I. Research Plan

a. Specific Aims

Cell surface carbohydrate structures are known to be characteristic markers for different types of cells (1-3). The transformations of normal to cancerous cells are also often associated with the alteration of cell surface carbohydrates (4, 5). The expression or over-expression of certain carbohydrates, such as sialyl Lewis X (sLex), sialyl Lewis a (sLea), Lewis X (Lex) and Lewis Y (Ley), has been correlated with the development of certain cancers, particularly carcinomas (6-17). Therefore, these cell surface carbohydrates can be used for the cell-specific identification and targeting of certain cancers. The long-term goal of this project is the development of artificial receptors which can recognize characteristic carbohydrate structures with high selectivity and affinity. Such receptors could be used for the development of fluorescent tags for cell-specific identification, tissue-specific imaging (such as MRI), and targeted delivery of therapeutic agents. In this project, we will use sLex as the model carbohydrate and use colon cancer as the model biological system for our studies. This selection was made because the expression of sLex is often associated with the development and the metastatic potential of colon cancer (16-31). The short-term objective of this R21/R33 application is the development of tissue-specific fluorescent tags (sensors) which can recognize sLex with high affinity and selectivity. Such fluorescent tags could be used for the pathological analysis and identification of colon cancer cells over-expressing sLex. The design takes advantage of the high affinity interaction of boronic acid with vicinal diol structures commonly found in carbohydrates for the preparation of high affinity receptors. It is understood that different cell surface carbohydrates will have different numbers and spatial orientations of vicinal diol moieties. For example, sLex, sLea, Lex, and Ley all have different numbers and/or different spatial orientations of vicinal diol structures. A number of boronic acid moieties arranged in a special three-dimensional orientation could bind the complementary carbohydrate moiety with high selectivity and affinity.

Specifically, we are interested in developing artificial chemical receptors that have boronic acid moieties with certain spatial arrangements for the recognition of sLex. For the construction of such fluorescent sensors, we will use an integrated approach combining template-directed synthesis, combinatorial chemistry, and computer molecular modeling aided design. This project will be divided into two phases. In the first phase (R21 phase) of the project, we will focus on evaluating the feasibility of the approach aimed at developing the chemical methodologies for the construction of such fluorescent tags. If successful, the second phase (R33 phase) will involve the optimization of the methods and sensors developed in the first phase for optimal affinity and selectivity for sLex. In the second phase these fluorescent tags will also be evaluated using colon cancer cell lines over-expressing sLex. The short term goals of the R21 phase will be accomplished by pursuing the following specific aims.

**Specific Aim 1.** Synthesis of a number of fluorescent and non-fluorescent arylboronic acid moieties.

**Specific Aim 2.** Template-directed combinatorial synthesis of potential sLex artificial receptors by cross-linking two boronic acid moieties using small libraries of linkers.

**Specific Aim 3.** Screening of the artificial receptors thus prepared for their binding to chemically synthesized sLex.

**Specific Aim 4.** Computer molecular modeling work to aid the further optimization of the fluorescent sensors.

**Specific Aim 5.** Selection and characterization of sLex-expressing and control cell lines.

The sLex-specific artificial receptors thus prepared have the potential to be used for cell identification, detection, and tagging for the purposes of localization, staging, tissue biopsy, and fluorescence-directed surgical removal of cancer cells. Although not specifically proposed in this application, the value of boron-11 in magnetic resonance imaging (MRI) (32-34), might allow these compounds to be used as specific MRI contrast agents for cancers over-expressing sLex. Such tissue-specific compounds could also serve as vehicles for the targeted delivery of cancer chemotherapeutic agents. Targeted delivery of these boron compounds could also be used to broaden the applications of boron neutron capture therapy (BNCT) of cancers, currently limited to brain tumors (35-41). Such small molecule sensors could also have the following advantages over antibody-based detection/delivery systems: (1) greater stability during storage and in vivo; (2) increased permeability through biological membranes and consequently enhanced target accessibilities; (3) intrinsic sensitivity to binding accompanied by significant fluorescence intensity changes, making detection and visualization easier and more suitable for high throughput screening, and (4) lower propensity to elicit undesirable immune responses. In addition, it is very likely that the production of such small organic molecular sensors will cost less than the production of antibodies and that the conjugation of such small molecule sensors to other therapeutic agents could conceivably be easier than the...
conjugation of antibodies to other therapeutic agents for targeted drug delivery. The same method, once successfully developed, could also be used for the construction of artificial fluorescent tags for cell-surface carbohydrates implicated in other cancers or diseases.

b. Background and Significance

Cancer is one of the leading causes of death. Early detection and intervention is one of the most important factors in determining the success rates in the treatment of many types of cancers. Despite the tremendous progress made in recent years, there are still many challenging issues concerning the early diagnosis, localization, staging, and treatment of such diseases. This sentiment is reflected in a recent PA requesting applications for the development of "innovative technologies for the molecular analysis of cancer." Such technologies include new methodologies for the "analysis of gene expression and gene products." Specifically, efforts in the development of "new chemistry or contrast agents" and "molecular detection systems" that can be used for the "molecular analysis either in vitro, in situ, or in vivo (by imaging or other methods) in the discovery process, as well as in pre-clinical models and clinical research" are within the scope of this PA. In response to this PA, we propose to develop a novel type of small-molecular organic compounds that can recognize cell-surface-specific carbohydrate structures for the fluorescent tagging of pathologically relevant cells. Such fluorescent tags, if successfully developed, could have far-reaching impact in the rapid diagnosis, prognosis, staging, and treatment of those cancers expressing the target carbohydrates.

It is well known that cell surface carbohydrate structures as part of glycosylated proteins and peptides are characteristic signatures of different cell types (2, 42). Such characteristic carbohydrate structures might be considered the zip code for cell addressing. Specific to this project, it is well known that altered cell surface carbohydrates, such as sialyl Lewis X (sLex), sialyl Lewis a (sLea), Lewis X (Lex) and Lewis Y (Ley), have been associated with the development and progression of many types of cancers (1, 4-17, 35, 37-41, 43-45). Therefore, these cell surface carbohydrates can be used for the cell-specific identification and targeting of certain cancers. Monitoring the expressions of these carbohydrates could help the diagnosis and prognosis of the disease. The development of molecular analysis probes targeted on these pathologically relevant carbohydrates will undoubtedly aid the diagnosis, prognosis, staging, and treatment of many cancers. We are interested in the development of small molecule fluorescent sensors that can recognize such cell surface carbohydrates with high affinity and specificity.

In this application, the initial focus will be on sLex using colon cancer as our model biological system. It is well known that the expression of sLex is often associated with the development and the metastatic potential of colon cancers (16-31). For example, multiple glycosyltransferase genes responsible for the further glycosylation of glycoconjugates were found to be up-regulated in many forms of colorectal cancer (18). Higher concentrations of sLex have been found in the sera of patients with colon cancer (24) and sLex has also been implicated in mediating the metastasis and invasiveness of colon cancer [(20, 24-26, 29-31) and see Specific Aim 5 for more details]. Therefore, fluorescent sensors that can bind to sLex with high affinity and specificity could be used for the in vitro pathological analysis of colon cancer cells or tissues, in vivo optical imaging, and fluorescence-directed surgical removal of cancerous cells. Although beyond the scope of the work proposed in this application, such molecular tags could also be used to target MRI contrast agents to those cells that over-express certain carbohydrates. They could also be used for the cell-specific delivery of cytotoxic agents for therapeutic applications. As will be discussed later, the compounds of our design are boron-containing compounds. Boron-11 itself has been used for MRI studies (32-37). Furthermore, boron compounds, including boronates, have been used in BNCT of certain cancers (35, 37, 39-41, 44, 46-48). However, one of the major problems to be overcome in this area is the targeted delivery of boron compounds. The boron-based sensors to be developed in this application could be used to target boron-containing compounds to specific cell types based on their cell surface carbohydrate structures for boron neutron capture therapy and boron-11 MRI studies.

To develop such sensors or molecular tags targeted on sLex, it is desirable to use recognition moieties that can recognize unique structural features on carbohydrates with high affinities. It has been known since the 1940's that a boronic acid moiety can bind to compounds with a vicinal diol (dihydroxyl) structural motif with high affinity (49-51), and such vicinal diol structures are commonly found in carbohydrates. However, when the vicinal diols are on a ring structure, as is often the case in carbohydrates, it is only the cis diols that can bind with boronic acid very tightly. Figure 1 shows the structures of Ley, Lex, and sLex with those vicinal diols that can interact tightly with boronic acid highlighted in circles. For sLex, either pair of the vicinal diols on the side chain could react with boronic acid to form a boronate. However, only one pair is highlighted. Although the drawing in Figure 1 does not specify the overall conformations and the spatial arrangements of each molecule, it is intuitive to expect that the spatial arrangements of these vicinal diol structural moieties are likely to be different among these carbohydrates due to their grossly different primary (connectivity) structural features. Conceivably, artificial receptors with a special
arrangement of two or more boronic acid moieties, which are complementary to the spatial arrangement of the vicinal diol structures of a particular carbohydrate, could be used as a selective and sensitive tag or sensor for that particular carbohydrate.

In the preparation of such carbohydrate-specific artificial receptors, it would be desirable if the binding would generate a detectable signal so that the binding event could be monitored conveniently. Recently, we have developed a new method of making fluorescent sensors for sugars through template-directed polymerization of boronic acid monomers (52). The sensors thus prepared are intrinsically sensitive to the binding event with significant fluorescence intensity increases upon binding with a model sugar, D-fructose (See Preliminary Results Section c-1-i and the reprint (52) provided in the Appendix for details). We plan to extend the application of such fluorescent boronic acid compounds for the preparation of small molecule fluorescent tags for sLex. The key to the project is the preparation of artificial receptors that have the correct spatial arrangement of the boronic acid moieties so that they have the necessary affinity and specificity for the target carbohydrate.

The construction of such receptors could take on several different approaches. The first one is de novo design. The de novo design approach, however, requires prior knowledge of the three-dimensional structure including the preferred conformation of the carbohydrate moiety, the de novo design of a complementary binding pocket, the synthesis of such a receptor, and the analysis of its binding. The development of computer molecular modeling in recent years has made such a task much more reliable and rational. However, conformational modeling of complex structures such as sLex is not a trivial task. Although there have been reports on the conformational studies of sLex using NMR (53, 54), the accurate depiction of the preferred conformation for such a flexible molecule in solution is very difficult. Even with the best technology available, this de novo design approach still requires much trial and error. The second approach is template-directed synthesis. It is well known that template-directed synthesis favors the formation of the desired compounds that best complement the template (55-63). In this case the target carbohydrate sLex can be used as the template to direct the synthesis (Figure 2). The template carbohydrate (1) could first react with the boronic acid compounds to give complex 2. The subsequent cross-linking of the two boronic acid moieties of 2 can be directed by the template effect of the carbohydrate moiety to give a compound (3) with the optimal complementary interactions with the target carbohydrate 1. After the extraction of the template carbohydrate from the final complex (3), the fluorescent diboronic acid sensor 4 can be evaluated for its ability to bind the target carbohydrate, the template (1). The third possible approach is through combinatorial chemistry, in which a large number of compounds can be synthesized randomly and screened to select the compound(s) with high affinity. A truly random approach relies on the sheer number of possible compounds generated to help ensure that a "hit" can be identified. However, this number could be so large that it would be impractical for an academic lab to take such a brute-force approach. Recently there have been a number of reports of using focused combinatorial library methods for the construction and selection of catalysts, artificial receptors and sensors (64-72).
In this application we plan to use an integrated approach by combining the power of template-directed synthesis, combinatorial chemistry, and computer molecular modeling. In this approach, the boronic acid-carbohydrate complex (2, Figure 2) will be prepared first. The combinatorial synthesis of potential artificial receptors will be carried out by template-directed cross-linking of the two immobilized boronic acid moieties of 2 with a small library of linkers (See Research Design and Methods, Specific Aim 2 for details). The template effect should favor the formation of the sensors that best complement the template molecule (55-63, 73). These libraries of potential fluorescent sensors will be screened using both a chemically synthesized target carbohydrate, sLex (See Specific Aim 3 for details), and in vitro cell culture systems. The selection and structural analysis of those with the best binding in a particular library will aid the design of the next generation of linkers to be used for improved affinity and selectivity of the fluorescent sensors. Computer molecular modeling will be used for the design of these linkers and the examination of the binding of these potential sensors with the target carbohydrate, sLex (See Specific Aim 4 for details). Such an approach will hopefully create a synergistic effect among three technologies in the search of fluorescent receptors for the target carbohydrate, sLex. Reiterations of the process of template-directed synthesis, screening, molecular modeling analysis, and refinement or redesign of the next generation linkers should allow us to better understand the structural requirements for optimal binding and further improve the affinity and specificity of the fluorescent sensors prepared. The computer molecular modeling work will be carried out in collaboration with Professors Stefan Franzen and Michael Prisant. Professor Franzen is an expert in studying macromolecular structures including their conformations and Professor Prisant is an expert in developing new algorithms in computer molecular modeling. Letters from Professors Franzen and Prisant are provided to indicate their willingness to collaborate. The evaluation of the binding of the fluorescent tags to cancer cells will be carried out in collaboration with Professor Brent W. Weston, a pediatric oncologist and an expert in studying the expression of cell surface carbohydrates, at the University of North Carolina at Chapel Hill. A letter from Professor Weston is provided to indicate his willingness to collaborate on this project.

It is true that no project is guaranteed to succeed. With an R21/R33 application aimed at exploring the feasibility of novel technologies, success seems to be more uncertain. Therefore, it is important to analyze the important issues of this project to see if there is a reasonable chance of its success. The use of boronic acid moieties is one of the key components of our design because of the known high affinity of a boronic acid moiety for carbohydrates through their interaction with vicinal diols. It has been reported that a single boronic acid moiety can have affinities for a sugar with $K_d$'s in the $10^{-4}$ M range (74, 75). Introduction of two boronic acids, which could bind to two pairs of vicinal diols on one molecule, could dramatically improve the binding affinity compared with compounds with only one boronic acid moiety (76-78). There are several examples showing that the incorporation of the second boronic acid binding moiety, in un-optimized situations, could increase the binding affinity for certain saccharides by factors up to about 200 (76-78). Furthermore, the use of properly positioned diboronic acid moieties could help improve the specificity for the target carbohydrate (76-82). Therefore, it would be reasonable to expect that further optimization of structures with two or more boronic acid moieties could afford selective receptors that could bind the target compounds with $K_d$'s in the sub-micromolar range, which would be approaching the affinities of enzymes and monoclonal antibodies for their respective ligands. Although the local concentration of the target carbohydrate on a cell surface is difficult to estimate, it is high enough to be recognized by relevant enzymes and antibodies. Therefore, artificial receptors that have affinities ($K_d$'s) for the target carbohydrates in the sub-micromolar range should be reasonable candidates for cell-specific fluorescent tags.

By design, the fluorescent tags thus prepared are expected to show a significant increase in fluorescence intensities upon binding with the target carbohydrate. This allows the easy monitoring of the binding of the fluorescent tags to the target carbohydrate and target cells. This feature should also help the adaptation of the technology, if successfully developed, to clinical applications. As described earlier, such fluorescent sensors could be used for (a) the pathological analyses of biopsy tissues in vitro; (b) correlation of the staging of the disease with changes in cell surface carbohydrate expression; (c) optical imaging, localization, and fluorescence-directed surgical removal of cancerous cells; (d) boron-11 MRI studies of cancers over-expressing certain carbohydrates; (e) cell-specific delivery of boron compounds for boron neutron capture therapy of certain cancers; and (f) delivery of therapeutic agents targeted on cancerous cells over-expressing certain carbohydrates. Some of these applications (d and e) are unique to these boron-containing compounds. The other applications, on the other hand, could conceivably be achieved with antibody-based targeting systems. However, the small molecule sensors proposed in this application have the following advantages over antibody-based systems. First, because these fluorescent tags are small molecule sensors, they are expected to be much more stable than antibodies and other protein-based tags both in vitro and in vivo and, therefore, should be easier to store and use without the stability problems of protein-based tags (83-87). Second, these small molecule sensors should be more permeable through biological membranes and...
consequently have enhanced target accessibilities (83-87). Third, the small molecule sensors should be much less likely to elicit undesirable immune responses compared with antibodies. Fourth, the significant fluorescence intensity changes intrinsic to the binding event should make the detection and visualization easier and more suitable for high throughput screening. Fifth, it is very likely that the production cost of such small organic molecular sensors will be less than that of antibodies and the conjugation of such small molecule sensors to other therapeutic agents could conceivably be easier than that of protein-based antibodies for targeted drug delivery, which should also make the clinical application of such sensors more practical. The sensors thus developed also could benefit the detection of other cancer cells expressing sLex. The same technologies developed also could be used for the preparation of fluorescent sensors of other cell-surface carbohydrates implicated in cancer and other diseases.

c. Preliminary Results

Our research team has extensive experience in the areas of research that will be involved in carrying out the proposed work. Collectively, we have extensive expertise in the areas of 1) molecular recognition and template-directed synthesis, 2) fluorescent sensor development, 3) drug delivery, 4) cell surface saccharide expressions related to cancer development and progression, 5) computer molecular modeling, and 6) solid phase chemistry and synthetic organic chemistry. Because of space limitations, only our work in the areas of sensor development, template-directed synthesis, and drug delivery is presented in this section to demonstrate our expertise in these areas.

In the fluorescent sensor development area (Section 1), we have developed a fluorescent sensor for D-fructose using template-directed polymerization of fluorescent boronic acid monomers. In this project, we have designed and synthesized a boronic acid-based fluorescent monomer (5, Figure 4), which can bind to the vicinal diol structures of a sugar moiety very tightly; as designed, such binding results in significant fluorescence intensity increase due to a suppression of the photoelectron transfer (PET) mechanism of the monomer [See below and (52) provided in the Appendix]. Also using template-directed polymerization, we have developed fluorescent receptors for L-tryptophan. Such receptors showed a significant increase in their fluorescence intensity upon binding of L-tryptophan in the presence of a fluorescence quencher. These polymeric receptors are sensitive to L-tryptophan at single digit μM concentrations (Figure 7, Section 1-ii) (88). Using similar methods, we have also prepared molecular transporters for D-tryptophan that can discriminate between enantiomers with about a three-fold difference (89). These studies were aimed at developing polymeric receptors for separation and molecular detection purposes. In the drug delivery area (Section 2) (90-92), we have developed a coumarin-based prodrug system, which allows for the bioreversible derivatization of amine drugs (93-96), peptides (97-101) and peptide mimetics (102-104). These cyclic prodrugs of peptides and peptide mimetics showed significantly improved membrane permeabilities and/or oral bioavailabilities compared with their corresponding parent drugs. In addition, in the area of solid phase chemistry we have developed two novel resin linkers (105, 106) for solid phase peptide and organic synthesis. One reprint (105) is provided in the Appendix section to demonstrate our expertise in solid phase chemistry.

1. Sensor development and molecular recognition

i Developing fluorescent sensors using functional monomers that are sensitive to the binding of carbohydrates.

Custom-made fluorescent sensors for organic molecules have a wide range of potential applications (107-110). Traditionally, such sensors have been prepared through de novo design and synthesis (111-124). Recently, molecular imprinting or template polymerization has gained much attention as a convenient method for the construction of binding sites for different analytes. Such an approach does not require prior knowledge of the three-dimensional structure of the analyte and the de novo construction of the complementary binding site.
Molecular imprinting is a technique first demonstrated in the late 1940's by Dickey (125) . The preparation of imprinted polymers involves (Figure 3): 1) pre-arrangement of the print molecule (template) and the functional monomers at low temperature so that complementary intermolecular interactions among functional groups can develop; 2) polymerization of the monomers under conditions that cause minimal disturbance to the print molecule-monomer interactions; and 3) extraction of the print molecules from the polymers, which leaves behind "receptor sites" that are complementary to the templates or print molecules in terms of size, shape, and functional group orientations. This technique has been used for the preparation of selective recognition sites for a wide variety of molecules (126-129) . Naturally, such polymeric receptors also have the potential to be developed as fluorescent sensors (128, 130-134) . Conceivably, polymeric receptors could be constructed with a fluorescent tag built into them so that binding and dissociation of the analyte would change the fluorescence emission sufficiently for the binding event to be monitored conveniently. However, such efforts have been hampered by the lack of appropriate fluorescent tags which could respond to the binding event with an intrinsic fluorescence intensity change.

In our search for carbohydrate sensors, we were interested in designing and synthesizing a structural motif, which would respond to the binding event with an intrinsic fluorescence intensity change and could be incorporated into the imprinted polymer. There are ample literature precedents showing that boronic acid derivatives can bind to vicinal diols tightly through ester formation (49-51) and that boronic acid-containing monomers can be used for the preparation of selective binding sites for saccharides through molecular imprinting (135-137) . Furthermore, boronic acid moieties, when attached to certain fluorescent molecules, have been shown by Shinkai and others to significantly affect the fluorescence intensity of such fluorescent tags upon ester formation with vicinal diols (80, 82, 113, 117, 138) . By taking advantage of such known properties, we designed and synthesized an anthracene-boronic acid conjugate (5, Figure 4) with a methacrylate moiety attached to allow for its polymerization [See the reprint (52) provided in the Appendix for more details]. In the boronic acid form, the fluorescence of the anthracene moiety can be quenched by the lone pair electrons of the nitrogen through PET. However, upon ester formation (7), the increased acidity of the boron atom causes the lone pair electrons of the nitrogen to be donated to the open shell of the boron atom, which subsequently disrupts the fluorescence quenching mechanism, and consequently causes a significant fluorescence intensity increase. The utility of such a functional monomer (5) was tested by preparing a polymeric fluorescent sensor for D-fructose, which is known to bind to two boronic acid moieties (79) . As designed, the binding of D-fructose to the resulting fluorescent polymers showed significant fluorescence intensity changes (Figure 5) (52).

The functional monomer (5) was synthesized in five-steps starting from 10-(hydroxymethyl)-9-anthraldehyde (8) (139) (Scheme 1). The hydroxy aldehyde (8), upon treatment with tert-butylidemethylsilyl chloride (TBSCI) in the

![Figure 4. The effect of binding with D-fructose on the fluorescence intensity of 5.](image)

![Figure 5. A typical set of fluorescence emission spectra of D-fructose imprinted polymer at different concentrations of D-fructose (λ ex: 370 nm).](image)

**Scheme 1. Synthesis of fluorescent boronic acid monomer 5.**

<table>
<thead>
<tr>
<th>Step</th>
<th>Reaction Conditions</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>TBSCI/imidazole/DMF, 91%</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>i) NH2Me/MeOH, 4-5 h, ii) NaBH4/MeOH, 2 h, 89%</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>K2CO3/MeCN, reflux, 42%</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>TBAF/THF, 74%</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>Methacrylic anhydride, DMAP/TEA/DCM, 70%</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 5. A typical set of fluorescence emission spectra of D-fructose imprinted polymer at different concentrations of D-fructose (λ ex: 370 nm).](image)

![Figure 4. The effect of binding with D-fructose on the fluorescence intensity of 5.](image)
presence of imidazole in DMF at room temperature, yielded the silyl ether aldehyde in 91% yield \((140)\). Treatment of the silyl ether aldehyde with methylamine in methanol followed by reduction with sodium borohydride afforded the silyl ether amine \(9\) in 89% yield. The reaction of amine \(9\) with \(2,2\text{-dimethylpropane-1,3-diyl[\(o\text{-}(bromomethyl)phenyl]\}-boronate (12) (141)\ in the presence of potassium carbonate in acetonitrile gave the amine boronate intermediate \(10\) in 42% yield. The cleavage of the silyl protecting group with tetrabutylammonium fluoride (TBAF) in THF at room temperature followed by an aqueous work-up gave \(11\) in 74% yield. Reaction of \(11\) with methacrylic anhydride in anhydrous THF in the presence of 4-dimethylaminopyridine (DMAP) afforded the methacrylic ester \(5\) in 70% yield \((52)\).

The fluorescent sensor preparation started with the synthesis of the D-fructose-monomer complex \(7\) as the template for the polymerization (Figure 4). Then polymerization was carried out by following procedures previously reported using ethylene glycol dimethacrylate (EGDM) as the cross-linker at 65 °C under nitrogen using 2,2'-azobisiso(2-methylpropionitrile) (AIBN) as the free radical initiator \((88)\). The template molecule, D-fructose, was extracted by repeated washing and the polymer thus prepared was ground into small particles for the subsequent binding studies. Without the addition of D-fructose, the fluorescence intensity (I) of the polymeric particles was low (Figures 5,6). However, with the addition of D-fructose, the fluorescence intensity was enhanced significantly in a concentration-dependent fashion. Fluorescence intensity changes were observed at \(\mu\text{M}\) to mM fructose concentrations (Figures 5,6) \((52)\). Control experiments were conducted by studying the binding of the imprinted polymer with other sugars such as D-glucose and D-mannose and the binding of the control polymer with D-fructose to further verify that the imprinting process helped in the creation of binding sites which preferentially recognized the template molecule [See reprint \((52)\) provided in the Appendix for more details].

Our studies demonstrated that sensitive fluorescent sensors can be prepared using template-directed polymerization with properly designed recognition moieties. The fluorescent monomer \(5\) developed in this study could also be used for the preparation of fluorescent sensors of other carbohydrates and biologically important catecholamines, which also have a vicinal diol structural moiety. Similarly designed compounds could also be used for the preparation of small molecule fluorescent tags for other saccharides, including polysaccharides such as sLex, sLea, Ley, and Lex. It needs to be noted that the template-directed polymerization as described above gives polymeric sensors. However, for this proposed project, we do not plan to make polymeric sensors. Instead, we plan to make small molecule fluorescent sensors or tags because of the intrinsic limits of polymeric materials for biological applications.
ii. Developing enantioselective fluorescent polymeric sensors of L-tryptophan using molecular imprinting (template-directed polymerization) techniques.

The method described above indeed could be very useful for the creation of fluorescent sensors for a wide variety of carbohydrates and other compounds that have a vicinal diol structure moiety. However, it cannot be used for the preparation of fluorescent sensors for compounds that do not possess a vicinal diol structural moiety. To develop a method which is generally applicable for the preparation of fluorescent sensors for compounds with a wide variety of structural features, we also developed a method of using an external quencher for the study of the binding of the target compounds to the fluorescent sensor prepared through template polymerization [See the reprint (88) provided in the Appendix for more details]. The functional monomer (13) used in this case for the imprinted polymer preparation (Figure 3) was a dimethacrylic acid-based monomer with a known fluorescent probe, the dansyl moiety, attached.

The preparation of the imprint polymer followed similar procedures as described in the previous section using L-tryptophan as the template (88). The polymers prepared were tested for their ability to recognize L-tryptophan selectively. We studied the effect of several fluorescent quenchers on the fluorescent emission intensity of the imprinted polymers. It was found that p-nitrobenzaldehyde almost completely quenched the fluorescent emission of the imprinted polymer at 10 mM, presumably due to the passive non-specific diffusion of the fluorescent quencher into the polymeric cavities. It was postulated that if L-tryptophan could bind to the imprinted polymeric cavities with fairly high specificity, it should be able to displace the p-nitrobenzaldehyde in the polymeric cavity and consequently diminish the effect of the quencher on the fluorescent emission intensity of the imprinted polymer. This was indeed the case. Upon mixing with L-tryptophan, the fluorescent intensity of the imprinted polymer in the presence of 3 mM of the quencher, p-nitrobenzaldehyde, significantly increased over that of the control. This fluorescent change was sensitive to L-tryptophan at single digit µM concentrations (88) (Figure 7).

It was also found that the effect of the template compound (L-tryptophan) on the fluorescence intensity is greater than that of its enantiomer (D-tryptophan), which is in turn greater than that of L-phenylalanine showing the specificity of the imprinted polymeric receptors [Please see the reprint (88) provided in the Appendix for more details]. It should also be noted that under identical conditions L-tryptophan at 10 mM did not have much of an effect on the fluorescent emission intensity (less than 5%) (88) of the control polymer, which was prepared under identical conditions except in the absence of the template, L-tryptophan. Because making molecular receptors using the molecular imprinting technique does not require prior knowledge of the three-dimensional structural features of the receptor, it should be feasible to extend the method under discussion to the preparation of fluorescent sensors of other molecules for enantioselective fluorescent detections (109, 110, 142).

iii. Developing enantioselective polymeric transporters of D-tryptophan using molecular imprinting techniques.

In addition to the preparation of fluorescent sensors, we have also used template-directed polymerization for the
preparation of polymeric transporters for separation purposes (89). Developing new methods for the separation of enantiomers is of great current interest because of the importance, challenge, and high cost associated with such separations. This is particularly true in the pharmaceutical industry because of the requirement for the high purity, including enantiopurity, of the final drug products (143). Using similar methods, polymeric receptors for D-tryptophan were prepared. The ability of these polymers to enantioselectively transport D-tryptophan across an organic phase was examined by using a U-shaped tube (89). All polymers exhibited selective transport of the print molecule over its enantiomer. For example, polymers prepared at with D-tryptophan as the template were able to transport D-tryptophan at a rate that was about 3-fold higher than that of L-tryptophan (89). Our studies demonstrated that molecular imprinting can be used to prepare enantioselective molecular transporters which can be used for the separations of enantiomers without chromatography. Serial enantioselective transports could potentially be used as a practical approach for the large scale separation of enantiomeric mixtures.

2. Drug delivery research

For many potent biologically active compounds, undesirable physicochemical properties are often a major cause of their low bioavailability due to poor stability and/or low permeability through biological barriers, such as the BBB (blood brain barrier) and the intestinal barrier. Therefore, such undesirable physico-chemical properties hinder the development of these potent and biologically active compounds as clinically useful agents. Finding solutions to these problems is a very contemporary issue, particularly because of the rapid development of biotechnology and the discovery of increasing numbers of biologically important peptides, protein, and peptide mimetics, which tend to have these undesirable pharmaceutical and biopharmaceutical properties (83-87, 90, 91, 144). We have developed a novel prodrug strategy aimed at optimizing the physicochemical characteristics (size, charge, lipophilicity, hydrogen bonding potential) of a compound for improved passive transport of drugs through such membrane barriers (90-92, 145-147).

The design takes advantage of the known facile cyclization of coumarinic acid derivatives (such as 15 and 19) due to the cis geometry of the side chain double bond which brings the side chain carboxyl group (or its derivatives) within close proximity of the phenol hydroxyl group (Schemes 2,3) (148-153). Using this system, we were able to prepare esterase-sensitive prodrugs of amines (93-96), peptides (97-101), and peptide mimetics (91, 102-104) with improved physicochemical properties and membrane permeability and well-controlled bioreversibility.

In studying the applicability of this prodrug system for the preparation of amine prodrugs, we synthesized a series of model amine prodrugs (14) and examined the structure-activity relationship. We have found that as designed, esterase could catalyze the release reaction (93). Furthermore, the structural bulkiness of the R group of the ester part (14) had very little effect on the release rates (95). We have also found that the release rates can be manipulated through the introduction of substituents on the phenyl ring (96).

In the development of prodrugs for peptides, we targeted opioid peptides. This class of compounds has tremendous therapeutic potential as analgesic agents but is faced with poor bioavailability problems largely due to their poor permeation through biological membranes such as the BBB and the intestinal barrier (154-156). We have prepared the cyclic prodrugs (97, 98, 101) of two opioid peptides, Leu-enkephalin (H-Tyr-Phe-Gly-Gly-Leu-OH) and DADLE (H-Tyr-Phe-D-Ala-Gly-D-Leu-OH), a metabolically stable analog of Leu-enkephalin. The permeabilities of these two cyclic prodrugs 18 were evaluated using monolayers of Caco-2 cells, an in vitro cell culture model of the intestinal barrier. The cyclic prodrug of DADLE (18a) was found to be about 31 fold as permeable through monolayers of Caco-2 cells as the parent peptide, DADLE, while the cyclic prodrug of [Leu]-enkephalin (18b) was more than 665 fold as permeable as the parent compound, [Leu]-enkephalin (97, 99, 100).
In studying the application of this prodrug strategy for the delivery of peptide mimetics, we have prepared the prodrugs of several RGD (Arg-Gly-Asp) analogs (Figure 8, 21-23) (91, 102, 103). These RGD analogs have known anti-platelet activities and have tremendous potentials for the treatment of myocardial infarction [for a review, see (157)]. For example, MK-383 (the active ingredient of tirofiban, 23, Figure 8) is a drug by Merck that has been approved for clinical use as an iv dosage form (158-160). However, it is generally true that these RGD analogs, which show potent in vitro, ex vivo, and in vivo (when administered intravenously) anti-aggregation activities, are not permeable or have very low permeability through the intestinal barrier [for a review, see (157)] and consequently low oral bioavailability. However, the cyclic prodrugs of these RGD analogs showed dramatically improved membrane permeabilities (102, 104). The cyclic prodrug of MK-383 also showed promising in vivo biological activities after oral administration to dogs (102). All of these clearly demonstrated the application potential of this prodrug strategy. Further pharmacological evaluations of these cyclic prodrugs are underway in different animal models.

One major concern with the development of a prodrug system is the toxicity of the end product of the prodrug moiety after the drug is released. In this case, it is coumarin (16) (Schemes 2 and 3). It should be emphasized that coumarin (16) is a naturally occurring compound found in sweet clover, tonka beans, and citrus oils and is widely used in cosmetics and perfumes (161, 162). Consequently, the toxicity of coumarin (16) has been studied extensively, and coumarin has been found to be relatively safe in humans in numerous clinical trials (161, 163-170).

3. Selected publications in related areas with those provided in the Appendix indicated with an asterisk:

i. Sensor development and molecular recognition.

ii. Drug delivery:

iii. Solid phase chemistry:

iv. Cell surface carbohydrate studies:
The key to the feasibility phase of the project is the development of one or more lead compounds that can bind to Slex with high affinity and selectivity. As briefly stated in the Background and Significance Section, we will use boronic acid moieties as the recognition motif for the vicinal diol structural features of the target polysaccharide, Slex. It is desirable that the binding of the boronic acid compounds is associated with an output of a detectable signal. Therefore, we plan to use fluorescent boronic acid compounds, which show a significant fluorescence intensity change upon binding with a vicinal diol, as one of the basic building blocks of the target sensors. In constructing the sensors, we propose to use an integrated approach by combining the strengths of template-directed synthesis, combinatorial chemistry, and computer molecular modeling. Therefore, the project would involve the design and synthesis of fluorescent and non-fluorescent boronic acid compounds as the basic building blocks (Specific Aim 1), the synthesis of Slex as the template (Specific Aim 2), template-directed synthesis of small libraries of potential sensors by cross-linking boronic acid moieties immobilized on the template (Specific Aim 2), and screening and structural characterizations (Specific Aim 3) of the potential fluorescent sensors synthesized. Computer molecular modeling will be used to aid the analysis of the optimal structural requirements for the sensor construction and the design of the linkers and boronic acid compounds needed (Specific Aim 4). Such an integrated approach will hopefully create a synergistic effect of all three technologies to be used in search of fluorescent receptors for the target carbohydrate, Slex. The computer molecular modeling work will be carried out in collaboration with Professors Stefan Franzen and Michael Prisant. Again, Professor Franzen is an expert in studying macromolecular structures including conformations, and Professor Prisant is an expert in the development of new algorithms in molecular modeling. Letters from Professors Franzen and Prisant are provided to indicate their willingness to collaborate. The initial screening will be carried out in solution using chemically synthesized Slex. Once an appropriate lead compound has been identified using the chemical screening method, it will be further evaluated in a cell culture system using colon cancer cell lines expressing Slex. This will most likely occur at the end of the first phase (R21) and the beginning of the second phase (R33) of the project. This evaluation will be carried out in collaboration with Professor Brent W. Weston, a pediatric oncologist and an expert in studying the expression of cell surface carbohydrates, at the University of North Carolina at Chapel Hill. A letter from Professor Weston is provided to indicate his willingness to collaborate on this project. During the R21 phase of the project, Dr. Weston will work on the selection and characterization of appropriate cell lines that express Slex and control cell lines (Specific Aim 5).

Specific Aim 1. Synthesis of a number of fluorescent and non-fluorescent arylboronic acid moieties.

Design. In designing the appropriate fluorescent or non-fluorescent boronic acid building blocks, several factors need to be considered. First, a boronic acid moiety binds with a vicinal diol most tightly when the pH is greater than its pKₐ (49-51). However, the pKₐ of an unsubstituted aryl boronic acid moiety is around 9-10 (171, 172). As a result, the binding is best at pH far above the physiological pH. This problem can be overcome through the introduction of electron-withdrawing groups on the phenyl ring or appropriately positioned lone-pair electron donors (see below) (75-78, 172). For example, while the pKₐ of phenylboronic acid is about 8.9 (171), the pKₐ of 3-pyridylboronic acid is 4.0 (171, 173, 174), about 3 pH units lower than the physiological pH. Therefore, boronic acid moieties with different pKₐ's can be used to fine-tune the binding affinity. Second, it has been reported that...
boronic acid moieties, when attached to certain fluorescent molecules, can significantly affect the fluorescence intensity of such fluorescent tags upon ester formation with vicinal diols \((80, 82, 113, 117, 138)\). One specific example is the conjugate of an anthracene moiety and a benzylamine moiety with the boronic acid attached at the ortho position \((141)\). Such a conjugate is known to show significant fluorescence intensity changes upon ester formation with vicinal diols \((77, 78, 175)\). As shown in Figure 4 (Preliminary Results Section c-1-i) in the free boronic acid form \((5)\), the nitrogen lone pair electrons can quench the fluorescence of the anthracene moiety through PET. However, upon ester formation \((7)\), there is an increased tendency for the nitrogen lone pair electrons to be donated to the open shell of the boron atom, which in turn suppresses the fluorescence quenching mechanism and results in a significant fluorescence intensity increase. It is also known that this appropriate positioning of a nitrogen, which can readily donate lone-pair electrons to the open shell of boron, also lowers the apparent pKₐ of an arylboronic acid to about 5.2 \((80, 136, 176)\). Consequently, binding of boronic acid \(5\) with vicinal diols can be readily achieved at physiological pH.

Building on our experience with fluorescent boronic compounds \((52)\), we first designed two types of boronic acid compounds which should show increased fluorescence upon binding with vicinal diols. The first one \((24a, Figure 9)\), as has been discussed in the Preliminary Results section, takes advantage of the reduced quenching of the anthracene fluorescence upon ester formation due to the masking of the nitrogen lone pair electrons \((52, 77, 78, 175)\). For the cross linking of the boronic acid moieties during template-directed synthesis, we plan to use imine formation followed by reduction (See details for Specific Aim 2) because imine formation is reversible and allows for an equilibrium-driven process, which should favor the synthesis of the target compounds that best complement the template \((55, 73)\). Therefore, an aromatic aldehyde is built into the arylboronic acid compounds as the handle for the reductive amination reaction (Figure 9). Substituents will be introduced at different positions for the synthesis of analogs (Figure 9). These substituents will help to fine-tune the apparent pKₐ's as well as the conformations of these boronic acid compounds, and provide diversity in the sensor construction.

The synthesis of compounds \(24a-j\) (Figure 9) can be accomplished by following well established procedures. Compounds \(25a-j\) \((X = -Br or -I)\) are all commercially available for the synthesis of \(24a-j\) (Scheme 4). The introduction of the boronic acid moiety to give \(26\) can be accomplished by following literature procedures using a Pd catalyst with a diboronic acid compound \((177, 178)\). Bromination at the benzylic position can be accomplished by following literature procedures \((141)\). It should be noted that we have conducted similar reactions in the preparation of other boronic acid compounds such as \(12\) \((141)\) (Scheme 1, Preliminary Results Section c-1-i). Then nucleophilic substitution with \(9\) (prepared in Scheme 1, Preliminary Results Section) should give \(28\). Deprotection of the TBS group using TBAF followed by MnO₂ oxidation should give the final product \(24\). It should also be noted that we have extensive experience in preparing similar boronates for a separate project, part of which has been described in the Preliminary Results section.
The design of the second fluorescent boronic acid compound (34) borrows heavily from a literature design that has a pyridylboronic acid moiety attached to anthracene (80). This compound is also known to have a very low pKa (about 4) and shows increased fluorescence intensity upon boronate formation (80). The synthesis can be accomplished by following the route described in Scheme 5. It should be noted that we have synthesized compound 33 for a different project and are one step away from the desired compound 34. These two types of fluorescent boronic acid compounds will serve as the “reporter” recognition site of the fluorescent sensors/tags.

Because only one fluorescent boronic acid moiety is needed for each diboronic fluorescent sensor, we also designed a series of non-fluorescent boronic acid compounds, which are easier to synthesize, to increase diversity. Since the relative spatial orientation of the boronic acid moieties is something that we would like to explore in the development of specific fluorescent tags, we have designed non-fluorescent arylboronic acid compounds (36-46, Figure 10) with different relative positioning of the boronic acid moiety and the linker attachment point (aldehyde). In this series, the aldehyde to be used for the reductive amination linkage is ortho to the boronic acid moiety. In this series, compounds (36-41) with an electron withdrawing substituent (-NO2) at different positions are designed. Again, it is known that the lowering of the pKa of the boronic acid moiety helps to increase its binding affinity for vicinal diols. Naturally, the introduction of an electron-withdrawing substituent, such as -NO2, will help to lower the pKa of the boronic acids. The design of the arylboronic acid compounds with the boronic acid moiety meta (42-43) and para (44-46) to the linking point followed similar principles.
For the synthesis of 36-46, either the substituted toluene compounds 47 or the corresponding benzoic acid analogs are commercially available. The methyl group of 47 can be easily oxidized using KMnO4 to give the corresponding benzoic acid compounds by following well-established literature procedures (179). The benzoic acids can then be converted to the ester 48 using anhydrous HCl/methanol. Reduction of the ester with DIBAL should give the aldehyde, which can be easily protected as an acetal. Similar to the preparation of other boronic acid compounds described in Scheme 5, the introduction of the boronic acid moiety can be accomplished by following well-established literature procedures using a Pd catalyst with a diboronate (177, 178) (Scheme 6). It should be noted that for a different project, we have synthesized many boronic acid compounds and we are very comfortable with the kind of chemistry proposed.

**Specific Aim 2.**  
Template-directed combinatorial synthesis of potential sLex artificial receptors by cross-linking two boronic acid moieties using small libraries of linkers.

**General design.** As briefly described in the Background and Significance section, we will use a template-directed combinatorial method for the construction of the potential fluorescent tags that can recognize sLex with high affinity and selectivity. Because sLex has two pairs of vicinal diols that can bind to boronic acid moieties, the sensor to be constructed will have two boronic acids linked together.
In executing the general idea, we plan to conduct the template-directed synthesis on a solid phase (see below for rationales). First we need to synthesize sLex on solid phase. Then, boronic acid moieties will be reacted with the immobilized sLex in a 2 to 1 ratio to give complex 50 (Figure 11). Then template-directed linking of two boronic acid moieties on the same sLex molecule will be carried out by initially using about 20 diamines as the potential linkers. There are many literature precedents demonstrating that using solid phase reactions will help to ensure that no intermolecular-cross linking, which would result in the undesirable cross-linking of two molecules of 50, will happen. Tour and coworkers have synthesized polymers on solid phase without the interference of intermolecular reactions (180). The initial linking reaction is the imine (51) formation. Because imine formation is a reversible process, this method will allow for the thermodynamic equilibrium of all possible combinations, however, it will favor the compounds that have the optimal complementary interactions with the target saccharide as suggested by Lehn and others (55, 73). In a non-equilibrium combinatorial library, the distribution of the library components depends on many factors such as the reaction pathway, the kinetic barriers, and the stability of the products. However, in an equilibrium combinatorial library, the stability of the products is the predominant determinant. In turn, the stability of the products depends on the complementary interaction of the products with the template. Therefore, such a design should intrinsically favor the generation of the compounds with the optimal complementary interactions with the template. Experimentally, template-directed syntheses of small organic compounds have numerous precedents that favor the products with the best complementary interactions with the template (56-63). After the equilibrium imine formation, the reversible linkage (imine) will be converted to an irreversible one (amine) through reduction to give 52. The two-step process constitutes a reductive amination. Hydrolysis will release the potential fluorescent sensor 53.

In building libraries of these compounds we need to (1) synthesize sLex on a solid support, (2) design potential linkers that would allow us to examine many potential distances and spatial orientations of the linked boronic acid moieties, and (3) design combinatorial approaches that will allow us to synthesize and screen as many compounds as we can within a reasonable period of time and yet allow us to identify and characterize the compounds that we consider “hits.”

1. Synthesis of sLex on a solid support.

sLex will be synthesized using both solution and solid phase approaches. The sLex immobilized on a solid phase will be used for template-directed synthesis of the target compounds. The compound prepared in solution phase will be used for the screening of the binding affinity (See below in Specific Aim 3).

Much work has been done by the groups of Wong, Danishefsky, and Schmidt in developing polysaccharide synthetic methods. For the solid phase synthesis of sLex, we will follow essentially identical literature procedures for the construction of these compounds. Because the reactions are all known literature reactions (181) with essentially identical starting materials and products, the synthetic part will only be described briefly with reference to the relevant literature procedures.

A linker moiety of Galβ(1-4)Glcβ(1-O) nonanoic acid common in Lex, Ley and sLex will be immobilized on the solid phase via an amide bond (57, Scheme 7). Starting with 54, 55 can be obtained in three steps following literature procedures (181). First the free hydroxyl groups of 54 can be acylated with benzoyl chloride (BzCl) in
pyridine (Pyr) in 95% yield. Then the hexyldimethylsilyl (TDS) will be cleaved by TBAF in THF at -20 °C in high yield. The resultant free hydroxyl group will be reacted with CC13CN and 1.8-diazabicyclo [5.4.0]undec-7-ene (DBU) to afford trichloroacetimidate 55 in quantitative yield with exclusive α configuration. The Galβ1-4Glcα1-O(=NH)CCl3 55 will be attached to the polystyrene-based resin with 56a in the presence of BF3•OEt2 at room temperature (RT) with highly controlled β configuration at the 1-position (57a). Compound 57b can be used for the solution phase synthesis of the target carbohydrate.

The synthesis of sLex will be carried out by following modified literature procedures (181, 182) (Scheme 8, next page). 1-O-Silyl-protected azidoglucose derivative 58 (183) will be reacted with 59 in the presence of TMSOTf (trimethylsilyl triflate) to give the corresponding disaccharide 60. The free hydroxyl group in 60 will be glycosylated with trichloroacetimidate 61. Deacetylation (NaOCH3/CH3OH) should give 62 quantitatively. The 3-OH with highest activity will be selectively glycosylated with sialyl donor 63 (184, 185) to afford tetrasaccharide 64. Compound 64 will be transformed into 66 in 4 steps [removal of the benzyl group and reduction of azide by hydrogenation, acetylation (Ac2O/pyridine), deprotection of TDS (TBAF), and acetimidation (CCl3CN/DBU)], then immobilized to solid-phase with 57a using BF3•OEt2 as catalyst. Deacetylation will be achieved by treatment with NaOCH3 in CH3OH to give the final product sLex. It should be noted that almost all reactions are essentially literature reactions and we anticipate minimal problems. The yield indicated in the scheme for each reaction is the literature reported yield for that reaction. At the end of the solid phase synthesis, a portion of the product will be cleaved, purified, and structurally characterized to ensure that the solid phase synthesis was successful.

Solution phase synthesis will also be carried out by following essentially the same procedure except 57b will be used instead of 57a. Purified sLex in solution will be used for the binding studies in solution (See Specific Aim 3). All compounds will be fully characterized with 1H-, 13C-NMR, MS, and elemental analysis.

2. Linker selections and template-directed synthesis.

The key to a successful tag is the appropriate positioning of two or more boronic acid moieties in the sensor molecule. Both the length of the linker and its conformational constraints will determine the spatial arrangement of the boronic acid moieties. The unique terminal structure of sLex has two pairs of vicinal diols (Figure 1). Therefore, we plan to use a template-directed combinatorial library approach in search of the optimal linker between the two boronic acid moieties. In such a strategy, the boronic acid compounds will first react with the target carbohydrates, sLex immobilized on a solid phase to give a boronate-carbohydrate complex (50, Figure 11, previous page). The cross-linking will initially use imine chemistry, because it is a reversible process, through the reaction with a small library of diamines to give complex 51 (Figure 11). Because of the template effect, the cross-linking should favor the formation of the compounds which best complement the target carbohydrate structures. The use of reversible imine formation chemistry will also help the formation of the compounds which best complement that template (55, 73).

In designing the libraries of linkers to be used, we will initially examine the optimal length of the linker by testing a library of diamines with different spacers in between. sLex has two pairs of vicinal diol structures in its unique terminal tetrasaccharide moiety. The distance between the two pairs of vicinal diols is about 15 atoms, depending on how it is counted. However, it is difficult to guess the relative orientations of these vicinal diols. Therefore, we chose linkers ranging from 2-14 atoms to start with (Figure 12, page 43). Each arylboronic acid moiety has 2-9
atoms between the boron atom and the aldehyde to be connected to the diamine linkers. This gives a range of possibilities with about 13-32 atoms separating the two boronic acid moieties in a molecule that contains at least one fluorescent boronic acid moiety. It is understood that this rough estimation serves only as a starting point for the random construction of compound libraries. When conformation and stereochemistry are considered, the linear distance in terms of atom numbers does not necessarily reflect the true picture. Further design of linkers will heavily rely on the screening results of the initial libraries. Furthermore, we will also use computer molecular modeling (Specific Aim 4) in combination with the screening results and available information on the conformation of sLex for the further design of linkers.

We recognize that the ultimate fluorescent diboronic acid receptor(s) probably will have a highly constrained spacer separating the two boronic acid moieties to maintain the desired three-dimensional structure(s). However, we need to approach this problem in a stepwise fashion. Therefore, at the beginning of the project, we plan to start with linear linkers with very little conformational constraints in search of the optimal length required for the linker moiety. Once a preference for a length is found, we will then impose more conformational constraints in search of the optimal spatial relationship of the two boronic acid moieties. For example, if the linker with two atoms separating the two amines were identified as the one with the best binding, we would use the conformationally restricted analogs 88-94 (Figure 13, next page) for further optimization. Compounds 95-99 represent some constrained diamines with a

Scheme 8. Synthesis of sLex.

a. (i) TMSOTf (0.01 eq.), Et₂O (85%); (ii) NaCNBH₃, HCl (80%); b. (i) TMSOTf, Et₂O (89%); (ii) NaOCH₃, CH₃OH (quan.); c. TMSOTf (0.02 eq.), CH₃CN, -40 °C (74%); d. (i) H₂/Pd/C, CH₃OH; (ii) Ac₂O, Pyr (93%); (iii) TBAF, THF, H⁺, -30 °C; (iv) CCl₃CN, DBU; e. 57a, BF₃·Et₂O, CH₂Cl₂/ n-hexane, -10°C; f. NaOCH₃, CH₃OH.

Sialyl Lewis X (sLex)

We recognize that the ultimate fluorescent diboronic acid receptor(s) probably will have a highly constrained spacer separating the two boronic acid moieties to maintain the desired three-dimensional structure(s). However, we need to approach this problem in a stepwise fashion. Therefore, at the beginning of the project, we plan to start with linear linkers with very little conformational constraints in search of the optimal length required for the linker moiety. Once a preference for a length is found, we will then impose more conformational constraints in search of the optimal spatial relationship of the two boronic acid moieties. For example, if the linker with two atoms separating the two amines were identified as the one with the best binding, we would use the conformationally restricted analogs 88-94 (Figure 13, next page) for further optimization. Compounds 95-99 represent some constrained diamines with a
three-carbon spacer and compounds 100-105 represent some constrained diamines with a four-carbon spacer. The number of possible conformational constraints becomes exponentially higher with the longer linkers. It is difficult to list all possibilities at present. There are many diamines that have been reported in the literature and there are many methods for the preparation of diamines that we could use in the future [e.g., (190)]. The linkers shown in Figures 12 and 13 can only be treated as examples, because further design of other linkers will rely heavily on the results generated from the first round of screening and our computer molecular modeling studies (Specific Aim 4). In Figures 9 and 10 and Schemes 4 and 5, there are 22 different boronic acid compounds, 11 fluorescent and 11 non-fluorescent. In Figures 12 and 13, there are about forty linkers. If in the initial phase, we will use one fluorescent and one non-fluorescent boronic acid in each library, this will give slightly fewer than 5000 possible combinations with each compound having at least one fluorescent boronic acid. Although 5,000 compounds would not be considered very large in a combinatorial library, this is a large number of compounds to be synthesized and tested individually. Therefore, we will have to build small libraries of compounds and be able to screen one library at a time so that we can examine this large number of possible compounds quickly.

![Figure 12. Commercially available linear amines.](image)

![Figure 13. Examples of constrained diamines.](image)

Specifically, for the synthesis of the diboronic acid compounds, one non-fluorescent and one fluorescent arylboronic acid compound will be added to a solution containing the solid beads bearing the sLex molecules in a 1:1:1 ratio. Although, we desire to have 1 fluorescent and 1 non-fluorescent boronic acid compound bind to each sLex molecule, there is more than one way for two boronic acid compounds to be attached to one sLex molecule. It is also understood that on a particular sLex molecule, there may not be a 1:1 ratio of the two different boronic acid moieties (the fluorescent one and the non-fluorescent one) as desired. There could be sLex molecules that contain two identical boronic acid compounds of either the fluorescent one or the non-fluorescent one. Upon boronate formation, the linear diamine linker compounds in Figure 12 (about 20) will be added into the solution with a ratio of
each linker to 50 (Figure 11, page 40) being 1:2, i.e., each linker compound has what is needed to convert 50% of the boronate complex (50) to 51 if that compound has the optimal structural features. This would result in a 10-fold excess of total amines. The solution will be equilibrated overnight in dry acetonitrile. It is known that imine formation with an aromatic aldehyde can be carried out readily under neutral conditions, as has been demonstrated in the synthesis of 5 (Scheme 1) (52). It should be noted that in a separate project, we have studied such imine formation reactions extensively and accumulated much first-hand experience. This equilibrium process should favor the formation of the di-imine compounds that best complement the template structures (55, 73). In performing such cross-linking reactions, there is also the concern of the cross-coupling of boronic acid compounds attached to different sLex molecules in an intermolecular process. The intermolecular coupling is unlikely to give compounds that can bind to sLex with high selectivity and affinity. However, it is known that by conducting the coupling reactions on a solid phase, the chance for intermolecular cross-coupling can be dramatically reduced due to its similarity to a very high dilution situation. The immobilization separates the sLex molecules and restricts their movement, which in turn makes it unlikely for two sLex molecules to get into the vicinity of one another. It has been demonstrated that at a loading degree of 1.0 mmol/g, polymers with a molecular weight of about 5,000 can be synthesized in a stepwise fashion on a solid support without the interference of intermolecular reactions (191). Then, reduction of the imine 51 (Figure 11, page 40) will be carried out with NaCNBH3 to give 52. After the reduction, the solvent will be drained and the resin will be washed with solvents such as CH2Cl2 or THF to remove the unreacted amines. The boronic acid compounds 53 will be released by washing with an acetic acid solution. It is known that acidic solutions facilitate the hydrolysis of boronates. This is also the method that we used in preparing other sensors involving boronate formation (52). After solvent evaporation, the mixture (53) will be screened for activity as outlined in Specific Aim 3. Computer modeling will also be used to analyze the screening results which could indicate a preference for an optimal linker length. It should be noted that we have extensive experience in doing solid phase chemistry through our work in developing novel resin linkers for solid phase peptide and organic synthesis (105, 106). Furthermore, we have recently started some molecular modeling work (Specific Aim 4). Results generated from the modeling work could result in modification of the design of the initial linker library so the combinatorial chemistry is more focused.

It should be noted that we recognize that we could also use two different fluorescent boronic acids for the library construction. However, we anticipate the synthesis of the fluorescent boronic acid compounds being more time-consuming than that of the non-fluorescent ones. Therefore, it would be more economical if we use the non-fluorescent boronic acids to generate diversity.

Specific Aim 3. Screening of the artificial receptors thus prepared for their binding to chemically synthesized sLex.

One of the keys of the project is the utilization of a screening method which allows us to quickly screen a large number of compounds and to be able to pick out the ones that have high affinity for sLex. To simplify the analysis, we will use small libraries. Each library will be the cross linking of one fluorescent and on non-fluorescent boronic acid. Because there are eleven fluorescent and eleven non-fluorescent boronic acid compounds, there will be 121 libraries. If we start with 20 dianimes to serve as the cross linkers, there will be 60 possible compounds per library. Of the 60 compounds, 20 will contain two non-fluorescent boronic acid structures and 40 will contain at least one fluorescent boronic acid moiety. Because we will screen the library by observing changes in fluorescence intensity upon binding sLex, only the 40 fluorescent compounds will be valuable to us. Across 121 libraries, there will be 2,640 unique fluorescent compounds for analysis. This number is smaller than 4,840 (121 \(\times\) 40) because the compounds with two identical fluorescent boronic acid structures will be common to all libraries containing that particular fluorescent boronic acid.

The initial screening will be carried out in solution phase one library at a time. A small portion of this reaction mixture from a particular library will be transferred to a cuvette for fluorescence measurements. The total concentration of the diboronic acid compounds will be kept at around 10^{-4} \text{M} which will be calculated by assuming a 100% reaction yield. In each small library, we will be screening about 40 fluorescent diboronic compounds as described above. With a total concentration of about 10^{-4} \text{M}, the concentration of each compound will be in the range of low \(\mu\text{M}\) assuming all compounds were formed in equal quantity. Assuming the template-directed synthesis favors the desired compounds, the compounds with high affinity will be formed in a larger quantity than those that have lower affinities. This will give at least a \(\mu\text{M}\) concentration of the compound(s) to be tested. Then sLex prepared through solution phase chemistry (Specific Aim 2) will be added and concentration-dependent (if any) changes of the fluorescence intensity will be examined. Libraries that exhibit low \(\mu\text{M}\) sensitivity toward sLex will
be kept for further analysis. If no library exhibits sensitivity at low μM concentration, the most promising ones will be analyzed.

The libraries exhibiting most promising results in the above assay will be further analyzed. The same mixture will be dissolved in a solution in the low μM concentration range and mixed with the resin beads bearing sLex. It is expected that the compounds that have sensitivities toward sLex at low μM concentrations will be bound to the immobilized sLex molecules. A fluorescence intensity change of the bead will be a good indicator of those beads that have target compounds bound to it. Those beads that strongly fluoresce after the binding will be separated and the boronic acid compounds will be released using an acidic solution. The released compounds will be analyzed with MS to identify the structures of the compounds, which are bound to the resin beads. The NCSU MS Facility has all the instruments needed for this study. Particularly, nanoflow rate electrospray mass spectrometry has the high sensitivity for molecular weight identification. With nanoflow electrospray, the sample is suspended in 0.5 to 1 mL of solvent. This volume will be electrosprayed into the mass spectrometer at approximately 50 nL/min. At this flow rate the sample will provide signal for more than 10 minutes. During this time instrument performance can be optimized, molecular weight can be determined and product ion spectra can be obtained. A letter from Dr. Carol Haney, Director of the NCSU MS Facility, is provided to indicate our capability and her willingness to help with the MS studies if needed.

Alternatively, HPLC analysis can be carried out to identify the major component. If the template-directed synthesis does favor the formation of the desired product, the major components should also have the best affinity for sLex. Then HPLC-MS can be used to identify the desired product. It should be noted that we have extensive experience with HPLC work in our drug delivery project (93, 95, 97). Once the molecular weight of a potential candidate is identified, the compound will be synthesized, and the pure form will be used for further study of its binding affinity for sLex in the solution phase assay described above. The K_d will be estimated by using the concentration that gives half of the maximum fluorescence intensity change.

Because each library only contains 40 fluorescent diboronic compounds, the identification of the most abundant components should not be very difficult. In the worst case scenario, if a library exhibited high sensitivity, but we were not able to identify the most abundant form using MS, HPLC-MS and other screening methods, the total synthesis of all forty possible fluorescent diboronic acid compounds through a one-step reductive amination should not take more than a few weeks. One of the reasons that we would like to work with very focused libraries is that the number of compounds to be analyzed would be very manageable.

Screening the libraries of initial design, coupled with our molecular modeling work (Specific Aim 4), hopefully will give the information needed with regard to the optimal linker length which will help to guide the further design of the next generation linker libraries.

**Specific Aim 4. Computer molecular modeling work to aid the further optimization of the fluorescent sensors.**

**General.** We are using template directed synthesis to prepare artificial receptors for specific carbohydrate moieties expressed in colon cancer. In this technique, the design of the artificial receptor begins with the target carbohydrate immobilized on the solid phase. The "jaws" of the receptor are formed with boronic acid moieties bound to vicinal diol structures on the carbohydrate. The spatial orientation of these moieties is determined by cross linking. This last step is accomplished through a combinatorial approach -- namely reacting the boronic acid complex with a library of cross linkers. Understanding the approximate spatial deployment of these moieties will help us focus the cross linker library used in this synthesis. Specifically, the design of constrained cross linkers for boronate sLex tetrasaccharide complex depends on understanding the relative geometric deployment of the two pairs of vicinal diol motifs in sLex. We will use molecular modeling techniques on this system to set semi-quantitative bounds on various signatures of the vicinal diol motif. The simplest signatures are simply the vicinal diol pair distances. These bounds can be used, in turn, to design more highly focused cross-linker libraries.

**Molecular Dynamics Studies.** Kollman and coworkers have described the use of molecular dynamics (192) (MD) in conformational searching. A short summary of the most relevant aspects of their work follows. MD studies augment the molecular mechanics studies in two critical ways: First they allow us to estimate the relative entropy of a given conformation. Second they allow us to embed the sLex system in a droplet of explicit water molecules as part of our search for conformational minima. Both of these factors are crucial in evaluating the relative free energy and hence relative likelihood of a given configuration.

There are two means of using MD to understand conformational likelihood. The first is to use an adapted MD conformational search procedure. The second is to simply run a long trajectory of the system of interest. Both approaches are described below -- the adapted search procedure in principle can sample conformational space more
efficiently. The straight trajectory approach provides a straightforward estimation of confirmation probability.

The basic idea of the adapted MD conformational search (AMDCS) is to run a series of relatively short, less than 0.5 ns, constant temperature MD trajectories of the molecular system of interest. Each trajectory is started with a different set of initial conditions. During each trajectory, structures are recorded at a time interval which is less than the average residence time that on the average the system will spend in a given local minimum at the temperature of the simulation. The recorded conformations are then further energy optimized in a straightforward manner to the minimum point of their local wells. These distilled system configurations can then be analyzed with respect to their potential energy and other properties. In particular duplicate structures can be eliminated by comparison of root mean squared atomic coordinate distances. The end result of this process is to determine a series of minimum energy structures and their relative energies. The entropy of each configuration can be estimated by considering the relative curvature of the local minimum basin associated with that conformation. This is accomplished with, e.g., the NMODE module of Amber that uses a rigid rotator harmonic oscillator approximation. The output from NMODE can be used to calculate the vibrational entropy that is the crucial contribution needed to estimate the relative population of a given configuration. It is difficult to estimate the entropy in the presence of solvent water. To allow an estimate of the vibrational entropy, the AMDCS approach can be run using an effective dielectric constant \( \varepsilon \) with no solvent (see below). The relative population of a given conformational minimum is given by:

\[
P(r_i) = k \exp(-F_i/RT) = k \exp((-E_i-TS_i)/RT)
\]

where \( k \) is a proportionality constant, \( R \) is the gas constant, \( F_i \) is the free energy of a configuration \( i \), \( E_i \) is the potential energy of the configuration, and \( S_i \) is the entropy of the configuration. We note that the last two quantities are specified by the simulation.

The straight trajectory approach (STA) simply performs a more extended, more than 2 ns, MD simulation on the system of interest. The time period of the straight trajectory simulations needs to be on the order of the approach to classical thermodynamic equilibrium. Configurations are sampled, minimized, and filtered as before. Now however, the frequency of occurrence of a given minimum configuration is a direct measure of its ensemble population. In their study of crown ether, Kollman finds that the population predictions of the AMDCS provide a reasonable accounting of observed configurational population in the straight trajectory. The AMDCS approach thus has an advantage over the STA in that the time required for the simulation to approach classical thermodynamic equilibrium may be quite long for larger systems in arbitrary starting configurations. However estimation of population probability is straightforward even in solvated systems.

Application to sLex. We will apply an AMDCS and STA -- similar to that of Kollman and coworkers -- to understanding the sLex minima and in particular the geometric deployment of the vicinal diol motifs. The AMDCS will treat a model system without explicit solvent where medium interaction is approximated by an effective dielectric constant \( \varepsilon \). In the STA the sLex system will be embedded in a droplet of explicit TIP3P water molecules.

Some notes regarding the system at hand follow: Woods and coworkers have developed the GLYCAM93 force field specifically for carbohydrate simulation in water (193). This force field is compatible with the AMBER suite of programs. Starting configurations for our studies can be obtained from two sources. First, Wong and co-workers (54) have reported 4 minimum energy structures GESA-A,B,C,D for sLex. Second, recent NMR work has postulated structures for sLex bound to different substrates (186).

A full description of the vicinal-diol relative geometry is given by the pair distance in conjunction with \( \theta_1 \) and \( \theta_2 \), the colatitudinal angles, and \( \Phi \), the relative azimuthal angle. In the AMDCS study we can simply consider the \textit{cis} diol pair geometry of selected minimum energy configurations along with their respective probabilities as a guide to selection of a spread of cross-linker molecular geometries. A comparable analysis can be accomplished with the results of the STA minimum energy structures. Moreover the STA can provide a probability surface \( P(r, \theta, \Phi) \) at a range of temperatures. This analysis would monitor these variables over the course of the entire trajectory and then bin their frequencies. Using the conformations of sLex thus obtained, similar studies can be carried out with the boronate complex (2, Figure 2), which will help the rational selection of possible linkers for the template-directed combinatorial chemistry work. Once a linker is known to bind with strong affinity for the sLex vicinal-diol pair, new classes of linkers based on more rigid structures can be tested using the modeling approach. This feedback will be used to choose new classes of linkers for combinatorial synthesis.

Specific Aim 5. Selection and characterization of sLex-expressing and control cell lines

Background. Selectin-mediated tumor cell adhesion is a biological process that has recently become a potential target for pharmacologic interruption of carcinoma invasion and metastasis. Tumor cell “rolling” has been modeled \textit{in vitro}, with carcinoma cells binding to cytokine-activated vascular endothelium (26, 194, 195). The fucosylated
ligand components sLex and sLea on the surface of circulating carcinoma cells have been shown to bind endothelial selectins (26, 194-201). The specificity of this interaction is supported by studies which show that anti-sLex (198) and anti-sLea (194) antibodies block adhesion of carcinoma cell lines to HUVEC (human umbilical vein endothelial cells), although the specificity of these antibodies is not absolute when compared to direct carbohydrate analyses.

Studies of carcinoma cell surface glycoproteins and glycolipids carrying sLex/sLea suggest more complexity in the definition of functional roles. Mannori et al (23) examined selectin-mediated adhesion of several mucinous and non-mucinous colon carcinoma cells that express sLex and/or sLea. COLO205 cells bind to recombinant purified E-, P-, and L-selectin; HT29 cells bind E-selectin but not P- or L-selectin. Treatment of cells with an O-sialoglycoprotease (which selectively cleaves mucin-type O-linked glycoproteins) decreases adhesion to purified P- and L-selectin. O-sialoglycoprotease did not significantly diminish adhesion of COLO205 or HT29 cells, suggesting that E-selectin is the more important receptor for their adherence. These findings also show that although E-selectin can recognize both mucinous (202, 203) and nonmucinous ligands, the latter group of fucosylated structures may be more critical. The role of carcinoma cell surface glycolipids modified by sLex/sLea is more complex (201), but these species are also closely associated with E-selectin interactions. When taken together, these data suggest that glycan expression on cells like COLO205 and HT29 could be targeted for inhibition, with widespread effects on complex ligand formation and on E-selectin mediated adhesion.

Carbohydrate-mediated interactions appear to be critically important in colon carcinoma metastasis to the liver (202, 203). Several recent experimental models support this hypothesis in the context of selectins and their ligands. Tail vein injections of transgenic mice constitutively expressing E-selectin demonstrate the primacy of sLex interactions to hepatic localization. In this model, B16F10 melanoma cells, which normally migrate to the lungs, were redirected to the liver when surface sLea expression was introduced (204). Neo-expression of fucosylated glycans alone (possibly without colon-specific surface proteins or lipids) appears to be capable of inducing high levels of adhesion to liver endothelial cells expressing E-selectin. Although constitutive hepatic expression of E-selectin is not physiologic, liver inflammation is a frequent occurrence in clinical oncology and forms a suitable microenvironment for carcinoma adherence (205-209).

Recent studies have replicated cytokine-induced liver inflammation in experimental models with nude mice and demonstrate selectin/ligand interactions in vivo (210). KM12-HX colon carcinoma cells, selected for high sLex expression, bind avidly to endothelial cells and colonize mouse liver more efficiently than KM12-LX cells, their unselected counterparts (30). Adhesion of KM12-HX cells is inhibited by antibodies specific for E-selectin. The treatment of KM12-HX cells with benzyl N-acetyl-a-D-galactosaminide, an inhibitor of O-linked carbohydrate chain extension, reduces the rate of adhesion by approximately 50%. Using the opposite approach, another colon cancer cell line was selected for high liver metastatic capacity (OCUC-LM1) and then found to express more surface sLea than its parental cell line (26). It has also been shown that intraperitoneal injection of monoclonal anti-sLea blocks carcinoma metastasis to the liver in nude mice (200). After examining several cell adhesion markers, expression of sLea was found to be the most significant predictor of hepatic metastasis in this model and others (26, 200, 210).

Clinical correlates demonstrate that sLea/sLex expression on carcinoma cells is associated with advanced stage disease and poor prognosis (26, 211-215). Greater surface expression of dimeric sLex, an aberrant difucosylated antigen (214), correlates with shorter survival time, presumably related to venous invasiveness and/or metastatic potential (214-216). Related glycans, including Lea, Lewis b (Leb), Lex, Ley, and VIM-2 (217-221), have not been studied as extensively in carcinoma cells (222, 223). In another report, metastatic liver specimens were shown to express higher levels of sLex and sLea when compared to their matched primary tumors, with highest expression on the advancing tumor edge (31, 222, 223). Pathologic studies have also correlated progression of colon carcinoma with surface expression of sLex, sLea, and aberrantly fucosylated species related to these epitopes (31, 214, 215, 222, 223). Increased expression of these glycans is independent of the epidermal growth factor receptor (EGF-R), carcinoembryonic antigen (CEA), and MUC2, which have all been reported to be over-expressed in colorectal liver metastases (26, 31, 224). A recent immuno-histochemical survey of human tumor specimens confirmed that cancers of the colon, pancreas, and stomach expressed high levels of sLex and sLea (31, 224, 225), regardless of underlying glycoprotein or glycolipid scaffolds, which differ in these carcinoma types.

Experimental Design. The proposal's immediate goal in the R 21 phase, as stated above, is to construct artificial fluorescent receptors to specifically identify these biologically important surface carbohydrates for future targeting experiments. In this Specific Aim, cell lines that express sLex, sLea, and other important fucosylated glycans will be screened to provide suitable assays for the synthetic products detailed in Aims 1-3. The prototypical molecule chosen for the R 21 phase is sLex. Colon cancer cells are chosen for study in this proposal. HT-29 cells express moderately high levels of sLex and
sLex (226). A variant cell line, HT-29LMM (214, 227, 228), expresses higher surface levels of fucosylated ligands. This cell line was derived from metastatic lesions resulting from intrasplenic injections of HT-29 cells in nude mice (227). Other colon carcinoma cell line variants selected for high metastatic potential have been reported to express more sLex and FUT3 transcript than parental cells (229), but much less is known about their genotype and phenotype. The invasive and tumorigenic colon carcinoma cell line COLO-205 expresses high levels of sLex/sLea on complex structures, modifying specific glycolipids as well as N-linked and O-linked glycoproteins (23, 230-234). Recent studies have shown that sLex- and sLea-containing structures have more functional significance than other glycans expressed on carcinoma cells (194, 195, 235) and are likely responsible for the invasive and tumorigenic phenotype of COLO-205. Several different glycoproteins and glycolipids have been associated with proliferation and adhesion of colon carcinoma cells (23, 230-234, 236-241). Our choice of target cell lines to assess the specificity of the artificial receptors synthesized in Aims 1-3 will thus include cells that express these scaffolds (e.g., HT-29 and COLO-205) and those that don’t (e.g., CHO transfectants).

### Table 1. Preliminary and expected FACs analysis results of surface glycan expression in stable transfectant cell lines constructed for this application.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>sLex</th>
<th>sLea</th>
<th>Lex</th>
<th>Ley</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHO-FUT1*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+++</td>
</tr>
<tr>
<td>CHO-FUT3</td>
<td>+++</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CHO-FUT1/3*</td>
<td>+++</td>
<td>-</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>CHO-FUT4</td>
<td>-</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CHO-FUT7*</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>B16F10-FUT3*</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Indicates cell lines to be constructed and expected results.

Table with cDNAs encoding the human α(1,2)-, α(1,3)-, and α(1,4)-fucosyltransferase (FUTs) as indicated. Methods for transfection, cloning, and selection of cells expressing the desired antigens are previously described (204, 242-246). FUT1, the H blood group α(1,2)-FUT, has been reported to be over-expressed in progressive colon carcinomas (247); functional significance for this observation is presumed to be dependent on Ley expression. Biosynthesis of Ley requires coordinate action of FUT1 to generate the H type II precursor Fucα1->2Galβ1->4GlcNac, which is then fucosylated on the GlcNac moiety in α(1,3)-linkage by FUT3. CHO-FUT1/3* would not be expected to make sLea, because there are no type I acceptors in parental CHO. Conversely, B16F10 doesn't appear to have Type II acceptors, so no sLex, Lex, or Ley is likely to be expressed when transfecting FUT3. These cell lines, in aggregate, should allow specificity of the artificial receptors to be assayed.

It is not known if CHO-FUT1/7* would make Ley or not. This cell line could have the added advantage of not making Lex in addition to the other type II molecules. FUT6 could be transfected if VIM-2 becomes a specificity problem with Ley or Lex in the FUT3 or FUT4 constructs, since it doesn't construct VIM-2 in CHO cells (245). Another approach to distinguish between Type II molecules would be to use the LEC11 (Lex and sLex +) and LEC12 (Lex + only) CHO mutant system; these spontaneously arising mutants have been well-characterized (248, 249). Other alternative methods that may be required include different expression vectors (204); additional antibody reagents to increase specificity (16, 195, 226, 242, 247); fucosyltransferase assays and RT-PCR to confirm transfectant phenotypes (242-246, 250); neuraminidase and/or fucosidase digestion followed by HPLC with NMR standards (201, 242) to confirm carbohydrate structures on transfected cells. If more detailed chain analyses are required, HPAEC with previously described reagents will be performed (250).

### Table 2. Expression of surface glycans in targeted carcinoma cell lines.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>sLex</th>
<th>sLea</th>
<th>Lex</th>
<th>Ley</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>HT-29</td>
<td>++</td>
<td>++</td>
<td>+/-</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>HT-29LMM*</td>
<td>+++</td>
<td>++++</td>
<td>+/-</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>COLO205*</td>
<td>++</td>
<td>+++</td>
<td>+/-</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>HEPG2*</td>
<td>+++</td>
<td>-</td>
<td>+</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Capan2</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

*Functional models exist for these cell lines.

### Protocol 1. Although little preliminary information regarding surface expression of sLex and related structures has been specifically developed for this proposal, our laboratory has considerable experience using recombinant glycosyltransferases to synthesize these glycans (242-245). Our first step is to modify existing stable transfectant cell lines that do not express carcinoma-specific glycoproteins or glycolipids:

#### Methods. CHO and B16F10 parental cell lines, which don't express any of these carbohydrate antigens or the fucosyltransferases responsible for their biosynthesis, are transfected with cDNAs encoding the human α(1,2)-, α(1,3)-, and α(1,4)-fucosyltransferase (FUTs) as indicated. Methods for transfection, cloning, and selection of cells expressing the desired antigens are previously described (204, 242-246). FUT1, the H blood group α(1,2)-FUT, has been reported to be over-expressed in progressive colon carcinomas (247); functional significance for this observation is presumed to be dependent on Ley expression. Biosynthesis of Ley requires coordinate action of FUT1 to generate the H type II precursor Fucα1->2Galβ1->4GlcNac, which is then fucosylated on the GlcNac moiety in α(1,3)-linkage by FUT3. CHO-FUT1/3* would not be expected to make sLea, because there are no type I acceptors in parental CHO. Conversely, B16F10 doesn't appear to have Type II acceptors, so no sLex, Lex, or Ley is likely to be expressed when transfecting FUT3. These cell lines, in aggregate, should allow specificity of the artificial receptors to be assayed.

It is not known if CHO-FUT1/7* would make Ley or not. This cell line could have the added advantage of not making Lex in addition to the other type II molecules. FUT6 could be transfected if VIM-2 becomes a specificity problem with Ley or Lex in the FUT3 or FUT4 constructs, since it doesn't construct VIM-2 in CHO cells (245).

Another approach to distinguish between Type II molecules would be to use the LEC11 (Lex and sLex +) and LEC12 (Lex + only) CHO mutant system; these spontaneously arising mutants have been well-characterized (248, 249).

Other alternative methods that may be required include different expression vectors (204); additional antibody reagents to increase specificity (16, 195, 226, 242, 247); fucosyltransferase assays and RT-PCR to confirm transfectant phenotypes (242-246, 250); neuraminidase and/or fucosidase digestion followed by HPLC with NMR standards (201, 242) to confirm carbohydrate structures on transfected cells. If more detailed chain analyses are required, HPAEC with previously described reagents will be performed (250).

### Protocol 2. Moving toward carcinoma cells that express sLex and related antigens in the context of complex glycoprotein and glycolipid scaffolds, additional characterization of cells for expression of α(1,2)-fucosylated structures is required:

#### Methods. The cell line HT-29LMM has recently been well-characterized for its expression of sLex/sLea, adhesion to HUVEC (human umbilical vein endothelial cells), and metastasis to the liver in a nude mouse model (243)
COLO205 cells have been studied for sLex-associated proliferation in vitro and in vivo (251). HEPG2, an hepatocellular carcinoma line, expresses sLeA alone. Capan2, a pancreatic carcinoma cell line (250), doesn't have type II acceptors (Weston, et al, unpublished data); it is useful as a sLex-negative sLeA-positive line. Glycoprotein and glycolipid scaffolds have been described for these lines as well (e.g., HT-29 express CEA, COLO-205 express CD43, HEPG2 express α-1-acid glycoprotein). The pattern of sLex and sLeA expression varies in these lines according to the underlying N-linked and O-linked structures, as well as glycolipid antigens (23, 230, 231).

Although expression of sLex and sLeA are well defined for these cells, including the FUTs responsible for their biosynthesis, these "native" carcinoma lines will be studied further with monoclonal antibodies directed at H type II structures (16, 247, 252). We will begin with previously described flow cytometric methods (253). Given previous studies showing limited specificity of anti-H and anti-Ley antibodies, multiple reagents will likely be required to confirm the presence or absence of these structures. In addition, neuraminidase and fucosidase treatment followed by HPLC with NMR-proven standards will be used to verify carbohydrates identified by immunologic procedures (as above).

II. Milestone

The goal of the first phase of this project is the demonstration of the feasibility. A measurable milestone of this phase is the identification of at least one good lead compound which has high affinity (sensitive at least in the low µM range) for the target saccharide, sLex. Then, in the second phase of this project (R33 phase), this compound will be further evaluated for its specificity for the target saccharide in the presence of other saccharides including other cell surface saccharides, such as Lex and Ley. Further optimization of such sensors will be carried out at the same time as the biological evaluation using cell culture systems.

III. R33 Phase

a. Specific Aims

After the synthesis and identification of the lead compounds that have high affinity for the target carbohydrate, sLex, it will be important to examine how well they work with cancer cells over-expressing sLex and whether they have the desired selectivity among other carbohydrates including cell-surface carbohydrates. To achieve a practical fluorescent tag, it will also be important to fine-tune the lead compound(s) to further improve their affinity and selectivity. We plan to achieve these goals by pursuing the following specific aims.

Specific Aim 1. Examination of the binding and fluorescent tagging of other carbohydrates commonly found in the blood, such as D-glucose, D-fructose, D-mannose, and sucrose.

Specific Aim 2. Examination of the binding and fluorescent tagging of other carbohydrates commonly found on cell surfaces, such as Lex and Ley.

Specific Aim 3. Examination of the binding and fluorescent tagging of cancer cells that express sLex.

Specific Aim 4. Further refining of the structures of the lead compound(s) including computer molecular modeling aided modification of the fluorescent tags.

Such studies will help to address the practical problems commonly encountered in real analysis; i.e., specificity and sensitivity in biological assays.

b. Background and Significance

In the developmental phase (R33) of the project it will be necessary to study the issues related to the practicality of the fluorescent tags thus prepared. It is also necessary to use such results to guide the further modification of the fluorescent tags for optimal sensitivity and specificity. Finding a fluorescent tag which has high affinity for the target carbohydrate is only the first step. It is equally important that the fluorescent tag(s) has the desired specificity so that other carbohydrates in a biological system would not interfere with the application of the technology. Possible interference may come from the carbohydrates in the blood and other cell surface carbohydrates. We plan to test the binding of the fluorescent tags with D-glucose, D-fructose, D-mannose, D-galactose, and sucrose (the binding to sucrose is expected to be weak because of the lack of cis-diols on the ring structures), which are commonly found in the blood (254, 255). In terms of carbohydrates found on cell surfaces, there are again many possibilities. However, we will initially focus on interference due to binding with Lex and Ley, which are also implicated in cancer.
development and other pathological processes (1, 2, 4, 14, 15, 256-262). To accomplish this part, we plan to chemically synthesize Lex and Ley and use them for binding studies. The same fluorescent tags will also be tested using cell-culture systems so that the practicality study will go beyond chemical systems. For the further modification of the fluorescent tags, we will rely on computer aided design for the synthesis of more selective fluorescent tags specifically for sLex. These studies will help evaluate the practicality of the fluorescent tags thus developed. Furthermore, the specificity information thus generated could also help us in the design of fluorescent tags for other carbohydrates implicated in other cancers or diseases.

d. Research Design and Methods

General. In considering the practical application of the sensors proposed in this application, one should consider several factors. The first factor to be considered is the specificity of the sensors for the target carbohydrate. Because boronic acid moieties are known to recognize carbohydrates through boronate formation and there are many different types of carbohydrates in the biological system, the interference in the binding by non-target carbohydrates will be a real issue in the application of the proposed sensors. There have been literature reports that proper positioning of two boronic acid moieties, reduced the non-specific binding of each monomeric boronic acid functionality resulting in improved specificity (76-78). Similarly, appropriately positioned diboronic acid moieties should also afford the opportunity for the construction of sensors with not only high affinity but also high specificity for our target carbohydrate. Therefore, it is critical that we examine this issue as we move forward in the developmental phase of the project. The interference of the binding by other carbohydrates can be divided into two classes. The first one is the interference by carbohydrates which come from food and supply energy to cells, mostly some common monosaccharides and disaccharides, such as glucose, fructose, galactose, sucrose, and lactose [e.g., (254, 255)]. The other class is cell surface carbohydrates as part of glycoproteins or glycopeptides (1-4). Examination of the interference by these carbohydrates in chemical systems will help the mechanistic understanding and the further design and optimization of the proposed fluorescent sensors (Specific Aims 1 and 2). Similar examinations using cell culture systems will help to test how such sensors will work in real biological systems (Specific Aim 3). We will conduct both. The information obtained from such binding studies will be used for the further optimization of the sensors (Specific Aim 4). Again, during the optimization process, we will continue to use an integrated approach of combining the power of template-directed synthesis, combinatorial chemistry, and computer molecular modeling.

Specific Aim 1. Examination of the binding and fluorescent tagging of other carbohydrates commonly found in the blood, such as glucose, fructose, sucrose, and mannose.

It is well known that there are many saccharides in the blood and inside cells [e.g., (254, 255)]. Some of these saccharides, such as D-glucose, could reach low mM concentration in the blood. Although D-glucose normally does not bind to boronic acid very well (77, 78), at high concentrations (mM) it could still be an important interfering factor for the in vivo application of this sensor and targeted drug delivery. D-fructose, also commonly found in the blood in high mM concentrations, could present a problem because most boronic acid compounds do bind D-fructose very well. However, the affinity of boronic acids for a particular monosaccharide can be manipulated through the introduction of a second boronic acid moiety in a molecule (76-78). It will be very important to see whether these non-target carbohydrates interfere with the sLex sensors selected in the R21 phase.

The binding studies will be carried out by following the same methods as described in Specific Aim 3 of the R21 phase. Briefly, the sensor selected in the R21 phase will be dissolved in acetonitrile at $10^{-5} \text{M}$. The concentration-dependent fluorescence intensity changes will be examined with D-glucose, D-fructose, D-mannose, lactose, and sucrose. The $K_d$ will be estimated by the concentration giving half of the maximum fluorescence intensity change.

Specific Aim 2. Examination of the binding and fluorescent tagging of other carbohydrates commonly found on cell surfaces, such as Lex and Ley.

For the examination of the binding interference by other cell surface carbohydrates using chemically synthesized carbohydrates, we selected Lex and Ley because of their implications in many biological and pathological processes (1, 2, 4, 14, 15, 256-262). We recognize that there are many other cell surface carbohydrates that could also interfere. In the evaluation with cell cultures, the interference by sLea will also be examined (Specific Aim 3). To accomplish the goals of this specific aim, we will first need to synthesize Lex and Ley.

Such synthesis can be accomplished by following literature procedures (181). The synthesis of Lex will start
with 106 (263). (Scheme 9). Regioselective O-benzoylation of 106 can be achieved by following literature procedures. The primary hydroxyl groups in 106 can be acylated with 2.2 eq. of benzoyl cyanide in the presence of triethylamine (TEA) at -40°C to give di-O-benzoyl derivative 107 in high yield. An additional equivalent of benzoyl cyanide/TEA at -40°C will give the tri-O-benzoyl product 108. Fucosylation with trichloroacetimidate 59 in the presence of ZnCl₂•OEt₂ as catalyst (0.25 eq.) at RT will provide trisaccharide 109 in 97% yield with the desired configuration. Following debenzylation and reduction of the azide group by hydrogenation using Pd/C as the catalyst in CH₃OH, acetylation with Ac₂O and pyridine followed by deprotection of the silyl (TDS) protective group in 109 with TBAF and acid at -30°C, the free hydroxyl group can be reacted with CCl₃CN and DBU to afford trichloroacetimidate 110. Compound 110 can be reacted with 57b in the presence of BF₃•OEt₂ in CH₂Cl₂ and n-hexane to give 111 in high yields. The intermediate 111 can be transformed by a high yielding two-step sequence (TFA-catalyzed removal of isopropylidene, and deacylation with NaOCH₃) into the target Lex. All yields given in the scheme are literature yields.

The synthesis of Ley (Scheme 10) will then start with the fucosylation of di-O-benzoyl derivative 107 with trichloroacetimidate 59 (2.3 eq.) in the presence of ZnCl₂•OEt₂ as catalyst at RT to give tetrasaccharide 112 in 93% yield with high stereoselectivity (18I). Debenzylation and reduction of azide by hydrogenation in MeOH followed by acetylation with Ac₂O and pyridine will give 113. After deprotection of the TDS group in 112 with TBAF and acid at -30°C, the free hydroxyl group will be reacted with CCl₃CN and DBU to afford trichloroacetimidate 110 in high yields. Then 110 will be reacted with 56 in the presence of BF₃•OEt₂ as catalyst in CH₂Cl₂ and n-hexane to give 114 in 69% yield. The intermediate 114 can be transformed by a high yielding two-step sequence (TFA-catalyzed removal of isopropylidene and deacylation with NaOCH₃) into the target Ley.

These carbohydrates can be used for the binding studies using the same procedures described in Specific Aim 1 of the R33 phase and Specific Aim 3 of the R21 phase.
Specific Aim 3. Examination of the binding and fluorescent tagging of cancer cells that express sLex.

General. Stable transfectant and "native" carcinoma cell lines that have been characterized according to the R21 studies above will be tagged with the optimized artificial receptor(s) directed at sLex. By this point, the affinity of the optimized fluorescent sensors for sLex and their non-specific affinity for Lex, Ley, and other mono- and disaccharides will have been examined with chemically synthesized carbohydrates. In this specific aim, the ability of these sensors to specifically tag sLex in a complex membrane setting will now be tested using cells expressing sLex alone, cells expressing other carbohydrates together with sLex, and cells that do not express sLex (Please see Tables 1,2 in Specific Aim 5 of the R21 phase).

Methods. Cells will be plated in duplicate using previously described conditions for HT-29LMM and allowed to grow overnight (243). The media will be replaced with serum-free staining media (242) and concentrations of fluorescent tag ranging from 0.1 nM to 50 µM. After practical concentrations of artificial receptors have been determined, incubation times will be optimized. Cells will be assayed for viability throughout (243). Control plates will include serum-free staining media without tag and FITC-conjugated monoclonal antibodies directed at sLex and related glycans. The cells will first be observed and photographed using fluorescent microscopy as previously described (264). Distribution and intensity will be noted. Fluorescence intensity will be further quantified by flow cytometry using modifications of previous protocols (242, 243, 253).

Future studies. Successful identification of sLex positive carcinoma cells will allow us to incorporate these tags into functional colon cancer models, which include selectin-mediated cell adhesion to endothelial cells, metastasis to the liver in nude mice, as well as our recently developed proliferation and tumorgenesis assays. Design of such experiments awaits validation of the tags outlined in the R 21 section of this proposal.

Specific Aim 4. Molecular modeling aided further optimization of the fluorescent sensors.

The molecular modeling work will be essentially the same as that described in Specific Aim 4 of the R21 phase. Because the future design of more compounds heavily relies on the results generated in the first phase, it would not be possible to identify specific structural characteristics for further optimization. However, we have extensive experience in the area of molecular recognition and sensor development and we are confident and stand ready to fully utilize the information gained by that point for the further design of libraries of linkers to achieve optimal results.

e. Human Subjects
No human subjects involved.

f. Vertebrate Animals

No vertebrate animal experiments involved.

g. References


54. Ichikawa, Y.; Lin, Y.C.; Dumas, D.P.; Shen, G.J.; Junceda, E.; Williams, M.A.; Bayer, R.; Ketcham,


180. Huang, S. and Tour, J.M., Rapid Solid-phase Synthesis of Oligo(1,4-phenyleneethynylene)2 by a


251. Hiller, K.M.; Mayben, J.P.; Benk, K.M.; Manousos, G.A.; Cameron, H.S.; and Weston, B.W., Transfection of α(1,3)Fucosyltransferase Antisense Sequences Impairs the Proliferative and...
h. Consortium/Contractual Arrangements.

As described in the application, the biological evaluation (Specific Aim 5 of the R21 phase and Specific Aim 3 of the R33 phase) of the artificial sensors to be synthesized will be subcontracted to Professor Brent Weston, University of North Carolina at Chapel Hill. Dr. Weston is a pediatric oncologist and an expert in studying cell surface expression of carbohydrates implicated in cancer development. A letter from Dr. Weston is provided to indicate his interest in this collaboration. An authorized official letter from the University of North Carolina is also provided to indicate the institutional commitment to this project.

i. Collaborators and Consultants

As described in the application, the computer molecular modeling work (Specific Aim 4 of the R21 phase and Specific Aim 4 of the R33 phase) will be conducted in collaboration with Professors Stefan Franzen and Michael Prisant of the Department of Chemistry, North Carolina State University. Professor Franzen is an expert in studying macromolecular structures, including their conformations, and Professor Prisant is an expert in developing new methodologies in computer molecular modeling. Their participation in this project will help ensure that the molecular modeling work will be conducted in the most expeditious fashion. Letters from Professors Franzen and Prisant are provided to indicate their interest in this collaboration.

The project also involves using MS for the identification/detection of potential fluorescent sensors. A letter from
Dr. Carol Haney, Director of the NCSU MS Facility, is provided to indicate the capability of the MS Facility and her willingness to help with the MS component.