Biomarkers and Biosensors for Studies of Blood Cancer Risks

S. M. Rappaport, Martyn Smith, Mark van der Laan, Luoping Zhang, Richard Mathies, Evan Williams, Bernhard Boser

University of California, Berkeley

The overall theme of Berkeley's Center for Exposure Biology is "Biomarkers and Biosensors for Studies of Blood Cancer Risks." This theme reflects the long-standing interests of our investigators in the etiology of human blood-borne cancers (leukemias, lymphomas, and myelomas) and our close associations with clinicians and epidemiologists who maintain repositories of stored tissues for molecular epidemiology studies of these diseases. Blood-borne cancers will cause the deaths of approximately 60,000 people in the U.S. this year and half-a-million Americans are living with these diseases. They account for over 10 percent of all cancer deaths and are the leading disease killers of children in the U.S. and of men under age 39. Our center also highlights particular research strengths of our investigators in the areas of laboratory-on-a-chip technologies, microsensing, mass spectrometry, toxicology, bioinformatics, and exposure biology. The center has the following three projects and goals:

Project 1 - Protein adducts as molecular signatures of carcinogen dose. The major goal of Project 1 is to demonstrate the viability of 'protein adductomics' as a true 'omics' approach, one of general importance to the GEI and of particular importance to Exposure Biology. We propose to enrich unknown cysteinyl adducts of human serum albumin and then to profile these adducts in the serum of lymphoma cases and controls. If successful, this project will identify unknown initiators of human lymphomas.

Project 2 - High throughput detection of genetic mutations as biomarkers of hematological cancer risk. This project will develop and apply laboratory-on-a-chip technologies to investigate mutation spectra in individual cells that are involved with human leukemia and lymphoma. By amplifying DNA from single cells, we will measure mutation spectra at critical loci in susceptible individual cells in a large background of normal cells. Then, by developing high throughput single cell PCR assays to detect low frequencies of mutations of importance, we will prepare a library of hot-spots for mutations in key genes or regions related to blood-borne cancers.

Project 3 - Advanced biosensing for molecular and cellular biomarkers. This project will develop biosensors that make it possible to perform immunoassays quickly and conveniently, either in a laboratory or in the field, using a single drop of blood. It will validate the biosensors by measuring blood protein adducts of polycyclic aromatic hydrocarbons in archived blood from asphalt workers. By developing the technology to utilize protein adducts as biomarkers of internal dose in large epidemiology studies, our Center will open up important new avenues of inquiry regarding the links between chemical exposures and human cancers.

Application of biomarkers in future studies. In our Center, we will take a first-step towards providing revolutionary new biomarkers for the detection of carcinogens and genetic mutations of relevance to blood-cancer risk in healthy individuals. Such biomarkers will be useful measures of individual dose and early effect in epidemiological studies of leukemia and lymphoma, which have long latency periods, and will provide tools for early detection for those individuals at risk.

Project 1. Protein adducts as molecular signatures of carcinogen dose. Participants: Stephen Rappaport (Leader), Evan Williams, and Mark van der Laan

Leukemia/lymphoma risk and environmental exposures. Current evidence points to a causal relationship between adult leukemias and exposure to benzene; however, evidence associating leukemias and lymphomas with other environmental exposures is far weaker. This is not surprising because information about chemical exposures in air, water, and food is almost universally lacking, and even the few available environmental measurements vary tremendously between subjects and within subjects over time. Furthermore, exposures to chemical carcinogens arise naturally from the diet as well as from endogenous inflammatory and metabolic processes. For example, the hematotoxicant 1,4-benzoquinone can be produced not only from metabolism of benzene, but also from metabolism of arbutin (a glycosylated hydroquinone present in many plant-based foods), and from metabolism of phenol from both exogenous sources (e.g., drugs such as PeptoBismol®) and endogenous sources (e.g., catabolism of aromatic amino acids). Thus, the role that chemical carcinogens play in the etiology of blood-borne cancers is obscured by the generally poor and, in any case, incomplete knowledge of internal exposures received prior to diagnosis.

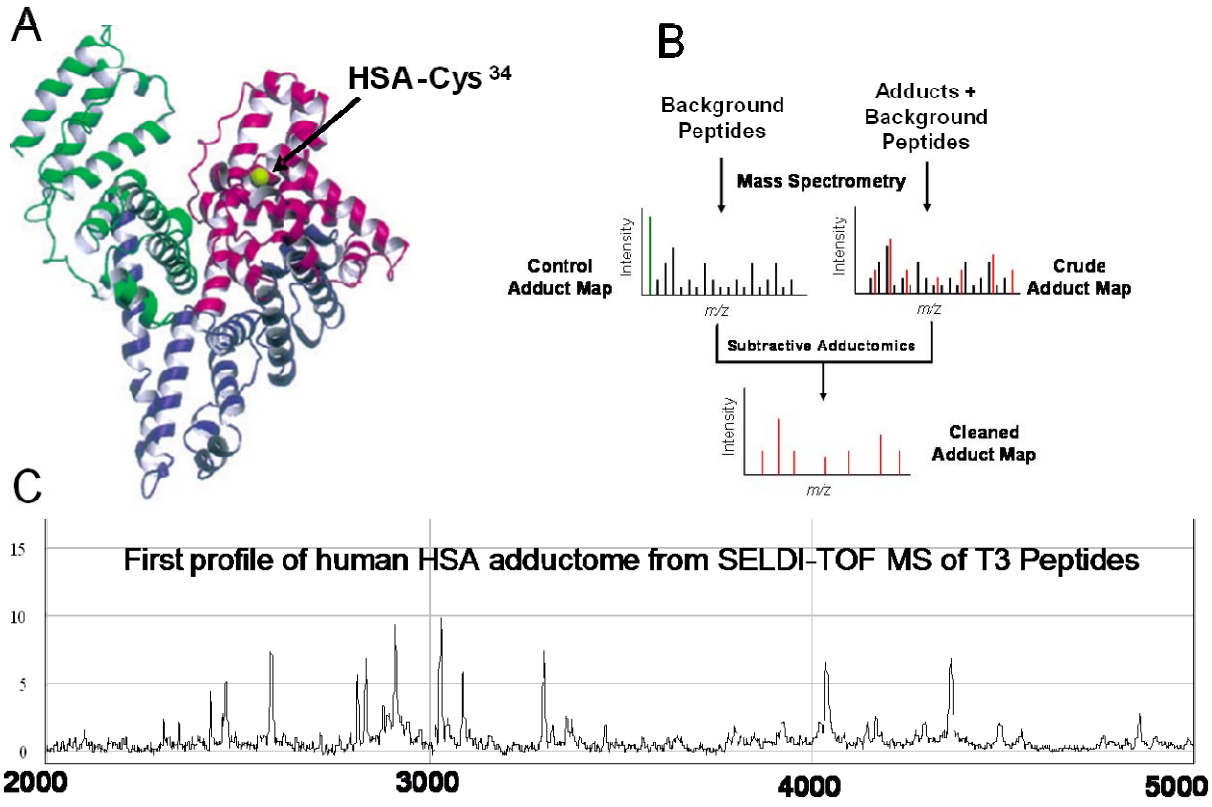
Protein adducts as biomarkers of carcinogen dose. Biomarkers of internal dose can be more accurate and precise surrogates for carcinogen exposures than environmental measurements per se. Yet, because chemical carcinogens are usually reactive electrophiles with very short life spans *in vivo*, it is rarely possible to measure them in target tissues. This has motivated the use of adducts of these electrophiles with abundant blood proteins, particularly hemoglobin (Hb) and human serum albumin (HSA), as measures of carcinogen dose. Electrophiles enter the blood from absorption in the lungs or gut (e.g., inhalation of ethylene oxide) or, more typically, via metabolism of procarcinogens in the liver or other tissues (e.g., production of benzene oxide by CYP metabolism of benzene). Once in the blood, electrophiles react at varying rates with all available nucleophiles to form adducts. Hemoglobin and HSA contain numerous nucleophilic loci capable of binding with electrophiles to form adducts. Since protein adducts are not repaired and are much more abundant than DNA adducts in blood, they are more useful measures of internal dose than DNA adducts, which have paradoxically received far more attention in this regard. Indeed, the kinetics of production and elimination of Hb and HSA adducts are sufficiently simple to permit straightforward estimates of systemic doses of carcinogens over the mean residence time of these proteins (28 d for HSA and 63 d for Hb in humans).

The protein adductome and carcinogen discovery. Although levels of targeted Hb and/or HSA adducts have been investigated in human blood for several environmental toxicants, only limited research has been conducted to characterize the full complement of adducts produced at a given locus on a protein. Yet, this 'protein adductome' (the set of all adducts on a given protein) is an entity of toxicogenomic importance and offers opportunities for true 'omics' experiments. Since the HSA and Hb adductomes reflect all stable products of nucleophilic reactions between major blood proteins and potential carcinogens during one to two months, they are particularly relevant to the discovery of carcinogens (via profiling experiments), arguably more so than the proteome or metabolome. **The goal of this project is to develop protein adductomics as a true 'omics approach for profiling possible carcinogens.**

Of the many loci available for nucleophilic substitution reactions with electrophiles, free thiols in Cys residues are important subadductomes because of their strong nucleophilicity. Since Cys subadductomes also offer the opportunity for selective enrichment via thiol-affinity chromatography, they are attractive candidates for adductomic experiments. Interestingly, human Hb and HSA each have only one free thiol available for adduction, i.e., Hb-Cys^{93}β and HSA-Cys^{34}. Because HSA-Cys^{34} is the most abundant serum thiol in human blood and is also inherently reactive towards electrophiles (much more so than Hb-Cys^{93}β), we focus upon adducts bound to HSA-Cys^{34}.

One method for selectively enriching HSA-Cys^{34} adducts takes advantage of the propensity of the free (unadducted) thiols to form disulfides with thiol-affinity resins. In our project, we remove unadducted Cys proteins and peptides in a series of steps involving thiol-affinity chromatography, and in doing so, create a pair of HSA peptide specimens, one representing background peptides ('control adducts') and one representing the unknown thiol-adducted peptides + background peptides ('crude adducts'). After mass spectrometry of these paired peptide specimens, we subtract the map representing control adducts from the spectrum representing crude adducts to produce a cleaned adduct map for each HSA sample. **Our goal is to compare cleaned adduct maps from biobanked specimens of HSA obtained from lymphoma cases and controls.** Adducts that are more abundant in cases than in controls suggest likely exposures to carcinogens of interest. In subsequent experiments we plan to identify unknown adducts that are associated with lymphoma status.

The figure below highlights some important features of our project.



A Structure of HSA, showing the lone free Cys residue (HSA-Cys^{34}) where adducts will be measured.

B Scheme for subtractive adductomics whereby the control adduct map from a given HSA-Cys[34] specimen is subtracted from the corresponding crude adduct map to pinpoint adducts of interest.
C Crude adduct map of HSA-Cys[34] adducts obtained from a specimen of commercial HSA.

Project 2. High throughput detection of genetic mutations as biomarkers of future hematological cancer risk. Participants: Richard Mathies (Leader), Martyn Smith, and Luoping Zhang

Biomarkers of early effect for blood-borne cancers. Carcinogenesis is a complex, multistage process which involves the accumulation of a variety of mutations within a particular cell and its progeny. Particular genes, chromosomal regions, or entire chromosomes are vulnerable to mutation at various points in carcinogenesis. This suggests that certain mutations play specific roles in determining the ability of a cell to survive and continue to the next step of this multi-step process, as well as potentially determining what the next mutation will be. These mutations, particularly early events, may provide biomarkers, which indicate genetic damage and potential cancer risk. Regarding blood-borne cancers, mutations in two specific genes, NRAS and nucleophosmin (NPM1) are common in myeloid leukemia. Cytogenetic studies of acute myeloid leukemia (AML) and myelodisplastic syndrome (MDS) have shown that most patients have acquired chromosome aberrations, including deletions of two specific regions on the q arms of chromosomes 5 and 7, namely, 5q31 and 7q22. These chromosome aberrations and accompanying mutations are known to cause AML, are found at higher frequency in chemically-induced leukemias associated with alkylating drugs and benzene exposure, and may be good biomarkers of early effect for myeloid leukemia. Chromosomal translocations affecting immunoglobulin genes are characteristic features of lymphoma, notably the t(14;18) translocation most commonly found in follicular lymphoma. Thus, t(14;18) may be a good biomarker of early effect for lymphoma.

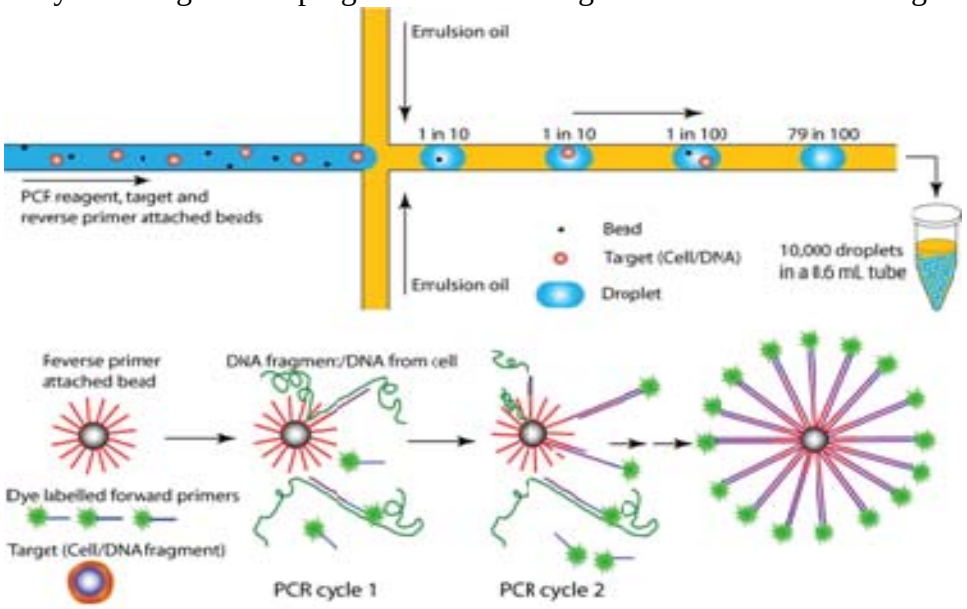
In this project we are developing methods for measuring mutations in NRAS, NPM1, deletions of 5q31 and 7q22, and t(14;18) in single cells as biomarkers of early effect for blood-borne cancers. These methods will be applied to individuals without these diseases using a laboratory-on-a-chip approach to detect infrequent mutational events at these critical loci in a large background of normal cells. This approach will take advantage of recent advances in single cell sorting and high throughput single cell genetic amplification and analysis.

Measuring mutations in single cells. Genetic damage at the chromosome level has traditionally been measured using cytogenetic analysis of metaphase spreads under a microscope. Indeed, classical chromosome aberrations have been shown to have predictive value for risks of cancer, particularly for blood-borne cancers. However, classical aberrations are a measure of overall chromosome damage, not of specific events on the causal pathways of particular diseases. In order to understand the mechanisms of exposure-related diseases, we need to measure specific events on the causal pathways of those diseases. Since these specific events are relatively rare among non-diseased populations, it is important to screen events among large populations or to examine large numbers of cells from each subject in order to attain sufficient statistical power. Although this is not possible with conventional cytogenetics, the use of PCR technology has vastly improved detection and identification of gene mutations. However, certain forms of

genetic and chromosomal mutation have been difficult to measure by real-time PCR and large amounts of genomic DNA or mRNA have been required for the analyses. Thus, the frequency and variability of many mutations have not yet been established in the normal population and PCR based studies of bulk cell averages have found limited applicability in molecular epidemiology studies.

There is also a desire to measure cancer-relevant mutations in critical target cells, such as stem and early progenitor cells, which are usually present in small numbers. Hence, methods are needed that can measure multiple forms of mutation in a given population of cells. One approach to achieving this goal is to develop PCR mutation assays that work at the single cell level and are sufficiently high throughput so that hundreds of thousands of cells can be analyzed in a relatively short period of time. The key to single cell genetic analysis is the idea of a PCR colony. The basic idea is to dilute a PCR template or cell to the level where only a single molecule or cell is present in a nanoliter reaction volume and then to amplify this template in that nanoliter volume. This amplification proceeds to produce a clonal colony of the target sequence that is determined by the primers. This pPCR colony or 'polony' can then be genetically analyzed or sequenced by using integrated high throughput microfluidic systems developed in the Mathies laboratory.

Fortunately, microfluidic and laboratory-on-a-chip technologies have now advanced to the point that single cell genetic analysis is feasible on a high-throughput scale. **Our goal is to analyze specific genetic targets in single human cells by wedding microscale PCR analysis in nanoliter volume emulsion droplets with laboratory-on-a-chip technologies**, which employ microfabricated microfluidics for sample transport, manipulation, and genetic analysis. Microfabricated microfluidic structures now provide a versatile set of platform technologies for performing electrophoretic separations, on-chip reactions, on-chip sample cleanup, and microfluidic manipulations of nanoliter volumes in fully integrated devices. In this project, we have already made significant progress towards this goal as illustrated in the figure below.



Schematic of our single target (Sheared genomic DNA/cell) analysis method. The target suspension and PCR reaction are mixed at a statistically dilute level producing 1% of the droplets with both a cell and a bead functionalized with reverse primer. Amplification in the droplet produces a large number of double stranded products that are linked to the bead by the

covalent reverse primer. In most cases the bead is fluorescently labeled because the forward primer carries a product specific label. Since, at least four distinctly different labels can be introduced; this gives the ability to perform multiplex identification if mixtures of target specific reverse primers are linked to the beads. After amplification the beads are isolated, pooled and run through a flow cytometer to determine the distribution of fluorescence on each bead.

Project 3. Advanced biosensing for molecular and cellular biomarkers. Participants: Bernhard Boser (Leader) and Richard Mathies

Measuring protein adducts in a single drop of blood. Protein adducts have several properties that make them ideal biomarkers of internal doses of carcinogens in epidemiology studies (see Project 1). However, the need to obtain venous blood samples has been a great impediment to the use of protein adducts and other biomarkers in studies involving hundreds or thousands of subjects. As a more acceptable alternative, blood samples can conveniently be collected as dried blood spots (DBS), using a finger lance to prepare single drops of blood that are absorbed on special specimen collection papers. However, there are no sensitive assays for protein adducts that can conveniently process samples from DBS in high throughput or point-of-analysis applications. In collaboration with Project 1, **we will measure protein adducts of polycyclic aromatic hydrocarbons (PAH), a ubiquitous class of environmental carcinogens, from DBS using an automated microscale ELISA assay**. Adducts of PAHs were chosen, in part, because PAHs are ubiquitous environmental carcinogens that are routinely detected in the general population due to smoking, as well as environmental sources (e.g., engine exhausts), and dietary sources (e.g., char-broiled foods). They are also found at higher levels in the blood of working populations exposed to PAH (e.g., asphalt workers). In this project we will measure PAH adducts bound to globin (Gb), which is isolated from hemoglobin (Hb) in a single DBS.

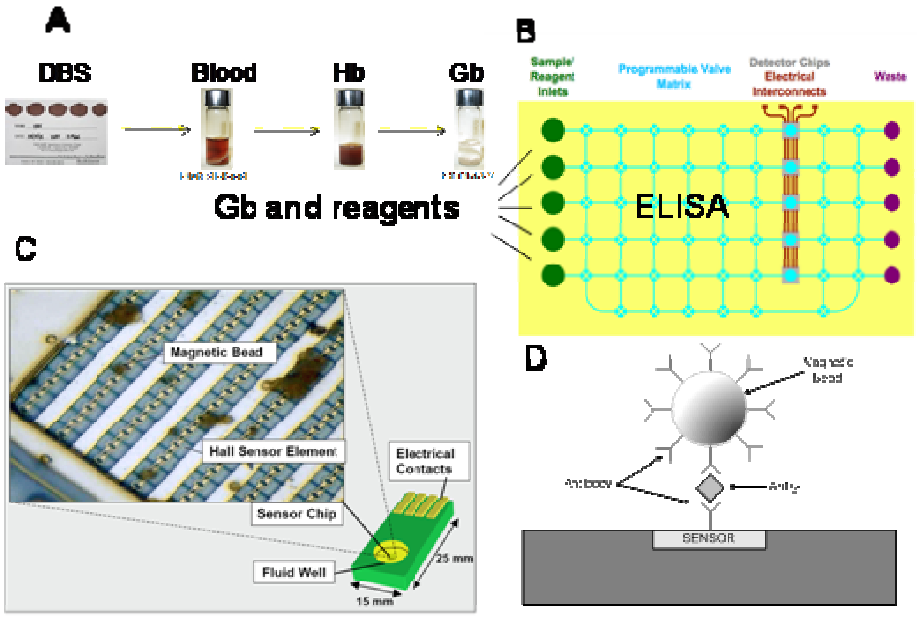
Many ELISA assays, such as those currently used to detect PAH-protein adducts, require a complex sequence of steps including sample extraction and concentration, mixing with reagents, washing, incubation, and detection. While the steps are similar for many assays, variations in the basic sequence and different reagents add further complications. To perform a large number of high throughput assays and to lower costs, a fully integrated microfluidic format is needed. The development of a simple, low cost, high throughput assay would advance our ability to screen blood samples of environmental interest and would advance laboratory studies to determine which adducts predict cancer risks. Furthermore, the development of assays and technologies that can be transitioned to point-of-analysis applications would enable real time evaluation of exposure levels to initiate strategies to reduce cancer risks.

Wedding laboratory-on-a-chip to biosensing technologies. Programmable laboratory-on-a-chip technology is ideally suited to automate assays for detecting protein adducts of interest. The small sizes of these devices add additional benefits including intrinsically small sample and reagent volumes (for low costs), fast reaction times, and the ability to measure several different adducts in a single device, possibly using a single DBS. Detection systems typically used in laboratory-on-a-chip devices have relied on fluorescence labeling and detection. Because fluorescence requires lasers and optical detectors many times larger than the microfluidic system itself, we use magnetic labeling to solve the integration and size problems. Magnetic labeling assays follow a standard ELISA 'sandwich-type' format but replace the fluorescent label with a small (0.1, …, 3 μm diameter) paramagnetic bead. These magnetic sensors produce an electrical

output that is easily processed. The capability to detect single beads translates into high sensitivity, potentially at the level of single molecules. A high dynamic range is achievable thanks to the electrical output of the sensor, which is proportional to the number of magnetic beads, and hence to target molecules bound to its surface. The possibility to manipulate magnetic beads is another attractive feature since it offers a means to automate assays.

In this project we will produce a cellular automaton consisting of a dense microfluidic array of valves, each of which functions as either a valve or a reactor, to process Gb specimens containing PAH adducts. This addressable array can perform a wide variety of programmable operations to execute different biochemical protocols on the same microdevice. The output from each microfluidic channel will enter a magnetic detector based on a CMOS Hall sensor array. This design realizes 1024 individual sensor elements along with a switch matrix and detection electronics on a single chip, yet requires only 4 interconnects thanks to the switching. This approach is, therefore, compatible with scaling up to very large number of detector elements and multiple chips, while still needing only a small number of interconnects.

The individual elements of the assay, the automaton, and the magnetic sensors are illustrated in the figure below.



A Isolation of globin (Gb) and the associated protein adducts from a dried blood spot (DBS).

B Microfluidic array of programmable valves for processing samples of Gb through an ELISA assay.

C Hall sensor for detecting magnetic beads. Sensors are connected to the microfluidic array via electrical interconnects (shown in B). Detector chips measure less than 1 mm on a side. At a cost of only pennies per chip, several can be integrated into the microfluidic device to detect several different antigens in a sample obtained from a single DBS.

D A single sensor element, structured for the ELISA format to detect antigens representing protein adducts or other analytes of interest.

Validation and future applications. The automaton developed in this project will be the first integrated immunological laboratory capable of performing complex assays automatically. We will validate the automaton by measuring PAH-Gb adducts in archived blood samples obtained from PAH-exposed asphalt workers and control workers. The simplicity for the user enables many new applications for ELISA assays, including more thorough investigation of environment-disease associations and in-the-field monitoring of a drop of whole blood or of a DBS. The rapid turnaround time associated with this technology can also motivate secondary screening of possible sources of exposure (while in the field) or interventions to reduce exposure levels.