Shining a New Light on RNA-Protein Interactions

Jeffry L. Corden[1,*]
[1]Department of Molecular biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe St., Baltimore, MD 21205, USA
*Correspondence: jcorden@jhmi.edu
DOI 10.1016/j.chembiol.2010.04.003

PAR-CLIP is a new technique for transcriptome-wide mapping of binding sites for RNA-binding proteins. In this technique, a photoactivatable nucleoside is incorporated into RNA and crosslinked to proteins in vivo with UV light. The crosslinked nucleoside causes a specific base change during processing for deep sequencing, allowing the precise identification of binding sites.

Eukaryotic transcripts can be differentially processed, transported, and translated in a myriad of combinations that generate complexity at the protein, cell, and organism level. Such posttranscriptional regulation of eukaryotic gene expression is orchestrated by hundreds of sequence-specific RNA-binding proteins (RBPs), and identifying transcriptome-wide interactions between RBPs and their target transcripts is essential in understanding the eukaryotic regulatory networks (Licatalosi and Darnell, 2010). In a recent paper published in *Cell*, Hafner et al. (2010) describe a robust new technique for identifying RBP binding sites in vivo.

A variety of powerful techniques have been used to identify RNA sequences bound by RBPs, but until now, all have had shortcomings. For example, in vitro RNA selection (SELEX) using purified Nova, a brain-specific RBP, yielded the consensus binding motif UCAY (Jensen et al., 2000), but such short sequences have limited value in predicting in vivo binding sites. Immunopurification of RBPs with associated RNA (Brooks and Rigby, 2000) has been used with some success but, under low stringency conditions such as low salt and the absence of detergent, high levels of nonspecific binding are observed, while increased stringency allows reassociation of RBPs after cell lysis, leading to artificial RNA-protein complexes. Treating cells with formaldehyde (Niranjanakumari et al., 2002) preserves in vivo RBP-RNA complexes but also results in protein-protein crosslinks that yield large ribonucleoprotein complexes that may contain RNA not directly or even functionally associated with the protein of interest.

Darnell and colleagues have used 254 nm UV irradiation of growing cells to covalently crosslink RNA to RBPs that are in direct contact in vivo. This approach (CLIP, in vivo crosslinking and immunoprecipitation) (Ule et al., 2003, 2005) has the advantage that RNA is covalently bound to the protein of interest but proteins are not crosslinked to each other. The covalent interaction allows stringent purification of the crosslinked RNA fragments. After trimming the attached RNA with ribonuclease to 70-100 nt, RNA-protein complexes can be purified using a variety of approaches, including SDS gel electrophoresis. Alternatively, the protein can be tagged and purified together with crosslinked RNA under denaturing conditions (Granneman et al., 2009). The purified RNA-protein complexes are digested with protease and the RNA fragments are ligated to 5′ and 3′ adaptors prior to reverse transcription and PCR amplification. High throughput sequencing results in a set of sequence reads representing RNA sequences bound to the RBP in vivo (see Figure 1 for the general outline of the workflow).

One drawback to this technique is the relatively inefficient formation of covalent bonds between RNA and protein. More intense 254 nm UV might improve crosslinking but can result in DNA damage and the induction of damage repair programs, which can include expression of novel RBPs (Sheikh, et al., 1997). Thus, the pattern of RBP binding may be influenced by the crosslinking protocol.

Tuschl and colleagues now report the next generation of high throughput in vivo RNA cross-linking. In their PAR-CLIP (photoactivatable ribonucleside-enhanced CLIP) approach, cells are grown in the presence of 4-thiouridine (4SU). This ribonucleoside is readily incorporated into RNA in vivo with no apparent alteration in the levels of transcripts as determined by microarray (Hafner et al., 2010). 4SU substituted RNA is 100–1,000-fold more efficiently crosslinked to protein than unsubstituted RNA, and crosslinking occurs at a longer wavelength (365nm) than used in previous CLIP approaches (Figure 1). Thus, PAR-CLIP has the advantage that crosslinking times can be much shorter and cells are less prone to UV damage.

Most importantly, Hafner et al. (2010) show that 4SU crosslinked to protein yields a signature "mark." Reverse transcriptase misincorporates G opposite the crosslinked 4SU base, yielding a characteristic T to C transition in the sequence corresponding to the bound RNA. This additional piece of information facilitates the identification of consensus RNA-binding sequences among clusters of bound RNA fragments. One limitation of 4SU crosslinking is the possible paucity of U residues in close proximity to the binding site. Hafner et al. (2010) provide a remedy for this by demonstrating that 6-thioguanosine gives very similar results to 4SU, including analogous G to A transitions in the cDNA.

In a tour de force demonstration of the power of PAR-CLIP, Hafner et al. (2010) present the transcriptome-wide distribution of five pre-mRNA processing proteins and seven proteins that mediate miRNA-guided mRNA recognition. The processing proteins provide "proof of principle" data sets that confirm known interactions and provide a wealth of new binding sites. Of particular note is the demonstration that insulin-like growth factor 2 mRNA-binding proteins recognize the consensus sequence CAUH (where H = U,C,A). Unambiguous identification of this low-information binding site was greatly facilitated by focusing on sequences close to the T to C transition in the sequence tags. Hafner et al. (2010) also applied

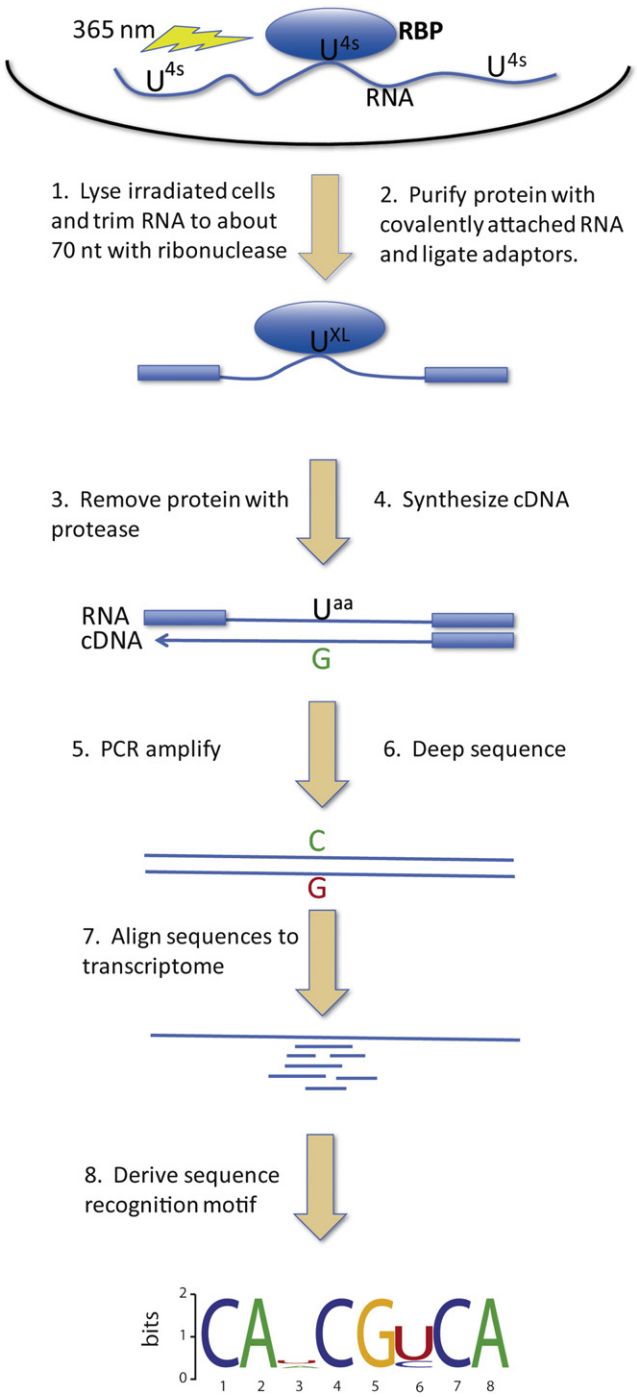

Figure 1. Illustration of the PAR-CLIP Technique

4SU-containing RNAs are crosslinked to RNAs in growing cells. After lysing cells, the RNA is trimmed with ribonuclease T1 and the protein of interest is purified by immunoprecipitation and SDS-PAGE. The protein-RNA complex is then cut out of the gel, and RNA is liberated by protease digestion. RNA adaptors are added to each end and the RNA is then converted to DNA with reverse transcriptase. At this step, the reverse transcriptase misincorporates G opposite the 4SU-amino acid adduct. PCR-amplified cDNA is subjected to deep sequencing and the resulting sequence reads are aligned to the genome. By comparing different clusters of sequence reads, a consensus binding site can be obtained.

PAR-CLIP to argonaute proteins and derived a comprehensive miRNA interaction map for human HEK293 cells. Interestingly, about 50% of the miRNA binding sites are located in coding regions and about 6% of these interactions require bulges or mismatches in the miRNA-mRNA seed-pairing region. Again, this analysis benefitted from the observation of prominent crosslinking sites in the mRNA targets 1-2 nt downstream of sequences complementary to the miRNA seed sequence. This would place the crosslink near the center of the AGO-miRNA-target RNA complex.

PAR-CLIP is more than an incremental improvement of existing technology. The opportunity to control the presence of the photoactivatable nucleoside opens up a range of experiments not possible with direct UV irradiation. For example, pulse chase experiments with 4SU could help unravel the temporal order of different RNA processing steps. Another potential application of PAR-CLIP is in cell-type specific crosslinking. Metazoan cells lack the enzyme uracil phosphoribosyltransferase (URPT) and are thus unable to salvage uracil. Hafner et al. (2010) skirt this limitation by labeling with 4-thiouridine. However, if cells are provided with the URPT gene from *Toxoplasma gondii*, they are able to incorporate 4-thiouracil into RNA (Cleary, 2008). Expressing URPT from a *Drosophila* neuronal-specific promoter, Miller et al. (2009) were able to use thio-biotin coupling to isolate cell-type specific RNAs from animals that were fed 4-thiouracil. This same approach could certainly be used to specifically crosslink RBPs in specific cell types. PAR-CLIP used in this way could lead to the elucidation of RNA regulatory networks in complex tissues. Given this new tool, our understanding of RNA regulatory networks is bound to expand at an increasing CLIP.

REFERENCES

Brooks, S.A., and Rigby, W.F. (2000). Nucleic Acids Res. *28*, E49.

Cleary, M.D. (2008). Methods Enzymol. *448*, 379–406.

Granneman, S., Kudla, G., Petfalski, E., and Tollervey, D. (2009). Proc. Natl. Acad. Sci. USA *106*, 9613–9618.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., et al. (2010). Cell *141*, 129–141.

Jensen, K.B., Musunuru, K., Lewis, H.A., Burley, S.K., and Darnell, R.B. (2000). Proc. Natl. Acad. Sci. USA *97*, 5740–5745.

Licatalosi, D.D., and Darnell, R.B. (2010). Nat. Rev. Genet. *11*, 75–87.

Miller, M.R., Robinson, K.J., Cleary, M.D., and Doe, C.Q. (2009). Nat. Methods *6*, 439–441.

Niranjanakumari, S., Lasda, E., Brazas, R., and Garcia-Blanco, M.A. (2002). Methods *26*, 182–190.

Sheikh, M.S., Carrier, F., Papathanasiou, M.A., Hollander, M.C., Zhan, Q., Yu, K., and Fornace, A.J., Jr. (1997). J. Biol. Chem. *272*, 26720–26726.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). Science *302*, 1212–1215.

Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). Methods *37*, 376–386.

SNAP-Shots of Hydrogen Peroxide in Cells

Amy E. Palmer[1,*] and Philip J. Dittmer[1]
[1]Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309, USA
*Correspondence: Amy.palmer@colorado.edu
DOI 10.1016/j.chembiol.2010.04.002

Long regarded as a toxic byproduct, hydrogen peroxide is increasingly recognized as an important cellular signal. Efforts at defining the spatiotemporal nature of hydrogen peroxide production recently got a boost by the development of a series of organelle-targeted fluorescent probes by Srikun et al. (2010).

Reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2) are generally thought of as unwanted and detrimental byproducts, produced accidentally as a result of cellular metabolism. There are many good reasons for this bad rap. After all, organisms have evolved exquisite mechanisms for detoxifying ROS through enzymes such as catalase, superoxide dismutase, glutathione peroxidases, and peroxiredoxins, which serve to limit the buildup of ROS and are generally thought of as cellular protective agents. Without question, excessive generation of ROS such as H_2O_2 are toxic to cells, leading to oxidative stress, apoptosis or necrosis, and cell senescence (Finkel and Holbrook, 2000). However, H_2O_2 can be produced deliberately and in a regulated manner by the NADPH oxidase (NOX) and Dual oxidase (Duox) family of enzymes (Bedard and Krause, 2007). A classic example is the oxidative burst utilized by professional phagocytes such as macrophage and neutrophils to protect a host against invasion by pathogens. While this is a specialized example, NOX family enzymes exist in a wide range of nonphagocytic cells, suggesting that deliberate production of H_2O_2 plays a fundamental role in cell biology (Bedard and Krause, 2007).

Mounting evidence suggests that H_2O_2, produced by NOX extracellularly, can act in both an autocrine and paracrine fashion (Figure 1). For autocrine signal transduction, H_2O_2 is widely becoming recognized as a bona fide second messenger. Bursts of H_2O_2 are produced in response to a variety of stimuli, including growth factors, cytokines, hormones, calcium, and neurotransmitters (Bedard and Krause, 2007). The primary action of H_2O_2 as a signaling molecule is the oxidation of proteins to modulate their function. H_2O_2 can oxidize cysteine residues to sulfenic acid (Cys-S-OH) that can be readily reversed by cellular reductants such as glutathione and thioredoxin. However, H_2O_2 does not specifically oxidize any Cys-containing protein because the Cys must be deprotonated at physiological pH, and hence have a low pKa. Thus, H_2O_2 acts on select sites, including those found in a number of transcription factors and protein tyrosine phosphatases (Rhee, 2006). H_2O_2 can also modify histidine and methionine residues. By modulating the function of intracellular protein targets, H_2O_2 has been found to affect gene transcription, cell proliferation, differentiation, metabolism, and migration (Bedard and Krause, 2007). Lastly, Niethammer et al. (2009) recently provided convincing evidence that H_2O_2 produced by Duox serves as a paracrine signal for recruitment of leukocytes to wounds in the vertebrate zebrafish.

A paradigm is emerging that when and where H_2O_2 is produced has a profound impact on downstream cellular consequences. The ability to monitor the spatiotemporal nature of H_2O_2 production and clearance in real time would be an invaluable tool in elucidating H_2O_2 biology. Toward this end, Srikun et al. (2010) have now generated a family of H_2O_2 sensitive fluorescent probes targeted to various cellular organelles. These localized probes should help provide insight into the spatial heterogeneity of H_2O_2 signaling.

To generate organelle-targeted probes, Srikun et al. (2010) combined the power of a small molecule fluorescent indicator, namely the Peroxy Green probe previously developed by the same research group (Miller et al., 2007), with the genetic targetability of the SNAP-tag technology pioneered by Keppler et al. (2004) (Figure 1). Peroxy Green consists of a boronate-modified Tokyo Green fluorophore; reaction with H_2O_2 liberates the boronate, resulting in an increase in fluorescence and hence a ''turn-on'' signal. To be compatible with the SNAP-tag technology, this basic probe was modified to incorporate a moiety that could serve as a substrate for AGT (O^6-alkylguanine-DNA alkyltransferase). Two different Peroxy Green probes were synthesized; one conjugated to the traditional benzylguanine substrate (referred to by the authors as SPG1) and another linked to a benzyl-2-chloro-6-aminopyrimidine substrate

