SYNTHESIS AND ANTIVIRAL EVALUATION OF NUCLEIC ACID BASED (NAB) LIBRARIES

The present invention provides for compositions and methods to generate a library of small molecules that mimic the repertoire of interactions that exist amongst nucleic acids, and proteins, as well as, that exist between proteins and nucleic acids. The designed library of nucleoside analogs are for use as competitive and non-competitive viral replication inhibitors or serve as pro-drugs of inhibitors in anti-viral drug discovery. A facile parallel solution-phase chemistry has been developed to carry out the combinatorial synthesis of nucleoside analogs as a source of pharmaco-core-based chemical diversity.
SYNTHESIS AND ANTIVIRAL EVALUATION OF 
NUCLEIC ACID BASED (NAB) LIBRARIES

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to the design and synthesis of focused nucleic acid based (NAB™) libraries with different diversity attributes. More specifically these libraries are useful for their antiviral activity and for rapidly identifying novel molecular targets for antiviral intervention.

2. Background

The discovery of safe and effective antiviral drugs presents a formidable challenge compared to bacterial and parasitic agents. Indeed, very few virus-specific molecular targets have been identified that can be specifically subjected to antiviral intervention, because viral metabolic processes closely resemble host cellular processes. Nevertheless, three virus-encoded enzymes have been the targets of most “small molecule-type” antiviral drugs - polymerases, proteases, and most recently, neuraminidase. However, the rapid emergence of resistance to antiviral drugs is a major problem and unwieldy “cocktail regimens” often need to be employed as a desperate measure. Clearly, there exists a substantial unmet clinical need for antiviral drugs with different structures and unique mechanisms of action.

Historically, antiviral drugs have been designed using both mechanism, and structure-based drug design approaches. The combinatorial methodology is emerging as a powerful contemporary drug discovery tool, and consists of two steps in which: (a) biological methods are used to select and validate molecular targets, and (b) structure-, and mechanism-guided drug design approaches are used, in which a library of compounds are synthesized and evaluated for their ability to interfere with the biosynthesis, structure and/or function of the target. When insufficient structure and/or function of a target is available, combinatorial methods such as “diversity oriented organic synthesis for therapeutic target validation” or alternatively “combinatorial target-guided ligand assembly” have been employed.
The growing understanding of the molecular biology and biochemistry of viral replication as well as of related protein structure has facilitated the development of potent and specific antiviral drugs. Many nucleoside analogs have been found to have effective and specific antiviral activity in the past two decades, and a number of compounds in this class have enjoyed considerable clinical and commercial success for the treatment of viral diseases. The primary mechanism for the antiviral drug action of the nucleoside analogs was by acting as the chain terminators or as competitive inhibitors of viral-related enzymes or both; recently an alternative strategy has been the development of agents that act in their own right at an allostatic site as noncompetitive inhibitors.

An approach that seems appropriate for antiviral drug discovery, is the use of structurally diverse compounds to modulate biological pathway without regard to specific molecular target. This allows simultaneous functional validation of a target, as well as, the discovery of a lead structure that modulates the function of the target. We describe here the application of this concept for antiviral drug discovery.

SUMMARY OF THE INVENTION

The present invention provides for compositions and methods to generate a library of small molecules that mimic the repertoire of interactions that exist amongst nucleic acids, and proteins, as well as, that which exist between proteins and nucleic acids. Indeed, a number of proteins contain nucleotide-binding domains defined by the topology of protein α-helices and β-sheets. More specifically, a nucleic acid-based (NAB™) scaffold is described into which such diversity attributes are engineered to potentially target “hot spots” in protein-protein, and protein-nucleic acid interaction surfaces. As an illustrative example of an embodiment of the invention, we report here the synthesis and antiviral evaluation of NAB™ libraries against herpes simplex virus (HSV-1) using cell-based assays.

The invention provides for drug discovery in anti infectious agent therapeutics, by constructing libraries of nucleoside analogs in which a variety of functional and structural elements are attached to the nucleic acid-based (NAB™) scaffold through proper linkers. This pharmacore-based chemical class is intended to act as the competitive or noncompetitive viral replication inhibitor or serve as a prodrug of inhibitors.
The advantages of the present system are many. For example, there are many novel diversity features associated with libraries built around a NAB™ scaffold: (a) the scaffold can be used to create variable spatial display of hydrogen-bonding, hydrophobic, charge-transfer, electrostatic and such other non-covalent interactions; (b) the scaffold can be conformationally rigid or flexible and by linking the individual scaffolds together, one can fashion diverse molecular topology into library members; indeed, such shape-defining motifs as circles, pseudoknots, bulges and stem loops can be incorporated into libraries to target “hot spots on the receptor; (c) From a synthetic perspective, the NAB™ libraries can be assembled using well established solid-phase or solution-phase methods in nucleic acids field.

The exquisite specificity of such interactions reside in topology-associated molecular recognition defined by local and global conformations of the ligand and receptor, and a network of hydrogen bonding, hydrophobic, ionic, and van der Waals interactions that facilitate the binding interactions.

Examples of suitable compounds to generate a compound library are disclosed. For example, a compound library comprising two or more compounds of the following Formula I or I':

wherein L₁ and L₂ are independently O or a linking group such as e.g. an amide, ester, diester or the like, or an optionally substituted alkylene (e.g. C₁₋₂₀ alkylene), optionally substituted alkenylene (e.g., C₂₋₂₀ alkenylene) or alkynylene (e.g., C₂₋₂₀ alkynylene) having such groups either as a chain member of pendant to the chain, and which may be optionally...
substituted with one or more substituents selected from a group consisting of O, S, Se, 
NR¹NR², CR¹CR², OR, SR and SeR, or an enzymatically reactive;

Q is carbon or a heteroatom such as O, S or N;

R¹, R², R³ are each independently a hydrogen or a hydroxyl group or an optionally 
substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally 
substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, 
optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an 
optionally substituted polynucleotide, or an optionally substituted heteroaromatic or 
heterocyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or 
S atoms;

B is optionally substituted adenine, optionally substituted thymidine, optionally 
substituted cytosine or an optionally substituted guanine, preferably where the optional 
substituents are alkyl, carbocyclic aryl, or heteroaromatic or heterocyclic group preferably 
having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic 
structure that is covalently linked to the sugar ring;

n=1 to 5;

and pharmaceutically acceptable salts thereof.

In another preferred embodiment, the compound library is comprised of two or more 
compounds of the following Formula II or II':

![Diagram]

wherein X and Y are each independently selected from a group consisting of O, S, Se, 
NR¹NR², CR¹CR², OR, SR and SeR, or one or both of X and Y are an enzymatically reactive 
moiety;

R is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, 
optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted
cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an
optionally substituted mononucleotide, an optionally substituted polynucleotide, or an
optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3
separate or fused ring and 1 to 3 N, O or S atoms;

\[ R^1, R^2 \text{ and } R^3 \text{ are each independently selected from a group as defined by } R; \]

B is optionally substituted adenine, optionally substituted thymidine, optionally
substituted cytosine or an optionally substituted guanine, preferably where the optional
substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably
having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic
structure that is covalently linked to the sugar ring;

and pharmaceutically acceptable salts thereof.

In another preferred embodiment, at least one of the compounds that comprise the
compound library, has at least one furanose ring in the C-2' endo conformation or has at
least one furanose ring in the C-3' endo conformation.

In another preferred embodiment, at least one of the compounds that comprise the
compound library, has at least one furanose ring in the C-2' endo conformation and at least
one furanose ring in the C-3' endo conformation.

In another preferred embodiment, the compound library is comprised of compounds
having at least about 1 to 50 nucleoside residues, more preferably at least about 1 to 10
nucleoside residues, most preferably at least about 1 to 5 nucleoside residues.

In other preferred embodiments, the compound library is comprised of compounds
having at least about 4 nucleoside residues, or at least about 3 nucleoside residues, or at least
about 2 nucleoside residues, or at least about 1 nucleoside residue.

In another preferred embodiment at least one compound of the following formula III
or III' of the compound library is comprised of:
wherein X and Y are each independently selected from a group consisting of O, S, Se, \( NR^1NR^2 \), \( CR^1CR^2 \), OR, SR and SeR, or one or both of X and Y are an enzymatically reactive moiety;

\( R \) is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

\( R^1, R^2 \) and \( R^3 \) are each independently selected from a group as defined by \( R \);

\( B \) is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring; and pharmaceutically acceptable salts thereof.

In one preferred embodiment, the compound library is comprised of at least one compound which has a sugar group in open chain form and wherein an enantiomerically enriched mixture of a compound is present.

A method is provided for construction of the compound library using solid-phase synthesis or solution-phase synthesis. In brief, one or more reagents are added to a reaction vessel capable of agitation and containing a resin reaction support material. The reaction
vessel is agitated during reaction of the reagents, and then centrifuging the reaction vessel to remove desired reaction materials therefrom. The library may also be constructed using an automated solution-phase synthesis.

In one preferred embodiment the compound library is used to identify a specific interacting partner for a nucleic acid or a protein. The nucleic acid can be for example, RNA or DNA. Examples of proteins are antibodies, receptors or ligands.

Compositions and methods are also provided for the treatment of cells or mammals infected with viruses, for example, cells infected with a herpes virus, cells infected with a cytomegalovirus; or cells infected with bacteria.

In a preferred embodiments, the compositions of the present invention are useful for identifying specific inhibitors for viral kinases, viral polymerases, identifying specific compounds which cause disruption of the association between a helicase-primase complex and a viral nucleic acid to which it is bound. The compositions are useful in treating a mammal suffering from or susceptible to a fungal, viral, or a bacterial infection.

Particularly preferred viral organisms causing human diseases according to the present invention include (but not restricted to) Herpes viruses, Hepatitisviruses, Retroviruses, Orthomyxoviruses, Paramyxoviruses, Togaviruses, Picornaviruses, Papovaviruses and Gastroenteritisviruses.

Particularly preferred bacteria causing serious human diseases are the Gram positive organisms: Staphylococcus aureus, Staphylococcus epidermidis, Enterococcus faecalis and E. faecium, Streptococcus pneumoniae and the Gram negative organisms: Pseudomonas aeruginosa, Burkholdia cepacia, Xanthomonas maltophilia, Escherichia coli, Enterobacter spp, Klebsiella pneumoniae and Salmonella spp.

Particularly preferred protozoan organisms causing human diseases according to the present invention include (but not restricted to) Malaria e.g. Plasmodium falciparum and M. ovale, Trypanosomiasis (sleeping sickness) e.g. Trypanosoma cruzei, Leishmaniasis e.g. Leishmania donovani, Amebiasis e.g. Entamoeba histolytica.
Particularly preferred fungi causing human diseases according to the present invention include (but not restricted to) *Candida albicans*, *Histoplasma neoformans*, *Coccidioides immitis* and *Penicillium marneffei*.

The methods of treatment comprise administering to a mammal susceptible to or suffering from an infectious disease, a therapeutically effective amount of a compound. The compound may be administered alone or in combination with a pharmaceutically acceptable carrier.

Preferred compounds comprising compositions used for the treatment of cells or mammals susceptible to or suffering from an infectious disease causing agent such as, for example, viruses such as herpes virus, or cytomegalovirus; bacteria, fungi and the like are, are preferably of the following Formula I or I':

![Diagram](image)

wherein $L^1$ and $L^2$ are independently O or a linking group such as e.g. an amide, ester, diester or the like, or an optionally substituted alkylene (e.g. $C_{1-20}$ alkylene), optionally substituted alkenylene (e.g., $C_{2-20}$ alkenylene) or alkynylene (e.g., $C_{2-20}$ alkynylene) having such groups either as a chain member of pendant to the chain, and which may be optionally
substituted with one or more substituents selected from a group consisting of O, S, Se, NRᵢNRᵢ, CRᵢCRᵢ, OR, SR and SeR, or an enzymatically reactive;

Q is carbon or a heteroatom such as O, S or N;

R¹, R², R³ are each independently a hydrogen or a hydroxyl group or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroarlicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroarlicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring; n=1 to 5; and pharmaceutically acceptable salts thereof.

The compound preferably is comprised of a sugar group, wherein the sugar group is in open chain form and an enantiomerically enriched mixture of a compound is present.

Other preferred compounds, include compounds of the following Formula II or II':

```
II

II'
```

HO --O-- B

R³

Y ----P---- X

R

OR³

HO --O-- B

Y ----P---- X

R

wherein X and Y are each independently selected from a group consisting of O, S, Se, NR\(^1\)NR\(^2\), CR\(^1\)CR\(^2\), OR, SR and SeR, or one or both of X and Y are an enzymatically reactive moiety;

R is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

R\(^1\), R\(^2\) and R\(^3\) are each independently selected from a group as defined by R;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring; and pharmaceutically acceptable salts thereof.

Other preferred compounds also include compounds of the following Formula III or III':

![Diagram](image)

wherein X and Y are each independently selected from a group consisting of O, S, Se, NR\(^1\)NR\(^2\), CR\(^1\)CR\(^2\), OR, SR and SeR, or one or both of X and Y are an enzymatically reactive moiety;
R is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

R¹, R² and R³ are each independently selected from a group as defined by R;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring; and pharmaceutically acceptable salts thereof.

The methods of the invention are preferably employed for treatment or prophylaxis against diseases caused by infectious agents, particularly for treatment of infections as may occur in tissues or organs of a subject. The methods of the invention also may be employed to treat systemic conditions such as viremia or septicemia. The methods of the invention are also preferably employed for treatment of diseases and disorders associated with viral infections or bacterial infections, as well as any other disorder caused by an infectious agent.

The methods of the invention are also preferably employed for identifying molecules that interfere with the life cycle or biological pathways of an infectious disease causing agent.

Other aspects of the invention are disclosed infra.

30 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates the general structure of the chemical library built around the NAB™ scaffold. Two types of functional variants P2 and P4 were attached to the furanose scaffold P1, providing divergent target recognition elements such as hydrogen bonding, lipophilicity, “shape-in-space”, and electronic properties. The P3 elements are served as structural variants to reorient the furanose scaffold into certain conformations, thereby enabling overall changes in three dimensional projection for the functional elements.
Figure 2 shows the general structure of the libraries 1 and 2, with the functional R groups attached to the 5'- and 3'-ends of the nucleoside respectively. For description of R, B, and Z, see Figure 6 and Table 7.

Figure 3 shows the energy-favorable conformations of the furanose puckers resulting from the 2'-substitutions. The changes in the furanose puckers cause dramatically different projections in three-dimensional space for the functional groups attached to the furanose scaffold.

Figure 4 shows the chemical structures of the representative members of library 2. The 2'-deoxynucleotide analogs are listed in Panel A, and the 2'-O-methyl nucleotide analogs are listed in Panel B.

Figure 5 shows the HPLC (C 18, at 260 nm, Panel A) and \(^{31}\)P NMR (in D\(_2\)O, Panel B) profiles of the representative members of library 2. The chemicals are presented as diastereomeric mixture of phosphorothioates. The purity of the products is in general higher than 95%.

Figure 6 is a schematic illustration of the generation of classes of libraries with varying diversities. (i) 1H-Tetrazole in CH\(_3\)CH; (ii) 3H-1,2-benzodithiole-3-one-1,1-dioxide in CH\(_3\)CN; (iii) partition between 2% NaHCO\(_3\) and EtOAc; (iv) 28% NH\(_4\)OH; (v) Dowex H\(^+\); (vi) partition between CHCl\(_3\) and H\(_2\)O.

Figure 7 shows the general structure of a simple NAB library depicting diversity elements.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides for a library that possesses key diversity attributes. The libraries are comprised of di-, tri- and tetranucleotides that carry modifications at the backbone, sugar, and nucleobase as illustrated in figure 1.

Important elements that contribute to diversity in the present invention include, but not limited to, backbone modifications such as for example, phosphorothioates and phosphoramidates which provide desirable metabolic stability to compounds when used in
cell-based assays - the former also could potentially participate in electrostatic interactions, while the latter could facilitate hydrophobic and hydrogen-bonding interactions with the target receptor; dominant furanose modifications, such as for example, substitution of a 2'-OMe group in place of a 2'-hydrogen in the deoxyribofuranoside ring, a 2'-substituent which acts as a "conformational switch", that transforms the furanose ring pucker from the 2'-endo to 3'-endo thereby affecting the global conformation of the individual library members; nucleobase modifications, which, for example can, include both the replacement of the parent heterocyclic moiety, as well as, substitution on the heterocycle, providing additional hydrophobicity to the library members; linkages between the different moieties are, for example, 3' to 5'. In this way a repertoire of diversity attributes can be captured in a representative NAB™ library with a molecular weight range of about 400 to about 1200.

There are many novel diversity features associated with libraries built around a NAB™ scaffold: (a) the scaffold can be used to create variable spatial display of hydrogen-bonding, hydrophobic, charge-transfer, electrostatic and such other non-covalent interactions; (b) the scaffold can be conformationally rigid or flexible and by linking the individual scaffolds together, one can fashion diverse molecular topology into library members; indeed, such shape-defining motifs as circles, pseudoknots, bulges and stem loops can be incorporated into libraries to target "hot spots on the receptor; (c) From a synthetic perspective, the NAB™ libraries can be assembled using well established solid-phase or solution-phase methods in nucleic acids field. An illustrative example of a small molecule-type NAB™ library is shown in figure 7, and represents multiple elements of diversity.

"Hot spots" as used herein refers to the high and low affinity binding sites on a receptor.

As used herein, "pseudoknots" refers to a three dimensional spatial configuration in the nucleic acid based scaffold that is formed by the inclusion of molecules wherein the molecular interactions, such as a network of hydrogen bonding, hydrophobic, ionic, and van der Waals interactions that facilitate the binding interactions between the molecules, resulting in a twist in the scaffold forming a "knot-like" structure.

Also described is a straightforward and facile solution-phase chemistry which has been developed to carry out the parallel combinatorial synthesis, from which libraries of 3'-
modified nucleoside analogs 2 were generated as individual compounds in high purity. See the Examples which follow.

As used herein, "nucleoside" includes the natural nucleosides, including 2'-deoxy and 2'-hydroxyl forms, e.g., as described in Kornberg and Baker, DNA Replication, 2nd Ed. (Freeman, San Francisco, 1992). For an excellent treatise on nucleic acids chemistry, see: Sanger, W. Principles of Nucleic Acids Structure. Springer-Verlag: New York, 1984.

bicycloarabinonucleosides (see e.g. N.K Christensen, et al, J. Am. Chem. Soc., 120: 5458-5463 (1998). Such analogs include synthetic nucleosides designed to enhance binding properties, e.g., duplex or triplex stability, specificity, or the like.

The term "stability" in reference to duplex or triplex formation generally designates how tightly an antisense oligonucleotide binds to its intended target sequence; more particularly, "stability" designates the free energy of formation of the duplex or triplex under physiological conditions. Melting temperature under a standard set of conditions, is a convenient measure of duplex and/or triplex stability.

A compound library comprised of a nucleic acid based (NAB™) scaffold typically refers to a library of small molecules that mimic the repertoire of interactions that exist amongst nucleic acids and proteins, as well as that which exist between proteins and nucleic acids. For example, a number of proteins contain nucleotide-binding domains defined by the topology of protein α-helices and β-sheets. For example, polymerases, topoisomerases, p53 protein and the like.

An example of molecules which can mimic nucleic acid interactions are DNA mimics which are polymers composed of subunits capable of specific, Watson-Crick-like
hybridization with DNA, or of specific hybridization with RNA. The nucleic acids can be modified at the base moiety, at the sugar moiety, or at the phosphate backbone. Exemplary DNA mimics include, e.g., phosphorothioates.

In accordance with the present invention, generation of compound libraries, described in detail in the material and methods and examples which follow, allow for the identification of interactions between nucleic acids and proteins, of an organism, for example, infectious disease causing agents and the biological pathways involved in replication of an infectious agent, inhibiting a virus-specific enzyme, such as viral thymidine kinase and reverse transcriptase, receptor-ligand interactions that allow an infectious agent to infect a cell or mammal; molecules that allow infection of mammals by disease causing agents; and the like. Based on the identification of the above-said interactions, drugs can be designed which inhibit the desired mechanism of pathogenesis of a disease causing agent.

As used herein, a biological pathway includes a collection of cellular constituents that influence one another through any biological mechanism, known or unknown, such as by a cell's synthetic, regulatory, homeostatic, or control networks. The influence of one cellular constituent on another can be, *inter alia*, by a synthetic transformation of the one cellular constituent into the other, by a direct physical interaction of the two cellular constituents, by an indirect interaction of the two cellular constituents mediated through intermediate biological events, or by other mechanisms. Further, certain pathways that are of particular interest in this invention can be said to originate at particular cellular constituents, which influence, but are not in turn influenced by, the other cellular constituents in the pathway and among such pathways, those without feedback loops are said to be hierarchical. A feedback loop in a biological pathway is a subset of cellular constituents of the pathway, each constituent of the feedback loop influences and also is influenced by other constituents of the feedback loop. Infectious disease causing agents may interfere with such biological pathways or make use of biological pathways for survival and replication in a host organism. For example, HIV stimulates the CD4+ T cell's production of certain cytokines.

Examples of biological pathways, as understood herein, are well known in the art. They depend on various biological mechanisms by which the cellular constituents influence one another. Biological pathways include well-known biochemical synthetic pathways in which, for example, molecules are broken down to provide cellular energy or built up to
provide cellular energy stores, or in which protein or nucleic acid precursors are synthesized. The cellular constituents of synthetic pathways include enzymes and the synthetic intermediates, and the influence of a precursor molecule on a successor molecule is by direct enzyme-mediated conversion. Biological pathways also include signaling and control pathways, many examples of which are also well known. Cellular constituents of these pathways include, typically, primary or intermediate signaling molecules, as well as the proteins participating in the signal or control cascades usually characterizing these pathways. In signaling pathways, binding of a signal molecule to a receptor usually directly influences the abundances of intermediate signaling molecules and indirectly influences on the degree of phosphorylation (or other modification) of pathway proteins. Both of these effects in turn influence activities of cellular proteins that are key effectors of the cellular processes initiated by the signal, for example, by affecting the transcriptional state of the cell. Control pathways, such as those controlling the timing and occurrence of the cell cycle, are similar. Here, multiple, often ongoing, cellular events are temporally coordinated, often with feedback control, to achieve a consistent outcome, such as cell division with chromosome segregation. This coordination is a consequence of functioning of the pathway, often mediated by mutual influences of proteins on each other's degree of phosphorylation (or other modification). Also, well known control pathways seek to maintain optimal levels of cellular metabolites in the face of a fluctuating environment. Further examples of cellular pathways operating according to understood mechanisms will be known to those of skill in the art. Infectious disease causing agents may influence the biological pathways in any number of ways, such as for example, downregulation or upregulation of certain constituents of the biological pathway, etc.

As used herein, the term "infectious agent" or "infectious disease causing agent", refers to an organism wherein growth/multiplication leads to pathogenic events in humans or animals. Examples of such agents are: bacteria, fungi, protozoa and viruses.

The compound libraries are also useful in identifying interactions between molecules involved in autoimmune diseases such as diabetes or autoimmune diseases caused by viruses which mimic molecules of the immune system of a mammal, thereby developing drugs which inhibit autoimmune interactions.
According to the invention, a preferred compound library comprises two or more compounds of the following Formula I or I:\n
\[ \text{Diagram of formulas I and I'} \]

wherein \( L^1 \) and \( L^2 \) are independently O or a linking group such as e.g. an amide, ester, diester or the like, or an optionally substituted alkyne (e.g. \( C_{1,20} \) alkyne), optionally substituted alkenylene (e.g., \( C_{2,20} \) alkenylene) or alkynylene (e.g., \( C_{2,20} \) alkynylene) having such groups either as a chain member of pendant to the chain, and which may be optionally substituted with one or more substituents selected from a group consisting of O, S, Se, \( NR^1NR^2 \), \( CR^1CR^2 \), OR, SR and SeR, or an enzymatically reactive;

\( Q \) is carbon or a heteroatom such as O, S or N;

\( R^1, R^2, R^3 \) are each independently a hydrogen or a hydroxyl group or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;
B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heterocyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring; n=1 to 5; and pharmaceutically acceptable salts thereof.

Sugar groups of the nucleic acid base scaffold may be natural or modified (e.g. synthetic) and in an open chain or ring form. Sugar groups may be comprised of mono-, di-, oligo- or polysaccharides wherein each monosaccharide unit comprises from 3 to about 8 carbons, preferably from 3 to about 6 carbons, containing polyhydroxy groups or polyhydroxy and amino groups. Non-limiting examples include glycerol, ribose, fructose, glucose, glucosamine, mannose, galactose, maltose, cellobiose, sucrose, starch, amylose, amylopectin, glycogen and cellulose. The hydroxyl and amino groups are present as free or protected groups containing e.g. hydrogens and/or halogens. Preferred protecting groups include acetonide, t-butoxy carbonyl groups, etc. Monosaccharide sugar groups may be of the L or D configuration and a cyclic monosaccharide unit may contain a 5 or 6 membered ring of the α or β conformation. Disaccharides may be comprised of two identical or two dissimilar monosaccharide units. Oligosaccharides may be comprised of from 2 to 10 monosaccharides and may be homopolymers, heteropolymers or cyclic polysugars.

Polysaccharides may be homoglycans or heteroglycans and may be branched or unbranched polymeric chains. The di-, oligo- and poly-saccharides may be comprised of 1 → 4, 1 → 6 or a mixture of 1 → 4 and 1 → 6 linkages. The sugar moiety may be attached to the link group through any of the hydroxyl or amino groups of the carbohydrate.

Preferred modifications to the sugar include modifications to the 2' position of the ribose moiety which include but are not limited to 2'-O-substituted with an -O-lower alkyl group containing 1-6 saturated or unsaturated carbon atoms, or with an -O-aryl, or allyl group having 2-6 carbon atoms wherein such -O-alkyl, aryl or allyl group may be unsubstituted or may be substituted (e.g., with halogen, hydroxy, trifluoromethyl, cyano, nitro, acyl, acyloxy, alkoxy, carboxy, carbalkoxy, or amino groups), or wherein the 2'-O-group is substituted by an amino, or halogen group. None of these substitutions are intended to exclude the native 2'-hydroxyl group in case of ribose or 2'-H- in the case of deoxyribose.
Heteroaliphatic moieties may be branched, unbranched or cyclic and include heterocycles such as morpholino, pyrrolidinyl, etc.

The term "heterocycle" as used herein refers to cyclic heteroaliphatic groups and preferably three to ten ring atoms total, includes, but is not limited to, oxetane, tetrahydrofuranyl, tetrahydropyranyl, aziridine, azetidine, pyrrolidine, piperidine, morpholine, piperazine and the like.

The terms "aryl" and "heteroaryl" as used herein refer to stable mono- or polycyclic, heterocyclic, polycyclic, and polyheterocyclic unsaturated moieties having 3-14 carbon atom which may be substituted or unsubstituted. Substituents include any of the previously mentioned substituents. Non-limiting examples of useful aryl ring groups include phenyl, halophenyl, alkoxyphenyl, diaalkoxyphenyl, trialkoxyphenyl, alkylendioxyphenyl, naphthyl, phenanthryl, anthryl, phenanthro and the like. Examples of typical heteroaryl rings include 5-membered monocyclic ring groups such as thiophenyl, pyrrolyl, imidazolyl, pyrazolyl, furyl, isothiazolyl, furazanyl, isoxazolyl, thiazolyl and the like; 6-membered monocyclic groups such as pyridyl, pyrazinyl, pyrimidinyl, pyridazinyl, triazinyl and the like; and polycyclic heterocyclic ring groups such as benzo[b]thienyl, naphtho[2,3-b]thienyl, thianthrenyl, isobenzofuranyl, chromenyl, xanthenyl, phenoxathiinyl, indolizinyln, isoindolyl, indolyln, indazolyl, purinyl, isoquinolyl, quinolyl, phthalazinyl, naphthyridinyl, quinoxalinyl, quinazolinyl, benzoisoxazolo, benzimidazolyl, tetrahydroquinoline cinnolinyl, pteridinyl, carbazolyl, beta-carbolinyl, phenanthridinyl, acridinyl, perimidinyl, phenanthrolinyl, phenazinyl, isothiazolyl, phenothiazinyl, phenoxazinyl, and the like (see e.g. Katritzky, Handbook of Heterocyclic Chemistry). The aryl or heteroaryl moieties may be substituted with one to five members selected from the group consisting of hydroxy, C1-C8 alkoxy, C1-C8 branched or straight-chain alkyl, acyloxy, carbamoyl, amino, N-acylamino, nitro, halo, trihalomethyl, cyano, and carboxyl. Aryl moieties thus include, e.g. phenyl; substituted phenyl bearing one or more substituents selected from groups including: halo such as chloro or fluoro, hydroxy, C1-C6 alkyl, acyl, acyloxy, C1-C6 alkoxy (such as methoxy or ethoxy, including among others dialkoxyphenyl moieties such as 2,3-, 2,4-, 2,5-, 3,4- or 3,5-dimethoxy or diethoxy phenyl or such as methylenedioxyphenyl, or 3-methoxy-5-ethoxyphenyl; or trisubstituted phenyl, such as trialkoxy (e.g., 3,4,5-trimethoxy or ethoxyphenyl), 3,5-dimethoxy-4-chloro-phenyl, etc.), amino, --SO2 NH2, --SO2NH(aliphatic),
--SO₂N(aliphatic)₂, --O-aliphatic--COOH, and --O-aliphatic-NH₂ (which may contain one or two N-aliphatic or N-acyl substituents).

According to the present invention it is also preferred that the library is comprised of an enantiomerically enriched mixture of at least about one type of a compound illustrated by formula I or I' is present.

Other preferred compound libraries comprise two or more compounds of the following Formula II or II':

wherein X and Y are each independently selected from a group consisting of O, S, Se, NR₁NR₂, CR₁CR₂, OR, SR and SeR, or one or both of X and Y are an enzymatically reactive moiety;

R is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

R¹, R² and R³ are each independently selected from a group as defined by R;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic
structure that is covalently linked to the sugar ring; and pharmaceutically acceptable salts thereof.

In accordance with the present invention, preferred compounds that comprise the compound library include at least about one compound of the following formula III or \( \text{III}' \):

\[
\begin{align*}
\text{III} & \quad \text{III}' \\
\end{align*}
\]

wherein \( X \) and \( Y \) are each independently selected from a group consisting of \( O, S, Se, NR^1NR^2, CR^1CR^2, OR, SR \) and \( SeR \), or one or both of \( X \) and \( Y \) are an enzymatically reactive moiety;

\( R \) is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heterocyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 \( N, O \) or \( S \) atoms;

\( R^1, R^2 \) and \( R^3 \) are each independently selected from a group as defined by \( R \);

\( B \) is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heterocyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 \( N, O \) or \( S \) atoms, or a heterocyclic structure that is covalently linked to the sugar ring; and pharmaceutically acceptable salts thereof.
According to the present invention, the library is comprised of at least about one compound having at least about one furanose ring in the C-2' endo conformation or at least about one furanose ring in the C-3' endo conformation. Also preferred is a library comprised of compounds having at least one about furanose ring in the C-2' endo conformation and at least about one furanose ring in the C-3' endo conformation.

Preferably the compound library is comprised of compounds having a length of at least about 1 to about 50 nucleoside residues, more preferably the compounds have a length of at least about 1 to 10 nucleoside residues, most preferably the compounds have a length of at least about 1 to 5 nucleoside residues.

According to the present invention the library may be comprised of compounds that have at least about 1 nucleoside residue, or at least about 2 nucleoside residues, or at least about 3 nucleoside residues, or at least about 4 nucleoside residues.


According to the present invention, a facile parallel solution-phase chemistry has been developed to carry out the combinatorial synthesis of nucleoside analogs as a source of pharmacophore-based chemical diversity. The reactions adapted in the synthetic cycle are performed at room temperature in high yields in a one-pot reaction system, except for the final deprotection steps. An efficient and convenient liquid/liquid extraction strategy partitioning between organic solvents and water has been tailored to successfully address the challenge of intermediate/product purification often associated with the application of solution-phase combinatorial chemistry. The library members thus obtained are present as individual compounds with purity higher than 90%.
As used herein, "pharmacophores-based diversity" are compounds representative of the same and/or closely related diversity based on physical and chemical properties of molecules with similar characteristics to a previously identified molecule or pharmacophore. The similarity principle requires that for any pair of molecules, differences in activity are related to differences in structure.

Various methods can be utilized to determine properties of compounds suitable for use in a compound library, for example US Patent No: 6,240,374, which is incorporated herein in its entirety. The reference makes use of the following measures: 2 dimensional and 3 dimensional measures. As used herein, "2 dimensional and 3 dimensional measures" shall mean a molecular representation which includes any terms which specifically incorporate information about the two dimensional and three dimensional features of the molecule. These measures take into account the geometric features of a molecule and also reflect the properties which are derivable from its topology; that is, the network of atoms connected by bonds.

As used herein, "2 dimensional fingerprints" means a 2D molecular measure in which a bit in a data string is set corresponding to the occurrence of a given 2-7 atom fragment in that molecule. Typically, strings of roughly 900 to 2400 bits are used. A particular bit may be set by many different fragments.

As used herein, a "combinatorial screening library" means a subset of molecules selected from a combinatorial accessible universe of molecules to be used for screening in an assay.

As used herein, "molecular structural descriptor" or "descriptor" means a quantitative representation of the physical and chemical properties determinative of the activity of a molecule. The term "metric" is synonymous with molecular structural descriptor and is used interchangeably throughout this application.

"Patterson plots", as used herein, means two dimensional scatter plots in which the distance between molecules in some metric is plotted on the X axis and the absolute difference in some biological activity for the same molecules is plotted on the Y axis.
"Sigmoid plots", as used herein, means two dimensional plots for which the proportion of molecular pairs in which the second molecule is also active is plotted on the Y axis and the pairwise Tanimoto similarity is plotted in intervals on the X axis.

As used herein, "topomeric alignment", means conformer alignment based on a set of alignment rules.

Any method known to those of skill in the art may be used in designing a compound library with molecular diversity. Preferred methods include structure activity relationships (SAR), see for example, Clark-Lewis et al., "Structure-Activity Relationships of Interleukin-8 Determined Using Chemically Synthesized Analogs", J. Biol. Chem. 266(34):23128-23134 (1991); and the method described in US Patent No: 6,240,374. Briefly, the latter method requires that any valid molecular structural descriptor must have a "neighborhood property". That is: the descriptor must meet the similarity principle's constraint that it measure the chemical universe in such a way that similar structures (as defined by the descriptor) have substantially similar biological properties. Or stated slightly differently: within some radius in descriptor space of any given molecule possessing some biological property, there should be a high probability that other molecules found within that radius will also have the same biological property. If a descriptor does not have the neighborhood property, it does not meet the similarity principle, and can not be valid.

According to the present invention, to quantitatively analyze whether any given metric obeys the neighborhood principle is preferably used for determining patterns for use in identifying novel molecular targets for antiviral intervention. Absolute values of biological activity have also been considered the dependent variable with the structural metric as the independent variable. This is the case for traditional QSARs (quantitative structure activity relationships), which is also a preferred method for use in identifying compounds in the present invention.

It must be noted, although small differences in structure should be associated with small differences in activity, the converse is not necessarily true; large differences in activity are not necessarily associated with large differences in structure. Thus, it is important to use differences in both measures: biological differences and structural (metric) differences. Thus,
instead of looking at the values assigned by the metric to each molecule, the absolute differences in the metric values for each pair of molecules are the independent variables and the absolute differences in biological activity for each pair of molecules are the dependent variables. The absolute value is used since it is the difference, not its sign, which is important. Thus identified library members are used in constructing combinatorial libraries.

The library members are synthesized as individual compounds using the parallel solution-phase chemistry (figure 6). The key precursor thiophosphotriesters 6 are synthesized from the commercially available 3'-phosphoramidites 4 in a one-pot reaction system at room temperature in good yield by stepwise addition of reactants. Reaction of alcohol 3 and the 3'-phosphoramidite 4 in CH$_3$CN in the presence of 1 H-tetrazole provided phosphonphite triester 5 in quantitative yield as indicated by TLC analysis. Without the necessity of isolating 5, 3H-

1, 2-benzodithiole-3-one (3H-BD)$^9$ is added to the reaction mixture to oxidatively sulfurize 5 to give thiophosphotriester 6. After a preliminary purification by liquid/liquid extraction, the crude triester 6 is converted to the final product phosphorothioate 2 quantitatively by sequential treatment with NH$_4$OH (28%, 55 °C, 4 hrs) and Dowex resin to remove the corresponding protecting groups.

One of the challenges in applying the parallel solution-phase chemistry to the synthesis of combinatorial libraries is to develop convenient and efficient isolation/purification techniques to eliminate the otherwise time-consuming purification procedures (e.g. chromatography) often associated with solution-phase chemistry. This problem is overcome by designing a convenient two-stage liquid/liquid extraction strategy partitioning between organic solvents and water (figure 6). The extraction strategy is based on the uniqueness that the desired phosphorothioate product 2 is hydrophilic (water soluble) but its immediate precursor thiophosphotriester 6 is hydrophobic (organic solvent soluble) and also compound 6 could be converted to compound 2 quantitatively. The first liquid/liquid extraction was applied after thiophosphotriester 6 was formed. By partitioning between ethyl acetate (EtOAc) and 2% aqueous NaHCO$_3$, the organic soluble 6 was retained in EtOAc, while the water-soluble byproducts (e.g. sulfurization byproducts 1-benzothiole-2-oxa-3-one sulfoxides, disopropyl ammonium tetrazolides) and other excess reagents from all previous reactions (such as 1 H-tetrazole) are removed by repetitive aqueous washing. The second extraction is applied after compound 2 is formed. By partitioning between CHCl$_3$ and H$_2$O, it is anticipated that the hydrophilic phosphorothioate 2 is retained in H$_2$O, while the
CHCl₃-soluble ingredients in the product mixtures - either generated from the deprotection reaction by ammonia or carried over with 6 from the previous reactions - are removed by repetitive CHCl₃ washing. After the second extraction, HPLC analysis⁴ (C-18, reversed phase, λ₂₆₀) shows that the desired product 2 is presented in the aqueous layer in purity higher than 90% (Figure 4).

Two approaches are used to remove the dimethoxytrityl group at the 5'-end of compound 7. The conventional approach involves treatment in 80% acetic acid for 1 to 2 hr¹¹; a recent report showed that Dowex resin (strong acid form, Dowex 500 WX 8-200, Aldrich) could be used efficiently in the detritylation (10 min).¹² Although both approaches can be used, the preferred method is using the Dowex resin capture procedure. The advantage of using Dowex resin capture procedure in the parallel library synthesis are: (a) ease and convenience of operation; (b) reduced time of detritylation (10 min of Dowex versus 1 - 2 hrs of acetic acid); (c) minimizing the depurination caused by prolonged exposure of nucleosides to acetic acid; (d) abolition of evaporation of the acrid acetic acid. After the ammonia treatment, the resulting mixture containing compound 7 is passed through a short column of Dowex H⁺ for the detritylation to give product 2.¹³

The final desalting step (C-18 cartridges, Gilson) is used to remove the inorganic salts, and the product 2, thus obtained, is of purity higher than 95%. Analysis by NMR and MS confirmed the structures of the library members. Spectral analysis of selected members revealed that no detectable base-modifications had occurred during the synthesis. The chemical structures of the representative library members are depicted in Figure 5 and their corresponding spectral analysis is displayed in Figure 4 and Table 6.

A preferred method for tethering of the 5'-end of a nucleoside with functional hydrophobic groups is achieved through phosphorothioate linkers using the parallel solid-phase and solution-phase combinatorial chemistry. The "drug-like" attributes of the 5'-modified nucleotide analogs 1 (figure 2) are evaluated by screening the library members against hepatitis- and herpes-virus replication and a number of "hits" are identified in cell-based assays. Such cell-based assays are well-know in the art and include for example, plaque reduction assays, etc. Libraries of the general structure 2 (figure 2) are constructed in which the 3'-end of a nucleoside carries a variety of functional hydrophobic groups.
The library can also be constructed using an automated solution-pH... synthesis.

The members of the library may be comprised of synthetic oligonucleotides. Preferred synthetic oligonucleotides comprise at least one, and preferably more than one, modification. Modifications include, for example, modifications of the internucleotide linkage, the base or the sugar moiety, capped ends and chimeric or hybrid oligonucleotides.

Synthetic oligonucleotides include chemically synthesized polymers of deoxyribonucleotide and/or ribonucleotide monomers connected by internucleotide linkages. Oligonucleotides may be constructed entirely of deoxyribonucleotides, entirely of ribonucleotides or of a combination of deoxyribonucleotides and ribonucleotides, including hybrid and inverted hybrid oligonucleotides. Hybrid oligonucleotides contain a core region of deoxyribonucleotides interposed between flanking regions of ribonucleotides. Inverted hybrids contain a core region of ribonucleotides interposed between flanking regions of deoxyribonucleotides.

Synthetic oligonucleotides of the invention may be connected by standard phosphodiester internucleotide linkages between the 5' group of one mononucleotide pentose ring and the 3' group of an adjacent mononucleotide. Such linkages could also be established using different sites of connection, including 5' to 5', 3' to 3', 2' to 5' and 2' to 2', or any combination thereof. In addition to phosphodiester linkages, the mononucleotides may also be connected by alkylphosphonate, phosphorothioate, phosphorodithioate, alkylphosphonothioate, phosphoramidate, carbamate, carbonate, phosphate triester, acetamidate, or carboxymethyl ester linkages, or any combination thereof. Preferably, an oligonucleotide of the invention comprises at least one phosphorothioate internucleotide linkage, more preferably, all linkages in the oligonucleotide are phosphorothioate internucleotide linkages.

Oligonucleotides of the invention may be constructed such that all mononucleotides are connected by the same type of internucleotide linkages or by combinations of different internucleotide linkages, including chimeric or inverted chimeric oligonucleotides. Chimeric oligonucleotides have a phosphorothioate core region interposed between methylphosphonate or phosphoramidate flanking regions. Inverted chimeric oligonucleotides have a nonionic
core region (e.g. alkylphosphonate and/or phosphoramidate and/or phosphotriester internucleoside linkage) interposed between phosphorothioate flanking regions.

Synthetic oligonucleotides of the invention may be constructed of adenine, cytosine, guanine, inosine, thymidine or uracil mononucleotides. Preferred oligonucleotides are constructed from mononucleotides which contain modifications to the base and/or sugar moiety of the mononucleotide. Modifications to the base or sugar include covalently attached substituents of alkyl, carbocyclic aryl, heteroaromatic or heteroalicyclic groups having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure.

Alkyl groups preferably contain from 1 to about 18 carbon atoms, more preferably from 1 to about 12 carbon atoms and most preferably from 1 to about 6 carbon atoms. Specific examples of alkyl groups include, for example, methyl, ethyl, n-propyl, isopropyl, n-butyl, t-butyl, isobutyl, pentyl, hexyl, heptyl, octyl, nonyl, decyl etc.

Aralkyl groups include the above-listed alkyl groups substituted by a carbocyclic aryl group having 6 or more carbons, for example, phenyl, naphthyl, phenanthryl, anthracyl, etc.

Cycloalkyl groups preferably have from 3 to about 8 ring carbon atoms, e.g. cyclopropyl, cyclopentyl, cyclohexyl, cycloheptyl, 1,4-methylenecyclohexane, adamantyl, cyclopentylmethyl, cyclohexylmethyl, 1- or 2-cyclohexylethyl and 1-, 2- or 3-cyclohexylpropyl, etc.

Exemplary heteroaromatic and heteroalicyclic group include pyridyl, pyrazinyl, pyrimidyl, furyl, pyrrolyl, thienyl, thiazolyl, oxazolyl, imidazolyl, indolyl, benzothiazolyl, tetrahydrofuranyl, tetrahydropyranyl, piperidinyl, morpholino and pyrrolidinyl.

For good classification of various modifications to a scaffold, see: Kaztritzky, A. R.; Kiely, J. S.; Hebert, N.; Chassaing, C. Definition of templates within combinatorial libraries. J. Comb. Chem. 2000, 2, 2-5.

Other modifications include those which are internal or are at the end(s) of the oligonucleotide molecule and include additions to the molecule at the internucleoside phosphate linkages, such as cholesteryl, cholesterol, or diamine compounds with varying
numbers of carbon residues between the two amino groups, and terminal ribose, deoxyribose and phosphate modifications which cleave, or crosslink to the opposite chains or to associated enzymes or other proteins which bind to the viral genome. Additional linkers including non-nucleoside linkers include, but are not limited to, polyethylene glycol of varying lengths, e.g., triethylene glycol, monoethylene glycol, hexaethylene glycol, (Ma et al. (1993) *Nucleic Acids Res.* 21: 2585-2589; Benseler et al. (1993) *J. Am. Chem. Soc.* 115: 8483-8484), hexylamine, and stilbene (Letsinger et al, (1995) *J. Am. Chem. Soc.* 117: 7323-7328) or any other commercially available linker including abasic linkers or commercially available asymmetric and symmetric linkers (CloneTech, Palo Alto, California) (e.g., Glen Research Product Catalog, Sterling, VA).

Additionally oligonucleotides capped with ribose at the 3' end of the oligonucleotide may be subjected to NaIO₄ oxidation/reductive amination. Amination may include but is not limited to the following moieties, spermine, spermidine, Tris(2-aminoethyl)amine (TAEA), DOPE, long chain alkyl amines, crownethers, coenzyme A, NAD, sugars, peptides, dendrimers.

Oligonucleotides may also be capped with a bulky substituent at their 3' and/or 5' end(s), or have a substitution in one or both nonbridging oxygens per nucleotide. Such modifications can be at some or all of the internucleoside linkages, as well as at either or both ends of the oligonucleotide and/or in the interior of the molecule (reviewed in Agrawal et al. (1992) *Trends Biotechnol.* 10: 152-158). Some non-limited examples of capped species include 3'-O-methyl, 5'-O-methyl, 2'-O-methyl, and any combination thereof.

Synthetic oligonucleotides of the invention can be prepared by art recognized methods. For example, nucleotides can be covalently linked using art-recognized techniques such as phosphoramidite, H-phosphonate chemistry, or methylphosphoramidite chemistry (see, e.g., Goodchild (1990) *Bioconjugate Chem.* 2: 165-187; Uhlmann et al. (1990) *Chem. Rev.* 90: 543-584; Caruthers et al. (1987) *Meth. Enzymol.* 154: 287-313; U.S. Patent No. 5,149,798) which can be carried out manually or by an automated synthesizer and then processed (reviewed in Agrawal et al. (1992) *Trends Biotechnol.* 10: 152-158). Oligonucleotides with phosphorothioate linkages can be prepared using methods well known in the field such as phosphoramidite (see, e.g., Agrawal et al. (1988) *Proc. Natl. Acad. Sci. (USA)* 85: 7079-7083) or H-phosphonate (see, e.g., Froehler (1986) *Tetrahedron Lett.* 27:

In other aspects, the invention provides a pharmaceutical composition. The pharmaceutical composition is a physical mixture of at least one, and preferably two or more classes of compound libraries.

15 As used herein, "classes of libraries" refers to those libraries generated which fall into categories of diversity based on their chemical or physical properties, discussed supra. The different compounds in the different classes of libraries may have different specificities for target molecules depending on the sequences of the scaffold, modification(s), and/or lengths. In some embodiments, this pharmaceutical formulation also includes a physiologically or pharmaceutically acceptable carrier. Specific embodiments include a therapeutic amount of a lipid carrier.

20 As used herein the terms, "an anti-fungal effective amount" or "an anti-viral effective amount", or "an anti-bacterial viral effective amount", are used interchangeably with "therapeutically effective amount" or "effective amount" means an amount of a drug or pharmacologically active agent that is nontoxic but sufficient to provide the desired local or systemic effect and performance at a reasonable benefit/risk ratio attending any medical treatment.

30 As used herein, a "pharmaceutically acceptable" component is one that is suitable for use with humans and/or animals without undue adverse side effects (such as toxicity, irritation, and allergic response) commensurate with a reasonable benefit/risk ratio.

35 The classes of libraries of the present invention are suitable for use as therapeutically active compounds, especially for use in the control or prevention of herpes simplex virus (HSV). In vitro assays (see example 5) indicate a significant inhibition of HSV plaque formation.
In this aspect of the invention, a therapeutic amount of a pharmaceutical composition containing classes of libraries is administered to a cell to inhibit virus replication. In a similar aspect, the classes of libraries of the present invention can be used for treating human or other mammals infected with a virus comprising the step of administering to an infected animal or cell a therapeutic amount of a pharmaceutical composition containing at least one class of library, and in some embodiments, at least two classes of libraries.

In all methods involving the administration of classes of libraries of the invention, at least one, and preferably two or more identical or different classes of libraries may be administered simultaneously or sequentially as a single treatment episode in the form of separate pharmaceutical compositions.

More specifically, the invention includes methods of treatment of a mammal susceptible to (prophylactic treatment) or suffering from a disease associated with viruses, such as HSV. Methods in the present invention comprise administration of a therapeutically effective amount of one or more compounds of the invention to virally infected cells, such as mammalian cells, particularly human cells.

Administration of compounds of the invention may be made by a variety of suitable routes including oral, topical (including transdermal, buccal or sublingual), nasal and parenteral (including intraperitoneal, subcutaneous, intravenous, intradermal or intramuscular injection) with oral or parenteral being generally preferred. It also will be appreciated that the preferred method of administration and dosage amount may vary with, for example, the condition and age of the recipient.

Compounds of the invention may be used in therapy in conjunction with other pharmaceutically active medicaments, such as another anti-viral agent, or an anti-cancer agent. Additionally, while one or more compounds of the invention may be administered alone, they also may be present as part of a pharmaceutical composition in mixture with conventional excipient, i.e., pharmaceutically acceptable organic or inorganic carrier substances suitable for parenteral, oral or other desired administration and which do not deleteriously react with the active compounds and are not deleterious to the recipient thereof. Suitable pharmaceutically acceptable carriers include but are not limited to water, salt
solutions, alcohol, vegetable oils, polyethylene glycols, gelatin, lactose, amyllose, magnesium stearate, talc, silicic acid, viscous paraffin, perfume oil, fatty acid monoglycerides and diglycerides, petroethral fatty acid esters, hydroxymethyl-cellulose, polyvinylpyrrolidone, etc. The pharmaceutical preparations can be sterilized and if desired mixed with auxiliary agents, e.g., lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, colorings, flavorings and/or aromatic substances and the like which do not deleteriously react with the active compounds.

For parenteral application, particularly suitable are solutions, preferably oily or aqueous solutions as well as suspensions, emulsions, or implants, including suppositories. Ampules are convenient unit dosages.

For enteral application, particularly suitable are tablets, dragees or capsules having talc and/or carbohydrate carrier binder or the like, the carrier preferably being lactose and/or corn starch and/or potato starch. A syrup, elixir or the like can be used wherein a sweetened vehicle is employed. Sustained release compositions can be formulated including those wherein the active component is protected with differentially degradable coatings, e.g., by microencapsulation, multiple coatings, etc.


It will be appreciated that the actual preferred amounts of active compounds used in a given therapy will vary according to the specific compound being utilized, the particular compositions formulated, the mode of application, the particular site of administration, etc. Optimal administration rates for a given protocol of administration can be readily ascertained by those skilled in the art using conventional dosage determination tests.

All documents mentioned herein are incorporated herein by reference.
The present invention is further illustrated by the following examples. These examples are provided to aid in the understanding of the invention and are not to be construed as limitations thereof.

EXAMPLES
MATERIALS AND METHODS

General Methods

\(^1\)H, \(^{13}\)C, \(^{31}\)P NMR spectra were collected on a 500 MHz NMR spectrometer (Bruker). MS was performed by electrospray in negative mode. The HPLC analysis was run on a Waters 600 system equipped with a photodiode-array UV detector 996, autosampler 717, and Millenium, using a Radial-Pak liquid chromatography cartridge (8mm I.D., 8NVCl8). The C-18 cartridges (100 mg) were purchased from Gilson, and the desalting was performed on Gilson solid-phase extraction system (Gilson, ASPEC XL). The QIAshredder spin column was purchased from Qiagen. The N-protected nucleoside phosphoramidites 4a-h were purchased from CruaChem (Aston, PA). The Dowex resin (strong acid form, Dowex 500 WX 8-200) was cleaned before use. All other reagents were obtained from commercial sources and used without purification.

General procedures for library assembly (see scheme1)

The general procedures for library assembly are illustrated in figure 6. Each alcohol 3, at a concentration of 30 \(\mu\)mol, and each of the nucleoside amidites 4, at a concentration of 20 \(\mu\)mol were added sequentially to a series of conical microtubes (2 mL, Ultradent Scientific) containing 1 H-tetrazole solution in CH\(_3\)CN (1 mL, 100 \(\mu\)mol) under argon. The mixture was shaken in a platform shaker at room temperature for 5 min. Then 3H-BD (40 \(\mu\)mol) in CH\(_3\)CN was evaporated in a Speed Vac. One mL of ethyl acetate (EtOAc) was then added, followed by 2% aqueous sodium bicarbonate (0.5 mL). Following thorough mixing of the phases, the organic layer containing the intermediate thiophosphate triester 6, was separated and evaporated to dryness. Aqueous NH\(_4\)OH (28%, 1 mL) was then added to the residue in each microtube. The tightly capped tubes were heated at 55°C for 4 hrs. The aqueous ammoniacal solution was concentrated to dryness in a Speed Vac. The contents were dissolved in water (0.5 mL), then added to a spin column (QIAshredder) containing Dowex H\(^+\) (50 mg). After shaking for 10 min, the column was placed in a receptacle vial and centrifuged.
The flow-through was collected, diluted to 1 ml with water, and extracted with chloroform (2 x 0.4ml). The aqueous layer was then passed through a C-18 cartridge. The cartridge was then washed with water, and the product 2, was eluted with CH$_3$CN/ H$_2$O (10/90 for 2 ml, 50/50 for 2 ml). The corresponding eluent was collected and evaporated to dryness in a Speed Vac to provide the product 2, as a white solid in high purity (>95%, C-18 HPLC analysis at $\lambda_{260}$). Quantitation was achieved on the basis of A$_{260}$ units, and the yields of products 1, were found to range from 75 to 90% starting from 4.

"Drug-like" attributes and chemical diversity of the designed library

The pharmaceutical relevance and chemical diversity of the libraries building around the nucleic acid based (NAB$^\text{TM}$) scaffold (figure 1) could be recognized by a number of factors: (a) the nucleoside pharmacophores, including P1 and P2 elements, could provide the primary target (nucleic acids or proteins) recognition element by their unique "shape-in-space" and hydrogen-bonding characteristics; (b) the P4 elements (R groups, figures 1 and 2 and Table 1) could serve as additional functional variants providing a handle to incorporate divergent hydrophobic elements to facilitate the specific "drug"-target interaction; (c) the negatively-charged element of phosphorothioate linker (figure 2) could provide ionic interaction of the nucleotide analog with the target, while imparting aqueous solubility to the compounds; (d) importantly, the 2'-substituents (P3 elements, figures 1 and 2) could be used as the structural variants to "lock" the furanosyl scaffold into significantly different conformational puckers (e.g. C$_{2}$-endo versus C$_{3}$-endo puckers, figure 3) thereby enabling overall changes in local and/or global projection of the functional groups (e.g. P2 and P4) attached to the furanosyl scaffold.

Alcohol building blocks

The alcohols shown in table 1 were used as the building blocks to incorporate the functional variants R groups (figure 2). The selection of alcohols for the library construction was based on the contribution of the alcohols to molecular diversity by variable spatial displays of shape-in-space, charge-transfer, and van der Waals interactions. Such functional divergency was intended to explore the specific interaction between a ligand and its receptor: (1) to provide hydrophobic interactions to facilitate high affinity binding between the nucleoside analogs and target receptors; and/or (2) to aid in the inter- and/or intracellular delivery of the nucleoside analogs;
Assembly of the library using parallel solution phase chemistry.

The strategy to generate the desired libraries of 3'-modified nucleoside analogs, the general structure of library 2 is illustrated in figure 2, was similar to the method used to assemble the libraries of the 5'-modified nucleoside analogs of library 1, the general structure illustrated in figure 2. The library members were synthesized as individual compounds using the parallel solution-phase chemistry depicted in figure 6. The key precursor thiophosphotriesters 6, could be synthesized from commercially available 3'-phosphoramidites 4, in a one-pot reaction system at room temperature in good yield by step-wise addition of reactants. Reaction of alcohol 3, and the 3'-phosphoramidite in CH$_3$CN in the presence of 1 H-tetrazole provided phosphorhite triester 5, in quantitative yields as indicated by TLC analysis. Without the necessity of isolating 5, 3H-1,2-benzodithiole-3-one (3H-BD) was added to the reaction mixture to oxidatively sulfurize 5, to produce thiophosphotriesters 6. After a preliminary purification by liquid/liquid extraction, the crude triester 6, was then converted to the final product phosphorothioate 2, quantitatively by sequential treatment with NH$_4$OH (28%, 55°C, 4 hrs) and Dowex resin to remove the corresponding protecting groups.

To eliminate the other wise time-consuming purification procedures (e.g. chromatography) often associated with solution-phase chemistry, a convenient two-stage liquid/liquid extraction strategy partitioning between organic solvents and water (figure 6) was designed. The extraction strategy was based on the uniqueness that the desired phosphorothioate product 2, is hydrophilic but its immediate precursor thiophosphotriesters 6, is hydrophobic (organic solvent soluble) and also compound 6, could be converted to compound 2, quantitatively. The first liquid/liquid extraction was applied after thiophosphotriesters 6, was formed. By partitioning between EtOAc and 2% aqueous NaHCO$_3$, the organic soluble 6, was retained in EtOAc, while the water-soluble byproducts, for example, sulfurization byproducts such as 1-benzothiole-2-oxa-3-one; sulfoxides such as diisopropyl ammonium tetrazolides and other excess reagents from all previous reactions such as 1 H-tetrazole, were removed by repetitive aqueous washing.

The second extraction was applied after compound 2, was formed. By partitioning between CHCl$_3$ and H$_2$O, it is anticipated that the hydrophilic phosphorothioate 2, was retained in H$_2$O, while the CHCl$_3$-soluble ingredients in the product mixtures – either generated from the deprotection reaction by ammonia or carried over with 6, from the
previous reactions – were removed by repetitive CHCl₃ washing. After the second extraction, HPLC analysis (C-18, reversed phase, λ₂₅₀) showed that the desired product 2, was presented in the aqueous layer in purity higher than 90% (figure 4).

Two approaches could be used to remove the dimethoxytrityl group at the 5′-end of compound 7. The conventional approach involves treatment in 80% acetic acid for 1 to 2 hrs. A recent report showed that Dowex resin (strong acid form, Dowex 500 WX 8-200, Aldrich) could be used efficiently in the detritylation (10 min).

Although both approaches can be used, the advantage of using Dowex resin capture in the parallel library synthesis are: (a) ease and convenience of operation; (b) reduced time of detritylation (10 min of Dowex versus 1-2 hrs of acetic acid); (c) minimizing the depurization caused by prolonged exposure of nucleosides to acetic acid; (d) abolition of evaporation of the acrid acetic acid. After the ammonia treatment, the resulting mixture containing compound 7, was then passed through a short column of Dowex H⁺ for the detritylation to give product 2.

The final desalting step (C-18 cartridges, Gilson was used to remove the inorganic salts, and the product 2, thus obtained, was of purity higher than 95%. Analysis by NMR and MS confirmed the structures of the library members. Spectral analysis of selected members revealed that no detectable base-modifications had occurred during the synthesis. The chemical structures of the representative library members are depicted in figure 5 and their corresponding spectral analysis is displayed in figure 4 and Table 6.

ES-MS analysis (negative mode)
A scan spectrum was acquired first for each sample in order to verify the presence of the analyte in-the sample. Then a MS/MS spectrum was acquired for the analyte to give structural information on the samples. The samples were dissolved in isopropanol/H₂O (lit) containing 20 mM of triethylamine.

Assembly of the Libraries 1-4
The library synthesis (5 to 15 µmol scale) was performed using standard automated DNA synthesis protocols (DMT-off). Oxidative sulfurization was effected using 3H-1,2-
hendzothiole-3-one-1,1-dioxide.\textsuperscript{13} After synthesis, the CPG was dried using N\textsubscript{2}, and
transferred to 5-mL centrifuge tubes. Aqueous ammonium hydroxide (28%, 4 mL) was
added and the mixture was heated at 55 °C for 3 - 6 h. The resulting suspension was cooled
and centrifuged. The solution was evaporated in a Speed Vac. Each resulting product was
dissolved in water (3 to 5 mL), and extracted with ethyl acetate (2 x 1 mL). The aqueous
layer was evaporated to remove excess ethyl acetate and the residue taken up in ultra pure
water and filtered through 0.2 μm filter. Lyophilization gave the products as white foam.

\textit{Assembly of the library 5}

The requisite dinucleotide H-phosphonates were assembled on solid support using
standard H-phosphonate chemistry. Following washing, the dry CPG was transferred to 5
mL centrifuge tubes and a solution of amine in CCl\textsubscript{4} (10%, 3-4 mL).\textsuperscript{14} The mixture was
shaken for 20 to 30 min. Following washing, the CPG was treated with 28% aqueous
NH\textsubscript{4}OH (55 °C, 3-6 h). The suspension was cooled and centrifuged. The solution was
evaporated to dryness in vacuo, the residue dissolved in H\textsubscript{2}O (5 mL), and extracted with ethyl
acetate (2 mL). The aqueous layer was evaporated; residue taken up in ultra pure water and
filtered.

\textit{Analysis of the library}

Reversed-phase HPLC analysis of tile libraries was performed on a Waters 600
system equipped with a photodiode-array UV detector 996, autosampler 717, and
Millennium® 2000 software, using a Radial-Pak® liquid chromatography cartridge [8 mm
I.D., 8NVC18]. Mobile phase: Buffer A: 0.1 M NH\textsubscript{4}OAc; Buffer B: 20% A/80% CH\textsubscript{3}CN,
v/v: Gradient: 100% A, 0-3 min; 40% A, 40 min; 100% B, 49 min; 100% B. Product purity
ranged from 85 to 95%. Yields were estimated to be 65 to 90% on the basis of A\textsubscript{260} units.

\textit{Spectral characterization of representative library members}

Before spectral acquisition, the representative library members (10% of the library
size) were further purified by ion-exchange column (DEAE-5PW Resin, Buffer A: H\textsubscript{2}O,
Buffer B: 0.5 M NaCl) followed by desalting (C18 column, Buffer A: H\textsubscript{2}O, Buffer B: 20%
CH\textsubscript{3}CN in H\textsubscript{2}O) to give individual library members of 95-99% purity as determined by
reversed-phase HPLC.
\[ ^{31}\text{P} \text{ NMR analysis of selected library members revealed clear signals at } \delta 58-59 \text{ ppm, and 13-14 ppm, characteristic of phosphorothioate and phosphoramidate linkages respectively. Other peaks in } ^{31}\text{P} \text{ NMR spectra constituted less than 3\% of the total area corresponding to the desired product peaks. Additionally, the ES-MS of selected library members were consistent with the expected molecular weights corresponding to the assigned structures.} \]

\textit{Antiviral assays against HSV-1}

Vero cells (African green monkey kidney cells) (ATCC) were infected with HSV-1 at an MOI of 0.005 in 96 well plates in Dulbecco’s modified Eagle’s (DMEM) medium. The plates were maintained at 37 °C for 3 h. The compounds were added at 25 micromolar concentration, following which the plates were incubated at 37 °C: for 24 h. The media was removed, and the cells were fixed with 10\% formal saline for 10 minutes. The cells were stained with 0.1\% crystal violet, incubated at room temperature for 30 minutes, and then washed with distilled water. The viral plaques were counted, and the antiviral effect estimated as a percent reduction in the number of plaques compared with untreated control. Acyclovir was used as the positive control (EC\textsubscript{50}, at 6 micromolar).

Example 1

\textit{Synthesis of Libraries}

\textit{Library I}

A 64-member dinucleotide phosphorothioate library I (Table 1) was assembled on solid support using commercially available deoxyribonucleoside and 2’-OMe ribonucleoside phosphoramidite building blocks in conjunction with the corresponding controlled-pore-glass (CPG)-linked nucleosides. The assembly was carried out in a parallel synthesis mode (DMT-off, 10 to 15 \textmu mol scale). Following the assembly, each solid support was treated with aqueous ammonium hydroxide (28\%, 55 °C) to remove nucleobase-, arid phosphate-protecting groups and to cleave the products off the support. The products were purified and evaluated as described in the experimental section.
Example 2

**Library 2**

Using eight of the commercially available monomers and nucleoside-bound-CPGs, potentially 512 trinucleotide phosphorothioates could be assembled as a mixture of $R_p$, $S_p$ diastereomers. We prepared a representative 64-member library (Table 2) using phosphoramidite chemistry. Following work-up and extraction, the products were obtained 85% to 95% pure as determined by reversed-phase HPLC.

Table 2. 3-trinucleoside phosphorothioate library

Example 3

**Library 3**

The parallel assembly of tetranucleotides posed a special challenge because a 4096-member tetranucleotide library (representing an 8x8x8x8 array) could be assembled. In order to have a library amenable to parallel synthesis, three nucleotide positions were fixed with the fourth position being degenerate. Thus, each of the 64 trinucleotides were first assembled on CPG by parallel synthesis, and each CPG-bound trinucleotide was reacted with an equimolar
mixture of dA, dC, dG, and T nucleoside phosphoramidites (scheme 1). In this way, 256-member library was assembled (Table 3).

Scheme 1

Table 3: Deoxynucleoside phosphorothioate tetranucleoside library

| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |
| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |
| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |
| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |
| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |
| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |
| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |
| TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG | TACG |

9: represent a mixture of dA, dC, dG, and T

Example 4

Library 4

A representative base-modified library of 24 members were assembled as di-, tri-, tetranucleoside phosphorothioates (Table 4) using the corresponding commercially available phosphoramidite monomers.

Table 4: Base modified phosphorothioate library

| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |
| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |
| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |
| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |
| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |
| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |
| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |
| TAC-  | TAC-  | TAC-  | TAC-  | TAC-  | TAC-  |

C<sup>5Me</sup>, 5-Methyl-dC; A<sup>7-Deaza</sup>, 7-Deaza-dA; U<sup>5F</sup>, 5F-dU; U<sup>5F</sup>, 5-Fluoro-dU. A, C, G, and T are deoxyribonucleosides; δ, C, G, and U represent 2'-OMe-ribonucleosides.
Example 5

Library 5

In the phosphoramidate dinucleotide library (Table 5), an additional diversity element was incorporated at the backbone. A 192-member dinucleotide phosphoramidate library was assembled by parallel synthesis using H-phosphonate chemistry in conjunction with a series of amines. The requisite dinucleoside H-phosphonates (5'-DMT-off) were assembled on CPG using the corresponding commercially available H-phosphonates and the CPG-bound nucleosides in conjunction with 1-Adamantanecarbonyl chloride as the activator. The CPG-bond nucleoside H-phosphonates were converted to phosphoramidates by reported procedure."

Table 5: Phosphoramidate Libraries

<table>
<thead>
<tr>
<th>TXXN</th>
<th>TXXCN</th>
<th>TXXAN</th>
<th>TXXYN</th>
<th>TXXMN</th>
<th>TXXRN</th>
<th>TXXSN</th>
<th>TXXTN</th>
<th>TXXNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>TACN</td>
<td>TACCN</td>
<td>TACAN</td>
<td>TACYN</td>
<td>TACMN</td>
<td>TACRN</td>
<td>TACSN</td>
<td>TACTN</td>
<td>TACNN</td>
</tr>
<tr>
<td>TAAN</td>
<td>TAACN</td>
<td>TAAAN</td>
<td>TAAYN</td>
<td>TAAAN</td>
<td>TAAAN</td>
<td>TAAAN</td>
<td>TAAAN</td>
<td>TAAAN</td>
</tr>
<tr>
<td>AACN</td>
<td>AACCN</td>
<td>AACAN</td>
<td>AACYN</td>
<td>AACMN</td>
<td>AACRN</td>
<td>AACSN</td>
<td>AACTN</td>
<td>AACNN</td>
</tr>
</tbody>
</table>

N corresponds to N, N<sub>2</sub>, N<sub>3</sub>, N<sub>4</sub>, N<sub>5</sub>, N<sub>6</sub> & N<sub>7</sub> corresponds to N<sub>8</sub>, N<sub>9</sub>, N<sub>10</sub>. A, C, G, T correspond to deoxyribonucleosides, Δ, Ε, Ζ, Η correspond to 2'-OMe-rA, rC, rG, rT nucleosides. All internucleoside linkages are phosphoramidates.

Example 5

Antiviral Evaluation

Different classes of NAB libraries were prepared and evaluated. The results demonstrate that the libraries represent biologically relevant chemical diversity for drug discovery.
The libraries 1-5 were evaluated in antiviral assays against HSV-1. A number of compounds, induced 30 to 60% inhibition of HSV-1-induced plaques at a dose of 25 micromolar when compared with acyclovir (EC₉₀ of 6 micromolar).

Example 6

Characteristics of a possible new HSV-1 protein target for NAB™ compounds

The results from preliminary screening of antiviral activity for NAB™ compounds by SDS-PAGE and Western blotting have indicated that the compounds interact either directly or indirectly with one of the unique viral proteins. This protein has the following characteristics: A molecular weight of about 90kDa; it is an immediate early (IE) or early (E) protein as it appears during the early phases after infection around 6hr post infection and is most likely one of the important regulatory or structural proteins; the protein is inhibited completely by the NAB™ compounds between 6 and 24 hrs post infection where after it reappears; the protein is stained by both Coomassie Blue and Silver Stains and is most likely a glycosylated protein as determined by tunicamycin treatment. The nature and the kinetics of the protein, its production and its role or function during the virus growth cycle is determined by pulse chase experiments, virus labeling with ^35S-methionine, endo H/F digestions.
The following specific references, also incorporated herein by reference, are indicated in the above discussion and examples by the corresponding number generally within parenthesis or brackets.


What is claimed is:

1. A compound library comprising two or more compounds of the following Formula I or I’:

wherein L^1 and L^2 are independently O or a linking group such as e.g. an amide, ester, diester or the like, or an optionally substituted alkylene (e.g. C_{1-20} alkylene), optionally substituted alkenylene (e.g., C_{2-20} alkenylene) or alkynylene (e.g., C_{2-20} alkynylene) having such groups either as a chain member of pendant to the chain, and which may be optionally substituted with one or more substituents selected from a group consisting of O, S, Se,

NR^1NR^2, CR^1CR^2, OR, SR and SeR, or an enzymatically reactive;

Q is carbon or a heteroatom such as O, S or N;

R^1, R^2, R^3 are each independently a hydrogen or a hydroxyl group or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl,

optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring;

n=1 to 5;

and pharmaceutically acceptable salts thereof.
2. The library of claim 1 wherein at least one compound has a sugar group is in open chain form.

3. The library of claim 1 wherein an enantiomerically enriched mixture of a compound is present.

4. A compound library comprising two or more compounds of the following Formula II or II':

\[
\begin{align*}
&\text{II} \\
&\text{II'} \\
\end{align*}
\]

wherein X and Y are each independently selected from a group consisting of O, S, Se, NR\textsuperscript{1}NR\textsuperscript{2}, CR\textsuperscript{1}CR\textsuperscript{2}, OR, SR, and SeR, or one or both of X and Y are an enzymatically reactive moiety;

R is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

R\textsuperscript{1}, R\textsuperscript{2} and R\textsuperscript{3} are each independently selected from a group as defined by R;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring;
and pharmaceutically acceptable salts thereof.

5. The library of claim 1 wherein at least one compound is of the following formula III or III':

![Chemical structures](image)

wherein $X$ and $Y$ are each independently selected from a group consisting of O, S, Se, NR, CR, OR, SR and SeR, or one or both of $X$ and $Y$ are an enzymatically reactive moiety;

15 $R$ is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 $N$, $O$ or $S$ atoms;

18 $R^1$, $R^2$ and $R^3$ are each independently selected from a group as defined by $R$;

$B$ is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 $N$, $O$ or $S$ atoms, or a heterocyclic structure that is covalently linked to the sugar ring;

and pharmaceutically acceptable salts thereof.

6. The library of any one of claims 1-5, wherein the compounds have at least one furanose ring in the C-2' endo conformation.
7. The library of any one of claims 1-5, wherein the compounds have at least one furanose ring in the C-3' endo conformation.

8. The library of any one of claims 1-5, wherein the compounds have at least one furanose ring in the C-2' endo conformation and at least one furanose ring in the C-3' endo conformation.

9. The library of any one of claims 1-5, wherein the compounds have from 1 to 50 nucleoside residues.

10. The library of any one of claims 1-5, wherein the compounds have from 1 to 10 nucleoside residues.

11. The library of any one of claims 1-5, wherein the compounds have from 1 to 5 nucleoside residues.

12. The library of any one of claims 1-5, wherein the compounds have 4 nucleoside residues.

13. The library of any one of claims 1-5, wherein the compounds have 3 nucleoside residues.

14. The library of any one of claims 1-5, wherein the compounds have 2 nucleoside residues.

15. The library of any one of claims 1-5, wherein the compounds have 1 nucleoside residue.

16. The library of any one of claims 1 through 15 wherein the library has been constructed using solid-phase synthesis.

17. The library of any one of claims 1 through 15 wherein the library has been constructed using solution-phase synthesis.
18. A library of any one of claims 1 through 15 wherein the library is obtainable by a process comprising:
adding one or more reagents to a reaction vessel capable of agitation and containing a resin reaction support material;
agitating the reaction vessel during reaction of the reagents; and
centrifuging the reaction vessel and removing desired reaction materials therefrom.

19. The library of claim 18 wherein the library has been constructed using an automated solution-phase synthesis.

20. Use of the library of any one of claims 1 through 15 to find a specific interacting partner for a nucleic acid.

21. Use of the library of any one of claims 1 through 15 to find a specific interacting partner for a protein.

22. The use of claim 20 wherein the nucleic acid is RNA or DNA.

23. The use of claim 21 wherein the protein is an antibody, receptor or ligand.

24. A compound of the following Formula I or I':

\[ \begin{align*}
R^3 &- \text{L}^1 \\
R^2 &- \text{L}^2 \\
\text{R}^1 &
\end{align*} \]

\[ \begin{align*}
R^3 &- \text{L}^1 \\
R^2 &- \text{L}^2 \\
\text{R}^1 &
\end{align*} \]
wherein L¹ and L² are independently O or a linking group such as e.g. an amide, ester, diester or the like, or an optionally substituted alkyne (e.g. C₁⁻²₀ alkyne), optionally substituted alkenylene (e.g., C₂⁻²₀ alkenylene) or alkynylene (e.g., C₂⁻²₀ alkynylene) having such groups either as a chain member of pendant to the chain, and which may be optionally substituted with one or more substituents selected from a group consisting of O, S, Se, NR¹NR², CR¹CR², OR, SR and SeR, or an enzymatically reactive;

Q is carbon or a heteroatom such as O, S or N;

R¹, R², R³ are each independently a hydrogen or a hydroxyl group or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heterocyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heterocyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring;

n=1 to 5;

and pharmaceutically acceptable salts thereof.

25. A compound of claim 24 wherein the sugar group is in open chain form.

26. A compound of claim 24 wherein an enantiomerically enriched mixture of a compound is present.

27. A compound of the following Formula II or II':
wherein X and Y are each independently selected from a group consisting of O, S, Se, NR\(^1\)NR\(^2\), CR\(^1\)CR\(^2\), OR, SR and SeR, or one or both of X and Y are an enzymatically reactive moiety;

R is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

R\(^1\), R\(^2\) and R\(^3\) are each independently selected from a group as defined by R;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring;

and pharmaceutically acceptable salts thereof.

25 28. A compound of the following Formula III or III':

---

**Notice:** The chemical structures and formulas are represented in a textual format. For a precise interpretation, a chemical editor capable of displaying complex molecular structures is recommended.
wherein X and Y are each independently selected from a group consisting of O, S, Se, NR₁NR₂, CR₁CR₂, OR, SR and SeR, or one or both of X and Y are an enzymatically reactive moiety;

R is hydrogen or an optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, optionally substituted aralkyl, optionally substituted cycloalkyl, optionally substituted cycloalkenyl, optionally substituted carbocyclic aryl, an optionally substituted mononucleotide, an optionally substituted polynucleotide, or an optionally substituted heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused ring and 1 to 3 N, O or S atoms;

R¹, R² and R³ are each independently selected from a group as defined by R;

B is optionally substituted adenine, optionally substituted thymidine, optionally substituted cytosine or an optionally substituted guanine, preferably where the optional substituents are alkyl, carbocyclic aryl, or heteroaromatic or heteroalicyclic group preferably having from 1 to 3 separate or fused rings and 1 to 3 N, O or S atoms, or a heterocyclic structure that is covalently linked to the sugar ring;

and pharmaceutically acceptable salts thereof.

29. A method for treating virally infected cells comprising administering to the cells an anti-viral effective amount of a compound of any one of claims 24 through 28.

30. The method of claim 29 wherein the cells are infected with a herpes virus.

31. The method of claim 29 wherein the cells are infected with a cytomegalovirus.
32. A method for treating bacterially infected cells comprising administering to the cells an anti-bacteria effective amount of a compound of any one of claims 24 through 28.

33. A method for treating a mammal suffering from or susceptible to a viral infection, comprising administering to the mammal an anti-viral effective amount of a compound of any one of claims 24 through 28.

34. The method of claim 33 wherein the mammal is suffering from a herpes infection.

35. The method of claim 33 wherein the mammal is suffering from a cytomegalovirus infection.

36. A method for treating a mammal suffering from or susceptible to a bacterial infection, comprising administering to the mammal an anti-bacterial effective amount of a compound of any one of claims 24 through 28.

37. A pharmaceutical composition comprising a compound of any one of claims 24 through 28 and a pharmaceutically acceptable carrier.

38. A method for synthesis of a compound library, comprising:
adding one or more reagents to a reaction vessel capable of agitation and containing a resin reaction support material;
agitating the reaction vessel during reaction of the reagents; and
centrifuging the reaction vessel and removing desired reaction materials therefrom.

39. Use of the library or compound of any one of claims 1 through 15 or claims 24 through 28 to find a specific inhibitor for a viral kinase.

40. Use of the library or compound of any one of claims 1 through 15 or claims 24 through 28 to find a specific inhibitor for a viral polymerase.
41. Use of the library or compound of any one of claims 1 through 15 or claims 24 through 28 to find a specific compound which causes disruption of the association between a helicase-primase complex and a viral nucleic acid to which it is bound.

42. A method for treating a mammal suffering from or susceptible to a fungal infection, comprising administering to the mammal an anti-fungal effective amount of a compound of any one of claims 24 through 28.

43. A method for generating desired libraries of 3'-modified nucleoside analogs and 5'-modified nucleoside analogs, said method comprising:

synthesis of a desired library using a parallel solution-phase chemistry; and,

subjecting said library to a two-stage liquid/liquid extraction; wherein,

a first liquid/liquid extraction is applied to remove hydrophobic compounds by partitioning between ethyl acetate and at least about 2% aqueous NaHCO₃; and,

a second extraction is applied by partitioning between CHCl₃ and H₂O; wherein,

hydrophilic compounds are retained in the H₂O and CHCl₃-soluble compounds are removed by repetitive CHCl₃ washing.
Figure 1.
Figure 2
Figure 3.
Figure 5.
a: Z = H, B = adenine;  e: Z = OCH₃, B = adenine;  
b: Z = H, B = cytosine;  f: Z = OCH₃, B = cytosine;  
c: Z = H, B = guanine;  g: Z = OCH₃, B = guanine;  
d: Z = H, B = thymine;  h: Z = OCH₃, B = uracil.

Prt

Prf = Aᵇz, Cᵇz, Cⁱᵇu, T, U;

for description of R see Table 1.

Figure 6.
FIGURE 7.
Table 6. Yields and spectroscopic characterization of the representative library members 2a-g.
<table>
<thead>
<tr>
<th>T</th>
<th>H</th>
<th>xxxv</th>
<th>2d-xxxv</th>
<th>79%</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>OMe</td>
<td>iv</td>
<td>2e-iv</td>
<td>82%</td>
</tr>
<tr>
<td>A</td>
<td>OMe</td>
<td>v</td>
<td>2e-v</td>
<td>85%</td>
</tr>
<tr>
<td>A</td>
<td>OMe</td>
<td>xvii</td>
<td>2e-xvii</td>
<td>77%</td>
</tr>
<tr>
<td>A</td>
<td>OMe</td>
<td>xvii</td>
<td>2e-xvii</td>
<td>89%</td>
</tr>
<tr>
<td>A</td>
<td>OMe</td>
<td>xvi</td>
<td>2e-xvi</td>
<td>82%</td>
</tr>
<tr>
<td>A</td>
<td>OMe</td>
<td>xvi</td>
<td>2e-xvi</td>
<td>87%</td>
</tr>
<tr>
<td>C</td>
<td>OMe</td>
<td>xii</td>
<td>2f-xii</td>
<td>84%</td>
</tr>
<tr>
<td>C</td>
<td>OMe</td>
<td>xii</td>
<td>2f-xii</td>
<td>88%</td>
</tr>
<tr>
<td>C</td>
<td>OMe</td>
<td>xxi</td>
<td>2f-xxi</td>
<td>80%</td>
</tr>
<tr>
<td>G</td>
<td>OMe</td>
<td>i</td>
<td>2g-i</td>
<td>90%</td>
</tr>
<tr>
<td>G</td>
<td>OMe</td>
<td>viii</td>
<td>2g-viii</td>
<td>87%</td>
</tr>
<tr>
<td>G</td>
<td>OMe</td>
<td>xvi</td>
<td>2g-xvi</td>
<td>76%</td>
</tr>
</tbody>
</table>

Table 6 (continued). Yields and spectroscopic characterization of the representative library members 2a-g.
Table 7. Chemical structures of the alcohols used to introduce R groups in the prototypic library of 2.