Anal. Chem. **2001,** *73,* 571−581

MS/NMR: A Structure-Based Approach for Discovering Protein Ligands and for Drug Design by Coupling Size Exclusion Chromatography, Mass Spectrometry, and Nuclear Magnetic Resonance Spectroscopy

Franklin J. Moy, Kevin Haraki, Dominick Mobilio, Gary Walker, and Robert Powers*

Department of Biological Chemistry, Wyeth Research, 85 Bolton Street, Cambridge, Massachusetts 02140

Keiko Tabei, Hui Tong, and Marshall M. Siegel*

Discovery Analytical Chemistry, Wyeth Research, 401 North Middletown Road, Pearl River, New York 10965

A protocol is described for rapidly screening small organic molecules for their ability to bind a target protein while obtaining structure-related information as part of a structure-based drug discovery and design program. The methodology takes advantage of and combines the inherent strengths of size exclusion gel chromatography, mass spectrometry, and NMR to identify bound complexes in a relatively universal high-throughput screening approach. Size exclusion gel chromatography in the spin column format provides the high-speed separation of a protein−ligand complex from free ligands. The spin column eluent is then analyzed under denaturing conditions by electrospray ionization mass spectrometry (MS) for the presence of small molecular weight compounds formerly bound to the protein. Hits identified by MS are then individually assayed by chemical shift perturbations in a 2D $^{1}H-^{15}N$ HSQC NMR spectrum to verify specific interactions of the compound with the protein and identification of the binding site on the protein. The utility of the MS/NMR assay is demonstrated with the use of the catalytic fragment of human fibroblast collagenase (MMP-1) as a target protein and the screening of a library consisting of ∼32 000 compounds for the identification of molecules that exhibit specific binding to the RGS4 protein.

A well-established approach for drug discovery is the utilization of a biological assay to screen a large library of compounds (>100 000) to identify initial leads that effect the activity of target protein(s) in the assay (for reviews, see refs 1−6). The resulting identification of lead compounds from this high-throughput screening (HTS) effort initiates an iterative approach to optimizing the activity of the small molecules from feedback obtained from structural and biological activity data. A major drawback of this method is the typical requirement that the biological assay be completely redesigned with the identification of each new protein target. This effectively requires a large commitment of resources and time before new drug discovery projects can be initiated. Besides the difficulty associated with the design of a biological assay to properly screen the chemical library for the desired activity, there exists a number of other limitations that may hinder the analysis and utility of the assay. These are usually a result of the necessary complexity of the assay to reasonably mimic the cellular function of the target protein and to monitor changes in its activity. It is not uncommon for a biological assay to contain multiple proteins, to be a membrane-based assay, or to even be a cell-based assay. The consequence of these complex assays is the ambiguous nature of a positive hit since the molecular interaction between a target protein and a small molecule is not readily correlated to an observed biological response. As a result, these assays greatly limit a structure-based approach to drug optimization while making it extremely difficult to decipher a structure−activity relationship (SAR) from the initial chemical leads.

NMR has been extensively used to evaluate ligand binding with an obvious utility in structure-based drug discovery and design.[7−10] The "SAR by NMR" method, previously described by Hajduk et al., illustrates the utility of NMR to screen small molecules for their ability to bind proteins from observed chemical shift perturbation in 2D $^{1}H-^{15}N$ HSQC spectra[11−13] and 2D $^{1}H-^{13}C$ HSQC spectra.[14] In addition to determining whether the small

* Corresponding authors: (R.P.) powersr@war.wyeth.com; (M.M.S.) siegelm@war.wyeth.com.

(1) Armstrong, J. W. *Am. Biotechnol. Lab.* **1999**, *17*, 26−28.
(2) Gonzalez, J. E.; Negulescu, P. A. *Curr. Opin. Biotechnol.* **1998**, *9*, 624−631.
(3) Oldenburg, K. R. *Annu. Rep. Med. Chem* **1998**, *33*, 301−311.
(4) Fernandes, P. B. *Curr. Opin. Chem. Biol.* **1998**, *2*, 597−603.
(5) Kenny, B. A.; Bushfield, M.; Parry-Smith, D. J.; Fogarty, S.; Treherne, J. M. *Prog. Drug Res* **1998**, *51*, 245−269.
(6) Silverman, L.; Campbell, R.; Broach, J. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 397−403.
(7) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.
(8) Otting, G. *Curr. Opin. Struct. Biol.* **1993**, *3*, 760−768.
(9) Whittle, P. J.; Blundell, T. L. *Annu. Rev. Biophys. Biomol. Struct* **1994**, *23*, 349−375.
(10) Blundell, T. L. *Nature* **1996**, *384* (Suppl.), 23−26.

10.1021/ac0006270 CCC: $20.00 © 2001 American Chemical Society
Published on Web 01/06/2001

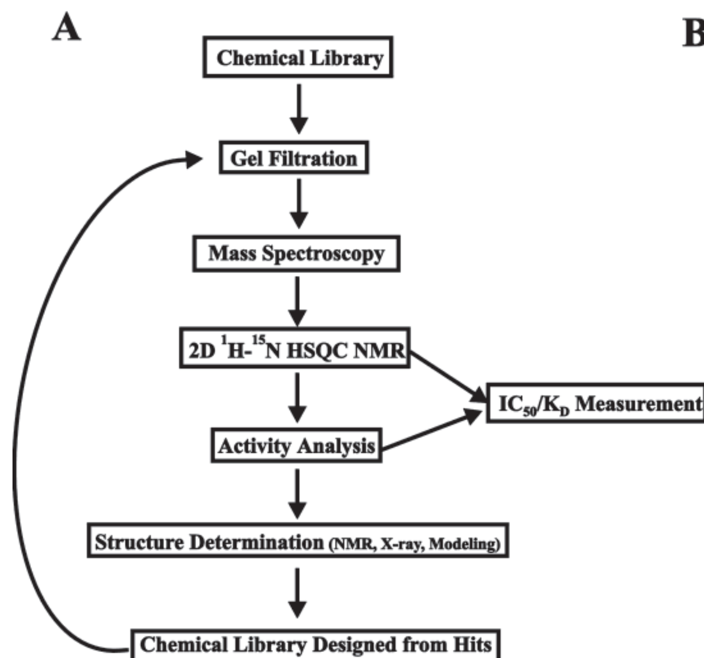
molecule binds the protein, the observed chemical shift perturbations also allow for the identification of the binding site on the protein surface. Nevertheless, the use of NMR as a primary screen has some significant obstacles that may limit its use in a high-throughput format. Mainly, the relatively low sensitivity of NMR requires significant quantities of isotope-enriched protein (>0.2 mM) and data acquisition time (>10 min) per sample, which drastically impacts the number of compounds that can be screened.[15,16] A response to these problems has been the utilization of mixtures, but this then requires deconvolution of the positive hits, which incurs a further commitment of sample supply and instrument resources. Furthermore, the utilization of mixtures may limit a compound's solubility below the concentration required by NMR while further complicating the necessity of maintaining consistent buffer conditions (pH, ionic strength) between samples. Additionally, the need to optimize the NMR data collection throughput usually results in a compromise between data quality and acquisition time.

Other attempts to minimize resource and sample requirements have focused on the application of 1D NMR techniques, particularly diffusion-edited measurements and transfer NOEs,[17−21] and the utilization of a Shapes compound library.[22] These 1D NMR experiments eliminate the need for labeled protein while minimizing sample quantities and data acquisition time. The Shapes library uses a very small set of molecular scaffolds (132) to represent a larger library where hits are used for virtual screening of the corporate compound collection. Again, the end result is to minimize both the sample requirement and experiment time. Unfortunately, these 1D NMR experiments do not provide information on the location of the binding site and the use of a small compound library reduces the chances of identifying an initial hit. Additionally, the utilization of mixtures is more difficult because of spectral overlap while also requiring a more complicated method for automated data analysis. Recently developed NMR cryoprobes and flow-through probes may provide some solutions to these issues since they may provide a 3−4-fold increase in sensitivity and a method for increased throughput, respectively.[17] Nevertheless, the real issue may be whether NMR is the best choice for the initial stage of the screening process since typical NMR experiments are time consuming and resource intensive. Given the observation that most assays have a hit rate on the order of 0.1−1%, which means that >99% of the data collected is negative information, it appears to be a more logical approach to eliminate a majority of the compounds before the NMR analysis stage.

A number of methods using only mass spectrometric detection have been proposed for the screening of drug candidates by evaluating noncovalent complexes between the ligand and a targeted protein. One general approach utilizes only electrospray mass spectrometry to directly produce ions of noncovalent complexes from a condensed-phase system under native conditions and then detects them in the gas phase.[23−27] The underlying assumption is that the gas-phase system mimics the condensed-phase system, which is "frozen" upon spraying into the gas phase. A second general approach utilizes an ancillary condensed-phase separation technique, such as spin column GPC,[28,29] microconcentration (centrifugation with ultrafiltration),[29−32] pulsed ultrafiltration,[33] dialysis, affinity chromatography,[34−37] frontal affinity chromatography,[38,39] in conjunction with detection by electrospray MS. The advantage of this approach is that the drug−protein complex is prepared by the ancillary technique so that detection of the drug by ESI-MS is all that is necessary to verify the formation of the drug−protein complex. The main disadvantage of these approaches for high-throughput screening is the inability of the mass spectrometric method to discriminate between specific and nonspecific binding of the drug to the targeted protein. Additionally, screening by MS does not provide any direct information regarding the binding site of the ligand or the structure of the protein−ligand complex, which are fundamental strengths of the "SAR by NMR" protocol.

Fundamentally, biological assays attempt to address two critical questions related to the potential utility of a small molecule: does the small molecule bind the protein target of interest and does this binding result in an observed perturbation in the activity of

(11) Hajduk, P. J.; Dinges, J.; Miknis, G. F.; Merlock, M.; Middleton, T.; Kempf, D. J.; Egan, D. A.; Walter, K. A.; Robins, T. S.; Shuker, S. B.; Holzman, T. F.; Fesik, S. W. *J. Med. Chem.* **1997**, *40*, 3144−3150.
(12) Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M., Jr.; Marcotte, P. A.; Severin, J.; Walter, K.; Smith, H.; Gubbins, E.; Simmer, R.; Holzman, T. F.; Morgan, D. W.; Davidsen, S. K.; Summers, J. B.; Fesik, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 5818−5827.
(13) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534.
(14) Hajduk, P. J.; Augeri, D. J.; Mack, J.; Mendoza, R.; Yang, J.; Betz, S. F.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122*, 7898−7904.
(15) Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663−5.
(16) Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S. J.; Sorensen, O. W.; Griesinger, C. *J. Biomol. NMR* **1994**, *4*, 301−306.
(17) Shapiro, M. J.; Wareing, J. R. *Curr. Opin. Drug Discovery Dev.* **1999**, *2*, 396−400.
(18) Meyer, B.; Weimar, T.; Peters, T. *Eur. J. Biochem.* **1997**, *246*, 705−709.
(19) Lin, M.; Shapiro, M. J.; Wareing, J. R. *J. Am. Chem. Soc.* **1997**, *119*, 5249−5250.
(20) Lin, M.; Shapiro, M. J.; Wareing, J. R. *J. Org. Chem.* **1997**, *62*, 8930−8931.
(21) Hajduk, P. J.; Olejniczak, E. T.; Fesik, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 12257−12261.
(22) Fejzo, J.; Lepre, C. A.; Peng, J. W.; Bemis, G. W.; Ajay; Murcko, M. A.; Moore, J. M. *Chem. Biol.* **1999**, *6*, 755−769.
(23) Smith, R. D.; Light-Wahl, K. J. *Biol. Mass Spectrom.* **1993**, *22*, 493−501.
(24) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1−23.
(25) Gao, J.; Cheng, X.; Chen, R.; Sigal, G. B.; Bruce, J. E.; Schwartz, B. L.; Hofstadler, S. A.; Anderson, G. A.; Smith, R. D.; Whitesides, G. M. *J. Med. Chem.* **1996**, *39*, 1949−1955.
(26) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10129−10133.
(27) Hofstadler, S. A.; Griffey, R. H. *Curr. Opin. Drug Discovery Dev.* **2000**, *3*, 423−431.
(28) Dunayevskiy, Y. M.; Lai, J.-J.; Quinn, C.; Talley, F.; Vouros, P. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1178−1184.
(29) Siegel, M. M.; Tabei, K.; Bebernitz, G. A.; Baum, E. Z. *J. Mass Spectrom.* **1998**, *33*, 264−273.
(30) Macht, M.; Fiedler, W.; Kuerzinger, K.; Przybylski, M. *Biochemistry* **1996**, *35*, 15633−15639.
(31) Linhares, M. C.; Kissinger, P. T. *Anal. Chem.* **1992**, *64*, 2831−2835.
(32) Wieboldt, R.; Zweigenbaum, J.; Henion, J. *Anal. Chem.* **1997**, *69*, 1683−1691.
(33) van Breemen, R. B.; Huang, C.-R.; Nikolic, D.; Woodbury, C. P.; Zhao, Y.-Z.; Venton, D. L. *Anal. Chem.* **1997**, *69*, 2159−2164.
(34) Nedved, M. L.; Habibi-Goudarzi, S.; Ganem, B.; Henion, J. D. *Anal. Chem.* **1996**, *68*, 4228−4236.
(35) Cai, J.; Henion, J. *Anal. Chem.* **1996**, *68*, 72−78.
(36) Papac, D. I.; Hoyes, J.; Tomer, K. B. *Anal. Chem.* **1994**, *66*, 2609−2613.
(37) Papac, D. I.; Hoyes, J.; Tomer, K. B. *Protein Sci.* **1994**, *3*, 1485−1492.
(38) Schriemer, D. C.; Hindsgaul, O. *Comb. Chem. High Throughput Screening* **1998**, *1*, 155−170.
(39) Schriemer, D. C.; Bundle, D. R.; Li, L.; Hindsgaul, O. *Angew. Chem., Int. Ed.* **1999**, *37*, 3383−3387.



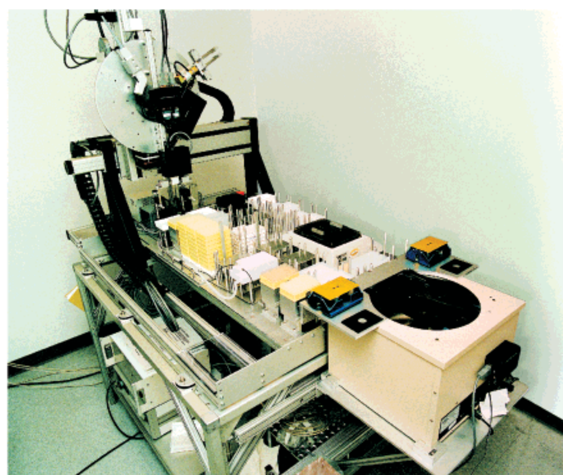


Figure 1. (A) Flow diagram of the MS/NMR structure-based assay. (B) The high-throughput screening robot used for the gel filtration component of the MS/NMR assay. Features of note are the centrifuge, shaker, sample reservoirs, and robot arm with 96-well pipettor.

the protein target? An assay protocol that attempts to answer these questions in separate stages of the procedure may result in a more efficient approach with a higher information content that is directly useable in a structure-driven drug discovery and design program. Additionally, the assay may provide a relatively universal HTS approach that is independent of the protein target. Toward this end, we report the development of the MS/NMR assay that takes advantage of and combines the inherent strengths of size exclusion gel chromatography, mass spectrometry, and NMR to identify bound complexes in a relatively universal high-throughput screening approach (Figure 1A). Size exclusion gel chromatography in the spin column format would provide the high-speed separation of a protein−ligand complex from free ligands, where either individual or compound mixtures are incubated with the protein target. The eluent from the size exclusion gel chromatography is then analyzed by electrospray ionization mass spectrometry for the presence of the low molecular weight compound. Hits identified by MS are then individually assayed by NMR where chemical shift perturbations in a 2D $^{1}H-^{15}N$ HSQC spectrum verify specific binding of the compound while identifying the small molecule binding site on the protein. The details of the MS/NMR assay are described below including applications with the screening for small molecules that specifically bind to the proteins MMP-1 and RGS4.

EXPERIMENTAL SECTION

Materials. Uniformly (>95%) ^{15}N-labeled human recombinant MMP-1 (18.7 kDa) and ^{15}N-labeled recombinant RGS4 (19.0 kDa) were expressed in *Escherichia coli* and purified as described previously.[40−42] All low molecular weight compounds used in the MS/NMR screen were synthesized in-house. The compounds were initially dissolved in DMSO from which mixtures were prepared and aliquots were removed for protein−ligand studies. The resins used in the size exclusion gel filtration chromatography were Sephadex G25 (Pharmacia) and polyacrylamide Bio-Gel P6 (BioRad) with a molecular weight fractionation range of 1000−5000 and 1000−6000, respectively. The gels were soaked and exhaustively washed (with 50 mM NH_4Ac, pH 7 for Sephadex G25, pH 4 for Bio Gel P6) to remove all soluble components which interfered with the mass spectral analysis. All buffers and solvents were purchased (Aldrich) and used as received.

MMP-1 Activity Assay. Studies on the effect of inhibitors on MMP-1 activity were performed using 10 mM Bis-Tris buffer with 5 mM $CaCl_2$, 100 mM NaCl, 0.1 mM $ZnCl_2$, and 2 mM NaN_3, at pH 7.0. Enzyme activity was assessed in a kinetic assay using nonlabeled recombinant MMP-1 and a peptide substrate.[43]

Compound Library for Screening. A key component of the MS/NMR assay is the appropriate design of the compound library such that the mixtures, which currently consist of 10 compounds, contain molecules that are not reactive and have unique molecular weights. The molecular mass of each compound in the mixture should differ by at least 3 Da to allow for clear identification of each component by the mass spectrometer. The molecular weight of the compound provides an effective means to identify "hits" from the MS/NMR assay while eliminating any requirement for deconvolution. The molecular weight effectively becomes an identification tag for each compound screened in the assay. Additional considerations in the selection of compounds for a screening library are solubility, structural diversity, and druglike characteristics.[44] It is also advantageous to design mixtures with

(40) Spurlino, J. C.; Smallwood, A. M.; Carlton, D. D.; Banks, T. M.; Vavra, K. J.; Johnson, J. S.; Cook, E. R.; Falvo, J.; Wahl, R. C.; Pulvino, T. A.; Wendoloski, J. J.; Smith, D. L. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 98−109.
(41) Moy, F. J.; Pisano, M. R.; Chanda, P. K.; Urbano, C.; Killar, L. M.; Sung, M.-L.; Powers, R. *J. Biomol. NMR* **1997**, *10*, 9−19.
(42) Moy, F.; Chanda, P. K.; Cockett, M. I.; Edris, W.; Jones, P. G.; Powers, R. *J. Biomol. NMR* **1999**, *15*, 339−340.
(43) Weingarten, H.; Feder, J. *Anal. Biochem.* **1985**, *147*, 437−440.
(44) Martin, E. J.; Critchlow, R. E.; Spellmeyer, D. C.; Rosenberg, S.; Spear, K. L.; Blaney, J. M. *Pharmacochem. Libr.* **1998**, *29*, 133−146.

a reasonable balance of acidic and basic molecules to avoid potentially drastic pH changes upon addition to the protein. The ~32 000 compounds used in the MS/NMR screen to identify binders to RGS4 corresponded to a subset of the corporate chemical library. The compounds were chosen based on these consideration with particular attention paid to minimizing reactivity and maintaining a unique molecular weight for each compound in a particular mixture.

The major advantage of using mixtures is to increase the throughput of the assay. The inherent limitation of the number of compounds used in a mixture is based on solubility issues and difficulties associated with finding appropriate groupings of compounds without reactivity issues and overlapping molecular weights. These were the primary factors in determining the size of the ~32 000 compound library. The choice of 10 compounds per mixture used in the examples described in the text does not necessarily represent an optimal number of compounds per mixture for the MS/NMR assay but clearly represents a viable approach.

Spin Column Size Exclusion Gel Filtration. The MMP-1 protein samples were prepared for the spin column size exclusion gel filtration experiments with MMP-1 at a concentration of 0.1 mM in a buffer consisting of 20 mM Tris, 100 mM NaCl, 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 2 mM NaN_3, and 3.5 mM DTT at pH 7.0 incubated with a single compound or mixtures consisting of 10 compounds at an approximate concentration of 1 mM for each compound. The RGS4 protein samples were screened against a ~32 000 compound library and were prepared for the spin column gel filtration size exclusion experiments with RGS4 at a concentration of 0.1 mM in a buffer consisting of 50 mM potassium phosphate, 2 mM NaN_3, and 50 mM DTT at pH 7.0 incubated with mixtures consisting of 10 compounds at an approximate concentration of 1 mM for each compound. The MMP-1 and RGS4 compound mixtures were incubated at room temperature for ~30 min. The final concentration of DMSO in the MMP-1 and RGS4 compound mixtures is 5%. The sample volume loaded on to and eluted from the gel consisted of ~10% of the gel volume. A total volume of 25 μL of the MMP 1-compound mixture is loaded on a Sephedex G25 column, consisting of ~300 μL of gel, in a Millipore multiscreen filtration system composed of a 0.65-μm hydrophilic Durapore filter. A total volume of 25 μL of the RGS4-compound mixture is loaded on a polyacrylamide Bio-Gel P6 column, consisting of ~300 μL of gel, in a Millipore multiscreen filtration system composed of a 0.65-μm hydrophilic Durapore filter. The samples were eluted using centrifugation (775*g* for a duration needed to elute 25 μL, corresponding to ~3 min total centrifugation time). Vacuum aspiration did not produce reliable spin column results with the Millipore multiscreen filtration system in an HTS format. Protein recovery through the spin columns has been found to be ~70%.

Sample Preparation Robot. A robot (Figure 1B) was designed to prepare samples for incubation, to perform gel filtration size exclusion chromatography in standard 96-well plate spin column format, and to collect the eluates in the 96-well plate format for delivery for mass spectral analysis. The robot simultaneously handles two 96-well plates through the entire sample preparation period. The total sample preparation time for the 33 plates used in the RGS4 screen was ~24 h. This time period included a 30-min incubation period for RGS4 with the compound mixture that occurred concurrently with the preparation of the gel filtration plates used in the subsequent chromatography step.

Mass Spectrometry. The gel filtration size exclusion chromatography eluates (25 μL) in the 96-well plates were each diluted with 25 μL water. A 5-μL aliquot from each sample was analyzed by mass spectrometry using automated ESI/MS methods in both positive and negative ionization modes with a Micromass LCT time-of-flight or a Quattro I triple quadrupole mass spectrometer each equipped with a Gilson 215 liquid handler and a Gilson 841 microinjector. A Hewlett-Packard 1100 HPLC system was used at a flow rate of 50 μL/min to introduce the samples into the mass spectrometer with a carrier solvent of 1:1 water−acetonitrile with 0.025% formic acid. The temperatures of the electrospray source, desolvation gas (N_2), and nebulizing gas (N_2) were maintained at 80 °C, 120 °C and ambient temperature, respectively. The nozzle−skimmer voltage was maintained at 20 V. Mass spectral data were acquired over the m/z range of 100−1200. Typically, mass spectra were acquired for about 0.75−1.0 min. The individual scans were combined, background subtracted, smoothed, baseline subtracted, and centroided using in-house MS processing software.[45] The total time for the analysis of each well was ~2.5 min, corresponding to ~4 h/96-well plate, resulting in a total analysis time for the 33 plates used in the RGS4 screen of ~6 days. Using in-house MS interpretation software which incorporates a smart background subtraction algorithm,[46] the processed spectra were automatically and efficiently interpreted for the presence of $[M + H]^+$, $[M + 2H]^{2+}$, and $[M + NH_4]^+$ ions consistent with a compound in the mixture. The program ranked the observed hits by a weighted signal-to-noise ratio scale and listed the results in an Excel spreadsheet.

NMR Spectroscopy. All NMR spectra were recorded on a Bruker DRX 600 spectrometer equipped with a triple-resonance gradient probe. For MMP-1, the gradient-enhanced 2D $^1H-^{15}N$ HSQC spectra were collected on a 0.3 mM ^{15}N-MMP-1 in a buffer consisting of 20 mM Tris, 100 mM NaCl, 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 2 mM NaN_3, and 3.5 mM DTT in 90% H_2O and 10% D_2O at pH 6.5 and 35 °C where the compound concentration ranged from 0.3 to 3.0 mM.[47] For RGS4, the gradient-enhanced 2D $^1H-^{15}N$ HSQC spectra were collected on 0.1 mM ^{15}N-RGS4 in a buffer consisting of 50 mM potassium phosphate, 2 mM NaN_3, and 50 mM DTT in 90% H_2O and 10% D_2O at pH 6.0 and 35 °C where the compound concentration was 1.0 mM. In all cases, 5% (v/v) DMSO was used to aid in compound solubility. In general, the 2D $^1H-^{15}N$ HSQC spectra were recorded with 256 complex points in t1, 2048 real points in t2, and 16 scans per increment. Spectra windows for t1 and t2 were 1723.7 and 8064.5 Hz, respectively, with the carrier at 4.75 and 115.2 ppm, respectively. Total acquisition time ranged from 15 min to 3 h per spectra. Data were processed and analyzed using NMRPipe, NMRWish,[48] and PIPP[49] on either a Sun Ultra 10 or SGI Octane workstation.

(45) Tong, H.; Bell, D.; Tabei, K.; Siegel, M. M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1174−1187.

(46) Tong, H.; Tabei, K.; Moy, F.; Powers, R.; Siegel, M. M. *Proceedings of the 48th ASMS Conference on Mass Spectrometry and Allied Topics*, Long Beach, CA, June 11−15, 2000; pp 935−936.

(47) Sklenar, V.; Piotto, M.; Leppik, R.; Saudek, V. *J. Magn. Reson., Ser. A* **1993**, *102*, 241−245.

(48) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.

Structure Determination of MMP-1−Inhibitor Complex. The determination of a solution structure of an inhibitor complexed to MMP-1 utilizing X-nucleus filtered multidimensional NMR experiments has been described previously in detail.[50] Briefly, the NMR chemical shift assignments and structural data in combination with the structure of the inhibitor-free MMP-1 is used as the starting point for the MMP-1−inhibitor complex. A minimal set of NMR experiments is used to assign the MMP-1 resonances that incur perturbations in the complex. The resonance assignments and bound conformation of the inhibitor in the complex are based on the 2D $^{12}C/^{12}C$-filtered NOESY,[51,52] 2D $^{12}C/^{12}C$-filtered TOCSY,[51,52] and $^{12}C/^{12}C$-filtered COSY experiments.[53] Intermolecular NOEs between MMP-1 and the inhibitor are obtained from the 3D ^{15}N-edited NOESY[54,55] and 3D ^{13}C-edited/^{12}C-filtered NOESY experiments.[56] The structure of the MMP-1−inhibitor complex is then determined by amending the restraints used for the refinement of the inhibitor-free MMP-1 NMR structure[57] with the distance restraints observed between MMP-1 and the inhibitor. The inhibitor-free MMP-1 NMR restraints are modified as appropriate for residues in the vicinity of the inhibitor by either removing restraints inconsistent with the complex structure and/or adding new restraints observed in the complex.

RESULTS AND DISCUSSION

General Overview. The large amount of negative data expected during the first iteration of any chemical library-based screen emphasizes the need for a high-throughput method that requires a minimal amount of sample and time. For a structure-based approach, the first step of the assay should clearly provide evidence of ligand binding while utilizing a relatively sensitive and universal technique. The combination of size exclusion gel filtration and mass spectrometry appears to meet these necessary requirements. In addition, the use of mixtures becomes a more attractive option since deconvolution of positive hits from the assay is unnecessary.

The inherent strength of NMR spectroscopy in the evaluation of potential protein ligands is the ability to directly observe specific binding of the ligand with the protein while identifying the ligand binding site from chemical shift perturbation maps. This capability is unique to NMR and provides a perfect compliment to the information obtained from the gel filtration mass spectrometry component of the assay. Furthermore, NMR may also be used to identify the structure of the protein−inhibitor complex. Because of the relatively large sample and time commitment of NMR, the optimal utilization of the NMR approach is dependent on the analysis of ligands that have been previously identified to bind with the protein. Mass spectrometry would minimize sample usage compared to NMR since detection limits are in the range of femtomoles while data acquisition times may be reduced by >5-fold compared to NMR.[29,32,33,58,59] This was a fundamental consideration for the incorporation of gel filtration and mass spectroscopy with NMR in the design of the MS/NMR assay to effectively screen a large library of compounds based on structural information (Figure 1A).

Gel Filtration Mass Spectroscopy. The first step of the MS/NMR assay consists of rapidly passing preincubated protein−compound mixtures through a Sephedex G25, polyacrylamide Bio-Gel P6, or comparable size exclusion column by centrifugation. The Sephedex G25 and polyacrylamide Bio-Gel P6 have a high retention for low molecular weight compounds while the relatively large molecular weight protein target will readily pass through the column. Therefore, the only low molecular weight species that will pass through the column are essentially those compounds that bind to the protein target. Similarly, DMSO, DTT, and other buffer components are retained in the excluded volume and do not pass through the column. The presence of protein in the sample injected into the mass spectrometer does not pose a problem in terms of false hits. The majority of the protein peaks is distributed generally over a wide number of masses of very low abundance and is found over a mass range generally greater than that of the low molecular weight compounds. Mass spectral analysis is performed on the filtrate in the molecular weight range for the compounds in the mixture to identify the positive hits. Since the molecular weights are known for each compound in the mixture as a prerequisite of the library design, the observation of a molecular ion in the mass spectrum simultaneously identifies the compound and the presence of a hit. The ready identification of a hit based on the compound's molecular weight effectively eliminates any need to deconvolute positive hits obtained from mixtures. In-house software has been developed to automate both the processing and the interpretation of mass spectral data and to determine the identity of hits from a database containing structure and molecular weight information for each of the mixtures in the library.[45,46]

The gel filtration and mass spectrometry components of the assay have been adapted to a 96-well plate format using the Millipore multiscreen filtration system. A robot has been constructed to automate the gel filtration component (Figure 1B) while a Gilson 215 liquid handler is used to transfer the filtrate from the 96-well plates to the mass spectrometer. Initial results indicate that we expect to achieve a further >10× reduction in the current sample requirements for the MS/NMR assay by adapting the protocol to a 384-well plate system and by further optimization of the mass spectrometry detection scheme. Additionally, only a small aliquot (~5 μL) of the filtrate is used for detection by mass spectrometry allowing for recycling of a majority of the protein sample for the remainder of the assay if necessary. Figure 2 illustrates the ESI mass spectra for a variety of MMP-1 inhibitors with IC_{50}'s ranging from 9 nM to 100 μM, which pass through the size exclusion spin column in the presence of MMP-1

(49) Garrett, D. S.; Powers, R.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson.* **1991**, *95*, 214−220.
(50) Moy, F. J.; Chanda, P., K.; Chen, J., M.; Cosmi, S.; Edris, W.; Skotnicki, J. S.; Wilhelm, J.; Powers, R. *Biochemistry* **1999**, *38*, 7085−7096.
(51) Petros, A. M.; Kawai, M.; Luly, J. R.; Fesik, S. W. *FEBS Lett.* **1992**, *308*, 309−314.
(52) Gemmecker, G.; Olejniczak, E. T.; Fesik, S. W. *J. Magn. Reson.* **1992**, *96*, 199−204.
(53) Ikura, M.; Bax, A. *J. Am. Chem. Soc.* **1992**, *114*, 2433−2440.
(54) Marion, D.; Driscoll, P. C.; Kay, L. E.; Wingfield, P. T.; Bax, A.; Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1989**, *28*, 6150−6156.
(55) Zuiderweg, E. R. P.; Fesik, S. W. *Biochemistry* **1989**, *28*, 2387−2391.
(56) Lee, W.; Revington, M. J.; Arrowsmith, C.; Kay, L. E. *FEBS Lett.* **1994**, *350*, 87−90.
(57) Moy, F. J.; Chanda, P. K.; Cosmi, S.; Pisano, M. R.; Urbano, C.; Wilhelm, J.; Powers, R. *Biochemistry* **1998**, *37*, 1495−1504.
(58) Davis, R. G.; Anderegg, R. J.; Blanchard, S. G. *Tetrahedon* **1999**, *55*, 11653−11667.
(59) Kaur, S.; McGuire, L.; Tang, D.; Dollinger, G.; Huebner, V. *J. Protein Chem.* **1997**, *16*, 505−511.

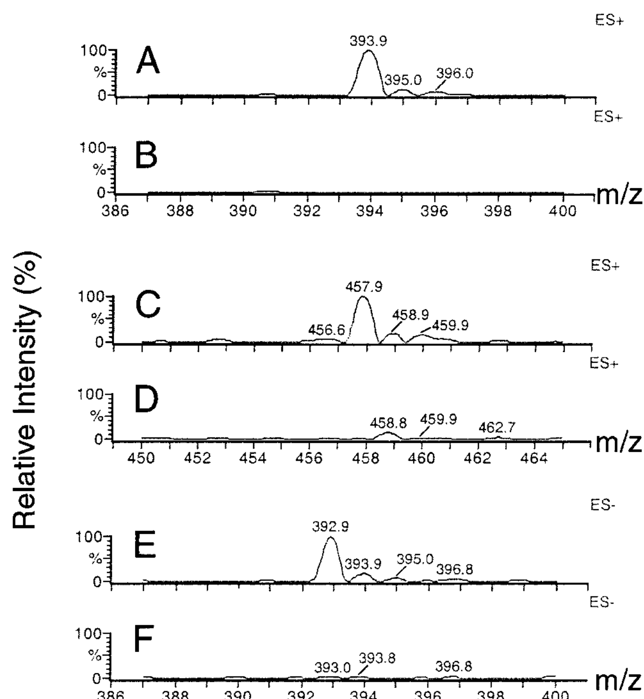


Figure 2. ESI mass spectral analysis of the filtrate after passing MMP-1 inhibitors through Sephadex G-25 columns in the presence and absence of MMP-1. (A) **1** (MW 393) and MMP-1, (B) **1** alone, (C) **2** (MW 457) and MMP-1, (D) **2** alone, (E) **3** (MW 394) and MMP-1, and (F) **3** alone. Same absolute intensity scale for (A) and (B), (C) and (D), (E) and (F), respectively. (A–D) in the ESI positive ionization mode; (E) and (F) in the ESI negative ionization mode.

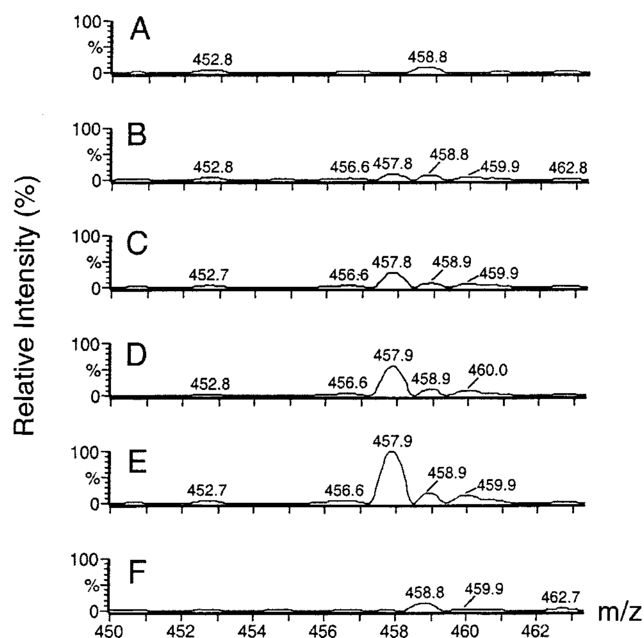


Figure 3. ESI (positive ionization mode) mass spectral analysis of the filtrate from the gel filtration titration of **2** (MW 457) with MMP-1 (A) MMP-1 alone at 50 μM and (F) **2** alone at 250 μM, respectively; (B–E) increasing amounts of MMP-1 (B) 20, (C) 30, (D) 40, and (E) 50 μM and increasing amounts of **2** from (B) 100, (C) 150, (D) 200, and (E) 250 μM. Same absolute intensity scale for (A–F).

(Table 1, compounds **1**–**3**). In all cases, when the compounds are passed through the size exclusion spin column in the absence of MMP-1, no ions are observed corresponding to the molecular weights of the small molecules. The molecular ions of the low molecular weight compounds are only observable in the filtrates in the presence of MMP-1. Further support of the validity of the procedure is seen by the response of a titration experiment with one of the weaker MMP-1 inhibitors (Figure 3, **2**). As expected, the increases in the relative intensities of the $[M + H]^+$ ion at m/z 457.9 correlate with the increases in the MMP-1 concentrations.

Table 1. Inhibitors of MMP-1

Compound Number	Structure	IC_{50} (nM)	MW (monoisotopic)
1		9	393
2		9900	457
3		89000	394
4		410	426
5		46	494
6		140	399
7		760	406
8		1000	450
9		17	471
10		3400	370
11		1100	448
12		7100	384
13		540	490

Mixture Analysis. The use of mixtures in lieu of the analysis of single compounds raises issues regarding mixture effects, i.e., suppression, competition, solubility, and synergistic effects, with the low molecular weight compounds and protein stability in the presence of the mixtures. To address these concerns, two different experiments were performed with MMP-1 and its inhibitors. Mixtures of 10 compounds were made containing only a single known inhibitor, and a mixture was made containing 10 known

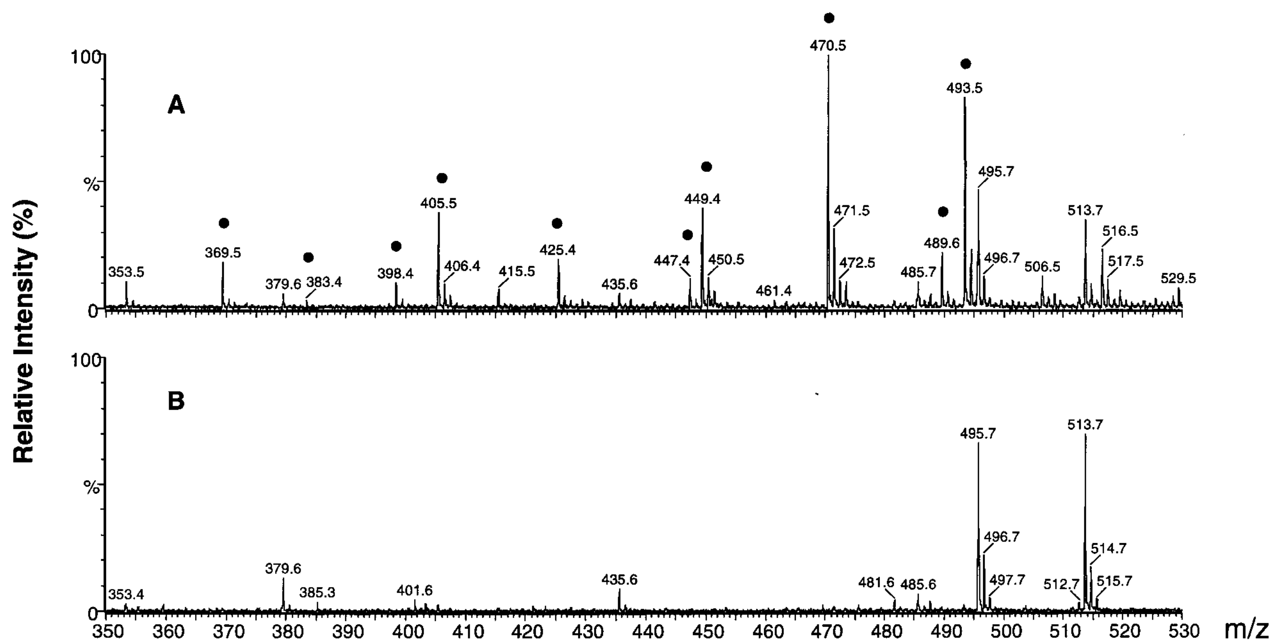


Figure 4. ESI (negative ionization mode) mass spectral analysis of the filtrate from the gel filtration analysis of a mixture containing 10 known MMP-1 inhibitors (A) with MMP-1 and (B) without MMP-1. The $[M - H]^-$ ions for the 10 compounds are indicated by solid circles (●) on the spectra. The mixture is composed of **4**−**13** listed in Table 1. Same absolute intensity scale for (A) and (B).

inhibitors with an IC_{50} range of 17 nM to 7100 nM. There was no significant effect on the response of the molecular ion of a single inhibitor (**1**−**3**) in a mixture of 10 compounds (data not shown). These results suggest that the presence of a compound in a mixture does not impede or impact on the ability of a binder to interact with the target protein. Of course, it is impossible to explore all possible combinations of mixtures and protein targets to exclusively rule out any detrimental effect on detecting an inhibitor in a mixture, but our limited analysis implies that any effect is probably minimal.

In the case of a mixture containing 10 known inhibitors **4**−**13** in the presence of MMP-1, all 10 compounds were detected by mass spectral analysis of the filtrate (Figure 4). The mixture of 10 hits is not a realistic representation of an expected result from a normal screen situation where <1% of the compounds would be hits. Nevertheless, this example provides valuable information regarding the effects of mixtures on the detection limits for the methodology. This particular examples illustrates that mixture effects may have a minimal contribution in the sensitivity of detecting positive hits since the gel filtration size exclusion chromatography does not exclusively select for the highest affinity binder in the mixture. The detection of a 7100 nM inhibitor in the presence of a 17 nM inhibitor may suggest that the off rate of the compound may be the most significant parameter contributing to the detection of the compound in the complex since the gel filtration size exclusion chromatography in the spin column format probably occurs under nonequilibrium conditions and favors dissociation of the protein−ligand complex.[60,61] If this is an accurate assessment of the results seen with the mixture of 10 known MMP-1 inhibitors, then this suggests that the off rates may be comparable for the different MMP-1 inhibitors despite the large range of IC_{50} values.

Robustness of the Gel Filtration Size Exclusion Chromatography. The success of the gel filtration size exclusion spin column methodology rests upon the selective retention of the low molecular weight compounds and the transmission of the protein−ligand complexes through the spin column. To verify that the detection of a compound in the eluent from the spin column is the result of a complex formed between the low molecular weight compounds and protein, the compound alone is analyzed in the absence of the protein using the gel filtration size exclusion spin column procedure. If the compound passes through the column and is detected in the eluent by mass spectral analysis, then the previous result obtained in the presence of the target protein is a false positive result. In large random libraries, very few materials were found to yield false positive results. Nevertheless, it was found that with Sephedex G25 resin material quaternary amines and compounds containing more than one carboxylic acid group tended to yield false positive results. To some extent, increasing the salt concentration in the sample buffer minimized the occurrences of these false positives. But an increase in the salt concentration may pose additional problems such as a potential effect on solubility of the compounds and affinity with the protein. The use of polyacrylamide Bio-Gel P6 in place of Sephedex G25 appears to have alleviated or at least minimized this problem without the introduction of different liabilities.

NMR Spectroscopy. The next step in the MS/NMR assay is the evaluation of the hits from the mass spectrometry analysis of the gel filtration filtrate by 2D $^1H-^{15}N$ HSQC NMR experiments. Since the NMR experiments are performed on identified hits, the data collection is conducted on single compounds. Additionally, since the number of compounds to be tested by NMR has been significantly reduced, greater care can be taken to maximize the quality of the data and greater attention can be applied in the analysis of the data. This implies that a weak binding compound

(60) Blom, K. F.; Larsen, B. S.; McEwen, C. N. *J. Comb. Chem.* **1999**, *1*, 82−90.
(61) Patapoff, T. W.; Mrsny, R. J.; Lee, W. A. *Anal. Biochem.* **1993**, *212*, 71−78.

that may induce a minimal number of modest chemical shift perturbations has a less likelihood of being missed and that false positives resulting from pH or buffer changes may be eliminated. An example of the 2D $^{1}H-^{15}N$ HSQC NMR spectra for the MMP-1 inhibitors exhibiting a positive response in the gel filtration mass spectrometry step of the MS/NMR assay (Figure 2) is shown in Figure 5A for **1**. Similar NMR spectra were observed for **2** and **3** that demonstrated a positive response in the gel filtration mass spectrometry experiment. All three compounds (**1**−**3**) induce chemical shift perturbations for residues in the vicinity of the catalytic Zn and the S1′ pocket in the MMP-1 active site. Particularly, residues 80−83, 114−119, and 136−142 exhibited the largest chemical shift changes in the presence of these inhibitors. The magnitude of the chemical shift perturbations and the number of residues exhibiting a chemical shift change are inversely related to the observed IC_{50} for each of the compounds; i.e., in general, a low IC_{50} correlates with a significant number of residues with a corresponding large chemical shift perturbation. The NMR experiment clearly corroborates the results from the gel filtration mass spectroscopy assay step while providing additional structural information concerning these protein−compound complexes.

In conjunction with previously determined NMR assignments and solution structure of MMP-1,[41,57] it is a straightforward procedure to map the amino acid residues exhibiting chemical shift perturbations onto the MMP-1 molecular surface to define the binding site of an inhibitor (Figure 5B). An observed clustering of amino acid residues in the same region of the protein surface also suggests a level of confidence that the inhibitor is binding specifically to the protein. An automated approach for the analysis of the library of collected HSQC spectra has been employed by using Tcl/Tk scripts written for the software program NMRWish[48] in combination with the GRASP[62] surface visualization software.

The NMR chemical shift perturbation analysis may be limited by the molecular weight of the protein to about 35 000−45 000,[63,64] but chemical shift assignments have been obtained on a 64 000 Trp repressor−DNA complex suggesting a potentially higher molecular weight upper limit.[65,66] NMR studies on "large" (>45 000) molecular weight protein−ligand complexes, where chemical shift assignments are not obtainable, may still provide direct evidence for binding and stoichiometry of low molecular weight molecules as well as the conformation of the bound ligand. Additionally, chemical shift perturbations in the absence of the resonance assignments may still be useful in clustering hits based on distinct perturbation patterns. Also, comparison of chemical shift perturbation patterns between a known ligand with a defined binding site and a new binder may provide information on the binding site of the new compound. This information would be invaluable for modeling the complex. Essentially, the NMR data are complimentary to the mass spectrometry results since the mass spectral analysis cannot distinguish between nonspecific and specific binders and does not provide detailed structural information about the complex.

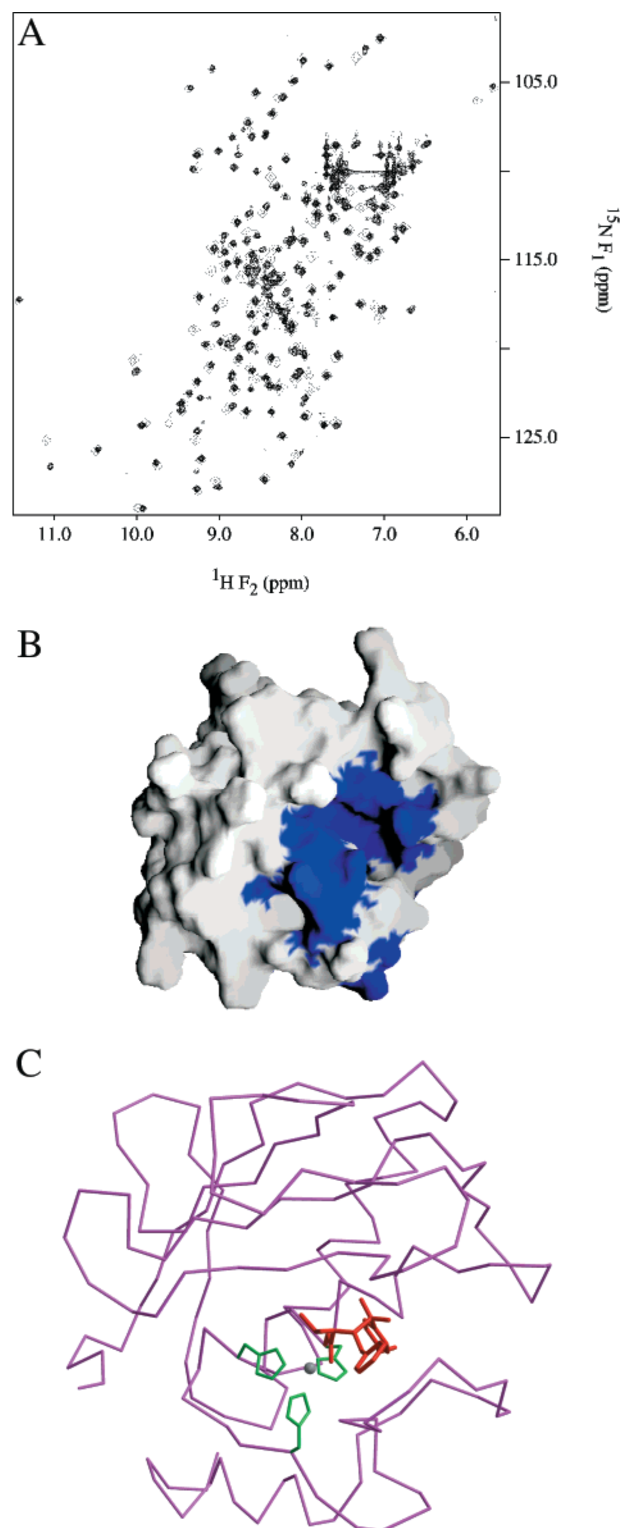


Figure 5. (A) 2D $^{1}H-^{15}N$ HSQC spectra of free MMP-1 (multiple contours) overlaid with MMP-1 complexed with **1**. (B) A GRASP[62] surface of the NMR solution structure of MMP-1 where residues that incurred a perturbation in the $^{1}H-^{15}N$ HSQC spectra in the MMP-1−**1** complex are blue, indicating the location of the ligand interaction with the protein. (C) NMR structure of the MMP-1−**1** complex. **1** is shown in red, the side chains for histidine are shown in green, the active site Zn is shown as a gray Van der Waals sphere, and the backbone atoms for MMP-1 are purple.

(62) Nicholls, A.; Sharp, K.; Honig, B. *Proteins: Struct., Func., Genet.* **1991**, *11*, 281ff.
(63) Gardner, K. H.; Kay, L. E. *Annu. Rev. Biophys. Biomol. Struct* **1998**, *27*, 357−406.
(64) Pervushin, K. V.; Wider, G.; Wuthrich, K. *J. Biomol. NMR* **1998**, *12*, 345−348.
(65) Shan, X.; Gardner, K. H.; Muhandiram, D. R.; Rao, N. S.; Arrowsmith, C. H.; Kay, L. E. *J. Am. Chem. Soc.* **1996**, *118*, 6570−6579.
(66) Shan, X.; Gardner, K. H.; Muhandiram, D. R.; Kay, L. E.; Arrowsmith, C. H. *J. Biomol. NMR* **1998**, *11*, 307−318.

Inhibitor Activity and Structure Determination. The confirmation by mass spectrometry and NMR that these compounds bind specifically to MMP-1 justifies obtaining biological data correlating the observed protein binding with a biological activity. Typically, IC_{50} values are obtained for each ligand from a biological assay that provides an initial ranking of the effectiveness of the chemical leads (Table 1). As a follow-up, K_D values can be obtained from NMR titration data or a variety of other analytical techniques.[8,67] Effectively, the high-throughput MS/NMR assay has inverted the typical steps involved in lead discovery. The MS/NMR method eliminates the need to convert a standard biological assay to a high-throughput format since the MS/NMR technique has become the primary screen. The small number of hits resulting from an MS/NMR screen could easily be handled by the standard biological assay without converting the assay to a high-throughput format.

After verifying that the compounds bind to the protein and effect the protein activity, the structure of the protein−ligand complex is elucidated by NMR, X-ray, and/or modeling. An example of a refined NMR structure of MMP-1 complexed with **1** is shown in Figure 5C.[50] Finally, the assay protocol is amenable to an iterative approach where a library of structural analogues, based on the initial hits, can be used to optimize further the affinity and activity of the ligand.

MS/NMR Screening of a 32 000 Compound Library with RGS4 Protein. Regulators of G-protein signaling (RGS) act as attenuators of the G-protein signal cascade by binding to the Gα subunit of G-proteins and inducing a 30-fold increase in the intrinsic Gα GTPase activity (for reviews, see refs 68−72). The regulation of the RGS4 protein suggests an adaptive response in the brain signal transduction pathway to compensate for desensitization and sensitization of G-protein-coupled receptor function, implying potential therapeutic roles in depression and epilepsy.[73] Previously, we presented the nearly complete NMR resonance assignments and high-resolution solution structure of free RGS4 as part of a structure-based drug design program.[42,74] The MS/NMR procedures described above for the robotic preparation of the gel filtration size exclusion spin column plates and analysis by MS and NMR have been applied for the identification of compounds that specifically bind to RGS4 using a ∼32 000 compound library. The goal of the MS/NMR screen is to identify compounds that bind RGS4, inhibit the interaction of RGS4 with Gα and prevent the induced GTPase activity. Furthermore, structure-based design, in conjunction with additional screens using directed libraries, will be employed to further develop the therapeutic utility of these inhibitors. To the best of our knowledge, and prior to our efforts, there have been no low molecular weight molecules identified that are active against RGS4. Mixtures of 10 compounds were analyzed for a total of ∼3200 mixtures, corresponding to 33 96-well plates in the MS/NMR assay. The top 50 hits identified by MS, based on a weighted signal-to-noise ratio score, were selected for further analysis by NMR. 2D $^{1}H-^{15}N$ HSQC spectra were collected for RGS4 in the presence of each of the top 50 compounds reported as MS hits. Analysis of the 2D $^{1}H-^{15}N$ HSQC spectra for chemical shift perturbations identified one compound that bound specifically to RGS4. A unique binding site on RGS4, which suggests a potential mechanism for the inhibition of the binding of RGS4 with Gα, was identified for this molecule from the chemical shift perturbations. Follow-up biological assays have demonstrated that this molecule inhibits RGS4 binding with Gα. Further development and analysis of the utility of the RGS4 inhibitor and the corresponding structural analogues are currently in progress.

A second compound was identified that effectively denatured RGS4 without precipitating the protein. The remaining compounds exhibited no effect on the RGS4 NMR spectra, indicating the lack of a specific interaction with the protein. Potential causes for a lack of an effect on the RGS4 NMR spectra, even though the compounds yielded a positive response in the MS analysis, may be nonspecific binding to RGS4 or flow-through of the gel filtration size exclusion column without binding RGS4. Follow-up analysis of the top 50 hits identified by MS determined that none of the individual compound passed through the gel filtration size exclusion column in the absence of RGS4, indicating that none of these compounds represent true false positives. These remaining compounds exhibit weak nonspecific binding to RGS4, which would not be expected to induce chemical shift perturbations in the RGS4 NMR spectra since NMR monitors an *average* structure. While these compounds are not interesting leads and provide no benefit for the drug development program, they cannot be accurately classified as false positives since they do effectively bind RGS4. These results also illustrate a unique feature of the MS/NMR assay. Since a compound is required to demonstrate direct binding to the protein target by both MS and NMR, the MS/NMR assay directly identifies and eliminates any false positives *before* further effort is expanded on the analysis of these compounds.

Comparison of MS/NMR Assay with Traditional Biological HTS. The results described above for the identification of MMP-1 and RGS4 inhibitors from the MS/NMR assay provide initial support for the utility of the MS/NMR assay as an alternative to traditional biological high-throughput screening approaches. Given these promising results for the MS/NMR assay, it is valuable to compare the benefits and potential liabilities relative to traditional biological high-throughput screens. A primary concern for any HTS assay is the efficiency of the screen. In our experience, the current throughput for a standard biological screen ranges from 1 to 6 months to screen approximately 230 000−260 000 compounds. For the RGS4 MS/NMR screen, ∼32 000 compounds were screened by gel filtration size exclusion chromatography and MS in ∼1 week. The NMR analysis of the resulting MS hits required ∼2.5 h/sample or an additional week for 50 compounds. There is a tradeoff in the NMR data collection approach between minimizing sample usage or experiment acquisition time. In the case of RGS4, it was decided that a protein concentration of 0.1 mM was judicious to minimize sample requirements resulting in a longer experiment time to acquire high-quality data. A 3-fold increase in sample concentration would reduce the experiment

(67) Ni, F. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 517−606.
(68) Dohlman, H. G.; Thorner, J. *J. Biol. Chem.* **1997**, *272*, 3871−3874.
(69) Kehrl, J. H. *Immunity* **1998**, *8*, 1−10.
(70) Arshavsky, V. Y.; Pugh, E. N., Jr. *Neuron* **1998**, *20*, 11−14.
(71) Zerangue, N.; Jan, L. Y. *Curr. Biol.* **1998**, *8*, R313−R316.
(72) De Vries, L.; Farquhar, M. G. *Trends Cell Biol.* **1999**, *9*, 138−144.
(73) Gold, S. J.; Ni, Y. G.; Dohlman, H. G.; Nestler, E. *J. Neurosci.* **1997**, *17*, 8024−8037.
(74) Moy, F. J.; Chanda, P. K.; Cockett, M. I.; Edris, W.; Jones, P. G.; Mason, K.; Semus, S.; Powers, R. *Biochemistry* **2000**, *39*, 7063−7073.

time to ~15 min/sample or ~1.5 days. The recent development of NMR cryoprobes may increase the speed of the NMR assay by a factor of from 4 to 16 times.[17] Likewise, with the advent of parallel acquisition and high-speed processing of electrospray spectra,[75,76] the throughput of the mass spectral side of the MS/NMR screening process could be improved by up to a factor of 8 times. The resulting throughput obtained for the MS/NMR assay is comparable to the upper range observed for traditional HTS assay. Adapting the MS/NMR assay to include recent NMR and MS technology developments will result in competitive throughput of the MS/NMR assay relative to traditional HTS assays.

Another consideration, in addition to throughput of the assay, is the preliminary effort required in implementing the screen. For a traditional biological HTS assay, this requires both conversion of an assay to an acceptable HTS format and a preliminary pilot run to verify the proper operation of the components of the robotic screen. Accurately predicting the time requirement to accomplish these goals is very difficult and will have a large range of possible outcomes (months to years) depending on the particular assay. The preliminary development of the HTS is eliminated with the MS/NMR assay since the only necessary modification is changing the target protein that is screened. The MS/NMR assay does require the availability of a protein structure and the corresponding NMR assignments to take full advantage of the analysis of the NMR component of the assay. It is important to note that this information is not required to actually execute the screen, only to take complete advantage of the analysis. Typically, the protein backbone resonance assignments required for interpretation of the chemical shift perturbations observed in the $^{1}H-^{15}N$ HSQC spectra may be obtained in 2−4 weeks. On the other hand, obtaining a quality protein structure by either NMR or X-ray is variable and may range from months to years. It is also plausible that the protein structure may already exist or a reasonable homology model could be readily obtained. Regardless, a structure-based drug design program is inherently dependent on the availability of the protein structure; therefore, the time commitment for obtaining the protein structure equally impacts both the utilities of the MS/NMR assay and a traditional HTS assay. Similarly, the availability of a supply of the protein of interest and an assay demonstrating the protein's activity are required prerequisites for any HTS effort.

The comparison of the efficiency between the MS/NMR assay and traditional approaches becomes more difficult when other factors are considered. Again, from our own experience, about 0.1−1% of the compounds in a traditional HTS are hits of which 60−80% are confirmed while in a complex assay only ~25% may be confirmed. Furthermore, it is very difficult to determine, in a general sense, the number of confirmed hits from traditional HTS that become viable leads and are developed into therapeutically active reagents. Similarly, it is difficult to infer any conclusions from our limited experience with the MS/NMR assay as to what percentage of hits will lead to biologically active compounds. However, the results with MMP-1 and RGS4 clearly establish that the MS/NMR assay does identify biologically active compounds. The inherent advantage of the MS/NMR assay compared to traditional HTS does infer a higher likelihood of identifying a compound with pertinent biological activity. A true hit from the MS/NMR assay identifies a compound that actually binds to the protein of interest. In general, this appears to be a rare occurrence for a traditional HTS assay since numerous alternative mechanisms can occur that result in a diminished activity that are unrelated and indistinguishable from a binding event with the target protein. Also, the nature of the MS/NMR assay directly eliminates any false positives since the compound has to demonstrate direct binding to the protein by both NMR and MS followed by exhibiting activity in the biological assay. The result is a focused chemistry effort on compounds that demonstrate the desired binding and biological activity.

Routine analysis by NMR of compounds that have been identified as confirmed hits in traditional HTS demonstrates a preponderance of problems including aggregate formation, low solubility, destabilization of the protein, nonstoichiometric binding, or no observable binding. Only a small number of compounds actually exhibit specific binding to the protein to justify further structural effort. Again, this number varies greatly depending on the nature of the HTS and the target protein. This point is illustrated by comparison of the MMP-1 and RGS4 structures. MMP-1 is an enzyme with a well-defined ligand binding pocket. Conversely, the activity of RGS4 is mediated through a protein−protein interaction where RGS4 presents a minimal binding pocket relative to MMP-1. Clearly, the MMP-1 structure is more amenable to the ready identification of a variety of ligands through an HTS assay, whereas finding compounds that bind RGS4 is a more formidable challenge. The end result is a higher expectation of false positives for RGS4 from a traditional HTS effort. Therefore, a higher success rate or efficiency with MMP-1 as the protein target compared to the lower results for RGS4 provides little utility in evaluating a HTS approach. Given the numerous complications outlined, it is extremely difficult to draw a definitive conclusion comparing the efficiency of the MS/NMR assay relative to traditional HTS. For the MS/NMR assay, the nature of the target protein and the compound library will dramatically impact the efficiency of the assay. This is also true for a standard HTS assay, where the methodology used in the screen may effect the efficiency of the assay to an even greater extent. Regardless, the real goal of either style assay is to identify lead compounds for a drug design program, which has been the case with traditional HTS approaches and which has now been demonstrated for the MS/NMR assay.

As stated previously, the exciting utility of the MS/NMR assay is its potential as a universal replacement for traditional biological assays as a primary screen to identify potential lead compounds. In addition to providing direct structural information of the binding interaction of compounds with the target protein, it is unnecessary to redesign the MS/NMR assay for each new protein target. This is typically the major obstacle for the development of biological assays for HTS. For the RGS4 project, the biological assays have not been readily amenable to a HTS approach and the MS/NMR assay has been providing invaluable information for this project. Toward this end, the use of the MS/NMR assay has identified a low molecular weight compound that has demonstrated direct binding to RGS4 and inhibition of the interaction of RGS4 with

(75) De Biasil, V.; Haskins, N.; Organ, A.; Bateman, R.; Giles, K.; Jarvis, S. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1165−1168.
(76) Bateman, R.; Jarvis, S.; Giles, K.; Organ, A.; De Biasi, V.; Haskins, N. *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics*, Dallas, TX, June 13−18, 1999; pp 1865−1866.

Gα. To the best of our knowledge, this represents the identification of the first small molecule inhibitor of RGS4.

CONCLUSION

The MS/NMR methodology appears readily adaptable to a wide range of protein targets and may prove to be a valuable approach for drug development. In addition, as described in the "SAR by NMR" protocol,[13] the MS/NMR approach is equally amenable to the concept of identifying and linking molecular substructures into a single lead compound.[77–80] The MS/NMR structure-based assay describes a novel protocol for drug discovery by linking the inherent strengths of a number of analytical techniques in a process that can be readily automated to screen a large library of chemical compounds while minimizing time and optimizing the use of resources. The MS/NMR assay has been successfully applied to identify a novel inhibitor of RGS4 and verify the interaction of known inhibitors to MMP-1.

Abbreviations: HTS, high-throughput screening; MS, mass spectroscopy; NMR, nuclear magnetic resonance spectroscopy; MMP-1, matrix metalloproteinase-1; RGS4, regulators of G-protein signaling-4; DTT, DL-1,4-dithiothreitol; 1D, one-dimensional; 2D, two-dimensional; HSQC, heteronuclear single-quantum coherence spectroscopy.

(77) Goodford, P. J. *J. Med. Chem.* **1984**, *27*, 557–64.
(78) Boehm, H. J. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 61–78.
(79) Boehm, H.-J. *Perspect. Drug Discovery Des.* **1995**, *3*, 21–33.
(80) Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, O. M. J.; Rayner, M. M. *Science* **1994**, *263*, 380–4.

ACKNOWLEDGMENT

The authors would like to thank Frank Deliglio for adaption of the NMRwish software to automate the analysis of NMR spectra for the presence of chemical shift perturbations; John Morin for enlightening discussions on HTS; Michelle Catino, Pranab K. Chanda, Scott Cosmi, Wade Edris, Laura Lin, Karl Malakian, Cheng Song, and Jim Wilhelm for providing the MMP-1 and RGS4 samples; Mei-Li Sung, Rebecca Cowling, and Loran M. Killar for providing the MMP-1 IC_{50} data. J. D. Albright, Efren G. Delos Santos, John W. Ellingboe, George T. Grosu, Yan Guo, Jeremy I. Levin, Jerauld S. Skotnicki, A. M. Venkatesan, and Arie Zask for chemical synthesis; Guang-Yi Xu for providing RGS4 NMR data; Kathleen Young and David Shuey for providing RGS4 activity data; and Gary Baisley, Duncan H. Bell, William F. Boulay, Leonard J. Candela, Barry A. Hausner, Jeanette M. Keating, Theodore J. Kornas, Frank W. Lambert, Miguel A. Maccio, Phil C. Pantliano, Marc Papi, Richard W. Scott, Mike S. Witek, Warren G. Yeisley, and Bart Zoltan for contributions in the development of the MS/NMR robot.

Received for review June 1, 2000. Accepted November 6, 2000.

AC0006270