



(22) Date de dépôt/Filing Date: 1996/11/13

(41) Mise à la disp. pub./Open to Public Insp.: 1997/05/22

(62) Demande originale/Original Application: 2 237 336

(30) Priorités/Priorities: 1995/11/14 (US08/558,633);  
1996/07/12 (US08/678,903); 1996/10/31 (US08/744,701)

(51) Cl.Int./Int.Cl. *G01N 24/08* (2006.01),  
*C40B 30/04* (2006.01), *G01N 33/53* (2006.01),  
*G06F 19/00* (2006.01), *G01N 33/68* (2006.01)

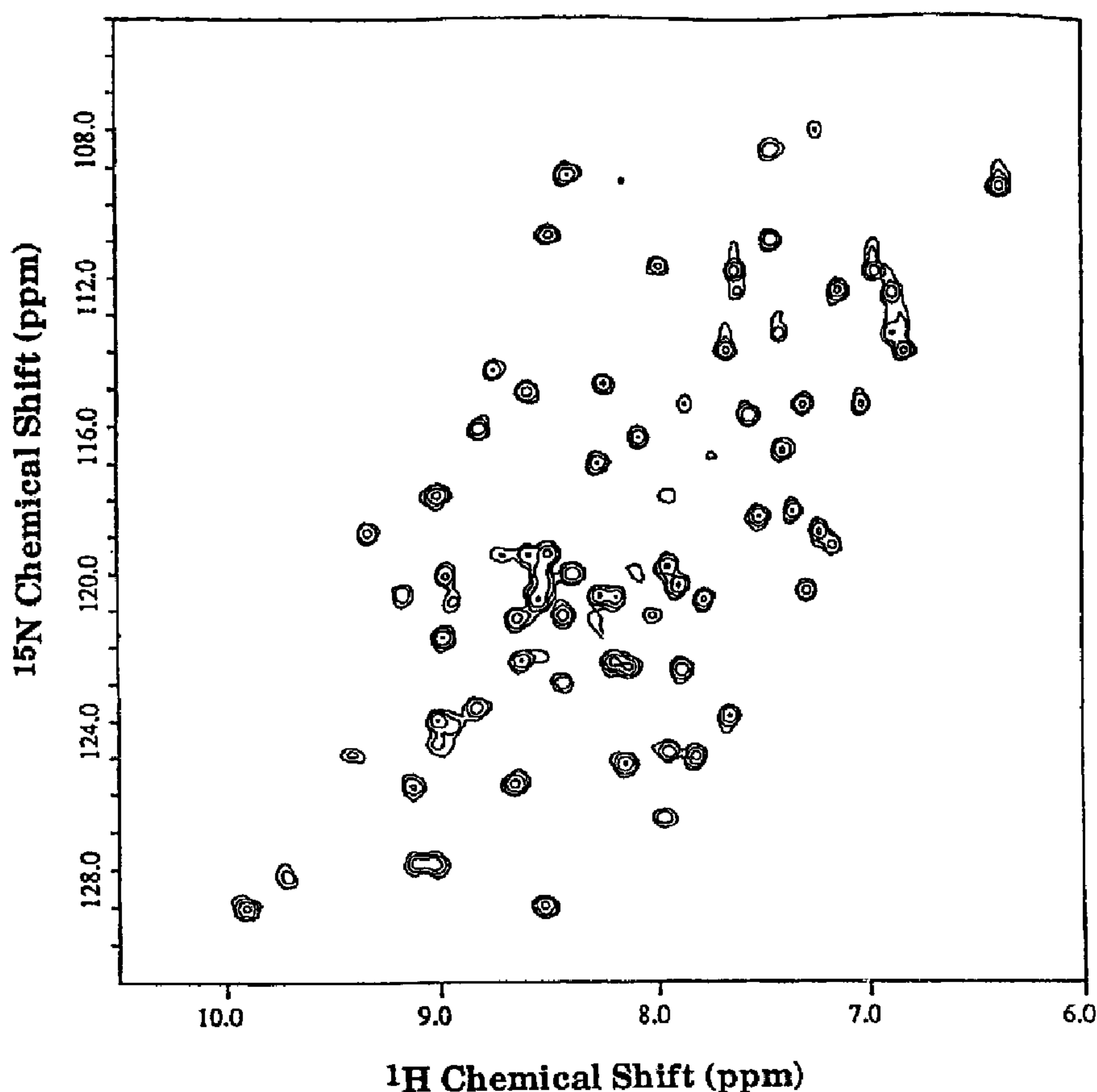
(71) Demandeur/Applicant:  
ABBOTT LABORATORIES, US

(72) Inventeurs/Inventors:  
FESIK, STEPHEN W., US;  
HAJDUK, PHILIP J., US;  
OLEJNICZAK, EDWARD T., US

(74) Agent: OGILVY RENAULT LLP/S.E.N.C.R.L.,S.R.L.

(54) Titre : METHODE D'IDENTIFICATION DE LIGANDS

(54) Title: METHOD FOR IDENTIFYING LIGANDS



(57) Abrégé/Abstract:

The present invention provides a process of designing compounds which bind to a specific target molecule. The process includes the steps of a) identifying a first ligand to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; b)



(57) **Abrégé(suite)/Abstract(continued):**

identifying a second ligand to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; c) forming a ternary complex by binding the first and second ligands to the target molecule; d) determining the three-dimensional structure of the ternary complex and thus the spatial orientation of the first and second ligands on the target molecule; and e) linking the first and second ligands to form the drug, wherein the spatial orientation of step (d) is maintained.



**ABSTRACT**

The present invention provides a process of designing compounds which bind to a specific target molecule. The process includes the steps of a) identifying a first ligand to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; b) 5 identifying a second ligand to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; c) forming a ternary complex by binding the first and second ligands to the target molecule; d) determining the three-dimensional structure of the ternary complex and thus the spatial orientation of the first and second ligands on the target molecule; and e) linking the first and second ligands to form the drug, wherein the 10 spatial orientation of step (d) is maintained.



## METHOD FOR IDENTIFYING LIGANDS

This application is a Divisional of Application No. 2,237,336 filed November 13, 1996.

### Technical Field of the Invention

The present invention pertains to a method for the use of two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectral analysis to design ligands that bind to a target biomolecule.

### 5 Background of the Invention

One of the most powerful tools for discovering new drug leads is random screening of synthetic chemical and natural product databases to discover compounds that bind to a particular target molecule (*i.e.*, the identification of ligands of that target). Using this method, ligands may be identified by their ability to form a physical association with a target molecule  
10 or by their ability to alter a function of a target molecule.

When physical binding is sought, a target molecule is typically exposed to one or more compounds suspected of being ligands and assays are performed to determine if complexes between the target molecule and one or more of those compounds are formed. Such assays, as is well known in the art, test for gross changes in the target molecule (*e.g.*,  
15 changes in size, charge, mobility) that indicate complex formation.

Where functional changes are measured, assay conditions are established that allow for measurement of a biological or chemical event related to the target molecule (*e.g.*, enzyme catalyzed reaction, receptor-mediated enzyme activation). To identify an alteration, the function of the target molecule is determined before and after exposure to the test compounds.

20 Existing physical and functional assays have been used successfully to identify new drug leads for use in designing therapeutic compounds. There are, however, limitations inherent to those assays that compromise their accuracy, reliability and efficiency.

A major shortcoming of existing assays relates to the problem of "false positives". In a typical functional assay, a "false positive" is a compound that triggers the assay but which  
25 compound is not effective in eliciting the desired physiological response. In a typical physical assay, a "false positive" is a compound that, for example, attaches itself to the target but in a non-specific manner (*e.g.*, non-specific binding). False positives are particularly prevalent and problematic when screening higher concentrations of putative ligands because many compounds have non-specific effects at those concentrations.

30 In a similar fashion, existing assays are plagued by the problem of "false negatives", which result when a compound gives a negative response in the assay but which compound is actually a ligand for the target. False negatives typically occur in assays that use concentrations of test compounds that are either too high (resulting in toxicity) or too low relative to the binding or dissociation constant of the compound to the target.

35 Another major shortcoming of existing assays is the limited amount of information provided by the assay itself. While the assay may correctly identify compounds that attach to



or elicit a response from the target molecule, those assays typically do not provide any information about either specific binding sites on the target molecule or structure activity relationships between the compound being tested and the target molecule. The inability to provide any such information is particularly problematic where the screening assay is  
5 being used to identify leads for further study.

It has recently been suggested that X-ray crystallography can be used to identify the binding sites of organic solvents on macromolecules. However, this method cannot determine the relative binding affinities at different sites on the target. It is only applicable to very stable target proteins that do not denature in the presence of high  
10 concentrations of organic solvents. Moreover, this approach is not a screening method for rapidly testing many compounds that are chemically diverse, but is limited to mapping the binding sites of only a few organic solvents due to the long time needed to determine the individual crystal structures.

Compounds are screened to identify leads that can be used in the design of new  
15 drugs that alter the function of the target biomolecule. Those new drugs can be structural analogs of identified leads or can be conjugates of one or more such lead compounds. Because of the problems inherent to existing screening methods, those methods are often of little help in designing new drugs.

There continues to be a need to provide new, rapid, efficient, accurate and reliable  
20 means of screening compounds to identify and design ligands that specifically bind to a particular target.

#### Brief Summary of the Invention

In accordance with the invention there is provided a method for discovering  
25 ligands to target molecules using structure-activity relationships obtained from nuclear magnetic resonance, comprising:

i) screening compounds having a molecular weight of 100-300 which bind to a subsite 1 of a given target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR to measure binding affinity;



- 2a -

ii) screening analogs prepared from binding results obtained in step i) to optimize binding of a first fragment to the target molecule;

iii) screening for compounds and corresponding analogs which bind to a nearby binding site, subsite 2, of the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR to measure binding affinity to optimize binding of a second fragment to the target molecule; and

iv) combining the first and second fragments to design a ligand.

In a principal aspect, the present invention provides a process for the design and identification of compounds which bind to a given target biomolecule. That process comprises the steps of: a) identifying a first ligand to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; b) identifying a second ligand to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; c) forming a ternary complex by binding the first and second ligands to the target molecule; d) determining the three dimensional structure of the ternary complex and thus the spatial orientation of the first and second ligands on the target molecule; e) linking the first and second ligands to form the drug, wherein the spatial orientation of step (d) is maintained.

This aspect of the present invention uses the two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopic screening process as set forth below to identify a first and subsequent ligands that bind to the target molecule. A complex of the target molecule and two or more ligands is formed and the three-dimensional structure of that complex is determined preferably using NMR spectroscopy or X-ray crystallography. That three-dimensional



- 3 -

structure is used to determine the spatial orientation of the ligands relative to each other and to the target molecule.

Based on the spatial orientation, the ligands are linked together to form the drug. The selection of an appropriate linking group is made by maintaining the spatial orientation of the ligands to one another and to the target molecule based upon principles of bond angle and bond length information well known in the organic chemical art.

Thus, the molecular design aspect of the present invention comprises identifying a first ligand moiety to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; identifying subsequent ligand moieties to the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy; forming a complex of the first and subsequent ligand moieties to the target molecule; determining the three dimensional structure of the complex and, thus, the spatial orientation of the first and subsequent ligand moieties on the target molecule; and linking the first and subsequent ligand moieties to form the drug to maintain the spatial orientation of the ligand moieties.

The identification of subsequent ligand moieties can be performed in the absence or presence of the first ligand (*e.g.*, the target molecule can be bound to the first ligand before being exposed to the test compounds for identification of the second ligand).

The present invention further contemplates a drug designed by the design process of this invention.

Chemical compounds can be screened for binding to a given target biomolecule by a process involving the steps of a) first generating a first two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of a  $^{15}\text{N}$ -labeled target molecule; b) exposing the labeled target molecule to one or a mixture of chemical compounds; c) next, generating a second two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of the labeled target molecule that has been exposed to one or a mixture of compounds in step (b); and d) comparing said first and second two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectra to determine differences between said first and said second spectra, the differences identifying the presence of one or more compounds that are ligands which have bound to the target molecule.

Where the process screens more than one compound in step (b), that is, a mixture of compounds, and where a difference between the first spectrum generated from the target molecule alone and that generated from the target molecule in the presence of the mixture, additional steps are performed to identify which specific compound or compounds contained in the mixture is binding to the target molecule. Those additional steps comprise the steps of e) exposing the  $^{15}\text{N}$ -labeled target molecule individually to each compound of the mixture, f) generating a two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of the labeled target molecule that has been individually exposed to each compound; and g) comparing each spectrum generated in step f) to the first spectrum generated from the target molecule alone to



- 4 -

determine differences in any of those compared spectra, the differences identifying the presence of a compound that is a ligand which has bound to the target molecule.

Because the chemical shift values of the particular  $^{15}\text{N}/^1\text{H}$  signals in the two-dimensional correlation spectrum correspond to known specific locations of atomic groupings in the target molecule (*e.g.*, the N-H atoms of the amide or peptide link of a particular amino acid residue in a polypeptide), the screening process allows not only for the identification of which compound(s) bind to a particular target molecule, but also permit the determination of the particular binding site of the ligand on the target molecule.

The dissociation constant,  $K_D$ , for a given ligand and its target molecule can be determined by this process, if desired, by performing the steps of a) generating a first two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of a  $^{15}\text{N}$ -labeled target molecule; b) exposing the labeled target molecule to various concentrations of a ligand; c) generating a two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum at each concentration of ligand in step (b); d) comparing each spectrum from step (c) to the first spectrum from step (a); and e) calculating the dissociation constant between the target molecule and the ligand from those differences according to the equation:

$$K_D = \frac{([P]_0 - x)([L]_0 - x)}{x}$$

An advantageous capability of the screening method is its ability to determine the dissociation constant of one ligand of the target molecule in the presence of a second molecule already bound to the ligand. This is generally not possible with prior art methods which employ "wet chemical" analytical methods of determining binding of a ligand to a target molecule substrate.

The process of determining the dissociation constant of a ligand can be performed in the presence of a second bound ligand. Accordingly, the  $^{15}\text{N}$ -labeled target molecule is bound to that second ligand before exposing that target to the test compounds.

The ability of the screening method to determine not only the existence of binding between one ligand and the target molecule, but also the particular site of binding in the presence of a second bound ligand, permits the capability to design a drug that comprises two or more linked moieties made up of the ligands.

In a preferred embodiment of the present invention, the target molecule used in the molecular design process is a polypeptide. The polypeptide target is preferably produced in recombinant form from a host cell transformed with an expression vector that contains a polynucleotide that encodes the polypeptide, by culturing the transformed host cell in a medium that contains an assimilable source of  $^{15}\text{N}$  such that the recombinantly produced polypeptide is labeled with  $^{15}\text{N}$ .



### Brief Description of the Drawings

In the drawings which form a portion of the specification:

FIG. 1 shows a  $^{15}\text{N}/^1\text{H}$  correlation spectrum of the DNA binding domain of uniformly  $^{15}\text{N}$ -labeled human papillomavirus E2. The spectrum (80 complex points, 4 scans/fid) was acquired on a 0.5 mM sample of E2 in 20 mM phosphate (pH 6.5), 10 mM dithiothreitol (DTT) and 10% deuterium oxide ( $\text{D}_2\text{O}$ ).

FIG. 2 shows  $^{15}\text{N}/^1\text{H}$  correlation spectra of the DNA binding domain of uniformly  $^{15}\text{N}$ -labeled human papillomavirus E2 before (thin multiple contours) and after (thick single contours) addition of a final test compound. The final concentration of compound was 1.0 mM. All other conditions are as stated in FIG. 1. Selected residues that show significant changes upon binding are indicated.

FIG. 3 shows  $^{15}\text{N}/^1\text{H}$ -correlation spectra of the DNA binding domain of uniformly  $^{15}\text{N}$ -labeled human papillomavirus E2 before (thin multiple contours) and after (thick single contours) addition of a second test compound. The final concentration of compound was 1.0 mM. All other conditions are as stated in FIG. 1. Selected residues that show significant changes upon binding are indicated.

FIG. 4 shows  $^{15}\text{N}/^1\text{H}$  correlation spectra of the catalytic domain of uniformly  $^{15}\text{N}$ -labeled stromelysin before (thin multiple contours) and after (thick single contours) addition of a test compound. The final concentration of compound was 1.0 mM. The spectra (80 complex points, 8 scans/fid) were acquired on a 0.3 mM sample of SCD in 20 mM TRIS (pH 7.0), 20 mM  $\text{CaCl}_2$  and 10%  $\text{D}_2\text{O}$ . Selected residues that show significant changes upon binding are indicated.

FIG. 5 shows  $^{15}\text{N}/^1\text{H}$  correlation spectra of the Ras-binding domain of uniformly  $^{15}\text{N}$ -labeled RAF peptide (residues 55-132) before (thin multiple contours) and after (thick single contours) addition of a test compound. The final concentration of compound was 1.0 mM. The spectra (80 complex points, 8 scans/fid) were acquired on a 0.3 mM sample of the RAF fragment in 20 mM phosphate (pH 7.0), 10 mM DTT and 10%  $\text{D}_2\text{O}$ . Selected residues that show significant changes upon binding are indicated.

FIG. 6 shows  $^{15}\text{N}/^1\text{H}$  correlation spectra of uniformly  $^{15}\text{N}$ -labeled FKBP before (thin multiple contours) and after (thick single contours) addition of a test compound. The final concentration of compound was 1.0 mM. The spectra (80 complex points, 4 scans/fid) was acquired on a 0.3 mM sample of FKBP in 50 mM phosphate (pH 6.5), 100 mM NaCl and 10%  $\text{D}_2\text{O}$ . Selected residues that show significant changes upon binding are indicated.

FIG. 7 shows a first depiction of the NMR-derived structure of the DNA-binding domain of E2. The two monomers of the symmetric dimer are oriented in a top-bottom fashion, and the N- and C-termini of each monomer are indicated (N and C for one monomer, N\* and C\* for the other). Shown in ribbons are the residues which exhibit significant



- 6 -

chemical shift changes ( $\Delta\delta(^1\text{H}) > 0.04$  ppm;  $\Delta\delta(^{15}\text{N}) > 0.1$  ppm) upon binding to a first test compound. These residues correspond to the DNA-recognition helix of E2. Selected residues are numbered for aid in visualization.

FIG. 8 shows a second depiction of the NMR-derived structure of the DNA-binding domain of E2. The two monomers of the symmetric dimer are oriented in a top-bottom fashion, and the N- and C-termini of each monomer are indicated (N and C for one monomer, N\* and C\* for the other). Shown in ribbons are the residues which exhibit significant chemical shift changes ( $\Delta\delta(^1\text{H}) > 0.04$  ppm;  $\Delta\delta(^{15}\text{N}) > 0.1$  ppm) upon binding to a second test compound. These residues are located primarily in the dimer interface region. Selected residues are numbered for aid in visualization.

FIG. 9 shows a depiction of the NMR-derived structure of the catalytic domain of stromelysin. The N- and C-termini are indicated. Shown in ribbons are the residues which exhibit significant chemical shift changes ( $\Delta\delta(^1\text{H}) > 0.04$  ppm;  $\Delta\delta(^{15}\text{N}) > 0.1$  ppm) upon binding to a test compound. These either form part of the S1' binding site or are spatially proximal to this site. Selected residues are numbered for aid in visualization.

FIG. 10 shows a ribbon plot of a ternary complex of first and second ligands bound to the catalytic domain of stromelysin.

FIG. 11 shows the correlation between the NMR binding data and a view of the NMR-derived three-dimensional structure of FKBP.

FIG. 12 shows a ribbon plot of a ternary complex involving FKBP, a fragment analog of ascomycin, and a benzanilide compound.

### Detailed Description of the Invention

The present invention provides a rapid and efficient method for designing ligands that bind to therapeutic target molecules.

Ligands are identified by testing the binding of molecules to a target molecule (*e.g.*, protein, nucleic acid, etc.) by following, with nuclear magnetic resonance (NMR) spectroscopy, the changes in chemical shifts of the target molecule upon the addition of the ligand compounds in the database.

From an analysis of the chemical shift changes of the target molecule as a function of ligand concentration, the binding affinities of ligands for biomolecules are also determined.

The location of the binding site for each ligand is determined from an analysis of the chemical shifts of the biomolecule that change upon the addition of the ligand and from nuclear Overhauser effects (NOEs) between the ligand and biomolecule.

Information about the structure/activity relationships between ligands identified by such a process can then be used to design new drugs that serve as ligands to the target molecule. By way of example, where two or more ligands to a given target molecule are identified, a complex of those ligands and the target molecule is formed. The spatial



- 7 -

orientation of the ligands to each other as well as to the target molecule is derived from the three-dimensional structure. That spatial orientation defines the distance between the binding sites of the two ligands and the orientation of each ligand to those sites.

Using that spatial orientation data, the two or more ligands are then linked together to form a new ligand. Linking is accomplished in a manner that maintains the spatial orientation of the ligands to one another and to the target molecule.

There are numerous advantages to the NMR-based discovery and design processes of the present invention. First, because a process of the present invention identifies ligands by directly measuring binding to the target molecule, the problem of false positives is significantly reduced. Because the present process identifies specific binding sites to the target molecule, the problem of false positives resulting from the non-specific binding of compounds to the target molecule at high concentrations is eliminated.

Second, the problem of false negatives is significantly reduced because the present process can identify compounds that specifically bind to the target molecule with a wide range of dissociation constants. The dissociation or binding constant for compounds can actually be determined with the present process.

Other advantages of the present invention result from the variety and detailed data provided about each ligand from the discovery and design processes.

Because the location of the bound ligand can be determined from an analysis of the chemical shifts of the target molecule that change upon the addition of the ligand and from nuclear Overhauser effects (NOEs) between the ligand and biomolecule, the binding of a second ligand can be measured in the presence of a first ligand that is already bound to the target. The ability to simultaneously identify binding sites of different ligands allows a skilled artisan to 1) define negative and positive cooperative binding between ligands and 2) design new drugs by linking two or more ligands into a single compound while maintaining a proper orientation of the ligands to one another and to their binding sites.

Further, if multiple binding sites exist, the relative affinity of individual binding moieties for the different binding sites can be measured from an analysis of the chemical shift changes of the target molecule as a function of the added concentration of the ligand. By simultaneously screening numerous structural analogs of a given compound, detailed structure/activity relationships about ligands is provided.

In part, the present invention provides a process of screening compounds to identify ligands that bind to a specific target molecule. That process comprises the steps of: a) generating a first two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of a  $^{15}\text{N}$ -labeled target molecule; b) exposing the labeled target molecule to one or more compounds; c) generating a second two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of the labeled target molecule that has been exposed to the compounds of step (b); and d) comparing the first and second



- 8 -

spectra to determine whether differences in those two spectra exist, which differences indicate the presence of one or more ligands that have bound to the target molecule.

Where a process of the present invention screens more than one compound in step (b) and where a difference between spectra is observed, additional steps are performed to identify which specific compound is binding to the target molecules. Those additional steps comprise generating a two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum for each individual compound and comparing each spectrum to the first spectrum to determine whether differences in any of those compared spectra exist, which differences indicate the presence of a ligand that has bound to the target molecule.

Any  $^{15}\text{N}$ -labeled target molecule can be used in a process of the present invention. Because of the importance of proteins in medicinal chemistry, a preferred target molecule is a polypeptide. The target molecule can be labeled with  $^{15}\text{N}$  using any means well known in the art. In a preferred embodiment, the target molecule is prepared in recombinant form using transformed host cells. In an especially preferred embodiment, the target molecule is a polypeptide. Any polypeptide that gives a high resolution NMR spectrum and can be partially or uniformly labeled with  $^{15}\text{N}$  can be used. The preparation of uniformly  $^{15}\text{N}$ -labeled exemplary polypeptide target molecules is set forth hereinafter in the Examples.

A preferred means of preparing adequate quantities of uniformly  $^{15}\text{N}$ -labeled polypeptides is to transform a host cell with an expression vector that contains a polynucleotide that encodes that polypeptide and culture the transformed cell in a culture medium that contains assimilable sources of  $^{15}\text{N}$ . Assimilable sources of  $^{15}\text{N}$  are well known in the art. A preferred such source is  $^{15}\text{NH}_4\text{Cl}$ .

Means for preparing expression vectors that contain polynucleotides encoding specific polypeptides are well known in the art. In a similar manner, means for transforming host cells with those vectors and means for culturing those transformed cells so that the polypeptide is expressed are also well known in the art.

The screening process begins with the generation or acquisition of a two-dimensional  $^{15}\text{N}/^1\text{H}$  correlation spectrum of the labeled target molecule. Means for generating two-dimensional  $^{15}\text{N}/^1\text{H}$  correlation spectra are well known in the art [See, *e.g.*, D. A. Egan, et al., *Biochemistry*, 32:8, pgs. 1920-1927 (1993); Bax, A., Grzesiek, S., *Acc. Chem. Res.*, 26:4, pgs. 131-138 (1993)].

The NMR spectra that are typically recorded in the screening procedure of the present invention are two-dimensional  $^{15}\text{N}/^1\text{H}$  heteronuclear single quantum correlation (HSQC) spectra. Because the  $^{15}\text{N}/^1\text{H}$  signals corresponding to the backbone amides of the proteins are usually well-resolved, the chemical shift changes for the individual amides are readily monitored.

In generating such spectra, the large water signal is suppressed by spoiling gradients. To facilitate the acquisition of NMR data on a large number of compounds (*e.g.*, a database



- 9 -

of synthetic or naturally occurring small organic compounds), a sample changer is employed. Using the sample changer, a total of 60 samples can be run unattended. Thus, using the typical acquisition parameters (4 scans per free induction decay (fid), 100-120 HSQC spectra can be acquired in a 24 hour period.

5 To facilitate processing of the NMR data, computer programs are used to transfer and automatically process the multiple two-dimensional NMR data sets, including a routine to automatically phase the two-dimensional NMR data. The analysis of the data can be facilitated by formatting the data so that the individual HSQC spectra are rapidly viewed and compared to the HSQC spectrum of the control sample containing only the vehicle for the  
10 added compound (DMSO), but no added compound. Detailed descriptions of means of generating such two-dimensional  $^{15}\text{N}/^1\text{H}$  correlation spectra are set forth hereinafter in the Examples.

A representative two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of an  $^{15}\text{N}$ -labeled target molecule (polypeptide) is shown in FIG. 1 (the DNA-binding domain of the E2  
15 protein).

Following acquisition of the first spectrum, the labeled target molecule is exposed to one or more test compounds. Where more than one test compound is to be tested simultaneously, it is preferred to use a database of compounds such as a plurality of small molecules. Such molecules are typically dissolved in perdeuterated dimethylsulfoxide. The  
20 compounds in the database can be purchased from vendors or created according to desired needs.

Individual compounds can be selected *inter alia* on the basis of size (molecular weight = 100-300) and molecular diversity. Compounds in the collection can have different shapes (e.g., flat aromatic rings(s), puckered aliphatic rings(s), straight and branched chain aliphatics  
25 with single, double, or triple bonds) and diverse functional groups (e.g., carboxylic acids, esters, ethers, amines, aldehydes, ketones, and various heterocyclic rings) for maximizing the possibility of discovering compounds that interact with widely diverse binding sites.

The NMR screening process utilizes ligand concentrations ranging from about 0.1 to about 10.0 mM. At these concentrations, compounds which are acidic or basic can  
30 significantly change the pH of buffered protein solutions. Chemical shifts are sensitive to pH changes as well as direct binding interactions, and "false positive" chemical shift changes, which are not the result of ligand binding but of changes in pH, can therefore be observed. It is thus necessary to ensure that the pH of the buffered solution does not change upon addition of the ligand. One means of controlling pH is set forth below.

35 Compounds are stored at 263°K as 1.0 and 0.1 M stock solutions in dimethylsulfoxide (DMSO). This is necessary because of the limited solubility of the ligands in aqueous solution. It is not possible to directly adjust the pH of the DMSO solution. In



- 10 -

addition, HCl and NaOH form insoluble salts in DMSO, so alternative acids and bases must be used. The following approach has been found to result in stable pH.

The 1.0 M stock solutions in DMSO are diluted 1:10 in 50 mM phosphate, pH 7.0. The pH of that diluted aliquot solution is measured. If the pH of the aliquot is unchanged (i.e., remains at 7.0), a working solution is made by diluting the DMSO stock solution 1:10 to make a 0.1 M solution and that solution is stored.

If the pH of the diluted aliquot is less than 7.0, ethanolamine is added to the 1.0 M stock DMSO solution, that stock solution is then diluted 1:10 with phosphate buffer to make another aliquot, and the pH of the aliquot rechecked.

If the pH of the diluted aliquot is greater than 7.0, acetic acid is added to the 1.0 M stock DMSO solution, that stock solution is then diluted 1:10 with phosphate buffer to make another aliquot, and the pH of the aliquot rechecked.

Ethanolamine and acetic acid are soluble in DMSO, and the proper equivalents are added to ensure that upon transfer to aqueous buffer, the pH is unchanged. Adjusting the pH is an interactive process, repeated until the desired result is obtained.

Note that this procedure is performed on 1:10 dilutions of 1.0 M stock solutions (100 mM ligand) to ensure that no pH changes are observed at the lower concentrations used in the experiments (0.1 to 10 mM) or in different/weaker buffer systems.

Following exposure of the  $^{15}\text{N}$ -labeled target molecule to one or more test compounds, a second two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum is generated. That second spectrum is generated in the same manner as set forth above. The first and second spectra are then compared to determine whether there are any differences between the two spectra. Differences in the two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectra that indicate the presence of a ligand correspond to  $^{15}\text{N}$ -labeled sites in the target molecule. Those differences are determined using standard procedures well known in the art.

By way of example, FIGs. 2, 3, 4, 5 and 6 show comparisons of correlation spectra before and after exposure of various target molecules to various test compounds. A detailed description of how these studies were performed can be found hereinafter in Examples 2 and 3.

Particular signals in a two-dimensional  $^{15}\text{N}/^1\text{H}$  correlation spectrum correspond to specific nitrogen and proton atoms in the target molecule (e.g., particular amides of the amino acid residues in the protein). By way of example, it can be seen from FIG. 2 that chemical shifts in a two-dimensional  $^{15}\text{N}/^1\text{H}$  correlation of the DNA-binding domain of E2 exposed to a test compound occurred at residue positions 15 (I15), 21 (Y21), 22 (R22) and 23 (L23).

It can be seen from FIG. 2 that the binding of the ligand involved the isoleucine (Ile) residue at position 15, the tyrosine (Tyr) residue at position 21, the arginine (Arg) residue at position 22 and the leucine (Leu) residue at position 23. Thus, a process of the present



invention can also be used to identify the specific binding site between a ligand and target molecule.

The region of the protein that is responsible for binding to the individual compounds is identified from the particular amide signals that change upon the addition of the compounds. These signals are assigned to the individual amide groups of the protein by standard procedures using a variety of well-established heteronuclear multi-dimensional NMR experiments.

To discover molecules that bind more tightly to the protein, molecules are selected for testing based on the structure/activity relationships from the initial screen and/or structural information on the initial leads when bound to the protein. By way of example, the initial screening may result in the identification of ligands, all of which contain an aromatic ring. The second round of screening would then use other aromatic molecules as the test compounds.

As set forth hereinafter in Example 2, an initial screening assay for binding to the catalytic domain of stromelysin identified two biaryl compounds as ligands. The second round of screening thus used a series of biaryl derivatives as the test compounds.

The second set of test compounds are initially screened at a concentration of 1 mM, and binding constants are measured for those that show affinity. Best leads that bind to the protein are then compared to the results obtained in a functional assay. Those compounds that are suitable leads are chemically modified to produce analogs with the goal of discovering a new pharmaceutical agent.

The present method also provides a process for determining the dissociation constant between a target molecule and a ligand that binds to that target molecule. That process comprises the steps of: a) generating a first two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum of a  $^{15}\text{N}$ -labeled target molecule; b) titrating the labeled target molecule with various concentrations of a ligand; c) generating a two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum at each concentration of ligand from step (b); d) comparing each spectrum from step (c) both to the first spectrum from step (a) and to all other spectra from step (c) to quantify differences in those spectra as a function of changes in ligand concentration; and e) calculating the dissociation constant ( $K_D$ ) between the target molecule and the ligand from those differences.

Because of their importance in medicinal chemistry, a preferred target molecule for use in such a process is a polypeptide. In one preferred embodiment, a process of determining the dissociation constant of a ligand can be performed in the presence of a second ligand. In accordance with this embodiment, the  $^{15}\text{N}$ -labeled target molecule is bound to that second ligand before exposing that target to the test compounds.

Binding or dissociation constants are measured by following the  $^{15}\text{N}/^1\text{H}$  chemical shifts of the protein as a function of ligand concentration. A known concentration ( $[P]_0$ ) of



- 12 -

the target molecule is mixed with a known concentration ( $[L]_0$ ) of a previously identified ligand and the two-dimensional  $^{15}\text{N}/^1\text{H}$  correlation spectrum was acquired. From this spectrum, observed chemical shift values ( $\delta_{\text{obs}}$ ) are obtained. The process is repeated for varying concentrations of the ligand to the point of saturation of the target molecule, when possible, in which case the limiting chemical shift value for saturation ( $\delta_{\text{sat}}$ ) is measured.

In those situations where saturation of the target molecule is achieved, the dissociation constant for the binding of a particular ligand to the target molecule is calculated using the formula:

$$K_D = \frac{([P]_0 - x)([L]_0 - x)}{x}$$

where  $[P]_0$  is the total molar concentration of target molecule;  $[L]_0$  is the total molar concentration of ligand; and  $x$  is the molar concentration of the bound species. The value of  $x$  is determined from the equation:

$$x = \frac{\delta_{\text{obs}} - \delta_{\text{free}}}{\Delta}$$

where  $\delta_{\text{free}}$  is the chemical shift of the free species;  $\delta_{\text{obs}}$  is the observed chemical shift; and  $\Delta$  is the difference between the limiting chemical shift value for saturation ( $\delta_{\text{sat}}$ ) and the chemical shift value of the target molecule free of ligand ( $\delta_{\text{free}}$ ).

The dissociation constant is then determined by varying its value until a best fit to the observed data is obtained using standard curve-fitting statistical methods. In the case where  $\delta_{\text{sat}}$  is not directly known, both  $K_D$  and  $\delta_{\text{sat}}$  are varied and subjected to the same curve-fitting procedure.

The use of the process described above to determine the dissociation or binding affinity of various ligands to various target molecules is set forth hereinafter in Examples 2 and 3.

Preferred target molecules, means for generating spectra, and means for comparing spectra are the same as set forth above.

In its principal aspect, the present invention provides a process of designing new ligands that bind to a specific target molecule by linking together two or more molecules that bind to the target molecule.

The initial step in the design process is the identification of two or more ligands that bind to the specific target molecule. The identification of such ligands is done using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectroscopy as set forth above.



- 13 -

Once two or more ligands are identified as binding to the target molecule at different sites, a complex between the target molecule and ligands is formed. Where there are two ligands, that complex is a ternary complex. Quaternary and other complexes are formed where there are three or more ligands.

5       Complexes are formed by mixing the target molecule simultaneously or sequentially with the various ligands under circumstances that allow those ligands to bind the target. Means for determining those conditions are well known in the art.

Once that complex is formed, its three-dimensional structure is determined. Any means of determining three-dimensional structure can be used. Such methods are well  
10       known in the art. Exemplary and preferred methods are NMR and X-ray crystallography. The use of three-dimensional double- and triple resonance NMR to determine the three-dimensional structure of two ligands bound to the catalytic domain of stromelysin is set forth in detail hereinafter in Example 4.

An analysis of the three-dimensional structure reveals the spatial orientation of the  
15       ligands relative to each other as well as to the conformation of the target molecule. First, the spatial orientation of each ligand to the target molecule allows for identification of those portions of the ligand directly involved in binding (i.e., those portions interacting with the target binding site) and those portions of each ligand that project away from the binding site and which portions can be used in subsequent linking procedures.

20       Second, the spatial orientation data is used to map the positions of each ligand relative to each other. In other words, discrete distances between the spatially oriented ligands can be calculated.

Third, the spatial orientation data also defines the three-dimensional relationships amongst the ligands and the target. Thus, in addition to calculating the absolute distances  
25       between ligands, the angular orientations of those ligands can also be determined.

Knowledge of the spatial orientations of the ligands and target is then used to select linkers to link two or more ligands together into a single entity that contains all of the ligands. The design of the linkers is based on the distances and angular orientation needed to maintain each of the ligand portions of the single entity in proper orientation to the target.

30       The three-dimensional conformation of suitable linkers is well known or readily ascertainable by one of ordinary skill in the art. While it is theoretically possible to link two or more ligands together over any range of distance and three-dimensional projection, in practice certain limitations of distance and projection are preferred. In a preferred embodiment, ligands are separated by a distance of less than about 15 Angstroms (Å), more  
35       preferably less than about 10 Å and, even more preferably less than about 5 Å.

Once a suitable linker group is identified, the ligands are linked with that linker. Means for linking ligands are well known in the art and depend upon the chemical structure of



- 14 -

the ligand and the linking group itself. Ligands are linked to one another using those portions of the ligand not directly involved in binding to the target molecule.

A detailed description of the design of a drug that inhibits the proteolytic activity of stromelysin, which drug was designed using a process of the present invention is set forth hereinafter in Example 4.

The following Examples illustrate preferred embodiments of the present invention and are not limiting of the specification and claims in any way.

#### Example 1

##### Preparation Of Uniformly <sup>15</sup>N-Labeled Target Molecules

##### A. Stromelysin

Human stromelysin is a 447-amino acid protein believed to be involved in proteolytic degradation of cartilage. Cartilage proteolysis is believed to result in degradative loss of joint cartilage and the resulting impairment of joint function observed in both osteoarthritis and rheumatoid arthritis. The protein possesses a series of domains including N-terminal latent and propeptide domains, a C-terminal domain homologous with homopexin, and an internal catalytic domain.

Studies have shown that removal of the N-terminal prosequence of approximately eighty amino acids occurs to convert the proenzyme to the 45 kDa mature enzyme. Furthermore, studies have shown that the C-terminal homopexin homologous domain is not required for proper folding of the catalytic domain or for interaction with an inhibitor. (See, e.g., A. I. Marcy, Biochemistry, 30: 6476-6483 (1991). Thus, the 81-256 amino acid residue internal segment of stromelysin was selected as the protein fragment for use in identifying compounds which bind to and have the potential as acting as inhibitors of stromelysin.

To employ the method of the present invention, it was necessary to prepare the 81-256 fragment (SEQ ID NO:1) of stromelysin in which the peptide backbone was isotopically enriched with and <sup>15</sup>N. This was done by inserting a plasmid which coded for the production of the protein fragment into an *E. coli* strain and growing the genetically-modified bacterial strain in a limiting culture medium enriched with <sup>15</sup>NH<sub>4</sub>Cl and <sup>13</sup>C-glucose.

The isotopically enriched protein fragment was isolated from the culture medium, purified, and subsequently used as the basis for evaluating the binding of test compounds. The procedures for these processes are described below.

Human skin fibroblasts (ATCC No. CRL 1507) were grown and induced using the procedure described by Clark et al., Archiv. Biochem. and Biophys., 241: 36-45 (1985). Total RNA was isolated from 1 g of cells using a Promega RNAgents® Total RNA Isolation System Kit (Cat.# Z5110, Promega Corp., 2800 Woods Hollow Road, Madison, WI 53711-5399) following the manufacturer's instructions. A 1 µg portion of the RNA was heat-



- 15 -

denatured at 80°C for five minutes and then subjected to reverse transcriptase PCR using a GeneAmp® RNA PCR kit (Cat.# N808-0017, Applied Biosystems/Perkin-Elmer, 761 Main Avenue, Norwalk, CT 06859-0156) following the manufacturer's instructions.

Nested PCR was performed using first primers (A) GAAATGAAGAGTC TTCAA  
5 (SEQ ID NO:3) and (B) GCGTCCCAGGTTCTGGAG (SEQ ID NO:4) and thirty-five cycles of 94°C, two minutes; 45°C, two minutes; and 72°C three minutes. This was followed by reamplification with internal primers (C) ATACCATGGCCTATCCAT TGGATGGAGC (SEQ ID NO:5) and (D) ATAGGATCCTTAGGTCTCAGGGGA GTCAGG (SEQ ID NO:6) using thirty cycles under the same conditions described immediately above to generate a DNA  
10 coding for amino acid residues 1-256 of human stromelysin.

The PCR fragment was then cloned into PCR cloning vector pT7Blue(R) (Novagen, Inc., 597 Science Drive, Madison, WI 53711) according to the manufacturer's instructions. The resulting plasmid was cut with NcoI and BamHI and the stromelysin fragment was subcloned into the Novagen expression vector pET3d (Novagen, Inc., 597 Science Drive,  
15 Madison, WI 53711), again using the manufacturer's instructions.

A mature stromelysin expression construct coding for amino acid residues 81-256 plus an initiating methionine was generated from the 1-256 expression construct by PCR amplification. The resulting PCR fragment was first cloned into the Novagen pT7Blue(R) vector and then subcloned into the Novagen pET3d vector, using the manufacturer's  
20 instructions in the manner described above, to produce plasmid (pETST-83-256). This final plasmid is identical to that described by Qi-Zhuang et al., Biochemistry, 31: 11231-11235 (1992) with the exception that the present codes for a peptide sequence beginning two amino acids earlier, at position 81 in the sequence of human stromelysin.

Plasmid pETST-83-256 was transformed into *E. coli* strain BL21(DE3)/pLysS  
25 (Novagen, Inc., 597 Science Drive, Madison, WI 53711) in accordance with the manufacturer's instructions to generate an expression strain, BL21(DE3)/pLysS/pETST-255-1.

A preculture medium was prepared by dissolving 1.698 g of Na<sub>2</sub>HP<sub>4</sub>•7H<sub>2</sub>O, 0.45 g of KH<sub>2</sub>PO<sub>4</sub>, 0.075 g NaCl, 0.150 g <sup>15</sup>NH<sub>4</sub>Cl, 0.300 13C-glucose, 300 µL of 1M aqueous  
30 MgSO<sub>4</sub> solution and 15 µL of aqueous CaCl<sub>2</sub> solution in 150 mL of deionized water.

The resulting solution of preculture medium was sterilized and transferred to a sterile 500 mL baffle flask. Immediately prior to inoculation of the preculture medium with the bacterial strain, 150 µL of a solution containing 34 mg/mL of chloramphenicol in 100% ethanol and 1.5 mL of a solution containing 20 mg/mL of ampicillin were added to the flask  
35 contents.

The flask contents were then inoculated with 1 mL of glycerol stock of genetically-modified *E. Coli*, strain BL21(DE3)/pLysS/pETST-255-1. The flask contents were shaken (225 rpm) at 37°C until an optical density of 0.65 was observed.



- 16 -

A fermentation nutrient medium was prepared by dissolving 113.28 g of  $\text{Na}_2\text{HP}_4 \cdot 7\text{H}_2\text{O}$ , 30 g of  $\text{KH}_2\text{PO}_4$ , 5 g  $\text{NaCl}$  and 10 mL of 1% DF-60 antifoam agent in 9604 mL of deionized water. This solution was placed in a New Brunswick Scientific Micros Fermenter (Edison, NJ) and sterilized at 121°C for 40 minutes.

5 Immediately prior to inoculation of the fermentation medium, the following pre-sterilized components were added to the fermentation vessel contents: 100 mL of a 10% aqueous solution of  $^{15}\text{NH}_4\text{Cl}$ , 100 mL of a 10% aqueous solution of  $^{13}\text{C}$ -glucose, 20 mL of an aqueous 1M solution of  $\text{MgSO}_4$ , 1 mL of an aqueous 1M  $\text{CaCl}_2$  solution, 5 mL of an aqueous solution of thiamin hydrochloride (10 mg/mL), 10 mL of a solution containing 34  
10 mg/mL of chloramphenicol in 100% ethanol and 1.9 g of ampicillin dissolved in the chloramphenicol solution. The pH of the resulting solution was adjusted to pH 7.00 by the addition of an aqueous solution of 4N  $\text{H}_2\text{SO}_4$ .

The preculture of *E. Coli*, strain BL21(DE3)/pLysS/pETST-255-1, from the shake-flask scale procedure described above was added to the fermentor contents and cell growth  
15 was allowed to proceed until an optical density of 0.48 was achieved. During this process, the fermenter contents were automatically maintained at pH 7.0 by the addition of 4N  $\text{H}_2\text{SO}_4$  or 4N KOH as needed. The dissolved oxygen content of the fermenter contents was maintained above 55% air saturation through a cascaded loop which increased agitation speed when the dissolved oxygen content dropped below 55%. Air was fed to the fermenter  
20 contents at 7 standard liters per minute (SLPM) and the culture temperature was maintained at 37°C throughout the process.

The cells were harvested by centrifugation at 17,000 x g for 10 minutes at 4°C and the resulting cell pellets were collected and stored at -85°C. The wet cell yield was 3.5 g/L. Analysis of the soluble and insoluble fractions of cell lysates by sodium dodecyl sulfate  
25 polyacrylamide gel electrophoresis (SDS-PAGE) revealed that approximately 50% of the  $^{15}\text{N}$ -stromelysin was found in the soluble phase.

The isotopically-labeled stromelysin fragment prepared as described above was purified employing a modification of the technique described by Ye, *et al.*, Biochemistry, 31: 11231-11235 (1992).

30 The harvested cells were suspended in 20 mM Tris-HCl buffer (pH 8.0) sodium azide solution containing 1 mM  $\text{MgCl}_2$ , 0.5 mM  $\text{ZnCl}_2$ , 25 units/mL of Benzonase® enzyme, and an inhibitor mixture made up of 4-(2-aminoethyl)-benzenesulfonyl fluoride ("AEBSF"), Leupeptin®, Aprotinin®, and Pepstatin® (all at concentrations of 1 µg/mL. AEBSF, Leupeptin®, Aprotinin®, and Pepstatin® are available from American International  
35 Chemical, 17 Strathmore Road, Natick, MA 01760.)

The resulting mixture was gently stirred for one hour and then cooled to 4°C. The cells were then sonically disrupted using a 50% duty cycle. The resulting lysate was



- 17 -

centrifuged at 14,000 rpm for 30 minutes and the pellet of insoluble fraction frozen at -80°C for subsequent processing (see below).

Solid ammonium sulfate was added to the supernatant to the point of 20% of saturation and the resulting solution loaded onto a 700 mL phenyl sepharose fast flow ("Q-Sepharose FF") column (Pharmacia Biotech., 800 Centennial Ave., P. O. Box 1327, Piscataway, NJ 08855). Prior to loading, the sepharose column was equilibrated with 50 mM Tris-HCl buffer (pH 7.6 at 4°C), 5 mM CaCl<sub>2</sub>, and 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The loaded column was eluted with a linear gradient of decreasing concentrations of aqueous (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (from 1 down to 0 M) and increasing concentrations of aqueous CaCl<sub>2</sub> (from 5 to 20 mM) in Tris-HCl buffer at pH 7.6.

The active fractions of eluate were collected and concentrated in an Amicon stirred cell (Amicon, Inc., 72 Cherry Hill Drive, Beverly, MA 01915). The concentrated sample was dialyzed overnight in the starting buffer used with the Q-Sepharose FF column, 50 mM Tris-HCl (pH 8.2 at 4°C) with 10 mM CaCl<sub>2</sub>.

The dialyzed sample was then loaded on the Q-Sepharose FF column and eluted with a linear gradient comprising the starting buffer and 200 mM NaCl. The purified soluble fraction of the isotopically-labeled stromelysin fragment was concentrated and stored at 4°C.

The pellet was solubilized in 8M guanidine-HCl. The solution was centrifuged for 20 minutes at 20,000 rpm and the supernatant was added dropwise to a folding buffer comprising 50 mM Tris-HCl (pH 7.6), 10 mM CaCl<sub>2</sub> 0.5 mM ZnCl<sub>2</sub> and the inhibitor cocktail of AEBSF, Leupeptin®, Aprotinin®, and Pepstatin® (all at concentrations of 1 µg/mL). The volume of folding buffer was ten times that of the supernatant. The mixture of supernatant and folding buffer was centrifuged at 20,000 rpm for 30 minutes.

The supernatant from this centrifugation was stored at 4°C and the pellet was subjected twice to the steps described above of solubilization in guanidine-HCl, refolding in buffer, and centrifugation. The final supernatants from each of the three centrifugations were combined and solid ammonium sulfate was added to the point of 20% saturation. The resulting solution thus derived from the insoluble fraction was subjected to purification on phenyl Sepharose and Q-Sepharose as described above for the soluble fraction.

The purified soluble and insoluble fractions were combined to produce about 1.8 mg of purified isotopically-labeled stromelysin 81-256 fragment per gram of original cell paste.

#### B. Human papillomavirus (HPV) E2 Inhibitors

The papillomaviruses are a family of small DNA viruses that cause genital warts and cervical carcinomas. The E2 protein of HPV regulates viral transcription and is required for viral replication. Thus, molecules that block the binding of E2 to DNA may be useful therapeutic agents against HPV. The protein rather than the DNA was chosen as a target,



- 18 -

because it is expected that agents with greater selectivity would be found that bind to the protein rather than the DNA.

The DNA-binding domain of human papillomavirus E2 was cloned from the full length DNA that codes for E2 using PCR and overexpressed in bacteria using the T7 expression system. Uniformly  $^{15}\text{N}$ -labeled protein was isolated from bacteria grown on a minimal medium containing  $^{15}\text{N}$ -labeled protein was isolated from bacteria grown on a minimal medium containing  $^{15}\text{N}$ -labeled ammonium chloride. The protein was purified from the bacterial cell lysate using an S-sepharose FastFlow column pre-equilibrated with buffer (50 mM Tris, 100 mM NaCl, 1 mM EDTA, pH = 8.3).

The protein was eluted with a linear gradient of 100-500 mM NaCl in buffer, pooled, and applied to a Mono-S column at a pH = 7.0. The protein was eluted with a salt gradient (100-500 mM), concentrated to 0.3 mM, and exchanged into a TRIS (50 mM, pH = 7.0) buffered  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (9/1) solution containing sodium azide (0.5%).

#### 15 C. RAF

Uniformly  $^{15}\text{N}$ -labeled Ras-binding domain of the RAF protein was prepared as described in Emerson et al., Biochemistry, 34 (21): 6911-6918 (1995).

#### 20 D. FKBP

Uniformly  $^{15}\text{N}$ -labeled recombinant human FK binding protein (FKBP) was prepared as described in Logan, *et al.*, J. Mol. Biol., 236: 637-648 (1994).

### Example 2

#### 25 Screening Compounds Using Two-Dimensional $^{15}\text{N}/^1\text{H}$ NMR Correlation Spectral Analysis

The catalytic domain of stromelysin was prepared in accordance with the procedures of Example 1. The protein solutions used in the screening assay contained the uniformly  $^{15}\text{N}$ -labeled catalytic domain of stromelysin (0.3 mM), acetohydroxamic acid (500 mM),  $\text{CaCl}_2$  (20 mM), and sodium azide (0.5%) in a  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (9/1) TRIS buffered solution (50 mM, pH=7.0).

Two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR spectra were generated at 29°C on a Bruker AMX500 NMR spectrometer equipped with a triple resonance probe and Bruker sample changer. The  $^{15}\text{N}/^1\text{H}$  HSQC spectra were acquired as 80 x 1024 complex points using sweep widths of 2000 Hz ( $^{15}\text{N}$ ,  $t_1$ ) and 8333 Hz ( $^1\text{H}$ ,  $t_2$ ). A delay of 1 second between scans and 8 scans per free induction decay(fid) were employed in the data collection. All NMR spectra were processed and analyzed on Silicon Graphics computers using in-house-written software.



- 19 -

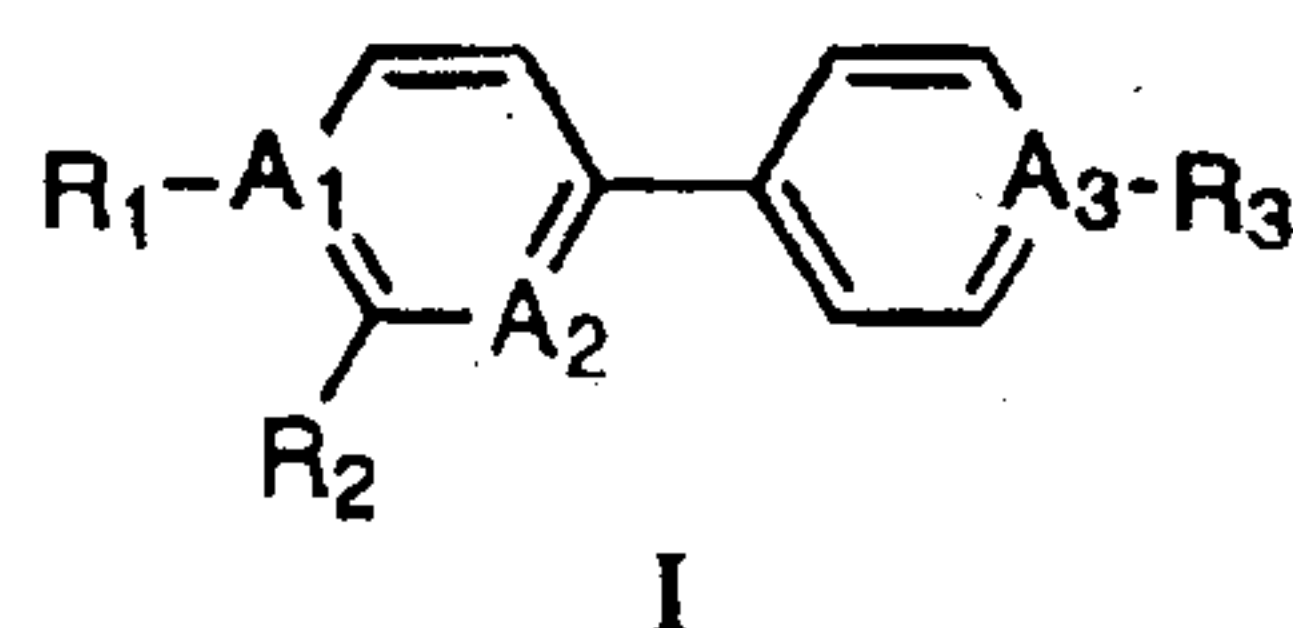
A first two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum was acquired for the  $^{15}\text{N}$ -labeled stromelysin target molecule as described above. The stromelysin target was then exposed to a database of test compounds. Stock solutions of the compounds were made at 100 mM and 1 M. In addition, a combination library was prepared that contained 8-10 compounds per sample at a concentration of 100 mM for each compound.

The pH of the 1 M stock solution was adjusted with acetic acid and ethanolamine so that no pH change was observed upon a 1/10 dilution with a 100 mM phosphate buffered solution (pH = 7.0). It is important to adjust the pH, because small changes in pH can alter the chemical shifts of the biomolecules and complicate the interpretation of the NMR data.

The compounds in the database were selected on the basis of size (molecular weight = 100-300) and molecular diversity. The molecules in the collection had different shapes (*e.g.*, flat aromatic rings(s), puckered aliphatic rings(s), straight and branched chain aliphatics with single, double, or triple bonds) and diverse functional groups (*e.g.*, carboxylic acids, esters, ethers, amines, aldehydes, ketones, and various heterocyclic rings) for maximizing the possibility of discovering compound that interact with widely diverse binding sites.

The NMR samples were prepared by adding 4  $\mu\text{l}$  of the DMSO stock solution of the compound mixtures that contained each compound at a concentration of 100 mM to 0.4 ml  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (9/1) buffered solution of the uniformly  $^{15}\text{N}$ -labeled protein. The final concentration of each of the compounds in the NMR sample was about 1 mM.

In an initial screen, two compounds were found that bind to the catalytic domain of stromelysin. Both of these compounds contain a biaryl moiety. Based on these initial hits, structurally similar compounds were tested against stromelysin. The structure of those biaryl compounds is represented by the structure I, below. (See Table 1 for definitions of  $\text{R}_1$ - $\text{R}_3$  and  $\text{A}_1$ - $\text{A}_3$ ).



In the second round of screening, binding was assayed both in the absence and in the presence of saturating amounts of acetohydroxamic acid (500 mM).

Many of the biaryl compounds were found to bind the catalytic domain of stromelysin. FIG. 4 shows a representative two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectrum before and after exposure of stromelysin to a biaryl test compound. It can be seen from FIG. 4 that the compound caused chemical shifts of  $^{15}\text{N}$ -sites such as those designated W124, T187, A199 and G204.

These sites correspond to a tryptophan (Trp) residue at position 124, a threonine (Thr) at position 187, an alanine (Ala) at position 199, and a glycine (Gly) at position 204 of SEQ ID NO. 1. FIG. 9 shows the correlation between the NMR binding data and a view of the



- 20 -

NMR-derived three-dimensional structure of the catalytic domain of stromelysin. The ability to locate the specific binding site of a particular ligand is an advantage of the present invention.

- Some compounds only bound to stromelysin in the presence of hydroxamic acid.
- 5 Thus, the binding affinity of some compounds was enhanced in the presence of the hydroxamic acid (i. e. cooperative). These results exemplify another important capability of the present screening assay: the ability to identify compounds that bind to the protein in the presence of other molecules.

Various biaryl compounds of structure I were tested for binding to stromelysin at  
10 differing concentrations. The  $^{15}\text{N}/^1\text{H}$  spectra generated at each concentration were evaluated to quantify differences in the spectra as a function of compound concentration. A binding or dissociation constant ( $K_D$ ) was calculated, using standard procedures well known in the art, from those differences. The results of this study are shown in Table 1. The values for R1-  
R3 and A1-A3 in Table 1 refer to the corresponding positions in the structure I, above.

15

Table 1

| Compound No. | R1                 | R2              | R3 | A1 | A2 | A3 | KD(mM) |
|--------------|--------------------|-----------------|----|----|----|----|--------|
| 1            | H                  | OH              | H  | C  | C  | C  | 1.1    |
| 2            | CH <sub>2</sub> OH | H               | H  | C  | C  | C  | 3.2    |
| 3            | Br                 | H               | OH | C  | C  | C  | 1.3    |
| 4            | H                  | H               | H  | N  | N  | C  | 1.6    |
| 5            | CHO                | H               | H  | C  | C  | C  | 1.7    |
| 6            | OCH <sub>3</sub>   | NH <sub>2</sub> | H  | C  | C  | C  | 0.4    |
| 7            | H                  | H               | H  | N  | C  | C  | 0.2    |



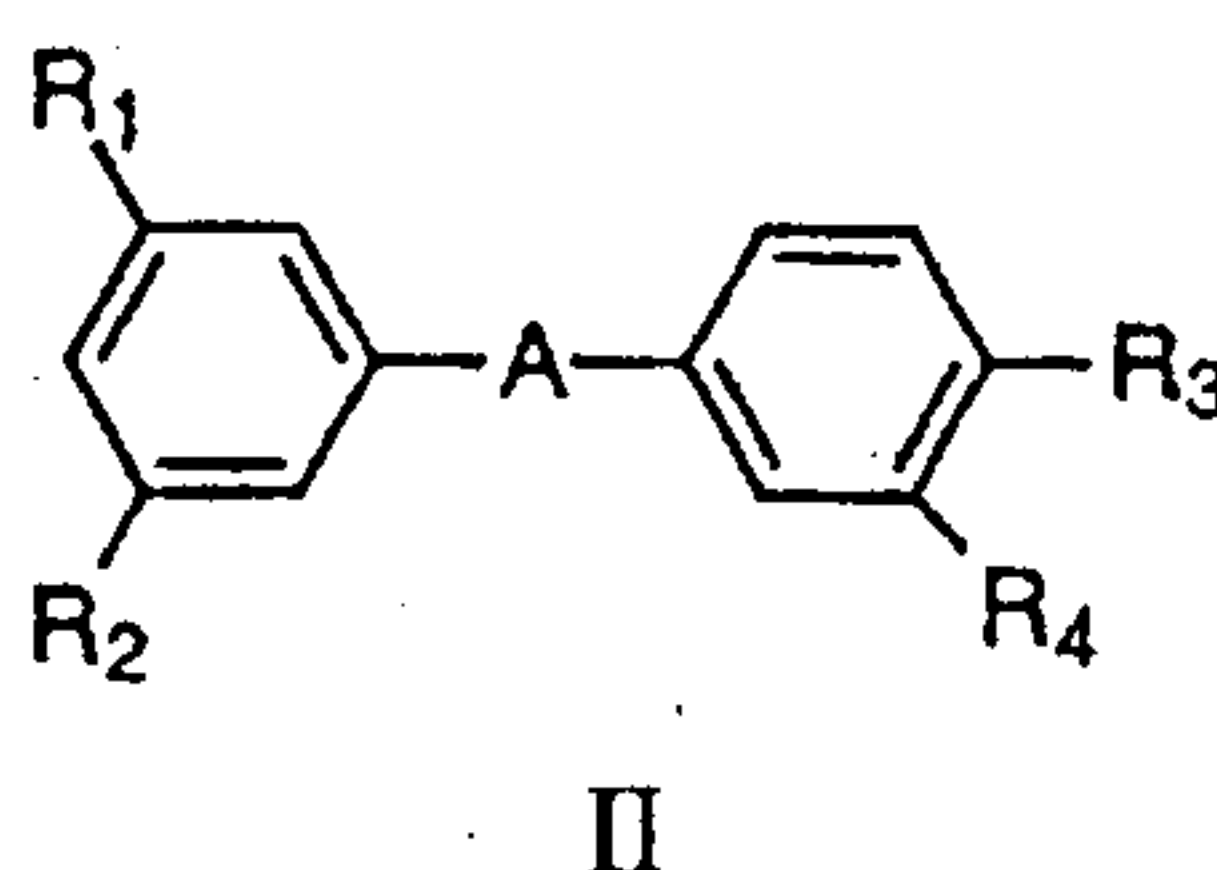
- 21 -  
Table 1 (Continued)

| Compound No. | R1                 | R2 | R3 | A1 | A2 | A3 | KD(mM) |
|--------------|--------------------|----|----|----|----|----|--------|
| 8            | OCOCH <sub>3</sub> | H  | H  | C  | C  | C  | 0.3    |
| 9            | OH                 | H  | OH | C  | C  | C  | 0.01   |
| 10           | H                  | H  | H  | N  | C  | N  | 0.4    |
| 11           | OH                 | H  | H  | C  | C  | C  | 0.3    |
| 12           | OH                 | H  | CN | C  | C  | C  | 0.01   |

The data in Table 1 show the utility of a process of the present invention in  
5 determining dissociation or binding constants between a ligand and a target molecule.

Another advantage of an NMR screening assay of the present invention is the  
ability to correlate observed chemical shifts from the two-dimensional <sup>15</sup>N/<sup>1</sup>H NMR  
correlation spectra with other spectra or projections of target molecule configuration. The  
results of a representative such correlation are shown in FIG. 9, which depicts regions within  
10 the polypeptide at which binding with the substrate molecule is most likely occurring. In this  
Figure, the apparent binding regions in stromelysin are shown for Compound 1 (from Table  
1).

Compounds from the database were screened in a similar manner for binding to the  
DNA-binding domain of the E2 protein. Those compounds had the structure II below, where  
15 R<sub>1</sub>-R<sub>4</sub> and A are defined in Table 2.



NMR experiments were performed at 29°C on a Bruker AMX500 NMR spectrometer  
equipped with a triple resonance probe and Bruker sample changer. The <sup>15</sup>N-/<sup>1</sup>H HSQC  
spectra were acquired as 80 x 1024 complex points using sweep widths of 2000 Hz (<sup>15</sup>N, t<sub>1</sub>)  
and 8333 Hz (<sup>1</sup>H, t<sub>2</sub>). A delay of 1 second between scans and 4 scans per free induction  
20 decay were employed in the data collection. All NMR spectra were processed and analyzed  
on Silicon Graphics computers.



- 22 -

FIGs. 2 and 3 show representative two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation spectra before and after exposure of the DNA-binding domain of E2 to a first and second test compound, respectively.

It can be seen from FIG. 2 that the first test compound caused chemical shifts of  $^{15}\text{N}$ -sites such as those designated I15, Y21, R22 and L23. Those sites correspond to an  
5 isoleucine (Ile) residue at position 15, a tyrosine residue (Tyr) at position 21, an arginine (Arg) residue at position 22 and a leucine (Leu) residue at position 23 of SEQ ID NO. 6.

It can be seen from FIG. 3 that the second test compound caused chemical shifts in the particular  $^{15}\text{N}$ -sites designated I6, G11, H38, and T52. Those sites correspond to an  
10 isoleucine (Ile) residue at position 6, a glycine (Gly) residue at position 11, a histidine (His) residue at position 38 and a threonine (Thr) at position 52 of SEQ ID NO. 6.

FIGs. 7 and 8 show the correlation between those NMR binding data and a view of the NMR-derived three-dimensional structure of the DNA-binding domain of E2.

Several structurally similar compounds caused chemical shift changes of the protein  
15 signals when screened at a concentration of 1 mM. Two distinct sets of amide resonances were found to change upon the addition of the compounds: one set of signals corresponding to amides located in the  $\beta$ -barrel formed between the two monomers and a second set corresponding to amides located near the DNA-binding site.

For example, compounds containing two phenyl rings with a carboxylic acid attached  
20 to the carbon linking the two rings only caused chemical shift changes to the amides in the DNA-binding site. In contrast, benzophenones and phenoxyphenyl-containing compounds only bound to the  $\beta$ -barrel. Other compounds caused chemical shift changes of both sets of signals but shifted the signals in each set by different amounts, suggesting the presence of two distinct binding sites.

25 By monitoring the chemical shift changes as a function of ligand concentration, binding constants for the two binding sites were also measured. The results of those studies are summarized below in Table 2.



- 23 -  
Table 2

| Comp. No. | A            | R <sub>1</sub> | R <sub>2</sub> | R <sub>3</sub>                                                               | R <sub>4</sub>     | DNA K <sub>D</sub> (mM) | $\beta$ -barrel K <sub>D</sub> (mM) | Filter binding assay |
|-----------|--------------|----------------|----------------|------------------------------------------------------------------------------|--------------------|-------------------------|-------------------------------------|----------------------|
| 13        | CO           | H              | H              | H                                                                            | OH                 | >50                     | 0.6                                 | -                    |
| 14        | O            | H              | H              | H                                                                            | CH <sub>2</sub> OH | >50                     | 2.0                                 | -                    |
| 15        | <sup>a</sup> | H              | H              | COO                                                                          | H                  | 2.0                     | >50                                 | +                    |
| 16        | <sup>a</sup> | Cl             | Cl             | COO                                                                          | H                  | 0.1                     | >50                                 | +                    |
| 17        | <sup>a</sup> | H              | H              | CH <sub>2</sub> COO                                                          | H                  | 4.2                     | 4.9                                 | +                    |
| 18        | <sup>a</sup> | H              | H              | CH=CHCOO                                                                     | H                  | 1.2                     | 6.2                                 | +                    |
| 19        | O            | H              | H              | CH <sub>2</sub> CH <sub>2</sub> CH(CH <sub>3</sub> )<br>-CH <sub>2</sub> COO | H                  | 0.5                     | 0.2                                 | +                    |
| 20        | O            | H              | H              | COCH <sub>2</sub> CH <sub>2</sub> COO                                        | H                  | 2.7                     | 4.8                                 | +                    |

<sup>a</sup> a dash (-) for A indicates no atom (*i.e.*, byphenyl linkage)

Uniformly <sup>15</sup>N-labeled Ras-binding domain of the RAF protein was prepared as described in Example 1 and screened using two-dimensional <sup>15</sup>N/<sup>1</sup>H NMR correlation spectral analysis in accordance with the NMR procedures described above. The results of a representative study are shown in FIG. 5, which depicts two-dimensional <sup>15</sup>N/<sup>1</sup>H NMR correlation spectra both before and after exposure to a test compound.

Uniformly <sup>15</sup>N-labeled FKBP was prepared as described in Example 1 and screened using two-dimensional <sup>15</sup>N/<sup>1</sup>H NMR correlation spectral analysis in accordance with the NMR procedures described above. The results of a representative study are shown in FIG. 6, which depicts two-dimensional <sup>15</sup>N/<sup>1</sup>H NMR correlation spectra both before and after exposure to a test compound.

15

### Example 3

#### Comparison of NMR, Enzymatic, Filter Binding and Gel Shift Screening Assays

Studies were performed to compare binding constants of ligands to various biomolecules, determined by the NMR method of the present invention, to similar results obtained from prior art methods.

20

In a first study, binding constants were determined, both by the NMR method of the present invention, and by a prior art enzymatic assay. The target molecule was the catalytic domain of stromelysin prepared in accordance with the procedures of Example 1. The NMR binding constants, K<sub>D</sub>, were derived using two-dimensional <sup>15</sup>N/<sup>1</sup>H NMR correlation



- 24 -

spectroscopy as described in Example 2. The  $K_D$  values so obtained were compared to an inhibition constant  $K_I$  as determined in an enzymatic assay.

The enzymatic assay measured the rate of cleavage of a fluorogenic substrate by following the fluorescence increase upon peptide cleavage which causes a separation between the fluorophore and quencher. Enzymatic activity was measured using a matrix of different concentrations of acetohydroxamic acid and biaryl compounds. The assay is a modification of the method described by H. Weingarten, *et al.* in Anal. Biochem., 147: 437-440 (1985) employing the fluorogenic substrate properties described by E. Matayoshi, *et al.* in Science: 247: 954-958 (1990).

Eight acetohydroxamic acid concentrations were used ranging from 0.0 to 1.0 M, and six compound concentrations were used, resulting in a total of 48 points. Individual compound concentration varied due to solubility and potency.

All NMR measurements were performed in the presence of 500 mM acetohydroxamic acid, except for the titration of acetohydroxamic acid itself. Dissociation constants were obtained from the dependence of the observed chemical shift changes upon added ligand. Inhibition constants were then obtained from the inhibition data using standard procedures.

The results of these studies are summarized below in Table 3, which shows the comparison of NMR-derived dissociation constants ( $K_D$ ) with inhibition constants measured in the enzyme assay ( $K_I$ ) using a fluorogenic substrate.

Table 3

| Compound No.         | NMR $K_D$ (mM) | Assay $K_I$ (mM) |
|----------------------|----------------|------------------|
| 4                    | 1.6            | 7.4              |
| 7                    | 0.17           | 0.32             |
| 9                    | 0.16           | 0.70             |
| 10                   | 0.40           | 1.8              |
| 12                   | 0.02           | 0.11             |
| Acetohydroxamic acid | 17.0           | 21.1             |



- 25 -

The data in Table 3 show that a NMR process of the present invention provides a rapid, efficient and accurate way of determining dissociation or binding constants of ligands to target biomolecules. Comparison of the binding constants determined by the two methods result in the same ranking of potencies of the compounds tested. That is, while the values for a given substrate as determined by the two methods are not equal, they are proportional to one another.

In a second study, the results for binding of the DNA-binding domain of E2 to its target DNA were obtained by prior art methods and compared with results obtained by the method of the present invention. The target was the DNA-binding domain of E2, prepared in accordance with the procedures of Example 1. NMR screening assays and NMR processes for determining ligand dissociation constants were performed as set forth above in Example 2.

The binding constant from the NMR process was compared to the results of a physical, filter binding assay that measured binding of DNA to the target. The high-throughput filter binding assay was performed using E2, prepared according to Example 2 above. The <sup>33</sup>P-labeled DNA construct comprised a 10,329 base pair plasmid formed by inserting the HPV-11 genome, containing three high affinity and one low affinity E2 binding sites, into the PSP-65 plasmid (Promega, Madison, WI).

The binding affinities at the different sites as determined by NMR were compared for a subset of the compounds to the inhibition of E2 binding to DNA as measured in the filter binding assay. As shown in Table 2 above, the activities determined in the filter binding assay correlated closely with the binding affinities calculated from the amides of the DNA-binding site but not to the affinities measured for the  $\beta$ -barrel site. This is consistent with the relative locations of each site.

In an alternative study, a comparison of the NMR-determined binding results was made with similar results obtained by a prior art gel-shift assay using techniques well known in the art. The gel-shift assay was performed using a GST fusion protein which contained full length E2 and a <sup>33</sup>P-labeled 62 base pair DNA fragment containing two E2 binding sites.

The method identified numerous compounds which gave positive results in the gel-shift assay. Some of these positive results, however, were believed to be due to binding to the DNA, since in these cases, no binding to the E2 protein was observed using the NMR method of this invention. These compounds were shown to indeed bind to DNA rather than to E2, as evidenced by changes in the chemical shifts of the DNA rather than the protein upon the addition of the compounds. These data show that yet another advantage of the present invention is the ability to minimize the occurrence of false positives.



- 26 -

## Example 4

Design of a potent, non-peptide inhibitor of stromelysin

Studies were performed to design new ligands that bound to the catalytic domain of stromelysin. Because stromelysin undergoes autolysis, an inhibitor was sought to block the degradation of stromelysin. That inhibitor would facilitate the screening of other potential ligands that bind to other sites on the enzyme.

The criteria used in selecting compounds in the screening for other binding sites was based primarily on the size of the ligand. The smallest ligand was sought that had enough solubility to saturate (>98% occupancy of enzyme) and inhibit the enzyme.

The cloning, expression, and purification of the catalytic domain of stromelysin was accomplished using the procedures set forth in Example 1. An initial step in the design of the new ligand was the identification of a first ligand that bound to the stromelysin target. Such identification was carried out in accordance with a two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR correlation screening process as disclosed above.

A variety of hydroxamic acids of the general formula  $\text{R}-(\text{CO})\text{NHOH}$  were screened for binding to stromelysin using the procedures set forth in Example 2. Of the compounds tested, acetohydroxamic acid [ $\text{CH}_3(\text{CO})\text{NHOH}$ ] best satisfied the selection criteria: it had a binding affinity for stromelysin of 17 mM and had good water solubility. At a concentration of 500 mM, acetohydroxamic acid inhibited the degradation of the enzyme, allowing the screening of other potential ligands.

The second step in the design process was the identification of a second ligand that bound to the target stromelysin at a site different from the binding site of acetohydroxamic acid. This was accomplished by screening compounds for their ability to bind stromelysin in the presence of saturating amounts of acetohydroxamic acid. Details of procedures and results of this second identification step are set forth above in Example 2.

The compound identified as a second ligand from these studies and used in subsequent design steps was the compound designated as Compound #4 in Table 1 (See Example 2).

The next step in the design process was to construct a ternary complex of the target stromelysin, the first ligand and the second ligand. This was accomplished by exposing the stromelysin target to the two ligands under conditions that resulted in complex formation. The three-dimensional structure of the ternary complex was then determined using NMR spectroscopy as described below.

The  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  backbone resonances of stromelysin in the ternary complex were assigned from an analysis of several 3D double- and triple-resonance NMR spectra (A. Bax, *et al.*, Acc. Chem. Res., 26: 131-138 (1993)). The  $\text{C}^\alpha$  resonances of adjacent spin systems were identified from an analysis of three-dimensional (3D) HNCA (L. Kay, *et al.*, J. Magn. Reson., 89: 496-514 (1990)) and HN(CO)CA (A. Bax, *et al.*, J. Bio. NMR, 1: 99



- 27 -

(1991)) spectra recorded with identical spectral widths of 1773 Hz (35.0 ppm), 3788 Hz (30.1 ppm), and 8333 Hz (16.67 ppm) in the  $F_1(^{15}\text{N})$ ,  $F_2(^{13}\text{C})$  and  $F_3(^1\text{H})$  dimensions, respectively.

The data matrix was 38( $t_1$ ) x 48( $t_2$ ) x 1024( $t_3$ ) complex points for the HNCA spectrum, and 32( $t_1$ ) x 40( $t_2$ ) x 1024( $t_3$ ) complex points for the HN(CO)CA spectrum. Both spectra were acquired with 16 scans per increment. A 3D CBCA(CO)NH spectrum (S. Grzesiek, *et al.*, *J. Am. Chem. Soc.*, **114**: 6261-6293 (1992)) was collected with 32( $t_1$ ,  $^{15}\text{N}$ ) x 48( $t_2$ ,  $^{13}\text{C}$ ) x 1024( $t_3$ ,  $^1\text{H}$ ) complex points and 32 scans per increment. Spectral widths were 1773 Hz (35.0 ppm), 7575.8 Hz (60.2 ppm), and 8333 Hz (16.67 ppm) in the  $^{15}\text{N}$ ,  $^{13}\text{C}$  and  $^1\text{H}$  dimensions, respectively.

For all three spectra, the  $^1\text{H}$  carrier frequency was set on the water resonance and the  $^{15}\text{N}$  carrier frequency was at 119.1 ppm. The  $^{13}\text{C}$  carrier frequency was set to 55.0 ppm in HNCA and HN(CO)CA experiments, and 46.0 ppm in the CBCA(CO)NH experiment.

The backbone assignments were confirmed from an analysis of the crosspeaks observed in an  $^{15}\text{N}$ -separated 3D NOESY-HSQC spectrum and a 3D HNHA-J spectrum. The  $^{15}\text{N}$ -separated 3D NOESY-HSQC spectrum (S. Fesik, *et al.*, *J. Magn. Reson.*, **87**: 588-593 (1988)); D. Marion, *et al.*, *J. Am. Chem. Soc.*, **111**: 1515-1517 (1989)) was collected with a mixing time of 80 ms. A total of 68( $t_1$ ,  $^{15}\text{N}$ ) x 96( $t_2$ ,  $^1\text{H}$ ) x 1024( $t_3$ ,  $^1\text{H}$ ) complex points with 16 scans per increment were collected, and the spectral widths were 1773 Hz (35.0 ppm) for the  $^{15}\text{N}$  dimension, 6666.6 Hz ( $t_2$ ,  $^1\text{H}$ , 13.3 ppm), and 8333 Hz (16.7 ppm) for the  $^1\text{H}$  dimension.

The 3D HNHA-J spectrum (G. Vuister, *et al.*, *J. Am. Chem. Soc.*, **115**: 7772-7777 (1993)), which was also used to obtain  $^3\text{J}_{\text{HNH}\alpha}$  coupling constants, was acquired with 35( $t_1$ ,  $^{15}\text{N}$ ) x 64( $t_2$ ,  $^1\text{H}$ ) x 1024( $t_3$ ,  $^1\text{H}$ ) complex points and 32 scans per increment. Spectral widths and carrier frequencies were identical to those of the  $^{15}\text{N}$ -separated NOESY-HSQC spectrum. Several of the  $\text{H}^\beta$  signals were assigned using the HNHB experiment. The sweep widths were the same as in the  $^{15}\text{N}$ -separated NOESY-HSQC spectrum that was acquired with 32( $t_1$ ,  $^{15}\text{N}$ ) x 96( $t_2$ ,  $^1\text{H}$ ) x 1024 ( $t_3$ ,  $^1\text{H}$ ) complex points.

The  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts were assigned for nearly all sidechain resonances. A 3D HCCH-TOCSY spectrum (L. Kay, *et al.*, *J. Magn. Reson.*, **101b**: 333-337 (1993)) was acquired with a mixing time of 13 ms using the DIPSI-2 sequence (S. Rucker, *et al.*, *Mol. Phys.*, **68**: 509 (1989)) for  $^{13}\text{C}$  isotropic mixing. A total of 96 ( $t_1$ ,  $^{13}\text{C}$ ) x 96( $t_2$ ,  $^1\text{H}$ ) x 1024( $t_3$ ,  $^1\text{H}$ ) complex data points were collected with 16 scans per increment using a spectral width of 10638 Hz (70.8 ppm,  $w_1$ ), 4000 Hz (6.67 ppm,  $w_2$ ), and 4844 (8.07 ppm,  $w_3$ ). Carrier positions were 40 ppm, 2.5 ppm, and at the water frequency for the  $^{13}\text{C}$ , indirectly detected  $^1\text{H}$ , and observed  $^1\text{H}$  dimensions, respectively.

Another 3D HCCH-TOCSY study was performed with the  $^{13}\text{C}$  carrier at 122.5 ppm to assign the aromatic residues. The spectra were collected with 36( $t_1$ ,  $^{13}\text{C}$ ) x 48( $t_2$ ,  $^1\text{H}$ ) x



- 28 -

1024 ( $t_3$ ,  $^1\text{H}$ ) complex points with spectral widths of 5263 Hz (35.0 ppm,  $w_1$ ), 3180 Hz (5.30 ppm,  $w_2$ ), and 10,000 (16.7 ppm,  $w_3$ ). Carrier positions were 122.5 ppm, 7.5 ppm, and at the water frequency for the  $^{13}\text{C}$ , indirectly detected  $^1\text{H}$ , and observed  $^1\text{H}$  dimensions, respectively.

5 A  $^{13}\text{C}$ -separated 3D NOESY-HMQC spectrum (S. Fesik, *et al.*, J. Magn. Reson., 87: 588-593 (1988)); D. Marion, *et al.*, J. Am. Chem. Soc., 111: 1515-1517 (1989)) was recorded using a mixing time of 75 ms. A total of 80 ( $t_1$ ,  $^{13}\text{C}$ ) x 72 ( $t_2$ ,  $^1\text{H}$ ) x 1024 ( $t_3$ ,  $^1\text{H}$ ) complex data points with 16 scans per increment were collected over spectral widths of 10638 Hz (70.49 ppm,  $w_1$ ), 6666.6 Hz (13.3 ppm,  $w_2$ ), and 8333.3 Hz (16.67 ppm,  $w_3$ ). The  
10  $^1\text{H}$  carrier frequencies were set to the water resonance, and the  $^{13}\text{C}$  carrier frequency was placed at 40.0 ppm.

Stereospecific assignments of methyl groups of the valine and leucine residues were obtained by using a biosynthetic approach (Neri *et al.*, Biochem., 28: 7510-7516 (1989)) on the basis of the  $^{13}\text{C}$ - $^{13}\text{C}$  one-bond coupling pattern observed in a high-resolution  $^1\text{H}$ ,  $^{13}\text{C}$ -  
15 HSQC spectrum (G. Bodenhausen, *et al.*, J. Chem. Phys. Lett., 69: 185-189 (1980)) of a fractionally  $^{13}\text{C}$ -labeled protein sample. The spectrum was acquired with 200( $^{13}\text{C}$ ,  $t_1$ ) x 2048( $^1\text{H}$ ,  $t_2$ ) complex points over spectral widths of 5000 Hz (39.8 ppm,  $^{13}\text{C}$ ) and 8333 Hz (16.7 ppm,  $^1\text{H}$ ). Carrier positions were 20.0 ppm for the  $^{13}\text{C}$  dimension, and at the water frequency for the  $^1\text{H}$  dimension.

20 To detect NOEs between the two ligands and the protein, a 3D  $^{12}\text{C}$ -filtered,  $^{13}\text{C}$ -edited NOESY spectrum was collected. The pulse scheme consisted of a double  $^{13}\text{C}$ -filter sequence (A. Gemmeker, *et al.*, J. Magn. Reson., 96: 199-204 (1992)) concatenated with a NOESY-HMQC sequence (S. Fesik, *et al.*, J. Magn. Reson., 87: 588-593 (1988)); D. Marion, *et al.*, J. Am. Chem. Soc., 111: 1515-1517 (1989)). The spectrum was recorded  
25 with a mixing time of 80 ms, and a total of 80 ( $t_1$ ,  $^{13}\text{C}$ ) x 80 ( $t_2$ ,  $^1\text{H}$ ) x 1024 ( $t_3$ ,  $^1\text{H}$ ) complex points with 16 scans per increment. Spectral widths were 8865 Hz (17.73 ppm,  $w_1$ ), 6667 Hz (13.33 ppm,  $w_2$ ), and 8333 Hz (16.67 ppm,  $w_3$ ), and the carrier positions were 40.0 ppm for the carbon dimension and at the water frequency for both proton dimensions.

30 To identify amide groups that exchanged slowly with the solvent, a series of  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra (G. Bodenhausen, *et al.*, J. Chem. Phys. Lett., 69: 185-189 (1980)) were recorded at 25°C at 2 hr intervals after the protein was exchanged into D<sub>2</sub>O. The acquisition of the first HSQC spectrum was started 2 hrs. after the addition of D<sub>2</sub>O.

35 All NMR spectra were recorded at 25°C on a Bruker AMX500 or AMX600 NMR spectrometer. The NMR data were processed and analyzed on Silicon Graphics computers. In all NMR experiments, pulsed field gradients were applied where appropriate as described (A. Bax, *et al.*, J. Magn. Reson., 99: 638 (1992)) to afford the suppression of the solvent signal and spectral artifacts. Quadrature detection in indirectly detected dimensions was



accomplished by using the States-TPPI method (D. Marion, *et al.*, J. Am. Chem. Soc., **111**: 1515-1517 (1989)). Linear prediction was employed as described (E. Olejniczak, *et al.*, J. Magn. Reson., **87**: 628-632 (1990)).

The derived three-dimensional structure of the ternary complex was then used to  
5 define the spatial orientation of the first and second ligands to each other as well as to the target stromelysin molecule.

Distance restraints derived from the NOE data were classified into six categories based on the NOE cross peak intensity and given a lower bound of 1.8 Å and upper bounds of 2.5 Å, 3.0 Å, 3.5 Å, 4.0 Å, 4.5 Å, and 5.0 Å, respectively. Restraints for  $\phi$  torsional angles  
10 were derived from  $^3J_{\text{HNH}\alpha}$  coupling constants measured from the 3D HNHA-J spectrum (G. Vuister, *et al.*, J. Am. Chem. Soc., **115**: 7772-7777 (1993)). The  $\phi$  angle was restrained to  $120\% \pm 40\%$  for  $^3J_{\text{HNH}\alpha} > 8.5$  Hz, and  $60\% \pm 40\%$  for  $^3J_{\text{HNH}\alpha} < 5$  Hz.

Hydrogen bonds, identified for slowly exchanging amides based on initial structures, were defined by two restraints: 1.8-2.5 Å for the H-O distance and 1.8-3.3 Å for the N-O  
15 distance. Structures were calculated with the X-PLOR 3.1 program (A. Brünger, "XPLOR 3.1 Manual," Yale University Press, New Haven, 1992) on Silicon Graphics computers using a hybrid distance geometry-simulated annealing approach (M. Nilges, *et al.*, FEBS Lett., **229**: 317-324 (1988)).

A total of 1032 approximate interproton distance restraints were derived from the  
20 NOE data. In addition, 21 unambiguous intermolecular distance restraints were derived from a 3D  $^{12}\text{C}$ -filtered,  $^{13}\text{C}$ -edited NOESY spectrum. Of the 1032 NOE restraints involving the protein, 341 were intra-residue, 410 were sequential or short-range between residues separated in the primary sequence by less than five amino acids, and 281 were long-range involving residues separated by at least five residues.

In addition to the NOE distance restraints, 14  $\phi$  dihedral angle restraints were included  
25 in the structure calculations that were derived from three-bond coupling constants ( $^3J_{\text{HNH}\alpha}$ ) determined from an HNHA-J spectrum (G. Viioster, *et al.*, J. Am. Chem. Soc., **115**: 7772-7777 (1993)). The experimental restraints also included 120 distance restraints corresponding to 60 hydrogen bonds. The amides involved in hydrogen bonds were  
30 identified based on their characteristically slow exchange rate, and the hydrogen bond partners from initial NMR structures calculated without the hydrogen bond restraints. The total number of non-redundant, experimentally-derived restraints was 1166.

The structures were in excellent agreement with the NMR experimental restraints. There were no distance violations greater than 0.4 Å, and no dihedral angle violations greater  
35 than 5 degrees. In addition, the simulated energy for the van der Waals repulsion term was small, indicating that the structures were devoid of bad inter-atomic contacts.

The NMR structures also exhibited good covalent bond geometry, as indicated by small bond-length and bond-angle deviations from the corresponding idealized parameters.



- 30 -

The average atomic root mean square deviation of the 8 structures for residues 93-247 from the mean coordinates was 0.93 Å for backbone atoms (C<sup>α</sup>, N, and C'), and 1.43 Å for all non-hydrogen atoms.

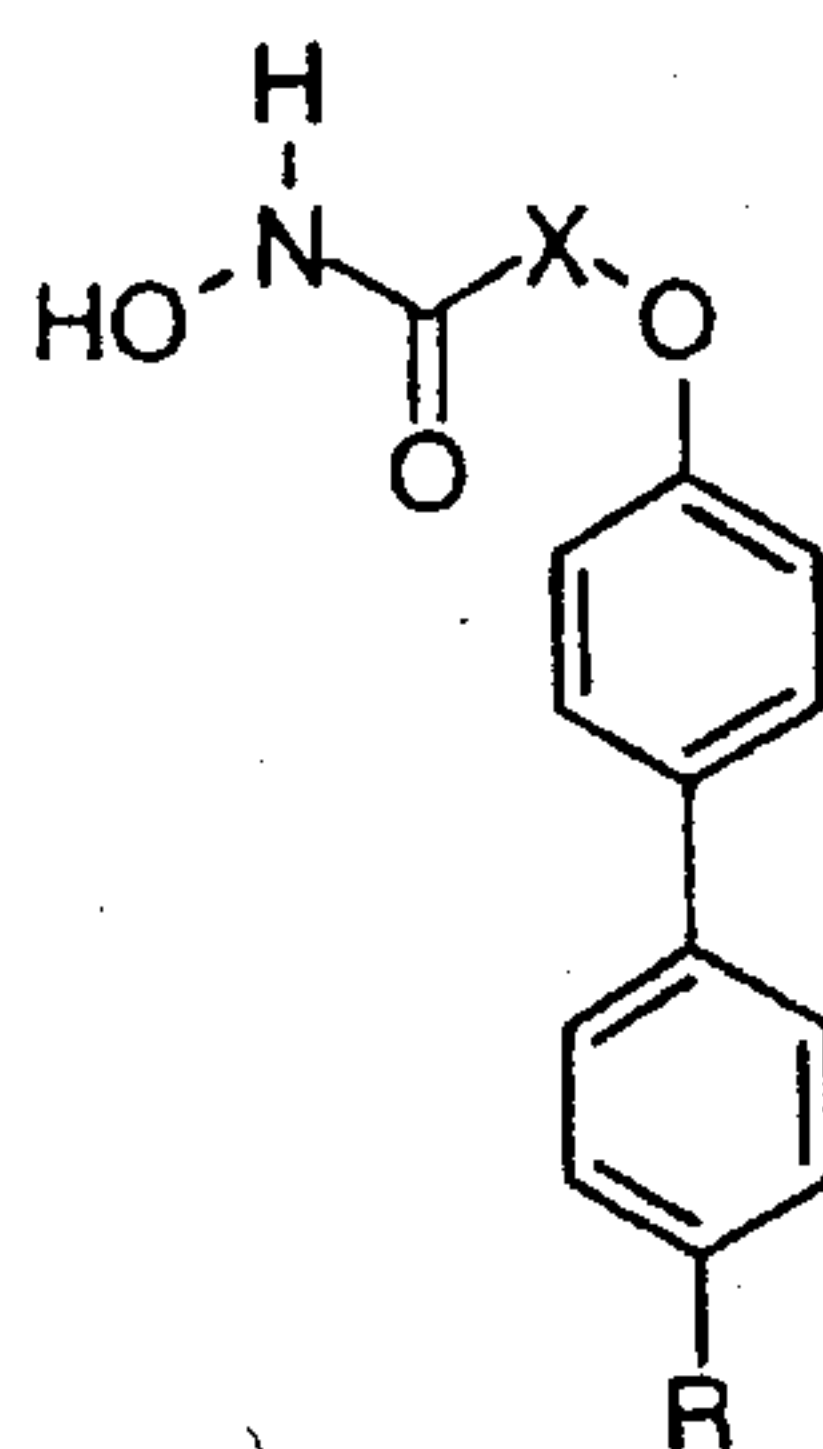
5 A ribbon plot of the ternary complex involving stromelysin, acetohydroxamic acid (the first ligand), and the second ligand is shown in Fig 10. The structure is very similar to the global fold of other matrix metalloproteinases and consists of a five-stranded β-sheet and three α-helices.

10 The catalytic zinc was located in the binding pocket. It was coordinated to three histidines and the two oxygen atom of acetohydroxamic acid. A biaryl group of the second ligand was located in the S1' pocket between the second helix and the loop formed from residues 218-223. This deep and narrow pocket is lined with hydrophobic residues which make favorable contacts with the ligand.

15 Based on the three-dimensional structure of the ternary complex as determined above and the structure/activity relationships observed for the binding to stromelysin of structural analogs of the second ligand (*i.e.*, other biaryl compounds), new molecules were designed that linked together the acetohydroxamic acid to biaryls.

20 As shown in Table 4 below, the initial biaryls chosen contained an oxygen linker and the absence or presence of CN para to the biaryl linkage. Initial linkers contained varying lengths of methylene units. Means for linking compounds with linkers having varying lengths of methylene units are well known in the art.

Table 4





- 31 -

| Compound Inhibition | X                               | R  | Stromelysin     |
|---------------------|---------------------------------|----|-----------------|
| 21                  | (CH <sub>2</sub> ) <sub>2</sub> | H  | 0.31 $\mu$ M    |
| 22                  | (CH <sub>2</sub> ) <sub>3</sub> | H  | 110 $\mu$ M     |
| 23                  | (CH <sub>2</sub> ) <sub>4</sub> | H  | 38%@100 $\mu$ M |
| 24                  | (CH <sub>2</sub> ) <sub>5</sub> | H  | 43%@100 $\mu$ M |
| 25                  | (CH <sub>2</sub> ) <sub>2</sub> | CN | 0.025 $\mu$ M   |
| 26                  | (CH <sub>2</sub> ) <sub>3</sub> | CN | 3.4 $\mu$ M     |
| 27                  | (CH <sub>2</sub> ) <sub>4</sub> | CN | 3.5 $\mu$ M     |
| 28                  | (CH <sub>2</sub> ) <sub>5</sub> | CN | 1.7 $\mu$ M     |

As expected based on the better binding of the CN substituted biaryls to stromelysin, the CN derivatives exhibited better stromelysin inhibition. The compound that exhibited the best inhibition of stromelysin contained a linker with two methylene units.

The present invention has been described with reference to preferred embodiments. Those embodiments are not limiting of the claims and specification in any way. One of ordinary skill in the art can readily envision changes, modifications and alterations to those embodiments that do not depart from the scope and spirit of the present invention.

#### Example 5

##### Design of potent, novel inhibitors of FKBP

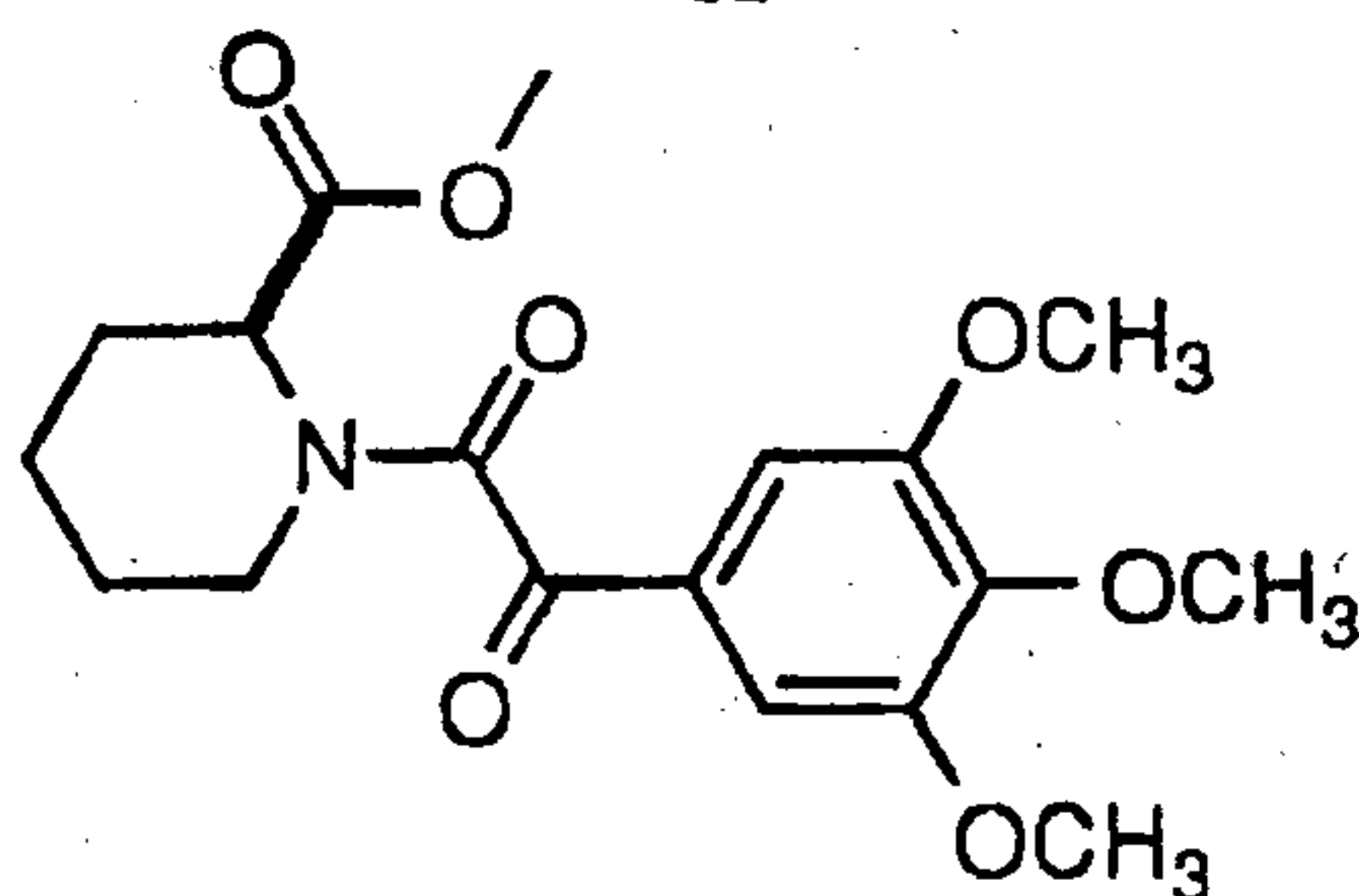
Studies were performed to design novel ligands that bound to FK-binding protein (FKBP).

The cloning, expression and purification of FKBP was accomplished as set forth in Example 1. An initial step in the design of the new ligand was the identification of a first ligand that bound to the FKBP target. Such identification was carried out in accordance with a two-dimensional <sup>15</sup>N/<sup>1</sup>H NMR correlation screening process as disclosed above.

A variety of low -molecular weight fragments and analogs of several known potent immunosuppressants (i.e. ascomycin, rapamycin) were screened for binding to FKBP using the procedures as set forth in example 2. Of the compounds tested, compound 29, below



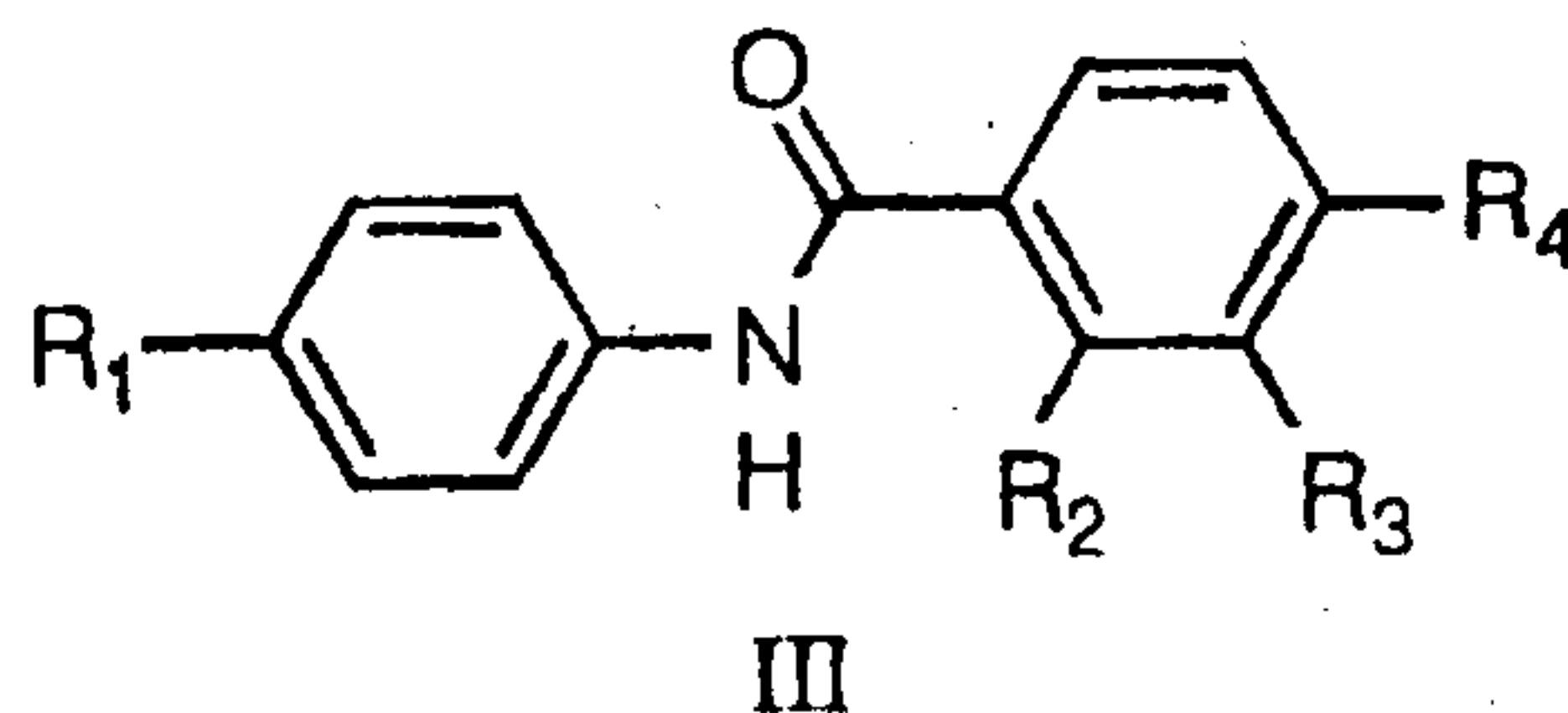
- 32 -



best satisfied the selection criteria: it had a binding affinity for FKBP of 2  $\mu$ M (measured by fluorescence by the methods known in the art) and saturated the protein (> 98% occupancy of the binding site) at ligand concentrations of 1 mM.

The second step in the design process was the identification of a second ligand that bound to the target FKBP at a site different from the binding site of compound 29. This was accomplished by screening compounds for their ability to bind to FKBP in the presence of saturating amounts of the ascomycin fragment analog (compound 29). Details of procedures for this second identification step are as set forth in example 2.

In an initial screen, a compound was found that contained a benzanilide moiety. Based on this initial hit, structurally similar compounds were obtained and tested against FKBP. The structure of these benzanilide compounds is represented by the structure III, below (see Table 5 for definitions of R<sub>1</sub>-R<sub>4</sub>).



In the second round of screening, binding was assayed both in the presence and in the absence of saturating amount of compound 29 (1 mM).

A structure-activity relationship was developed for these diphenyl amide compounds as set forth in Table A. Fig. 6 shows a representative two-dimensional <sup>15</sup>N/<sup>1</sup>H correlation spectrum before and after exposure of FKBP to a diphenyl amide test compound. It can be seen from Fig. 6 that the compound caused chemical shifts of <sup>15</sup>N sites such as those designated I50, Q53, E54, and V55. These sites correspond to an isoleucine (Ile) at position 50, a glutamine (Gln) at position 53, a glutamate (Glu) at position 54, and a valine (Val) at position 55 of SEQ ID NO # 7. Figure 11 shows the correlation between the NMR binding data and a view of the NMR-derived three-dimensional structure of FKBP. The ability to locate the specific binding site of a particular ligand is an advantage of the present invention.



- 33 -

Some compounds only bound to FKBP in the presence of compound 29. Thus the binding affinity of some compounds was enhanced in the presence of compound 29. These results exemplify yet another important capability of the present screening assay which is the ability to identify compounds that bind to the protein in the presence of other molecules.

5 Various benzanilide compounds were tested for binding to FKBP at multiple ligand concentrations. The  $^{15}\text{N}/^1\text{H}$  correlation spectra generated at each concentration were evaluated to quantify differences in the spectra as a function of compound concentration. A binding or dissociation constant ( $K_d$ ) was calculated, using standard procedures well known in the art, from those differences. The results of this study are shown in Table 5. The values for R1-R4  
10 in Table 5 refer to the corresponding positions in the structure III, above.

Table 5

| Compound No. | R1 | R2 | R3            | R4 | $K_d$ (mM) |
|--------------|----|----|---------------|----|------------|
| 30           | OH | OH | H             | H  | 0.8        |
| 31           | H  | H  | OH            | H  | 1.4        |
| 32           | H  | H  | H             | OH | 0.5        |
| 33           | OH | H  | H             | OH | 0.1        |
| 34           | OH | H  | H             | H  | 0.6        |
| 35           | OH | H  | $\text{CH}_3$ | OH | 0.5        |
| 36           | H  | H  | H             | H  | >5.0       |
| 37           | H  | OH | H             | H  | >5.0       |

15 The data in Table 5 show the utility of a process of the present invention in determining dissociation or binding constants between a ligand and a target molecule.

The next step in the design process was to construct a ternary complex of the target FKBP, the first ligand and the second ligand. This was accomplished by exposing the FKBP target to the two ligands under conditions that resulted in complex formation. The location  
20 and orientation of the ligands were then determined as described below.

The  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  resonances of FKBP in the ternary complex were assigned from an analysis of several 3D double- and triple-resonance NMR spectra. The assignment process was aided by known assignments of FKBP when complexed to ascomycin (R. Xu, et al., *Biopolymers*, 33: 535-550, 1993).  $^1\text{H}$  sidechain and  $^{15}\text{N}/^1\text{H}$  backbone resonances  
25 were identified from an analysis of three-dimensional (3D)  $\text{HC}(\text{CO})\text{NH}$  spectra recorded with spectral widths of 2000 Hz (39.5 ppm), 6250 Hz (12.5 ppm) and 8333 Hz (16.7 ppm) in the  $F_1(^{15}\text{N})$ ,  $F_2(^1\text{H})$  and  $F_3(^1\text{H})$  dimensions, respectively, and with a data matrix of  $46(\text{t1}) \times$



- 34 -

80(t2) x 1024(t3) complex points and 16 scans per increment.  $^1\text{H}$  and  $^{13}\text{C}$  sidechain and  $\text{C}\alpha$  resonances were identified from an analysis of 3D HCCH-TOCSY spectra (L. Kay, et al. J. Magn. Reson., 101b:333-337, 1993) recorded with spectral widths of 7541.5 Hz (60.0 ppm), 6250 Hz (12.5 ppm) and 8333 Hz (16.7 ppm) in the  $\text{F}_1(^{13}\text{C})$ ,  $\text{F}_2(^1\text{H})$  and  $\text{F}_3(^1\text{H})$  dimensions, respectively, and with a data matrix of 48(t1) x 64(t2) x 1024(t3) complex points and 16 scans per increment. Intermolecular NOEs between the ligand and FKBP were obtained from an analysis of a 3D  $^{12}\text{C}$ -filtered,  $^{13}\text{C}$  edited NOESY spectrum. The pulse scheme consisted of a double  $^{13}\text{C}$  filter sequence (A. Gemmecker, et al., J. Magn. Reson., 96:199-204, 1992) concatenated with a NOESY-HMQC sequence (S. Fesik, et al., J. AM. Chem. Soc., 111:1515-1517, 1989). The spectrum was recorded with a mixing time of 350 ms and a total of 46(t1,  $^{13}\text{C}$ ) x 64(t2,  $^1\text{H}$ ) x 1024(t3,  $^1\text{H}$ ) complex points and 16 scans per increment. Spectral widths of 7541.5 Hz (60.0 ppm), 6250 Hz (12.5 ppm) and 8333 Hz (16.7 ppm) were used in the  $\text{F}_1(^{13}\text{C})$ ,  $\text{F}_2(^1\text{H})$  and  $\text{F}_3(^1\text{H})$  dimensions, respectively.

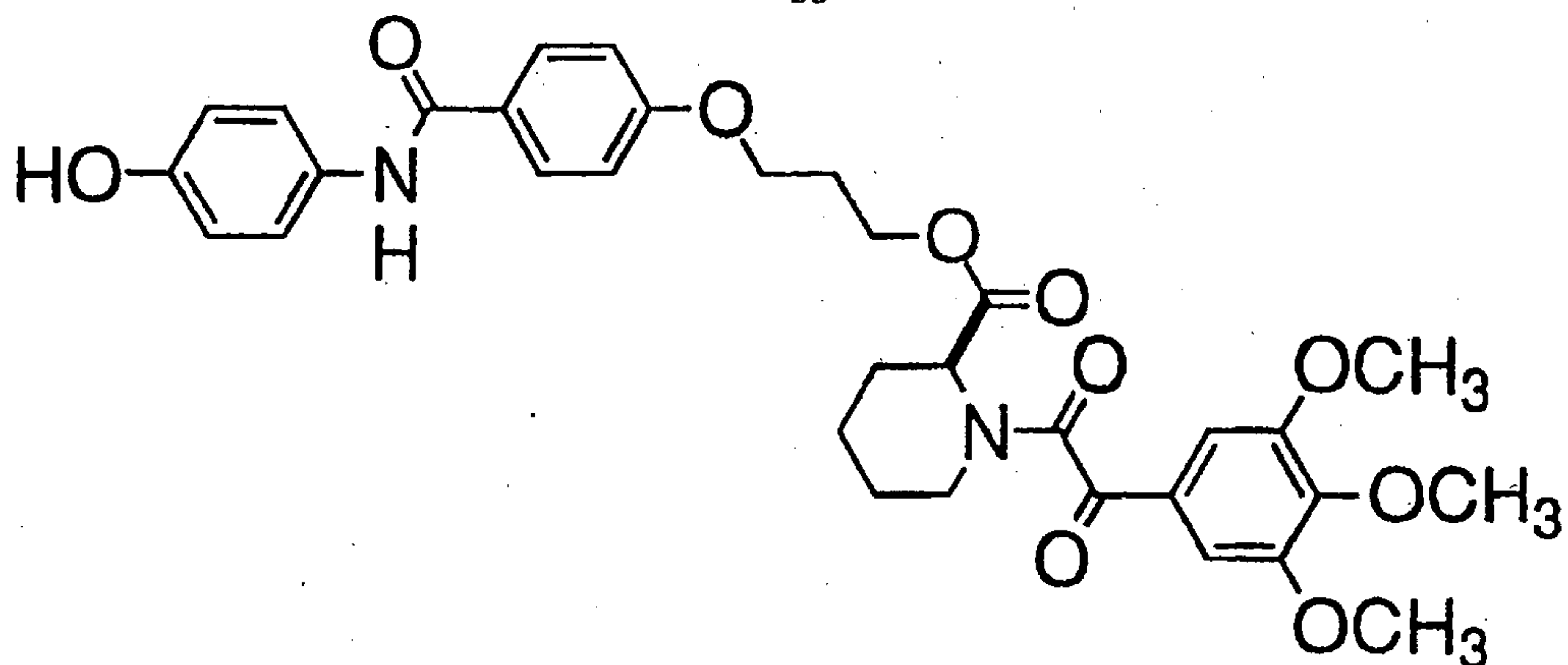
In all spectra, the  $^{15}\text{N}$  carrier frequency was set at 117.4 ppm, the  $^{13}\text{C}$  carrier frequency was set at 40.0 ppm, and the  $^1\text{H}$  carrier frequency was set on the water resonance. All spectra were recorded at 303K on a Bruker AMX500 NMR spectrometer. The NMR data were processed and analyzed on Silicon Graphics computers. In all NMR experiments, pulsed field gradients were applied where appropriate as described (A. Bax, et al., J. Magn. Reson., 99:638, 1992) to afford the suppression of the solvent signal and spectral artifacts. Quadrature detection in the indirectly detected dimension was accomplished by using the States-TPPI method (D. Marion, et al. J. Am. Chem. Soc., 87: 1515-1517, 1989). Linear prediction was employed as described (E. Olejniczak, et al., J. Magn. Reson., 87: 628-632, 1990).

Distance restraints derived from the NOE data were classified into three categories based on the NOE crosspeak intensity and were given a lower bound of 1.8 Å and upper bounds of 3.0, 4.0 and 5.0 Å. A total of 17 intermolecular distance restraints between the protein and compound 33 and 10 intermolecular distance restraints between the protein and compound 29 were used to define the location and orientation of the compounds when bound to FKBP using the known three-dimensional coordinates for the FKBP protein structure. A ribbon plot of the ternary complex involving FKBP, a fragment analog of ascomycin (compound 29), and a benzanilide compound (compound 33) is shown in Figure 12.

Based on the three-dimensional structure of the ternary complex as determined above and the structure activity relationships observed for the binding to FKBP of structural analogs of the second compound, a new molecule was designed that linked the ascomycin fragment analog to the benzanilide compound. This compound, shown below,



- 35 -



IV

has a 19 nM affinity for FKBP as determined by fluorescence titrations. This is a 100-fold  
 5 increase in potency over the ascomycin fragment analog (compound 29) alone ( $K_d = 2 \mu\text{M}$ ).

As shown by the above non-limiting examples, the present invention relates to a process for  
 designing a high-affinity ligand to a given target molecule, comprising

- a) identifying by the screening processes described herein at least two ligands  
 10 which bind to distinct binding sites on the target molecule using multidimensional NMR  
 spectroscopy;
- b) forming at least a ternary complex by exposing the at least two ligands to the  
 target molecule;
- c) determining the three dimensional structure of the complex and thus the spatial  
 15 orientation of the at least two ligands on the target molecule; and
- d) using the spatial orientation determined in step c) to design the high affinity  
 ligand which structurally resembles a combination of the at least two ligands which bind to  
 distinct sites on the target molecule. Preferably, the high-affinity ligand designed in the  
 above process serves as or is the basis for a drug which binds to a given target molecule and  
 20 performs, in vitro and in vivo, a targeted therapeutic effect in mammals including humans in  
 need of treatment thereof.

The process also relates to designing a high-affinity ligand to a given target molecule  
 comprising:

- a) identifying a first ligand to the target molecule using multidimensional NMR  
 25 spectroscopy;
- b) identifying a second ligand to the target molecule using multidimensional  
 NMR spectroscopy wherein the second ligand may be the same or different than the first  
 ligand and wherein the second ligand binds to a different site on the target molecule than the  
 first ligand;



- 36 -

- c) forming a ternary complex by binding the first and second ligands to the target molecule;
  - d) determining the three dimensional structure of the complex and thus the spatial orientation of the first ligand and the second ligand on the target molecule; and
  - 5 e) designing the ligand wherein the spatial orientation of step d) is maintained.
- In the process described above, the first and second ligands may have the identical molecular structure or formula wherein the moiety binds to at least two binding sites on the target molecule. The ligand that is based upon the structural combination of the first and second ligands then serves as a drug lead or drug upon actual synthesis of that combined compound and evaluation in the appropriate biological assays. The synthesis of the combined ligand, high-affinity ligand, drug lead or drug is achieved through synthetic or biological means. Conceptually, as indicated throughout the specification, the first and second ligands are linked (joined) together by carbon atoms, heteroatoms, or a combination thereof to form the ligand or drug lead. The processes described herein, of course, include syntheses of the high-
- 10 affinity ligand by linear or non-linear (convergent) means which ultimately produce the linked (combined) first, second or more ligands.

The first and second ligands may also have different molecular structures and either of the ligands may or may not bind to the other (distinct) binding site on the target molecule.

In more detail, the process of the invention also relates to designing a high affinity ligand to given target molecule, comprising

20 ligand to given target molecule, comprising

- a) preparing an isotopically-labeled target molecule wherein said molecule is enriched with an NMR detectable isotope;
  - b) generating a multidimensional NMR spectra of the isotopically-labelled target molecule;
  - 25 c) screening the isotopically-labeled target molecule by exposing the target molecule to a plurality of compounds to identify by multidimensional NMR spectroscopy at least a first and second ligand which bind to distinct sites on the target molecule ;
  - d) forming at least a ternary complex by exposing at least the first and second ligand to the isotopically-labeled target molecule;
  - 30 e) determining the spatial orientation of the at least first and second ligand on the isotopically-labeled target molecule;
  - f) using the spatial orientation determined in step e) to design the high affinity ligand based upon the combination of the at least first and second ligands. Of course, a plurality of ligands ( $1 + n$ ) can be combined to form a high affinity ligand which has the spatial orientation of the  $1 + n$  ( $n = 1 - \infty$ ) combined ligands. After the high-affinity ligand has been designed, the process may further include the step f) of making the high affinity ligand by synthetic or biological means. The at least two ligands (first and second ligand) may be linked by carbon atoms (e.g. by methylene or alkylene units) or by heteroatoms (e.g. by
- 35



- 37 -

nitrogen, oxygen, sulphur) or by other atoms which maintain or approximate the spatial orientation of the 1 + n ligands to the target molecule. Depending upon the ligands, the molecules may also be combined or joined (linked) directly to each other without intervening alkylene or heteroatom linker units. The high affinity ligand produced from the 1 + n combined ligands preferably shows an increase in binding potency to the target molecule in relation to any one of the 1 + n ligands. The present invention, therefore, includes high-affinity ligands designed by the processes shown herein wherein said high-affinity ligand has an increase in binding potency ( $K_d$ ) to the given target molecule over the at least two ligands which bind to distinct sites on the given target molecule.

10       The present invention also relates to a method for discovering high affinity ligands using structure-activity relationships obtained from nuclear magnetic resonance wherein said method comprises constructing a high-affinity ligand from ligands which bind to a subsite of a target molecule by ;

15       i)       screening low molecular weight ( < 450 MW) compounds which bind to a subsite 1 of the target molecule;

      ii)       screening analogs prepared from the initial results obtained in step i) to optimize binding to subsite 1;

20       iii)      screening for low molecular weight (< 450 mw) compounds and corresponding analogs which bind to a nearby binding site, subsite 2, of the target molecule using multidimensional NMR spectroscopy to measure binding affinity; wherein, after steps (i) - (iii), lead fragments are generated;

25       iv)      combining lead fragments generated from steps i) - iii) to design a high affinity ligand. Combining can be achieved by synthetic or biological means. Synthetic means includes organic synthesis of the combined ligand. Biological means includes fermentation or generation of the combined ligand through a cellular vehicle or system. Preferably, the target molecule is a polypeptide. The present invention also relates to the method as recited above wherein the combination of fragments produces a ligand with a higher binding potency ( $K_d$ ) than the individual fragments to the target molecule.



- 38 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Abbott Laboratories
- (ii) TITLE OF INVENTION: USE OF NUCLEAR MAGNETIC RESONANCE  
TO DESIGN LIGANDS TO TARGET BIOMOLECULES
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: SWABEY OGILVY RENAULT
  - (B) STREET: 1981 McGill College Avenue - Suite 1600
  - (C) CITY: Montreal
  - (D) STATE: Quebec
  - (E) COUNTRY: Canada
  - (F) ZIP: H3A 2Y3
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette
  - (B) COMPUTER: IBM Compatible
  - (C) OPERATING SYSTEM: Windows
  - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: 2,237,336
  - (B) FILING DATE: 13 November 1996
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/744,701
  - (B) FILING DATE: 31 October 1996
  - (A) APPLICATION NUMBER: 08/678,903
  - (B) FILING DATE: 12 July 1996
  - (A) APPLICATION NUMBER: 08/558,633
  - (B) FILING DATE: 14 November 1995
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Murphy, Kevin P.
  - (B) REGISTRATION NUMBER:
  - (C) REFERENCE/DOCKET NUMBER: 11899-486
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 514-845-7126
  - (B) TELEFAX: 514-288-8389
  - (C) TELEX:



- 39 -

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Arg | Thr | Phe | Pro | Gly | Ile | Pro | Lys | Trp | Arg | Lys | Thr | His | Leu | Thr | 1   | 5   | 10  | 15  |
| Tyr | Arg | Ile | Val | Asn | Tyr | Thr | Pro | Asp | Leu | Pro | Lys | Asp | Ala | Val | Asp | 20  | 25  | 30  |     |
| Ser | Ala | Val | Glu | Lys | Ala | Leu | Lys | Val | Trp | Glu | Glu | Val | Thr | Pro | Leu | 35  | 40  | 45  |     |
| Thr | Phe | Ser | Arg | Leu | Tyr | Glu | Gly | Glu | Ala | Asp | Ile | Met | Ile | Ser | Phe | 50  | 55  | 60  |     |
| Ala | Val | Arg | Glu | His | Gly | Asp | Phe | Tyr | Pro | Phe | Asp | Gly | Pro | Gly | Asn | 65  | 70  | 75  | 80  |
| Val | Leu | Ala | His | Ala | Tyr | Ala | Pro | Gly | Pro | Gly | Ile | Asn | Gly | Asp | Ala | 85  | 90  | 95  |     |
| His | Phe | Asp | Asp | Asp | Glu | Gln | Trp | Thr | Lys | Asp | Thr | Thr | Gly | Thr | Asn | 100 | 105 | 110 |     |
| Leu | Phe | Leu | Val | Ala | Ala | His | Glu | Ile | Gly | His | Ser | Leu | Gly | Leu | Phe | 115 | 120 | 125 |     |
| His | Ser | Ala | Asn | Thr | Glu | Ala | Leu | Met | Tyr | Pro | Leu | Tyr | His | Ser | Leu | 130 | 135 | 140 |     |
| Thr | Asp | Leu | Thr | Arg | Phe | Arg | Leu | Ser | Gln | Asp | Asp | Ile | Asn | Gly | Ile | 145 | 150 | 155 | 160 |
| Gln | Ser | Leu | Tyr | Gly | Pro | Pro | Pro | Asp | Ser | Pro | Glu | Thr | Pro |     |     | 165 | 170 |     |     |



- 40 -

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 83 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Ala Thr Thr Pro Ile Ile His Leu Lys Gly Asp Ala Asn Ile Leu
 1             5             10             15
Leu Cys Leu Arg Tyr Arg Leu Ser Lys Tyr Lys Gln Leu Tyr Glu Gln
             20             25             30
Val Ser Ser Thr Trp His Trp Thr Cys Thr Asp Gly Lys His Lys Asn
             35             40             45
Ala Ile Val Thr Leu Thr Tyr Ile Ser Thr Ser Gln Arg Asp Asp Phe
             50             55             60
Leu Asn Thr Val Lys Ile Pro Asn Thr Val Ser Val Ser Thr Gly Tyr
65             70             75             80
Met Thr Ile

```

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 18 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAAATGAAGA GTCTTCAA

18



- 41 -

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGTCCCAGG TTCTGGAG

18

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATACCATGGC CTATCCATTG GATGGAGC

28

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATAGGATCCT TAGGTCTCAG GGGAGTCAGG

30



- 42 -

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |    |    |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|
| Gly | Val | Gln | Val | Glu | Thr | Ile | Ser | Pro | Gly | Asp | Gly | Arg | Thr | Phe | Pro | 1   | 5   | 10 | 15 |
| Lys | Arg | Gly | Gln | Thr | Cys | Val | Val | His | Tyr | Thr | Gly | Met | Leu | Glu | Asp | 20  | 25  | 30 |    |
| Gly | Lys | Lys | Phe | Asp | Ser | Ser | Arg | Asp | Arg | Asn | Lys | Pro | Phe | Lys | Phe | 35  | 40  | 45 |    |
| Met | Leu | Gly | Lys | Gln | Glu | Val | Ile | Arg | Gly | Trp | Glu | Glu | Gly | Val | Ala | 50  | 55  | 60 |    |
| Gln | Met | Ser | Val | Gly | Gln | Arg | Ala | Lys | Leu | Thr | Ile | Ser | Pro | Asp | Tyr | 65  | 70  | 75 | 80 |
| Ala | Tyr | Gly | Ala | Thr | Gly | His | Pro | Gly | Ile | Ile | Pro | Pro | His | Ala | Thr | 85  | 90  | 95 |    |
| Leu | Val | Phe | Asp | Val | Glu | Leu | Leu | Lys | Leu | Glu |     |     |     |     |     | 100 | 105 |    |    |

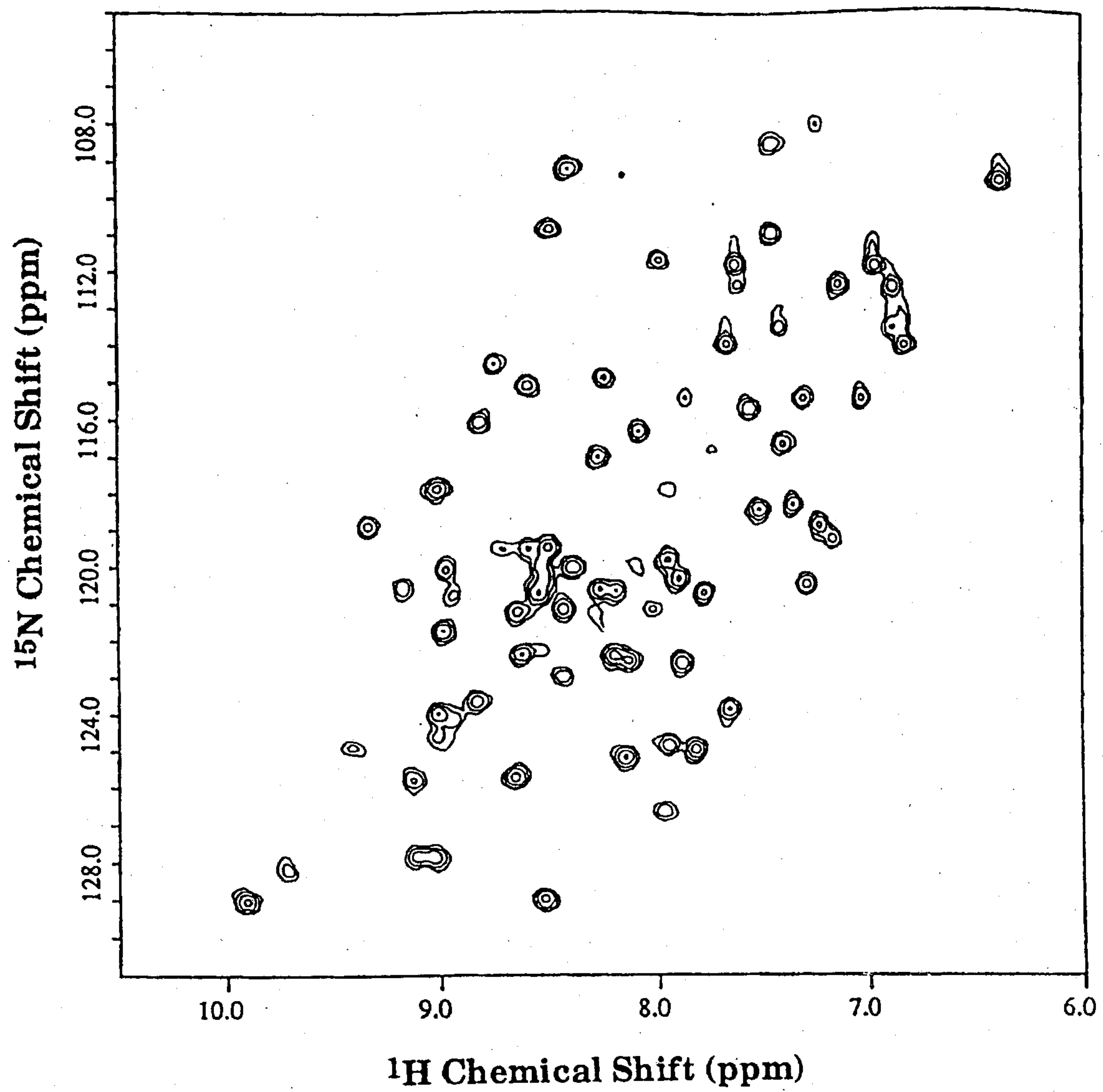


**CLAIMS:**

1. A method for discovering ligands to target molecules using structure-activity relationships obtained from nuclear magnetic resonance, comprising:
  - i) screening compounds having a molecular weight of 100-300 which bind to a subsite 1 of a given target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR to measure binding affinity;
  - ii) screening analogs prepared from binding results obtained in step i) to optimize binding of a first fragment to the target molecule;
  - iii) screening for compounds and corresponding analogs which bind to a nearby binding site, subsite 2, of the target molecule using two-dimensional  $^{15}\text{N}/^1\text{H}$  NMR to measure binding affinity to optimize binding of a second fragment to the target molecule; and
  - iv) combining the first and second fragments to design a ligand.
2. The method according to claim 1, further comprising, before step (iv):
  - A) forming at least a ternary complex by exposing the at least two fragments to the target molecule;
  - B) determining the three-dimensional structure of the complex and the spatial orientation of the at least two fragments on the target molecule; and then carrying out step (iv) by
  - C) using the spatial orientation determined in step B) to design the said ligand.
3. The method according to claim 1 or 2, wherein the target molecule is a protein.
4. The method according to claim 1 or 2, wherein the ligand has a higher binding potency to the target molecule than the fragments thereof.

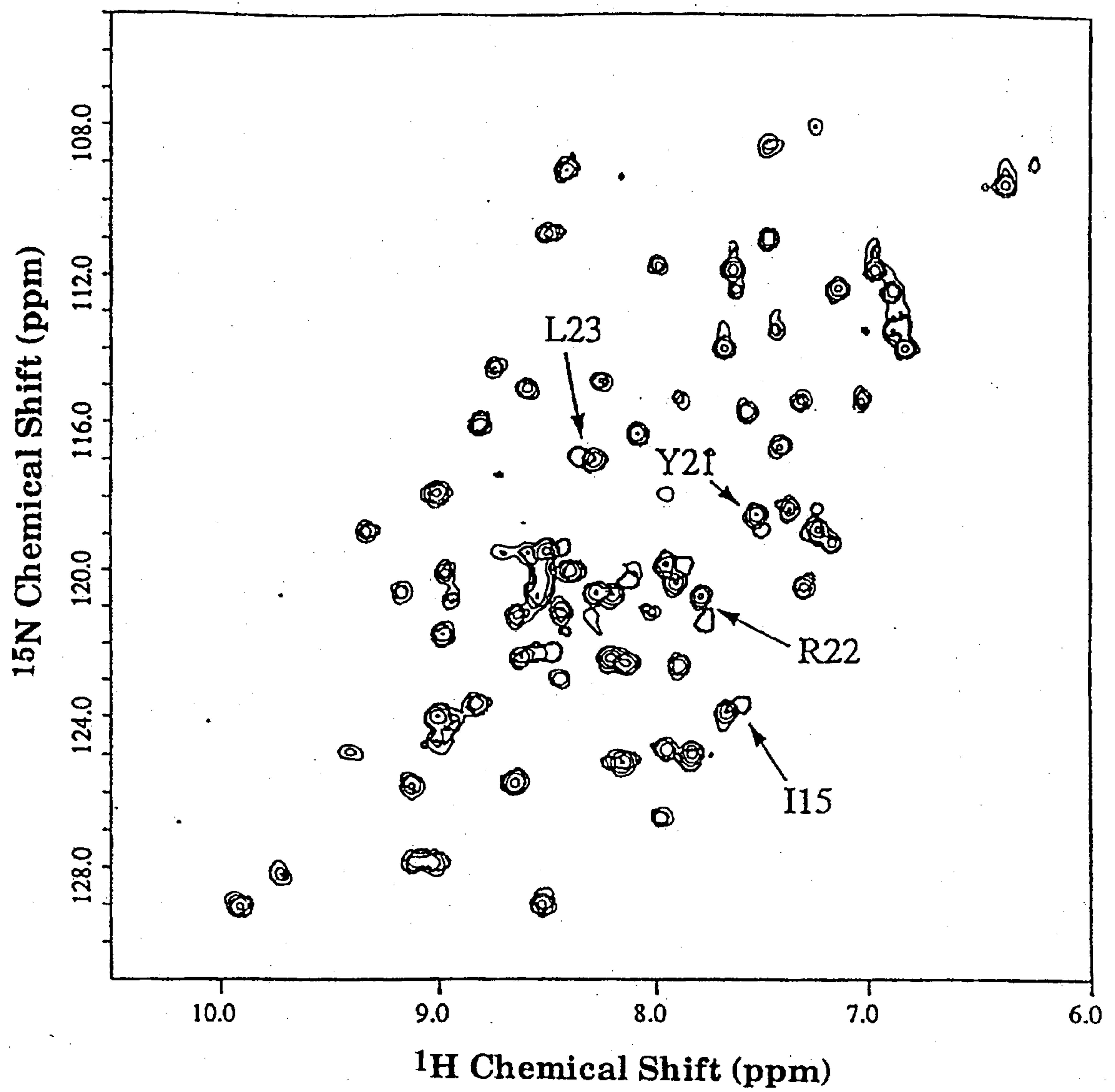


1/12

**Figure 1**

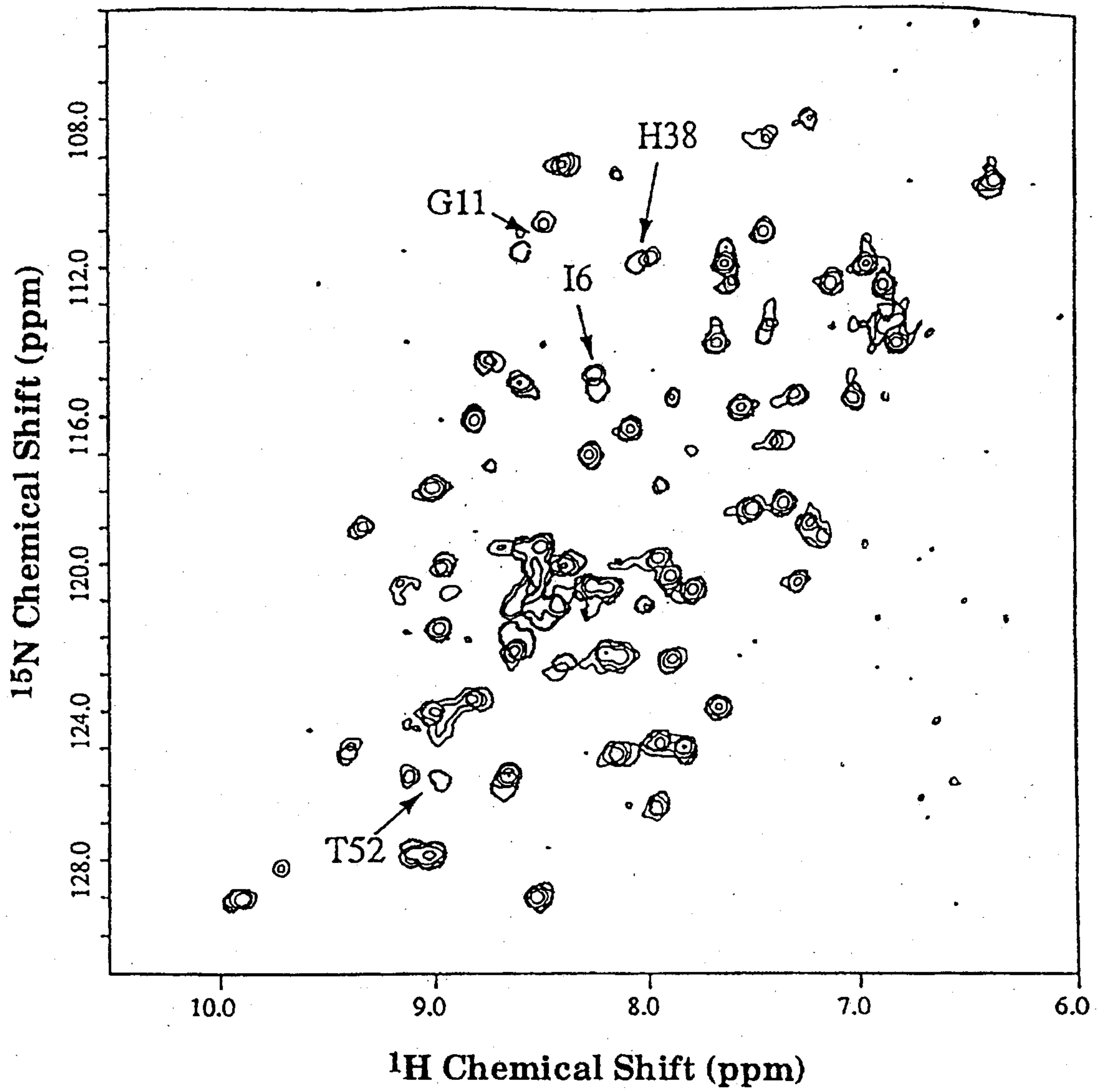


2/12

**Figure 2**

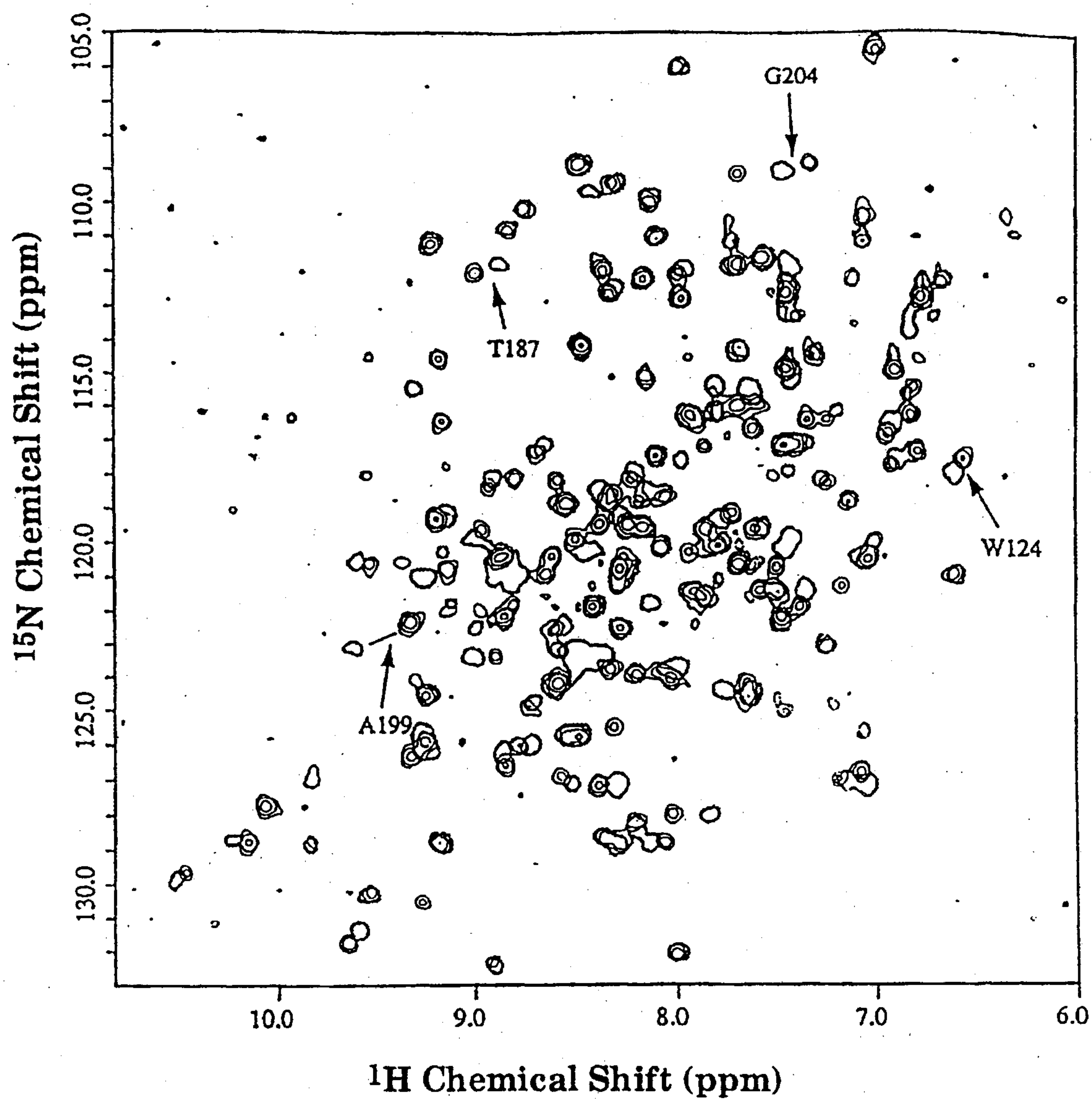


3/12

**Figure 3**

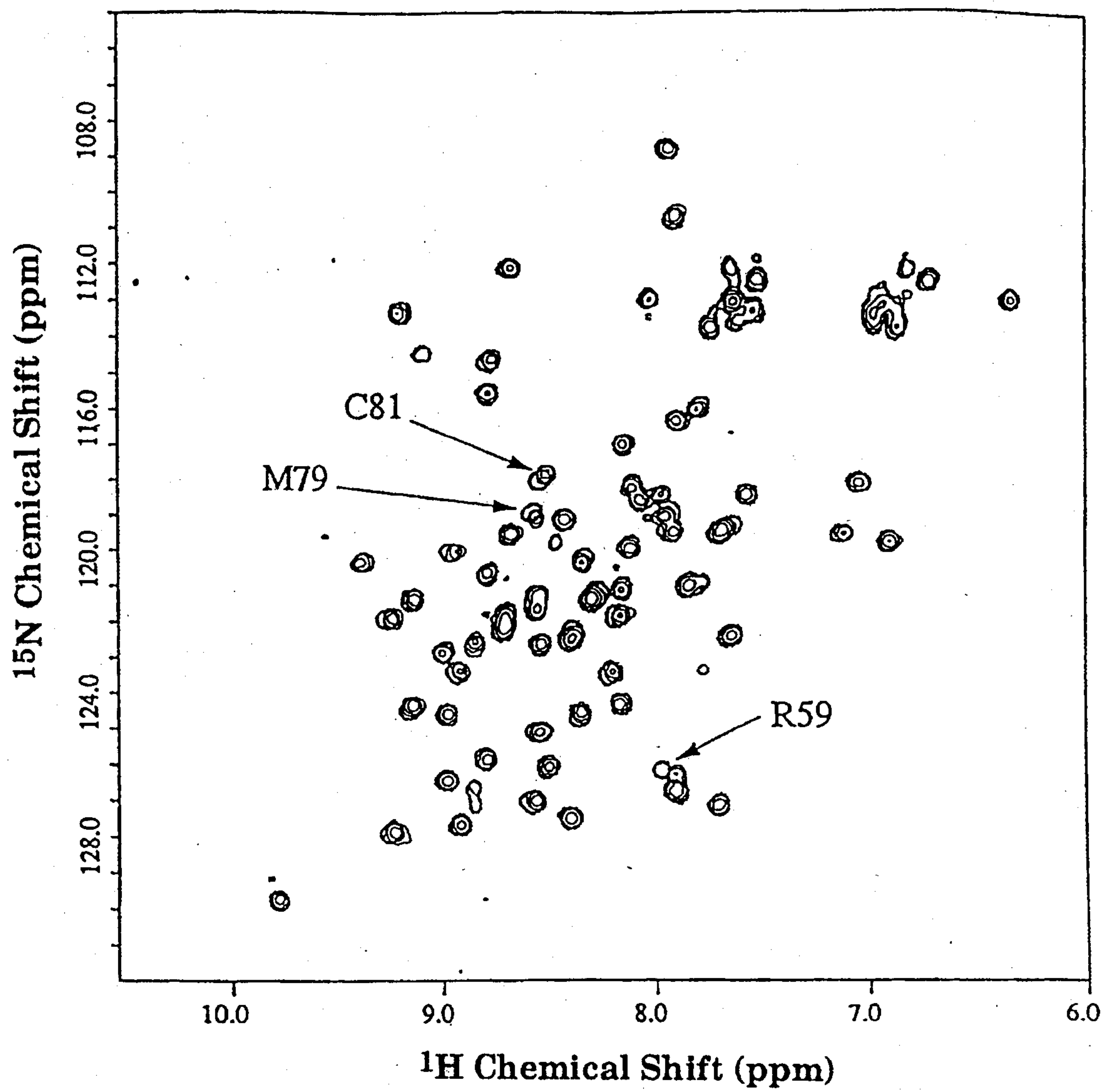


4/12

**Figure 4**

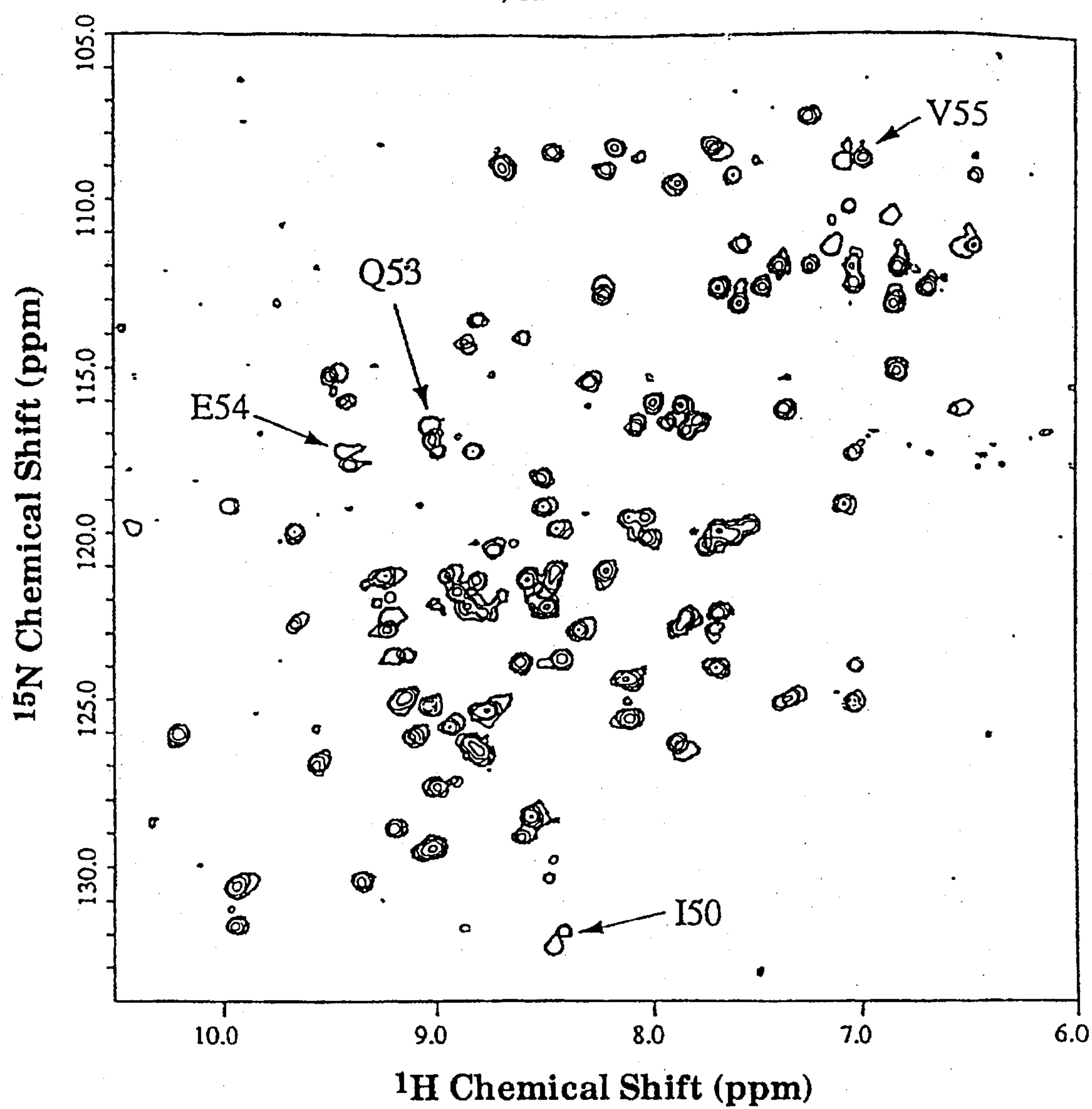


5/12

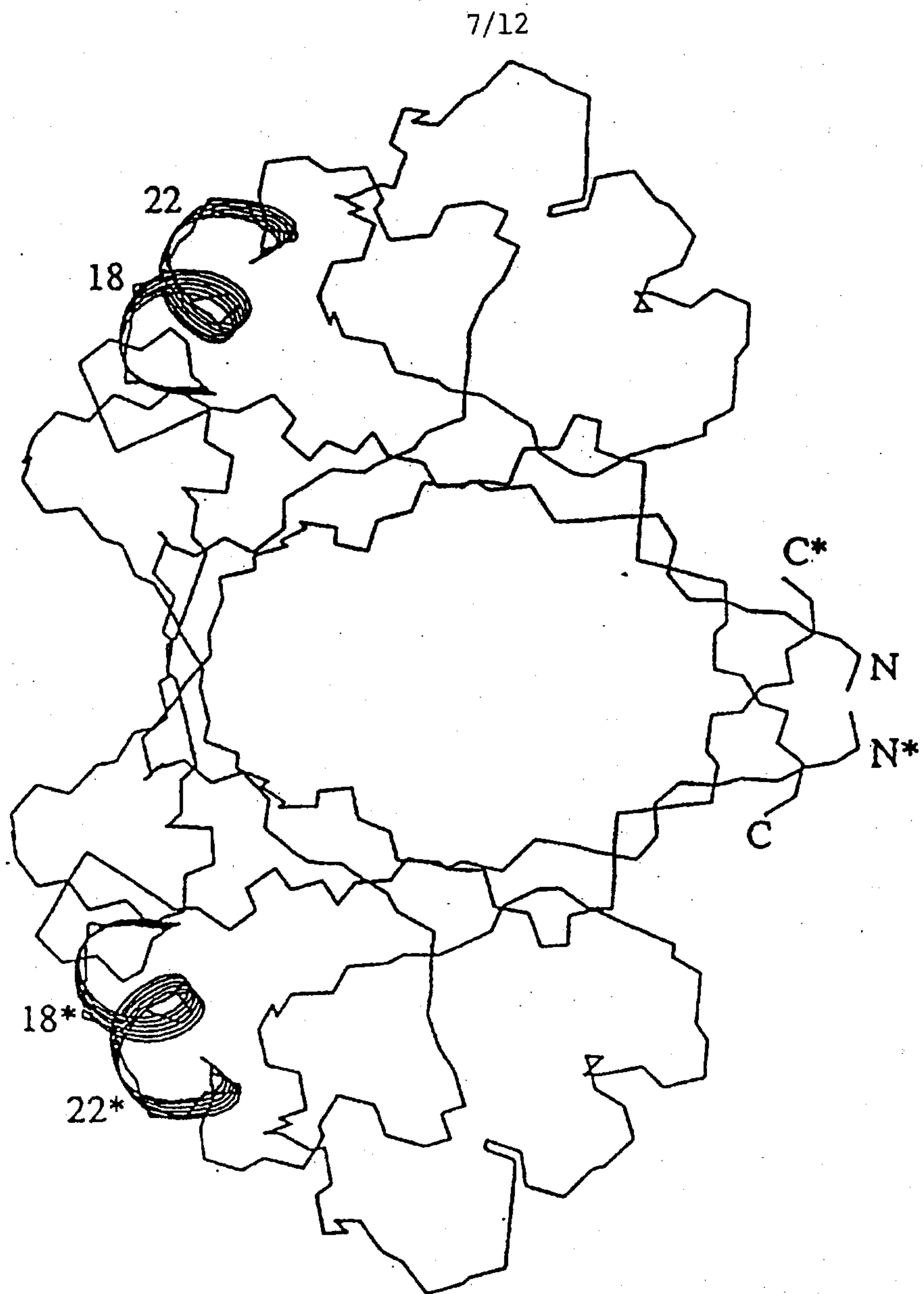
**Figure 5**



6/12

**Figure 6**

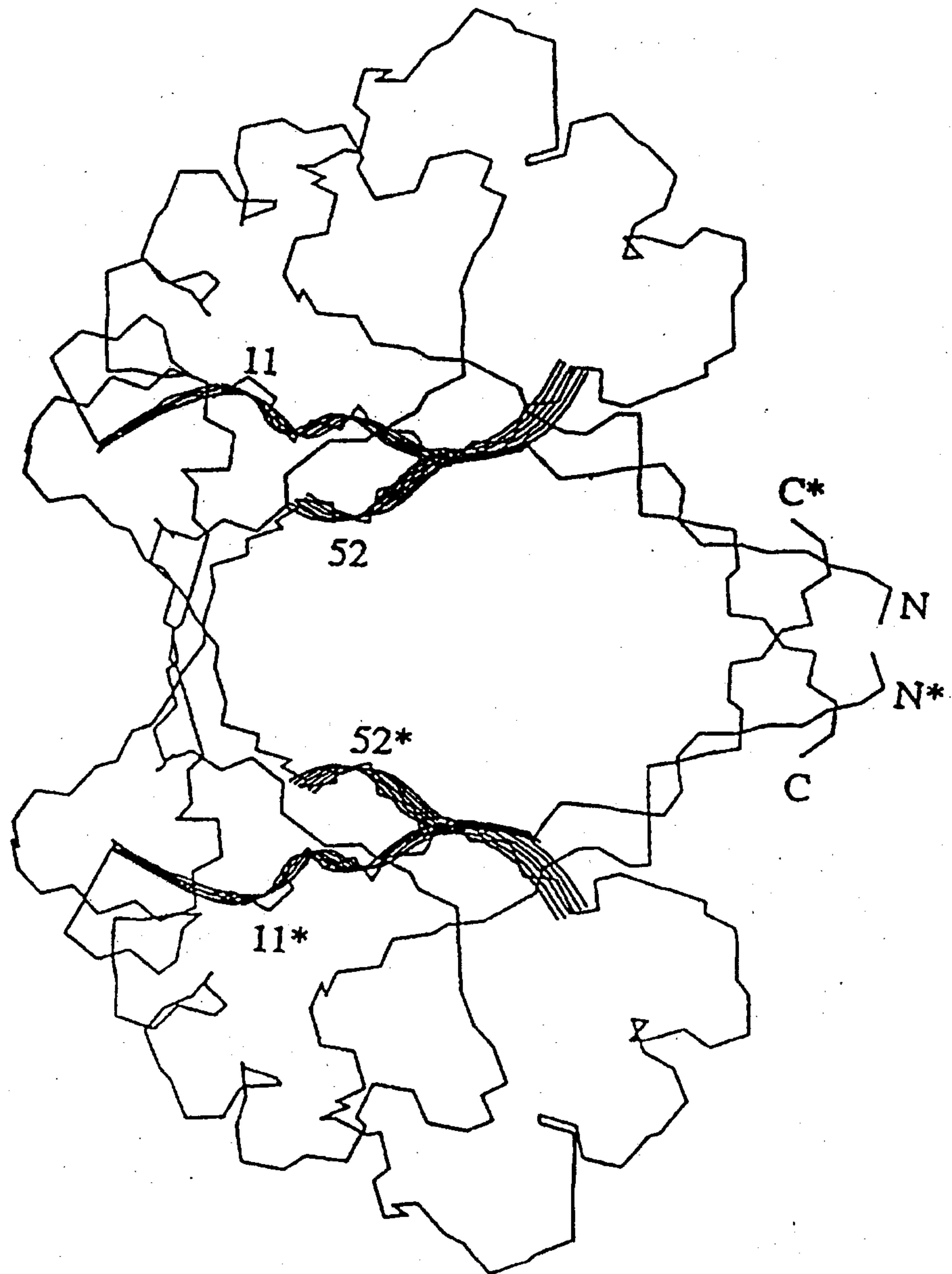




**Figure 7**



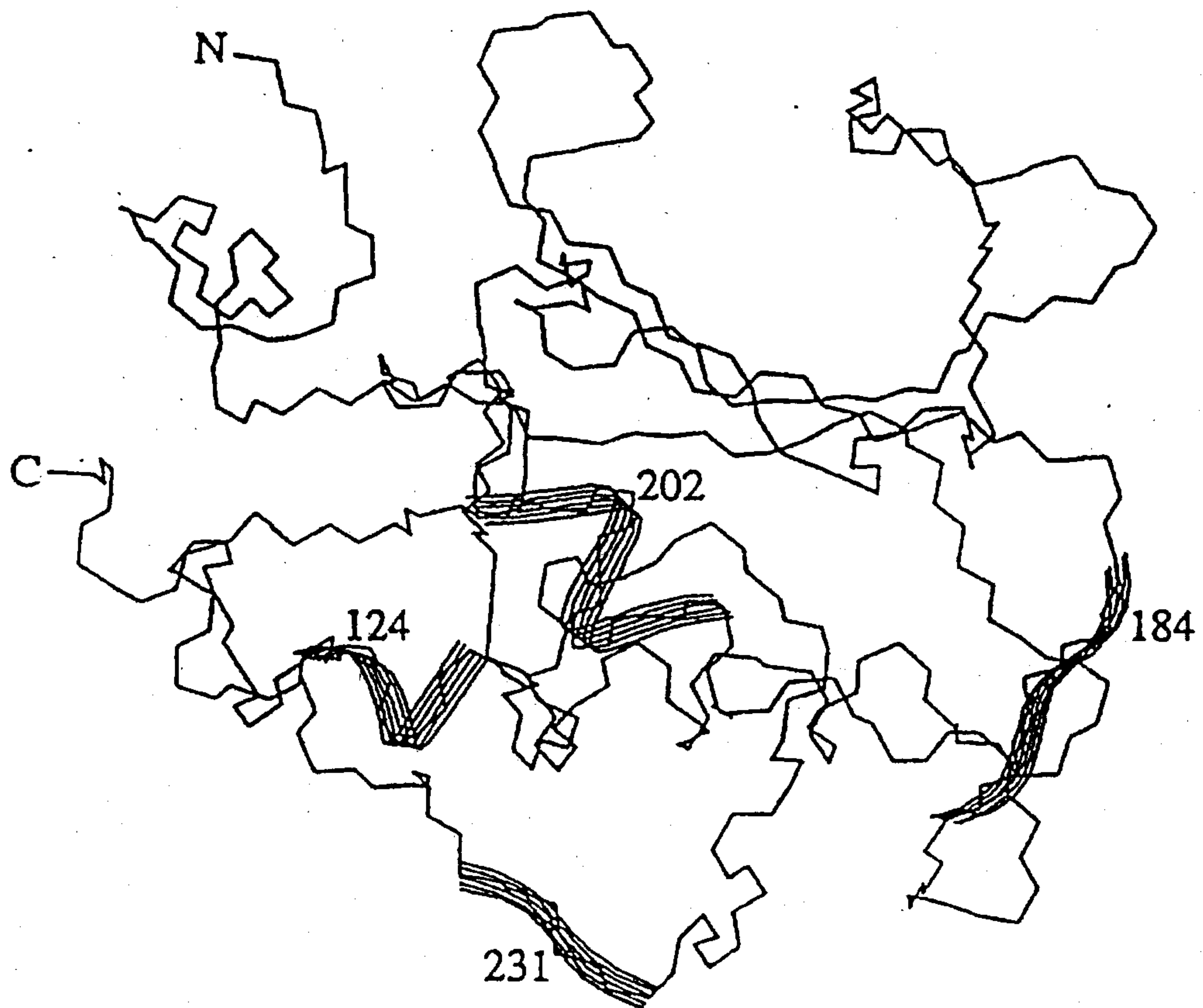
8/12



**Figure 8**



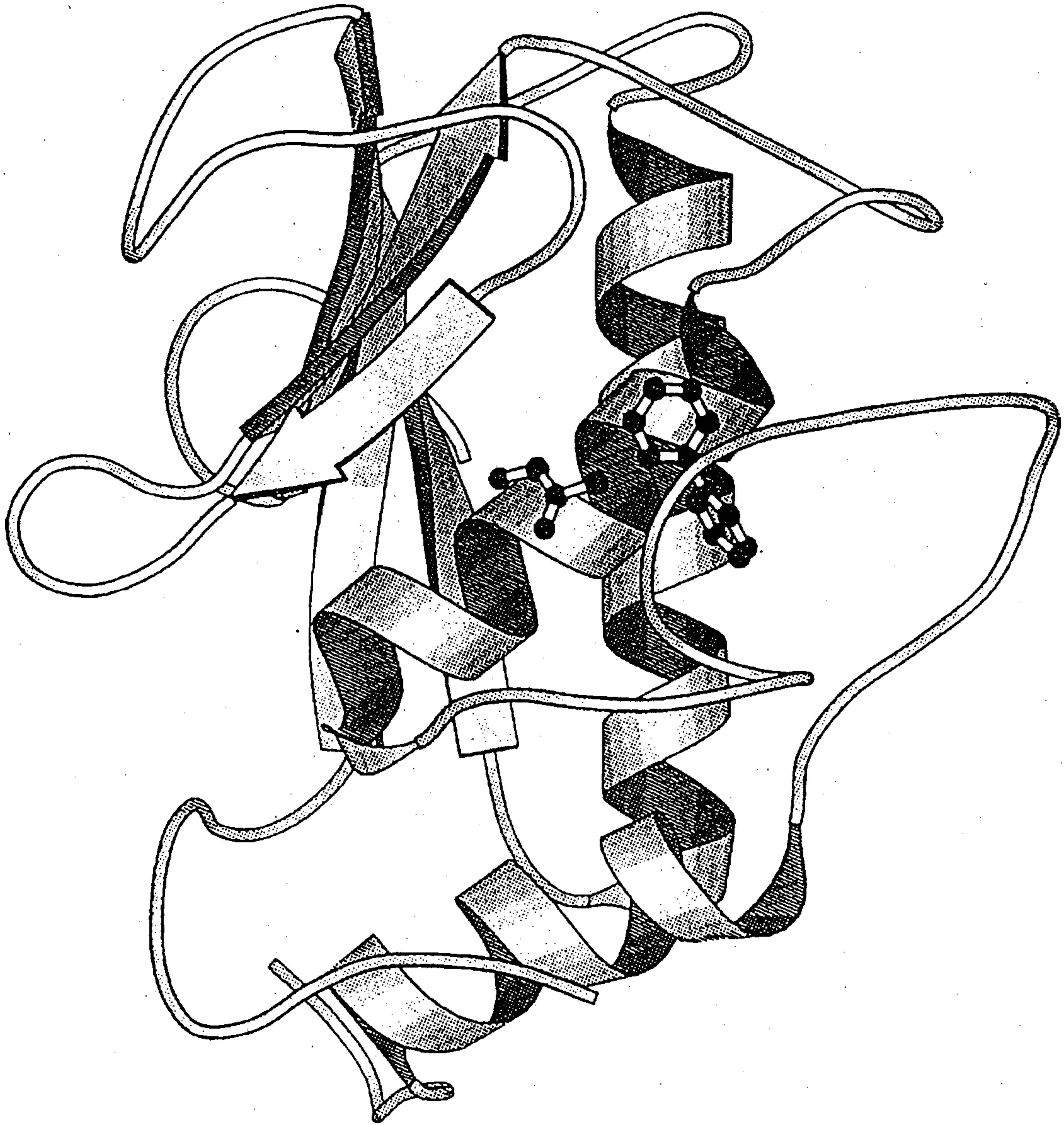
9/12



**Figure 9**



10/12



**Figure 10**



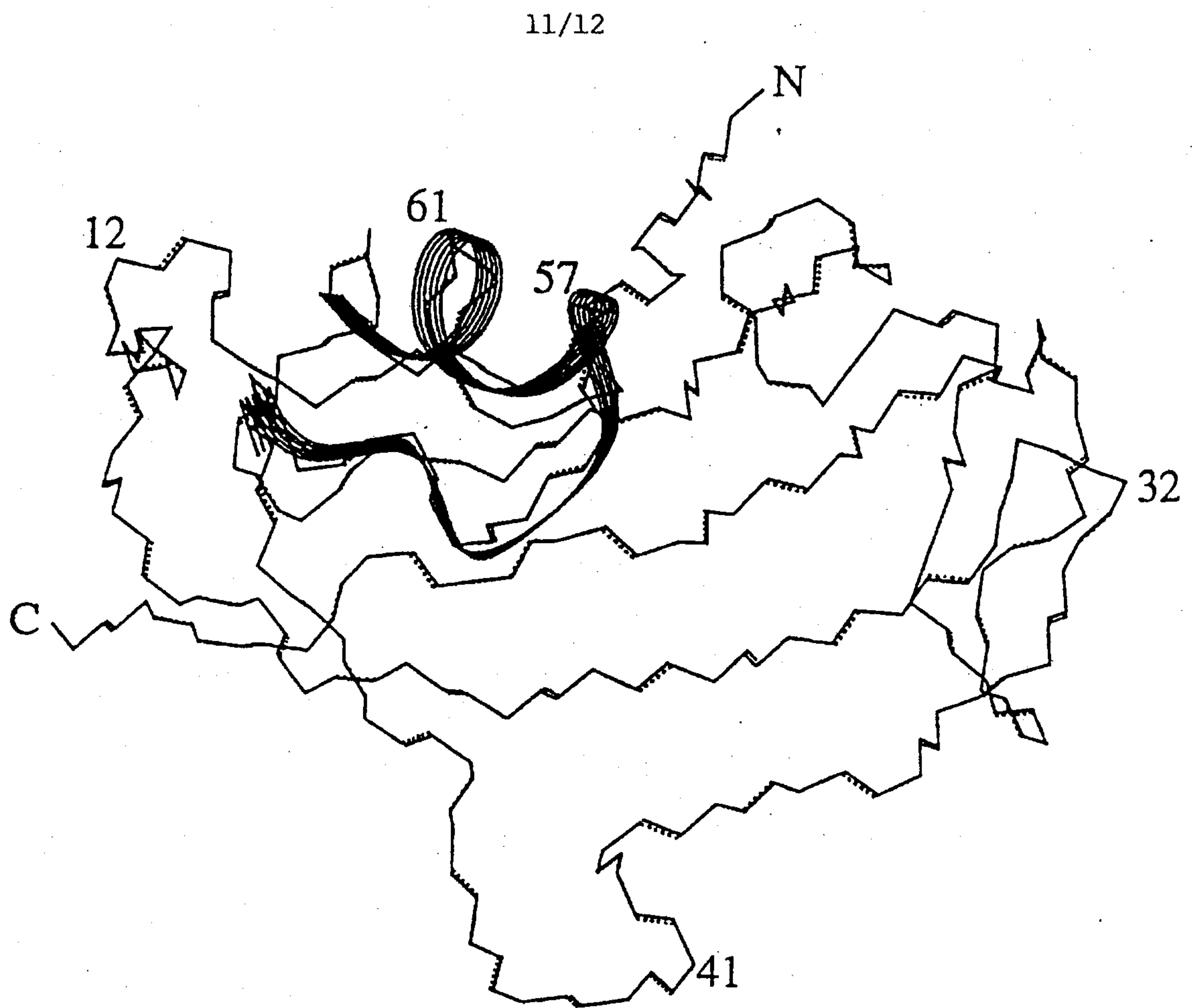


Figure 11



12/12

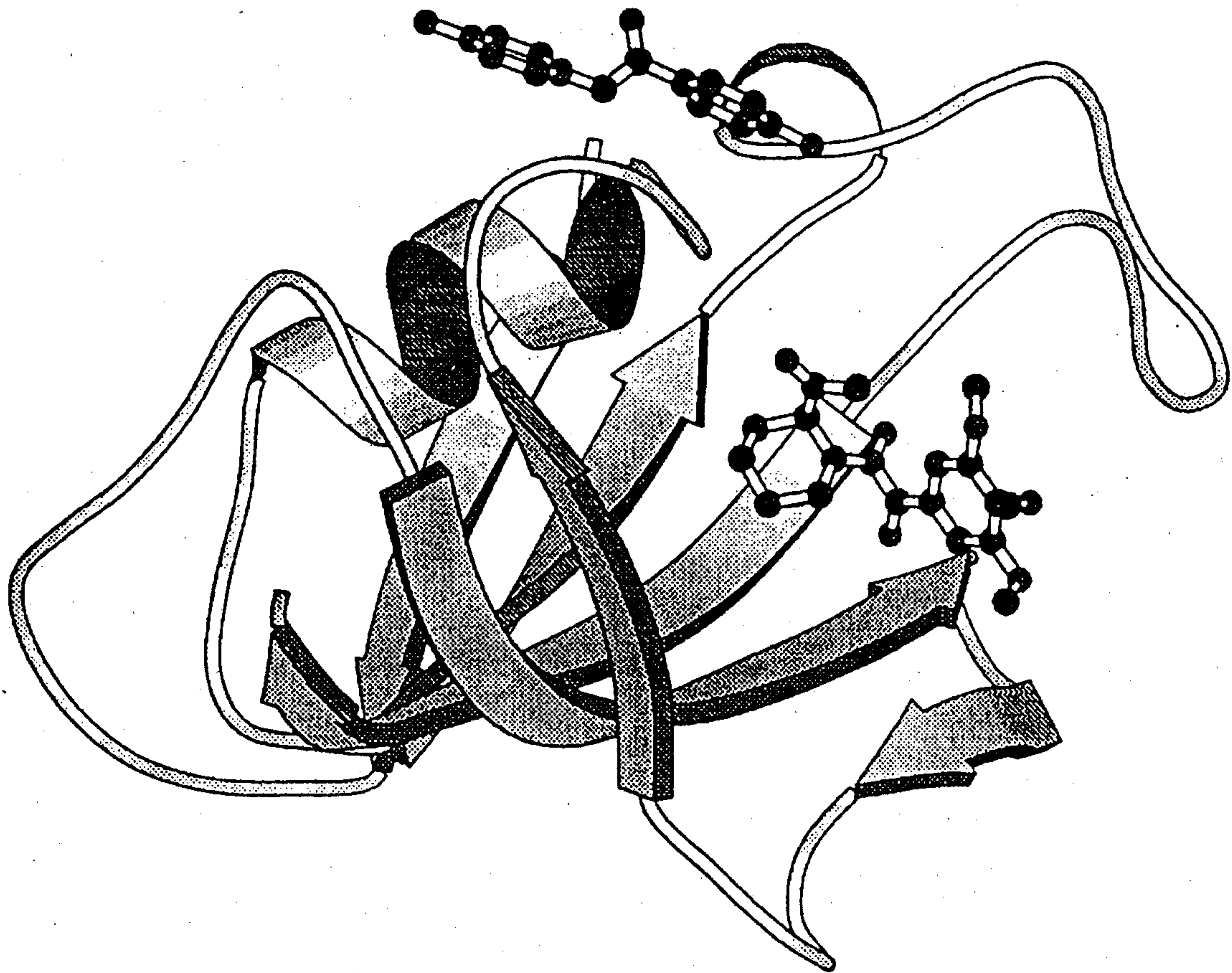


Figure 12



**$^{15}\text{N}$  Chemical Shift (ppm)**

108.0  
112.0  
116.0  
120.0  
124.0  
128.0

10.0

9.0

8.0

7.0

6.0

**$^1\text{H}$  Chemical Shift (ppm)**

