by incubating the eluate at 65°C overnight in the presence of Ribonuclease A (75 µg/ml) and 0.5 M NaCl. Elution from Streptavidin beads and reverse crosslink reaction were done overnight at 65°C in 10 mM Tris–HCl, pH 7.5, 1 mM EDTA, 1% SDS, 0.5 NaCl and Ribonuclease A (75 µg/ml). Next, the sample was incubated with 100 µg Proteinase K for 2 h at 37°C, after which DNA was purified by using QIAquick DNA cleanup system.

Figure 3. Tools for gene tagging by chromosomal integration. (A) Plasmids for PCR-based insertion of Avitag sequences at the C- and N- terminus of ORFs. The position of the upstream and downstream primers described in (B) are indicated. (B) Universal primer sequences for the tagging cassettes depicted in (A). To each primer 50 nt of yeast genomic sequence flanking the integration site should be fused. The reading frames are indicated for the appropriate primer sequences. (C) Plasmid map of pRS313-BirA-NLS. This represents a low-copy BirA-NLS expression plasmid bearing a HIS3 marker gene. The BirA-NLS expression cassette has also been inserted in low-copy plasmids bearing TRP1 (pRS314-BirA-NLS), LEU2 (pRS315-BirA-NLS), URA3 (pRS316-BirA-NLS).
(Qiagen), and analyzed by real-time PCR using a Chromo4 real-time detector system (MJ Research) and SYBR green mix consisting of 0.2 μM of each primer, 0.2 mM of each dNTP, 1:800000 dilution of SYBR green I (Molecular probes), 1x PCR Gold buffer, 4 mM MgCl₂ and 0.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems). All chromatin samples were analyzed by the same PCR program (10 min/95°C, and 40 cycles of 95°C/10 s, 58°C/10 s, 72°C/10 s) and checked for product specificity by DNA melting curve analysis. Immunoprecipitation efficiency for each sample was calculated from the amount of immunoprecipitated material divided over the input chromatin. In short, from a dilution series (1:10, 1:100, 1:1000 and 1:100000 dilutions) of input chromatin an exponential trendline was computed that was used for determination of the percentage of immunoprecipitated material. The mean and SD of each ChIP were calculated from at least three independent experiments, and two independent amplifications. Except for (Figure 2A and B), the input normalized ChIPs are presented as fold occupancy over either POL1 gene (Figure 2C and D) or HMR silent mating-type locus (Figure 5A–D) control regions. All oligonucleotide sequences of the primers used in ChIP detection are available upon request.

RESULTS AND DISCUSSION

To develop biotin-tagging in yeast we first constructed two episomal yeast high-copy expression plasmids for the BirA enzyme. Both the BirA and BirA-NLS constructs contain three copies of the HA-epitope at its N-termini. In addition, we added a nuclear localization signal (NLS) from the SV40 large T-antigen to obtain the BirA-NLS construct. In vivo biotinylation activity was tested in the EGY48 strain by coexpression of the DNA-binding domain of LexA fused to an NLS and either the Avitag (GLNDIFEAQEK; in bold the biotin acceptor side) (25) or the Biotag (ASSLRQILDSEQMEWRSNAGGS) (14) as biotin acceptors. Biotinylated proteins were detected using streptavidin–horseradish peroxidase (SA-HRP). Both LexA-Biotag and LexA-Avitag proteins were biotinylated dependent on BirA expression, but the NLS-derivative seemed to result in a slightly more efficient biotinylation of the tagged LexA (Figure 1A). Besides LexA, several endogenous yeast biotinylated proteins (most likely Acc1p, Arc1p, Hfa1p, Pyc1p and Pyc2p) were detected (26–29). This should not complicate ChIP assays as these proteins are not expected to crosslink to chromosomal DNA, because they localize to the cytoplasm or to the mitochondria (30). We analyzed biotinylation efficiency by incubating protein extracts with streptavidin prior to loading on a SDS–polyacrylamide (PAA) gel. Biotinylated proteins will bind streptavidin and display a reduced mobility in SDS–PAA gels (16). When cells were grown in SC-medium (containing 8 nM biotin), the LexA-Avitag fusion protein was biotinylated to a higher extent than the LexA-Biotag version (Figure 1B, compare lanes 4 and 10). Increasing biotin concentrations to 250 nM in the growth medium resulted in complete biotinylation of both proteins (Figure 1B, lanes 5 and 11). Previous biochemical analysis indicated that concentrations of 50 mM and higher are required for maximal BirA activity (31). Together, this indicates that the biotin substrate was limiting and that BirA expression was sufficient to achieve complete biotinylation of the target protein.

The EGY48 strain allows ChIP analyses of LexA proteins as it carries multiple LexA operator (LexAOp) sequences upstream of the LEU2 locus. To this end we isolated formaldehyde-crosslinked chromatin from EGY48 cells expressing the LexA-Avitag and BirA-NLS proteins. After incubation of chromatin extracts with streptavidin proteins we compared two different washing conditions (Materials and Methods). The ChIPs were analyzed by real-time PCR using primers for the LexAOp sequence and the POL1 open reading frame (ORF) as a non-binding control (32). About 20-fold enrichment of LexAOp signals over the POL1 control was observed with standard wash conditions (Figure 2A). No loss of LexAOp signal occurred during stringent washing (Figure 2A). However, the POL1 background signal was reduced (~5-fold) improving signal-to-noise ratios to above 100-fold (Figure 2A and C). For comparison we prepared crosslinked chromatin from EGY48 cells expressing a TAP-tagged LexA. In this case we observed a 10-fold enrichment of LexAOp signals (Figure 2B and C). As expected, stringent wash conditions resulted in loss of both signals (Figure 2B and C). As controls for specificity, we analyzed cells expressing BirA and untagged LexA or only LexA-Avitag. This yielded very low signals and no enrichment of LexAOp sequences (Figure 2D). In conclusion, using LexA as a model DNA-binding protein we found that the BirA-based in vivo biotinylation system performs remarkably well in ChIP experiments.

Next, we created generic tagging cassettes allowing chromosomal insertion of the Avitag (Figure 3A). The pUG6-Myc-C-Avitag plasmid was used to modify the Rpb3 gene, which encodes an essential subunit of RNA polymerase II. As shown in Figure 4, chromosomally expressed Rpb3-Avitag was efficiently biotinylated in the presence of a low-copy BirA-NLS expression plasmid (Figure 3C).

The Rpb3-Avitag strain was used to examine association of pol II to transcribed regions of genes expressed at high (PGK1 and PYK1), moderate (TUB1) or low levels (POL1) (33). The Rpb3-Avitag was compared with TAP-tagged Rpb3 and the 8WG16 antibody recognizing the C-terminal domain

![Figure 4](image-url)

Figure 4. Biotinylation efficiency of Rpb3-Avitag. Protein extract of Rpb3-Avitag strain was incubated with SA as indicated. Rpb3p was detected by immunoblotting using antibodies specific for the co-integrated Myc epitope.
of the Rpb1 subunit of pol II. The 8WG16 antibody is known to behave particularly well in ChIP assays, because it recognizes a heptad sequence (YSPTSPS), which is present in 26 copies in yeast. The transcriptionally silent HMR locus was used as a non-binding control. Under standard wash conditions comparable enrichment of PGK1 and PYK1 was observed (20– to 30-fold) with Rpb3-Avitag, Rpb3-TAP or 8WG16 (Figure 5A and B). However, stringent washing boosted signal-to-noise ratios for the Rpb3-Avitag to approximately 100-fold enrichments. Rpb3 binding could not be observed on TUB1 or POL1 with Rpb3-TAP or with 8WG16 antibodies. In contrast, the Rpb3-Avitag could be detected readily on the TUB1 and POL1 ORFs (Figure 5C and D). The ChIP signals with the Rpb3-Avitag were specific as no enrichment was observed in the absence of BirA-NLS expression. The different binding ratios detected with Rpb3-Avitag corresponded well to reported differences in expression levels of these genes tested (33).

Taken together, we have developed an efficient approach for affinity tagging of yeast proteins by in vivo biotinylation. This relies on coexpression of the bacterial BirA biotin ligase. Both the Avitag and Biotag sequences allow quantitative biotinylation of proteins in vivo. In comparison with other affinity tags (TAP, 170 residues; 3xHA, 36 residues; 13xMyc, 156 residues) the 15-amino acid Avitag is small, reducing the risk of interference with protein function.

We have now successfully modified multiple essential transcription factor genes using the Avitag. So far, no interference with protein function has been observed as determined by phenotypic analysis (F.J. van Werven and H.Th.M. Timmers, unpublished data). In all cases, quantitative biotinylation of the tagged protein was observed. In some cases (C-terminally tagged TBP) we noted that the biotin moiety may not be accessible in crosslinked protein–DNA complexes (F.J. van Werven and H.Th.M. Timmers, unpublished data). However, this has also been observed with other tagging approaches.

The applications of in vivo protein biotinylation are multiple. For ChIP assays we showed that capturing biotinylated transcription factors by streptavidin-coated beads allows stringent washing conditions. This resulted in high signal-to-noise ratios and improved detection of crosslinked protein–DNA complexes. Application of in vivo biotinylation to genome-wide localization studies should improve the quality of datasets. Combining in vivo biotinylation with another generic affinity tags would simplify sequential ChIP approaches to detect colocalization of proteins on DNA. In the mammalian system it has been shown that the biotinylation tag can also be used for complex purification and protein identification by mass spectrometry (5,14). In vivo biotinylation of yeast proteins can also form the basis for capturing multi-subunit complexes from relatively crude protein mixtures on streptavidin surfaces to enable structural studies by electron microscopy (34).

**ACKNOWLEDGEMENTS**

We thank J. Strouboulis for the CMV3xHAbirA, W.W.M. Pijnappel for the MV3xHAbirA NLS and A. Vosselman for the pLexA-NLS plasmids. We gratefully acknowledge our lab members for discussions and
K.W. Mulder, G.S. Winkler and F.C.P. Holstege for critical reading of this manuscript. This work was supported by grants 805.47.080 from the Earth and Life Sciences council (NWO-ALW) and 710.15.001 from the Chemical Sciences council (NWO-CW) of The Netherlands Organization for Scientific Research (NWO) and by The Netherlands Proteome Centre (NPC). Funding to pay the Open Access publication charges for this article was provided by NWO-ALW.

Conflict of interest statement. None declared.

REFERENCES