CCR frontiers IN SCIENCE

May 2004, Volume 3 Published by the National Cancer Institute's Center for Cancer Research, http://ccr.cancer.gov

Contents

NATIONAL CANCER INSTITUTE
NATIONAL INSTITUTES OF HEALTH
DEPARTMENT OF HEALTH AND HUMAN SERVICES

■ BIOTECHNOLOGY RESOURCES

RNA Silencing: A Mechanism, a Research Tool, and Perhaps More

Only 12 years have passed since an unexpected observation in petunias suggested the presence of a gene-silencing mechanism that acts at a posttranscriptional level, and it has been less than 3 years since a similar process was identified in mammalian cells. In this short time, these observations have profoundly changed our understanding of the role of RNA as an epigenetic regulator of gene expression, and we have seen the development of highly effective functional genomics tools that exploit this mechanism. CCR is establishing a new research and development section, the Gene Silencing Section, within the Office of Science and Technology Partnerships to develop and apply technologies based on RNA-mediated silencing of gene expression.

What Is RNA Silencing?

RNA silencing is a collective term for a group of related mechanisms, observed in nearly all organisms studied to date, that require a double-stranded RNA (dsRNA) structure at some point in the pathway to mediate a downregulation in gene expression through interaction of small RNA molecules (20–25 nucleotides) with the target RNA transcript. The term PTGS has generally been used to indicate epigenetic posttranscriptional gene silencing in plants and fungi, whereas the term RNA interference (RNAi) has usually been applied to a similar epigenetic gene-silencing effect observed in invertebrates and mammals. In somatic mammalian cells, RNAi is triggered by RNA duplexes of 21–23 nucleotides, called small interfering RNAs or siRNAs (Caplen N et al. *Proc Natl Acad Sci U S A* 98: 9742–7, 2001; Elbashir S et al. *Nature* 411: 494–8, 2001). SiRNAs can be introduced into mammalian cells in a number of forms, including duplexes of chemically synthesized RNA oligonucleotides and single-stranded RNA transcripts that make a hairpin structure, in which the stem

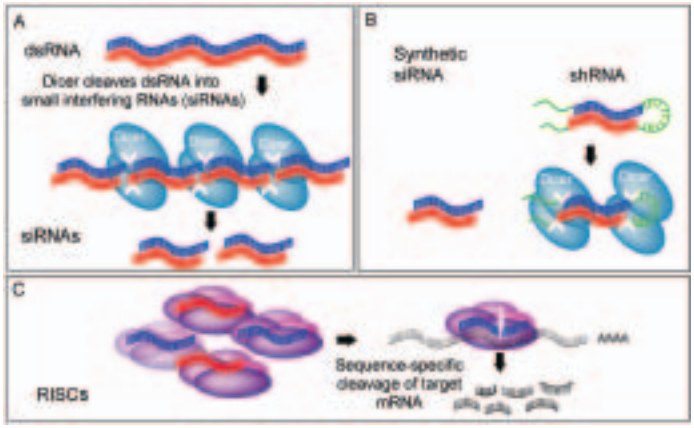

Figure 1. Schematic representations of aspects of the RNAi pathway. *A:* In invertebrates, double-stranded RNA (dsRNA) molecules are processed by Dicer (an RNase III enzyme) to generate small interfering RNAs (siRNAs). *B:* In mammalian cells, synthetic siRNAs or short hairpin RNAs (shRNAs) must be used to minimize the triggering of non-specific dsRNA responses. ShRNAs introduced as single-stranded RNA transcripts or expressed intracellularly from plasmid or viral vectors are processed by Dicer to generate siRNAs. *C:* The siRNA is unwound and incorporated into the RNA-induced silencing complex (RISC) that contains a number of proteins including several Argonaute family members. The siRNA single strand within RISC acts to guide the complex to the target RNA through sequence alignment. RISC then facilitates interaction with, and cleavage of, the target transcript, which is usually mRNA.

region corresponds to the siRNA sequence. Short hairpin RNAs (shRNAs), as these are known, have the advantage that they can be expressed from plasmid and viral vectors (Figure 1). RNA silencing has been implicated in a number of processes linked to the inhibition of gene expression, including heterochromatin formation, silencing of transposable elements, and as an antiviral response in plants, as well as in the posttranscriptional regulation of a number of endogenous genes.

RNAi as a Functional Genomics Tool

In mammalian cells, RNAi knockdowns have been generated in primary cells and cell lines and *in vivo* in mice. The inhibition of gene expression can be transient or long term depending on the form of the RNA trigger. Transcripts encoding proteins involved in a wide range of cellular processes have been targeted, including many relevant to the cancer process, and RNAi has also been used to block both the infection and replication of viruses, including HIV-1. RNAi has been rapidly adopted as a functional genomics tool in a wide range of species and has been adapted for high-throughput analysis of gene function in *Caenorhabditis elegans* and *Drosophila*.

RNAi knockdowns can be used to do the following:

- Improve our understanding of protein function by enabling studies of the effects of RNAi against the gene encoding the protein of interest;
- Translate bioinformatic, positional cloning, and cDNA array data, as only the transcript sequence is required to enable the generation of an RNAi effector molecule;
- Delineate pathways, for example, through the analysis of multiple genes simultaneously or consecutively—the effect of modulating expression by inducing different degrees of inhibition can also be studied;
- Generate new model systems, both cell lines and transgenic animals;
- Validate drug targets and identify drug-protein interactions that may enhance drug activity; and
- Identify new drug targets for which small molecules can be developed (or the trigger of RNAi itself can be developed).

CCR's RNAi Initiative

The Gene Silencing Section is the primary contact for the new RNAi initiative established within CCR. The RNAi initiative will focus on the following:

- Generation of key RNAi-based resources and standardized RNAi protocols including the development of an infrastructure for high-throughput RNAi analysis;
- Development of a phased program validating chemically synthesized siRNAs for RNAi against cancer-associated genes;
- Assessment and application of shRNA libraries;
- Generation of cell lines and mice with defined RNAi knockdowns; and
- Assessment of novel RNAi-based technologies with a potential for clinical application.

The RNAi initiative is building an infrastructure for high-throughput RNAi screening by using siRNA/shRNAs in cell culture. Initial studies will be based on 96- and 384-well–based assay systems, but we will also build on our recent experience developing novel RNAi microarray–based platforms (Mousses S et al. *Genome Res* 13: 2341–7, 2003). Collaborative research projects are under development with a number of CCR PIs and other CCR programs and initiatives. We welcome inquiries from CCR investigators who require general advice on the use of RNAi as a functional genomics tool or who wish to develop collaborative research projects using RNAi-based technologies. These might include projects that aim to validate or improve current RNAi methods, develop new applications for RNAi, and analyze the role of the RNAi mechanism itself as an epigenetic silencing process.

- **Natasha Caplen, PhD**
  Senior Scientist
  Gene Silencing Section
  Office of Science and Technology Partnerships
  Office of the Director
  NCI-Bethesda, Bldg. 37/Rm 3128B
  Tel: 301-451-1844
  Fax: 301-480-1717
  ncaplen@mail.nih.gov

TRANSLATIONAL RESEARCH

Harvesting the Blood's Secret Diagnostic Code

Liotta LA, Ferrari M, and Petricoin E. Written in blood. *Nature* 425: 905, 2003.

There is a great need to discover novel cancer biomarkers and translate them into routine clinical use. Conventional differential display technologies (gene array, two-dimensional polyacrylamide gel electrophoresis, etc.) followed by antibody production, validation, and enzyme-linked immunosorbent assay is an inherently costly and laborious process with long cycle times between discovery and clinical implementation. The paucity of new U.S. Food and Drug Administration–approved analytes for cancer detection is driving investigators to break out of this cycle. Mass spectroscopy serum proteomic pattern analysis can sort through tens of thousands of potential biomarkers in the time it takes you to read this sentence.

The recognition that biomarkers can be amplified by association with circulating carrier protein will likely lead to "designer" analyte harvesters.

Mass spectroscopy serum proteomic pattern diagnostics is a rapidly expanding field. The general hypothesis is that patterns of multiple low-molecular-mass (LMW) biomarkers in the blood specifically reflect the underlying pathologic state of an organ—even at a distance. Moreover, this pattern of features can achieve a higher accuracy and specificity than any single biomarker alone. Single analytes may show some discriminatory power for cancer detection in small study sets, but it is unclear whether any single analyte can detect cancer with high specificity across large heterogeneous populations. Mathematically, a pattern of multiple biomarkers will contain a higher level of discriminatory information, particularly across large heterogeneous patient populations and for complex multistage diseases such as cancer.

Because cancer cells themselves are deranged host cells, we may never find a true cancer-specific molecule. On the other hand, the complex proteomic signature of the tumor host microenvironment may be unique and may constitute a biomarker amplification cascade. The specificity of this unique microenvironment can be mirrored by a catalog of LMW proteins and peptides. These can include specific cleaved, or otherwise modified, proteins produced in enough abundance to be detected by current mass spectrometry platforms. Mass spectroscopy can discriminate clipped or modified versions of molecules with extremely high speed and resolution. If the biomarker were a cleaved version of a larger abundant protein, it would be nearly impossible to generate antibodies that recognize the cleaved version and do not cross-react with the much more abundant parent species. Consequently, mass spectroscopy is attractive for biomarker discovery as well as for routine high-throughput testing.

Knowing the identity of the proteins comprising the discriminatory ions could provide insights concerning their source and relationship to the underlying pathology. In fact, investigators are using mass spectrometry and enrichment strategies to identify the entire LMW region of the proteome. This region contains thousands of whole proteins and fragments derived from every class of cellular compartment—from transcription factors to oncogenes to membrane receptors and channels. In the future, we should be able to generate the ion patterns and then go directly to a list of the underlying identities in a database.

Under the assumptions that the LMW biomarkers contain important diagnostic information and that protein biomarkers useful in disease detection are of very low abundance, the search for biomarkers usually begins with a separation step to remove the abundant "contaminating" high-molecular-mass (HMW) proteins such as albumin, thyroglobulin, and immunoglobulins so that the analysis can focus on the LMW region. Researchers in the biomarker field could examine the proportion of the LMW species bound to the HMW fraction of the serum or plasma proteome. From a physiologic perspective, free-phase LMW molecules (less than 30 kDa) should be rapidly cleared through the kidney. Such rapid physiologic excretion may greatly reduce the concentration of free-phase LMW species to a level below detection. In the face of the vast excess of HMW serum proteins, low-abundance LMW species may tend to bind large carrier proteins. The abundant HMW carrier proteins exist above the cutoff for kidney clearance and hence possess a half-life many orders of magnitude larger than that of small molecules. Circulating carrier proteins may thus become the reservoir for the accumulation and amplification of bound LMW biomarkers, just as association with albumin is known to extend the half-life of short-lived proteins introduced into the circulation.

Most investigators are now using or developing methods by which the higher abundance proteins above 30 kDa are specifically subtracted from native serum or plasma before analysis.

We now recognize that this convention will dramatically diminish the chances of finding the important low-abundance LMW disease biomarkers. The failure to appreciate the partitioning of biomarkers with circulating carrier proteins may explain the paucity of new diagnostic markers entering the development pipeline.

The recognition that biomarkers can be amplified by association with circulating carrier protein will likely lead to "designer" analyte harvesters. Protein or nanoparticle harvesters could be engineered to circulate and collect low-abundance LMW molecules emanating from specific diseased tissue. Such an approach could dramatically improve the sensitivity and power of serum diagnostics.

■ **Lance A. Liotta, MD, PhD**
Chief, Laboratory of Pathology
NCI-Bethesda, Bldg. 10/Rm. B1B41
Tel: 301-496-2035
Fax: 301-480-0853
lance@helix.nih.gov

Spatial Proximity as a Novel Determinant in the Formation of Chromosomal Translocations in Lymphomas

Roix JJ, McQueen PG, Munson PJ, Parada LA, and Misteli T. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat Genet* 34: 287-91, 2003.

Blood-related cancers and lymphomas account for approximately 10 percent of new cancer cases and associated deaths in the United States; this year alone these malignancies will account for 57,000 deaths nationwide. A hallmark of many types of lymphomas is the presence of translocated chromosomes. Why and how these disease-causing illegitimate fusions between chromosomes occur in normal cells is only poorly understood. We recently began to study the cell biological aspects of chromosome translocations by asking whether the relative location of potential translocation partners influences their likelihood of forming translocations.

To address this question, we developed tools to precisely map the position of a gene in the cell nucleus. We implemented semi-automated microscopy tools to visualize specific genes using fluorescent probes and, in collaboration with Philip McQueen, PhD, at the Center for Information Technology, established statistical analysis methods to quantitatively describe the spatial positioning of any particular gene inside the mammalian cell nucleus (Figure 1, *top*). We then applied these tools to study genes involved in Burkitt's lymphoma. This type of lymphoma is very frequently associated with translocations between the MYC and immunoglobulin heavy chain (IGH) loci, less frequently between the MYC and immunoglobulin light chain lambda (IGL) loci, and sporadically between the MYC and the immunoglobulin light chain kappa (IGK) loci.

We first determined the positions of the involved genes within the three-dimensional space of the cell nucleus. Very surprisingly, none of these genes was completely randomly distributed. For example, the IGL gene was often found toward the center of the nucleus, whereas the IGH gene was often positioned about halfway between the nuclear center and the periphery (Figure 1, *bottom*). This observation has been confirmed for a number of randomly selected genes in the mouse and human genomes, suggesting that non-random spatial positioning is a general feature of mammalian genes. What this non-random distribution means for regulation and normal gene function is currently under investigation.

The differential frequencies of translocation of MYC with its three partners in Burkitt's lymphoma provided us with an ideal model system to ask how the relative position of potential translocation partners affects their likelihood of forming a translocation. When we measured the distance between the translocation-prone MYC and IGH loci within the space of the cell nucleus, these two loci were more frequently closer to each other than would be expected based on their distribution patterns. Furthermore, MYC was somewhat more distant from its next-frequent translocation partner IGL and was only rarely near IGK or a control gene that has never been reported to translocate with MYC.

We have found a similar correlation between proximity and translocation frequency for several other lymphomas. For instance, 446 clinical lymphomas have been associated with translocations involving the *CCND1* oncogene and IGH loci, whereas only 52 cases have been reported to involve the *B-cell lymphoma 6* (*BCL6*) gene and IGH. Similar to the situation in Burkitt's lymphoma, we observed that the IGH locus was closer to its most frequent translocation partner *CCND1* than to the *BCL6* locus, with which it translocates less frequently. These findings demonstrate a correlation between the clinically observed frequency of translocation between two loci and their spatial, three-dimensional arrangement in the cell nucleus.

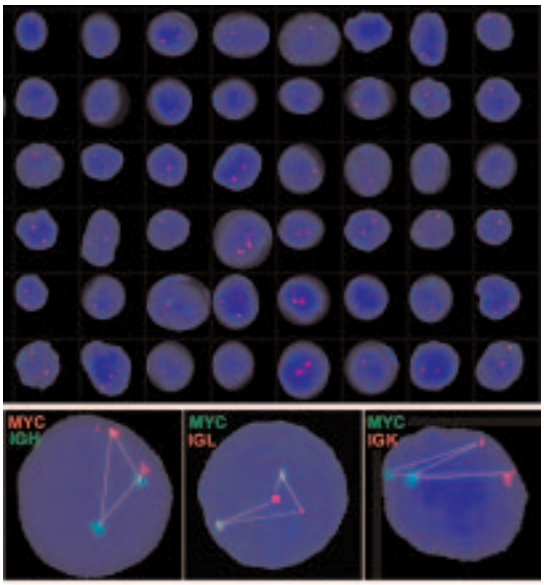

Figure 1. Mapping of gene loci in three-dimensional space. *Top:* The position of gene loci within the cell nucleus of B cells can be determined by fluorescence *in situ* hybridization and semi–high-throughput image analysis. Images of single cells are acquired using standard imaging methods, then combined into a "cell array" for efficient quantitative analysis by batch processing. Loci of interest are in red and green; the cell nucleus is in blue. Typically, two signals representing the two gene alleles are seen for both loci. *Bottom:* Distances between potential translocation partners are measured. Images of typical cell nuclei from a normal person show the primary translocation partners MYC and IGH (*left*) are frequently in spatial proximity. On average, variant IGL (*middle*) and IGK (*right*) translocation partner loci are found further from the MYC oncogene. A similar correlation between proximity and translocation frequency is observed for genes affected in other lymphomas.

We suggest that the spatial proximity of potential chromosome translocation partners is an important, previously unappreciated factor in the etiology of translocation formation. On the basis of observations in other systems, we believe the role of spatial positioning is not limited to lymphomas but might apply to many other tumor types in which translocations are frequent. Having established spatial positioning as a parameter in the

formation of cancerous translocations, we are now exploring the possibility of using spatial features of pre-cancerous genomes in diagnosis. To this end, we are combining the microscopy and analytical tools formulated in our laboratory with pattern recognition learning programs to enhance their sensitivity and reliability. We envision that our microscopy tools might be useful in optical high-throughput approaches, including cell-based assays in small-molecule screening. These efforts are a first step toward uncovering how the spatial organization of genomes in normal cells is related to the disease state and how we can use this information to improve the diagnosis and analysis of cancers.

■ **Jeffrey Roix**
Department of Biology
Massachusetts Institute of Technology
Cambridge, MA

■ **Tom Misteli, PhD**
Principal Investigator
Laboratory of Receptor Biology
and Gene Expression
NCI-Bethesda, Bldg. 41/Rm. B610
Tel: 301-402-3959
Fax: 301-496-4951
mistelit@mail.nih.gov

■ FROM THE DIRECTOR'S OFFICE

The National Cancer Institute: A Bridge for Genesis and Translation of New Ideas in Cancer Prevention and Treatment

The development of modern chemotherapy for cancer originated in a unique translational environment at the NIH more than 50 years ago. The opening of the Clinical Center in the 1950s allowed Roy Hertz, Griff Ross, and Mort Lipsett to demonstrate convincingly, and for the first time, that total cure of an advanced, potentially lethal cancer (in this case, choriocarcinoma) could be accomplished with chemotherapy. This pioneering achievement then led to the successful development of modern combination chemotherapy for leukemia and lymphoma in the 1960s and 1970s at the NCI. These were truly translational investigations, since the basic animal studies that developed these new combination therapies were initiated on the NIH campus by Lloyd Law's group and then translated for the first time into clinical reality by Tom Frei and Jay Freireich and their colleagues in the Clinical Center. These landmark collaborative studies represent a model for the cooperativity between basic science and clinical application that is still needed to achieve the NCI's goal of eliminating suffering and death due to cancer.

A generation later, however, it is clear that the same paradigms for the development and testing of cytotoxic agents that worked so well for the treatment of childhood leukemia will not suffice for the control of common epithelial malignancies, such as carcinoma of the lung, colon, pancreas, breast, and prostate, which collectively account for more than half of all cancer deaths. Therapy for leukemia is unfortunately a poor model for therapy for carcinoma, since leukemogenesis and carcinogenesis are very different processes.

We need to rethink our basic assumptions concerning the control of carcinoma and place more emphasis on control of disease in its earliest, pre-invasive stages. The disease itself is "carcinogenesis," an evolving process ultimately leading to the invasive state we call "cancer." This process is potentially more manageable in its earliest stages, before it becomes invasive. Yet, most efforts to control cancer are still focused on treatment of end-stage, invasive, and metastatic disease, rather than on its prevention.

Therefore, it will be important to develop new preventive strategies. These strategies will focus primarily on the following: 1) new molecular and cellular techniques for assessment of the risk of developing cancer, 2) the identification of new molecular targets for prevention, and 3) the development of new drugs to suppress the carcinogenic process, either at its very inception (inhibitors of initiation) or in its initiated but pre-invasive stages (inhibitors of promotion and progression). For the evaluation of the extent of risk in an individual, as well as the efficacy of a preventive drug that might be given to such a person, it will be necessary to develop new technologies, such as proteomics and nanotechnologies. It is essential to identify individuals at risk for cancer as early as possible in the pathogenesis of their disease, and then to evaluate the efficacy of a preventive intervention as quickly as possible.

Although such a preventive strategy holds great promise, it is intrinsically more complex, scientifically, than the classical clinical approach of waiting for relatively advanced disease to manifest itself, and then starting treatments that can be easily evaluated, because the patient is already symptomatic. The complexities of this new approach to prevention now provide the Center for Cancer Research at NCI with a unique opportunity to play a leadership role in a national effort to prevent cancer. The NCI has the unique ability to make the needed long-term commitments to a prevention effort. It has unique resources, both in terms of the scientific and clinical talents of its staff, as well as in terms of its world-class laboratories and new clinical facilities. In such a national effort, the NCI will act as a bridge whereby basic scientific knowledge can be applied in a translational manner to reach an ultimate goal.

■ **Michael B. Sporn, MD**
NCI Eminent Scholar
Dartmouth Medical School

X-ray Crystallography and Computational Analysis Illuminate the Molecular Evolution of Antibody Specificity and Affinity

Li Y, Li H, Yang F, Smith-Gill SJ, and Mariuzza RA. X-ray snapshots of the maturation of an antibody response to a protein antigen. *Nat Struct Biol* 10: 482-8, 2003.

Mohan S, Sinha N, and Smith-Gill SJ. Modeling the binding sites of anti-hen egg white lysozyme antibodies HyHEL-8 and HyHEL-26: an insight into the molecular basis of antibody cross-reactivity and specificity. *Biophys J* 85: 3221-36, 2003.

The humoral immune response includes a large repertoire of antibodies (Abs) with a broad range of antigen (Ag) affinities and specificities. It is generally held that early response Abs are less specific or are multispecific and evolve to become more conformationally constrained and specific. During this process, called affinity maturation, somatic mutations in the Ag-binding site increase an Ab's affinity for Ag by 10- to 100-fold, and the Ab response becomes more specific. Affinity maturation thus represents a prototypical example of molecular evolution. The structural changes mediating this functional evolution are only beginning to be understood.

Studies of Abs specific for small-molecule haptens suggest affinity maturation produces more polar amino acids, which increase intermolecular- and intramolecular-specific contacts and noncovalent bonds and thereby increase both Ab affinity and specificity. In addition, affinity maturation enhances hydrogen bond networks, which reduce the plasticity of the Ag-binding site and shape a "preconfigured" pocket with higher Ag complementarity. Affinity and flexibility are generally held to be inversely correlated.

Ab-protein complexes involve much larger contact interface areas than do Ab-hapten complexes. A decade ago we isolated hybridomas at different time intervals after immunization with hen egg white lysozyme (HEL) (Newman MA et al. *J Immunol* 149: 3260-72, 1992), including four highly homologous Abs (H26, H63, H10, and H8) which recognize nearly identical HEL epitopes (HyHEL-26, -63, -10, and -8, respectively). Their relative affinities and number of amino acid differences attributable to somatic hypermutation are the same: H26 < H63 < H10 ≤ H8. As such, these Abs represent different stages of Ab response maturation. However, unlike hapten-binding Abs, their cross-reactivities with mutant lysozymes are positively correlated with affinity.

X-ray crystallography of this set of Abs provided us the first structural view of affinity maturation to a protein Ag. The H8-protein complex has not yielded crystals suitable for x-ray analysis, so we used a chimeric Ab consisting of the H8 heavy chain and the H10 light chain (H8L10). The x-ray crystal structures of H26-, H63-, H10-, and H8L10-HEL showed that, in contrast to the hapten-binding Abs, the higher affinity binding does not correlate with number of contacts, number of hydrogen bonds, or buried surface area per se, but through increased complementarity and increased burial of apolar surface (the latter at the expense of polar surface). We also noted a progressive shift in the backbones and in the side chains of heavy chain hypervariable region 2. Within this loop, aromatic residues 53 and 58 were tyrosines in all the Abs except H8, where they had mutated to phenylalanine. This shift and the mutations are important in increasing both complementarity and interacting hydrophobic surfaces.

To date, interpretation of structure-function relationships for native H8 has relied on a homology-modeled structure based on the x-ray structure of the H10-HEL complex (Padlan EA et al. *Proc Natl Acad Sci U S A* 86: 5938-42, 1989), which was of lower resolution than the more recently obtained structures of H8L10 and H26. The H8 model predicted a large shift in the conformation of the heavy chain hypervariable region 2, like that seen in the H8L10 crystal structure. The H26 homology-modeled and x-ray crystal structures also were very similar in hydrogen bond formation and salt bridges (Sinha N et al. *Biophys J* 83: 2946-68, 2002). Thus, homology modeling combined with energy minimization of a lower resolution structure can predict the structure and atomic interactions of the same or a closely related protein and compare well to the higher resolution x-ray structure—a significant result, because high-resolution structures are not always available for proteins of interest.

The crystal structures may solve the unexpected cross-reactivity properties of the HEL Abs: fewer hydrogen bonds and contacts, more non-specific hydrophobic contacts, and no salt bridges confer flexibility to H8 and H8L10 protein-binding sites, allowing conformational adaptation and cross-reactivity with mutant Ags. In contrast, the specificity but lower affinity of H26 correlates with numerous H26-HEL hydrogen bonds and with salt bridges not attributable to somatic mutations, which likely restrain the Ag-binding site from assuming alternate conformations. Thus, although conformational flexibility is probably an important general mechanism associated with Ab cross-reactivity—allowing either induced fit or selection among conformational isomers by dissimilar Ags—higher affinity in protein-binding Abs is not necessarily achieved by less flexibility. Affinity maturation may select for high-affinity Abs with preconfigured complementarity or with preconfigured flexibility through modulation of Ag-binding site flexibility. In the case of H8, hydrophobic interactions may be particularly important in both affinity and flexibility.

Abs are increasingly used in therapy and diagnosis, so understanding Ab-Ag

association is of growing importance. The structural mechanisms we describe are likely to inform general molecular targeting efforts and be directly applicable to Ab design. The current dogma is that a protein and its complexes are a dynamic ensemble of multiple conformational states and either mechanism (preconfigured complementarity or preconfigured flexibility) may contribute to both complementarity and affinity. Thus, receptor and ligand flexibility must be considered in structure-based drug discovery.

■ **Sandra J. Smith-Gill, PhD**
Principal Investigator
Structural Biophysics Laboratory
NCI-Frederick, Bldg. 469/Rm. 206
Tel: 301-846-5203
Fax: 301-846-6326
smithgil@helix.nih.gov

■ CANCER AND CELL BIOLOGY

Chromatin Structure, HMGN Proteins, and DNA Repair

Birger Y, West KL, Postnikov YV, Lim J-H, Furusawa T, Wagner JP, Laufer CS, Kraemer KH, and Bustin M. Chromosomal protein HMGN1 enhances the rate of DNA repair in chromatin. *EMBO J* 22: 1665-75, 2003.

Preservation of the genetic information encoded in DNA requires efficient and correct repair of damage induced in it by intracellular and extracellular agents. Aberrations in DNA sequence due to incorrect repair result in mutations that disrupt the fidelity of gene expression and could lead to disease or death. Cells have developed several mechanisms to correct various types of damage to their DNA. All repair pathways must cope with the fact that nuclear DNA is associated with proteins and is stored and packed in a fibrous structure named chromatin, that looks like "beads on a string." The chromatin fiber is a dynamic structure that is continuously modified and that constantly rearranges, condenses, and de-condenses in response to external and internal stimuli. For example, during the progression of the cell cycle, the interphase chromatin fiber condenses and is tightly packed into metaphase chromosomes. During replication or at sites of intense transcriptional activity, chromatin de-condenses, and the DNA sequence is more accessible to regulatory factors and polymerases. The dynamic nature of the chromatin fiber is a key element in a cell's ability to execute cellular processes that require access to the DNA, such as transcription, replication, recombination, and repair.

Changes in chromatin condensation at specific sites are driven by the action of several regulatory factors, including a family of nuclear proteins named HMGN. HMGN are nonhistone proteins that bind to nucleosomes (the "beads" on the chromatin string), reduce the compaction of the chromatin fiber, and enhance DNA-dependent activities such as transcription and replication. We investigated whether these proteins also affect the rate of DNA repair and focused on the repair of DNA damage caused by ultraviolet (UV) irradiation, a major factor in skin cancer. Most of the UV-induced damage is repaired by an evolutionarily conserved pathway named nucleotide excision repair (NER). Faulty NER leads to several pathological conditions such as xeroderma pigmentosum, Cockayne's syndrome, and trichothiodystrophy.

The rate of repair of UV damage is faster in DNA than in chromatin, an indication that the chromatin fiber inhibits the action of the NER pathways. To test whether HMGN proteins affect the NER pathways, we shaved a small area on the back of genetically modified mice lacking the gene coding HMGN1 (*Hmgn1*$^{-/-}$ mice) and of their wild-type littermates (*Hmgn1*$^{+/+}$ mice) and exposed the areas to low UV doses. Histological examination of samples from the irradiated skins revealed thickening (acanthosis and hyperkeratosis) in the skin of *Hmgn1*$^{-/-}$

Continued on page 8…

■ ADMINISTRATIVE LINKS

NIH/NCI Clearance Guidelines for Publications/Presentations
Prior to submitting any manuscript to a peer-reviewed journal for publication, all CCR principal investigators *must* submit their manuscript to their Lab/Branch Chief for review and clearance. Formal written or oral presentations *must* also be reviewed/cleared by the Lab/Branch Chief. In addition, any collaborative publication or oral presentation by an NCI employee and a non-federal colleague reporting on work arising from or connected with a contract (including intramural support contracts), grant, or other award funded by the NCI *must* also be reviewed and cleared by the Chief.

To *document* the clearance of a manuscript (or presentation), staff members may use the NIH form 1616-1, Rev. 6-84 (NCI Manual issuance 1184), which is available at http://www1.od.nih.gov/OIR/sourcebook/oversight/pub-clear-form.htm. This form can be easily completed (online) and then printed out for filing purposes. For more information about the NIH clearance policy, visit http://www1.od.nih.gov/OIR/sourcebook/oversight/pub-clear.htm.

Reporting High-Impact Manuscripts
High-impact manuscripts should be reported to Tracy Thompson (thompstr@mail.nih.gov), Chief, CCR Office of Communication, as soon as possible **after acceptance but before publication.** Please include the anticipated publication date, an electronic or hard copy of the manuscript, and the journal name. "High-impact" manuscripts include but are not limited to papers that reflect a significant advance in your field or papers in any of the following areas: public health; tobacco-related issues; new technological advances; imaging; obesity, dietary fat, energy balance; nanotechnology; molecular targets; stem cells; angiogenesis; or combination therapies.

...continued from page 7

mice but not in the skin of their *Hmgn1*$^{+/+}$ littermates. This type of skin thickening is a typical response observed in sunlight-treated human skin. This response is exaggerated in the UV sensitivity disorder xeroderma pigmentosum, suggesting that loss of HMGN1 protein impaired the UV repair. Indeed, mouse embryonic fibroblasts (MEFs) derived from 13.5-day-old *Hmgn1*$^{-/-}$ embryos were much more sensitive to UV irradiation than were MEFs derived from *Hmgn1*$^{+/+}$ littermates.

The low cellular survival and the skin reaction after UV irradiation are directly linked to the ability of a cell to repair the damage induced in DNA by UV irradiation. The major photoproduct of UV irradiation is cyclobutane pyrimidine dimers (CPDs), and their removal from cellular chromatin reflects the cellular efficiency of the UV repair machinery. We found that the removal rate of CPDs in *Hmgn1*$^{-/-}$ MEFs was slower than that in *Hmgn1*$^{+/+}$ MEFs, an indication that HMGN1 protein affects the rate of damage removal from chromatin. To understand the mechanism whereby HMGN1 affects the UV repair process, we used cDNA microarrays to analyze the transcripts from irradiated and nonirradiated *Hmgn1*$^{-/-}$ and *Hmgn1*$^{+/+}$ cells. The analysis revealed that the expression of the various components of NER pathways was the same in both cell types, an indication that the reduced rate of UV repair is not due to changes in the NER complex itself. Indeed, host cell reactivation assays, a measure of the innate NER activity of a cell, indicate that both cell types have a fully functional NER pathway.

Next we reintroduced into *Hmgn1*$^{-/-}$ cells either wild-type HMGN1 proteins or mutant proteins that do not bind nucleosomes or do not unfold chromatin. The wild-type protein rescued the UV repair efficiency of the cells but the mutant proteins did not, an indication that the action of HMGN1 is linked to its ability to bind to and unfold chromatin. These results also suggest that HMGN1 is associated with chromatin in the regions undergoing UV repair. Indeed, chromatin immunoprecipitation assay (ChIP) revealed that HMGN1 is associated with the same genes that showed inefficient CPD removal. By micrococcal nuclease digestion, we demonstrated that the DNA of these genes in the chromatin isolated from *Hmgn1*$^{-/-}$ liver nuclei is less accessible than that isolated from wild-type *Hmgn1*$^{+/+}$ liver nuclei. Thus, we demonstrated that HMGN1 enhances the rate of DNA repair by reducing the compaction of the chromatin fiber and facilitating access of the repair components to the damage site.

The removal of CPDs, which is known to be affected by chromatin, has been subdivided into three major steps: In the first step, the damage is accessed. In the second step, the damage is repaired, and in the third step, the chromatin structure is restored to normal. We found that HMGN1 is involved in the first step of the NER process. It changes the compaction of the chromatin fiber and enhances the accessibility of the damaged DNA to the repair enzymes. The role of HMGN proteins extends beyond the DNA repair process. As ubiquitous abundant proteins that move dynamically in the nucleus, they play a part in maintaining the integrity of the genome and in optimizing gene expression during development and differentiation.

- **Yehudit Birger, PhD**
- **Kenneth H. Kraemer, MD**
- **Michael Bustin, PhD**

Laboratory of Metabolism
NCI-Bethesda, Bldg. 37/Rm. 3122B
Tel: 301-496-5234
Fax: 301-496-8419
bustin@helix.nih.gov