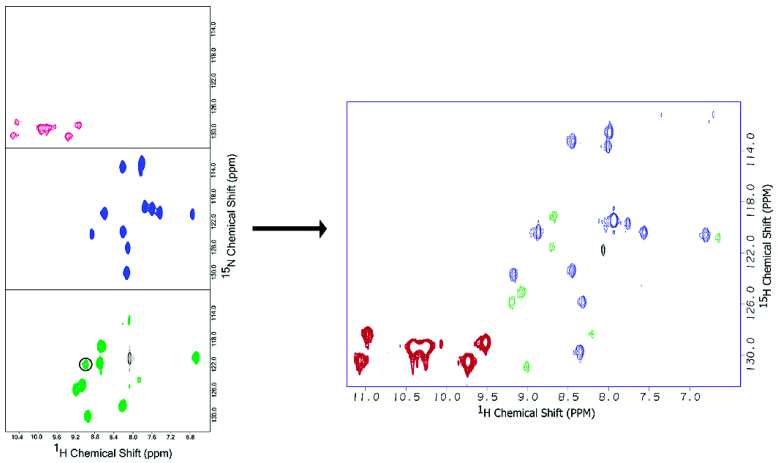
Article

RAMPED-UP NMR: Multiplexed NMR-Based Screening for Drug Discovery

Edward R. Zartler, Jeffrey Hanson, Bryan E. Jones, Allen D. Kline, George Martin, Huaping Mo, Michael J. Shapiro, Rong Wang, Haiping Wu, and Jiangli Yan

J. Am. Chem. Soc., **2003**, 125 (36), 10941-10946• DOI: 10.1021/ja0348593 • Publication Date (Web): 15 August 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 08/15/2003

RAMPED-UP NMR: Multiplexed NMR-Based Screening for Drug Discovery

Edward R. Zartler,[†] Jeffrey Hanson,[§] Bryan E. Jones,[‡] Allen D. Kline,[†] George Martin,[§] Huaping Mo,[†] Michael J. Shapiro,*[,†] Rong Wang,[‡] Haiping Wu,[§] and Jiangli Yan[†]

Contribution from the Discovery Chemistry Research and Technologies, Biology Research Technologies and Proteins, Lilly Research Labs, Lilly Corporate Center, Eli Lilly & Co., Indianapolis, Indiana 46285, and The Roche Protein Expression Group, Roche Diagnostics Corporation, 9115 Hague Road, Indianapolis, Indiana 46250

Received February 25, 2003; E-mail: shapiro_mike@lilly.com

Abstract: The crucial step in drug discovery is the identification of a lead compound from a vast chemical library by any number of screening techniques. NMR-based screening has the advantage of directly detecting binding of a compound to the target. The spectra resulting from these screens can also be very complex and difficult to analyze, making this an inefficient process. We present here a method, RAMPED-UP NMR, (**R**apid **A**nalysis and **M**ulti**p**lexing of **E**xperimentally **D**iscriminated **U**niquely Labeled **P**roteins using **NMR**) which generates simple spectra which are easy to interpret and allows several proteins to be screened simultaneously. In this method, the proteins to be screened are uniquely labeled with one amino acid type. There are several benefits derived from this unique labeling strategy: the spectra are greatly simplified, resonances that are most likely to be affected by binding are the only ones observed, and peaks that yield little or no information upon binding are eliminated, allowing the analysis of multiple proteins easily and simultaneously. We demonstrate the ability of three different proteins to be analyzed simultaneously for binding to two different ligands. This method will have significant impact in the use of NMR spectroscopy for both the lead generation and lead optimization phases of drug discovery by its ability to increase screening throughput and the ability to examine selectivity. To the best of our knowledge, this is the first time in any format that multiple proteins can be screened in one tube.

Introduction

A long-standing challenge in drug discovery is efficient delivery of lead compounds. Currently, drug discovery is performed primarily with high throughput functional screens. In this genomic era, the number of noncharacterized proteins is growing rapidly. Methods which do not rely on functional assays are becoming increasingly more important to this process. With the advent of fragment-based drug design, the identification of weak binding fragments is vital. NMR, one of many biophysical methods which detect binding,[1−3] easily detects weak binders (K_d < 10 mM) and is the only solution method that can also yield detailed structural data, indicating where on the protein the compound is binding.

NMR-based screening, whether ligand-observed or protein-observed, requires primary screens to determine the initial binding followed serially by experiments to verify and validate the initially determined binders. Validated binders are then screened in biophysical counterscreens and functional screens, if available, to determine selectivity to target and inhibition of antitargets. As practiced today, NMR screening is a serial process. Multiplexed screening can increase throughput; however, it is extremely difficult to multiplex functional screens. NMR screens are easily multiplexed: up to 100 compounds have been screened at once.[4] Protein-observed NMR screening using chemical shift perturbation, "SAR by NMR",[5] has shown its utility,[6−10] but cannot multiplex the proteins due to the inherent complexity of this experiment. We present here the

[†] Discovery Chemistry Research and Technologies, Lilly Research Labs.
[‡] Biology Research Technologies and Proteins, Lilly Research Labs.
[§] The Roche Protein Expression Group, Roche Diagnostics Corporation.

(1) Moy, F. J.; Haraki, K.; Mobilio, D.; Walker, G.; Powers, R.; Tabei, K.; Tong, H.; Siegel, M. M. *Anal. Chem.* **2001**, *73*, 571−581.
(2) Stockman, B. J.; Dalvit, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*, 187−231.
(3) Wong, C.-H.; Hendrix, M.; Manning, D. D.; Rosenbohm, C.; Greenberg, W. A. *J. Am. Chem. Soc.* **1998**, *120*, 8319−8327.
(4) Hajduk, P. J.; Gerfin, T.; Boehlen, J.-M.; Häberli, M.; Marek, D.; Fesik, S. W. *J. Med. Chem.* **1999**, *42*, 2315−2317.
(5) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534.
(6) 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.
(7) Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, J. G. M.; 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.
(8) Hajduk, P. J.; Zhou, M.-m.; Fesik, S. W. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2403−2406.
(9) Hajduk, P. J.; Boyd, S.; Nettesheim, D.; Nienaber, V.; Severin, J.; Smith, R.; Davidson, D.; Rockway, T.; Fesik, S. W. *J. Med. Chem.* **2000**, *43*, 3862−3866.
(10) Lesuisse, D.; Lange, G.; Deprez, P.; Benard, D.; Schoot, B.; Delettre, G.; Marquette, J.-P.; Broto, P.; Jean-Baptiste, V.; Bichet, P.; Sarubbi, E.; Mandine, E. *J. Med. Chem.* **2002**, *45*, 2379−2387.

basis for multiplexing protein-observed NMR screening with RAMPED-UP NMR.

Chemical Shift Perturbation

The output from chemical shift perturbation studies is a collection of $^1H-^{15}N$ correlation spectra. $^1H-^{15}N$ chemical shifts are environment sensitive and are reporters for ligand binding events. Ligand binding induced chemical shift perturbations are easily observed in amide resonances. Backbone amides in or near the site of ligand binding are particularly affected. Most of the $^1H-^{15}N$ resonances in a protein are the backbone amides, but side chains, such as Trp, Lys, His, Gln, Asn, or Arg, which are functionally important in a diverse range of proteins, also can be observed in $^1H-^{15}N$ spectra. Tryptophan side chains are especially good reporters due to their extreme downfield chemical shifts. The effect of binding on other side-chain resonances can vary widely. If the side-chain amide is not normally observed but upon ligand binding forms a hydrogen bond, it may become observable. Side-chain resonances which are directly involved in binding will be greatly perturbed by the binding event. Small or no perturbations are observed for those residues that are distant from the ligand binding site. If the binding is nonspecific or induces a conformational change, then the effect will be global, resulting in the shifting of a large majority of the resonances. This being said, the vast majority of peaks in a $^1H-^{15}N$ correlation spectrum are little affected upon ligand binding and add spectral complexity without providing substantive additional information. In the original "SAR by NMR" report, only 8 out of 107 resonances (~7.5%) were noted as showing significant perturbations upon compound binding.[5]

The complexity of $^1H-^{15}N$ correlation spectra increases quickly as the number of amino acids increases. For example, a single spectrum resulting from a moderate-size protein (30kDa) can have 300 $^1H-^{15}N$ correlation peaks. The advent of TROSY increased the size limit of protein NMR from 400 amino acids to more than 1000 amino acids for a deuterated protein[11] or 1000 observable resonances. For simple cases of small proteins, it is simple to detect peaks which have shifted. Detecting a subset of peaks, 30 of 500 for a large protein, which shift upon addition of compound can be difficult, especially if the spectrum is highly overlapped; doing this for tens or hundreds of spectra each with hundreds of peaks is onerous. A great deal of computational effort has been put forth to conquer this problem.[12] Reduced spectral complexity using methods such as SEA-TROSY[13] could speed up "SAR by NMR"; still only one protein can be screened at a time. Another disadvantage of TROSY-based methods is the requirement for protein deuteration (which can be costly or impossible to obtain) and high magnetic fields (≥800 MHz).

Reduced spectral complexity can also be achieved by isotopic labeling.[14] Specific labeling schemes utilizing ^{15}N amino acids, residue-specific ^{13}C-labeled methyl groups, ^{13}C-carbonyls, $^{13}C_\alpha$, or combinations of all of these have been developed. The quality of the proteins from all of these schemes can be improved by uniform deuteration of the background residues.[15] Recent reports use two labeled amino acids, one ^{15}N-labeled and one $^{13}C'$-labeled, generating $^{15}N-^{13}C'$ pairs which can be assigned by HNCO correlations experiments.[16−18] This allows assignment of some of the $^1H-^{15}N$ resonances and determination of the resonance perturbed by ligand. Typical labeling schemes use in vivo expression in *Escherichia coli* requiring often tedious and inefficient purification schemes. Scrambling of some amino acids (e.g., glycine and serine) is also frequent. Additionally, the incorporation of deuterium into the growth medium requires rigorous adaptation of the cells. Specialized media with a combination of uniform isotopic labels are commercially available, but are costly. These media cannot be used for specific labeling needs either. Last, a large percentage of pharmaceutically important proteins do not express well in *E. coli*, necessitating expression in yeast or insect cells. Isotopic labeling in insect cells or yeast is a nascent field and is not yet a reliable method for protein production.[19]

In vitro expression has many advantages over in vivo expression: expression levels are higher (>10 mg can be obtained from a 100 mL reaction overnight), labeling does not require changing growth conditions (especially important with regards to deuteration), there is no scrambling of labels, purification is more efficient (or even unnecessary[16]), and techniques to improve expression which affect viability of the cell are not an issue.[20−23] Many proteins which are problematic when expressed in *E. coli* are easily expressed using in vitro systems. For NMR purposes in vitro labeling can produce multimilligram quantities of protein with a limitless combination of labels more efficiently than in vivo expression. The quality of these proteins is higher than in vivo expressed protein and can be generated in a day. However, in vitro systems have their own problems. First, the protein must be able to be transcribed/translated efficiently by the *E. coli* machinery. Posttranslationally modified proteins will also not be easily expressed in vitro. With these caveats in mind, in vitro systems allow the production of a much greater array of uniquely labeled proteins than in vivo expressions systems.

Unique Labeling of Proteins

The essence of RAMPED-UP NMR is multiplexed screening of target and antitarget in the same tube. This method should not be confused with "SAR by NMR" times *N*. Although it could be used in that manner by extending the labeling to more amino acids per protein, it is really aimed at the question of selectivity and specificity. The advantage of RAMPED-UP NMR is that all proteins are screened simultaneously under identical conditions. To demonstrate this new methodology, we

(11) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1266−1271.
(12) Ross, A.; Schlotterbeck, G.; Klaus, W.; Senn, H. *J. Biol. NMR* **2000**, *16*, 139−146.
(13) Pellecchia, M.; Meininger, D.; Shen, A. L.; Jack, R.; Kasper, C. B.; Sem, D. S. *J. Am. Chem. Soc.* **2001**, *123*, 4633−4634.
(14) Marley, J.; Lu, M.; Bracken, C. *J. Biol. NMR* **2001**, *20*, 71−75.
(15) LeMaster, D. M. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 243−266.
(16) Weigelt, J.; van Dongen, M.; Uppenberg, J.; Schultz, J.; Wikstrom, M. *J. Am. Chem. Soc.* **2002**, *124*, 2446−2447.
(17) Guignard, L.; Ozawa, K.; Pursglove, S. E.; Otting, G.; Dixon, N. E. *FEBS Lett.* **2002**, *524*, 159−162.
(18) Yabuki, T.; Kigawa, T.; Dohmae, N.; Takio, K.; Terada, T.; Ito, Y.; Laue, E. D.; Cooper, J. A.; Kainosho, M.; Yokoyama, S. *J. Biol. NMR* **1998**, *11*, 295−306.
(19) Morgan, W. D.; Kragt, A.; Feeney, J. *J. Biol. NMR* **2000**, *17*, 337−347.
(20) Spirin, A. S.; Baranov, V. I.; Ryabova, L. A.; Ovodov, S. J.; Alakhov, J. B. *Science* **1988**, *242*, 1162−1164.
(21) Kigawa, T.; Muto, Y.; Yokoyama, S. *J. Biol. NMR* **1995**, *6*, 129−134.
(22) Kim, D.; Swartz, J. R. *Biotechnol. Bioeng.* **1999**, *66*, 180−188.
(23) Martin, G. A.; Kawaguchi, R.; Lam, Y.; DeGiovanni, A.; Fukushima, M.; Mutter, W. *BioTechniques* **2001**, *31*, 948−953.

chose three proteins: Target A, which has a known weak inhibitor,[24] Target B, and Target C, which have unknown responses to this inhibitor (Target A is phosopho tyrosine phosphatase 1b (PTP1b), Target B is k-RAS, and Target C is green fluorescent protein (GFP)). This experiment demonstrates target selectivity but could be used to determine target−antitarget selectivity, multiple target screening, or any combination of the above. A practical example of this experiment might be to study wild-type HIV-1 protease and mutants simultaneously. Two of the proteins studied have crystal structures: Target A (PDB:1C83[25,26]) and Target C (PDB:1EMA[25,27]). There is no crystal structure for Target B, but there is one for a highly homologous protein (PDB:1GNR[25]). Knowledge of the crystal structure guided the selection of appropriate amino acids for unique labeling, yet it was not required. For Target A, there is a Trp near the active site (Trp179). This fact and the distinctive chemical shift of the N^{ϵ} of tryptophan makes Target A an excellent choice for Trp labeling. Target B and C were labeled such that the label would provide a good sampling of the primary sequence, irrespective of the nature of the active site (see Supporting Information).

Target A was labeled using standard expression techniques in *E. coli* with ^{15}N-Trp$^{\epsilon}$. Target B and Target C were uniquely labeled using in vitro technology: Target B with ^{15}N-Ile and Target C with ^{15}N-Ala. For Target B, in the homologous protein h-Ras there are two isoleucines (21 and 24) on the helix that forms part of the binding site.[18,21] These two residues do not comprise part of the ligand binding site. These two isoleucines at a minimum could be expected to shift in response to ligand binding. The labels for Target B are not in the active site and thus will be distal reporters of the binding event. In Target C, the chromophore is formed by S65-Y66-G67, and the closest Ala to it is more than 19 amino acids away in primary sequence.

Methods

Protein Production. GFP and kRAS were provided by the Roche Protein Expression Group (RPEG), Roche Diagnostics Corporation, Indianapolis, IN. These two proteins were expressed and labeled according to a continuous exchange cell-free (CECF) protein expression protocol developed at RPEG. The proteins were purified by RPEG using a combination of standard and proprietary chromatography methods. A mixture of all 20 amino acids with either ^{15}N-amino acid (Ile or Ala) substitutions was added to the standard cell-free reaction mix to a final concentration of 2.1 mM. The cell-free protein expression reaction proceeded for 20 h at 30 °C. The solution was centrifuged at 20 000 rpm for 30 min at 4 °C. The supernatant was immediately purified using Ni^{2+}−chelate affinity chromatography (Qiagen). Purified GFP was then dialyzed into 20 mM HEPES, pH 7.5, 100 mM KCl, 1 mM $MgCl_2$, 24 mM CHAPS. k-RAS was dialyzed into the following: 20 mM HEPES, pH 7.5; 100 mM KCl; 1 mM $MgCl_2$; 24 mM CHAPS; 1 mM GDP. GFP was dialyzed into 20 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT-d_{10} (Cambridge Isotope Labs), and 10%−20% D_2O (Cambridge Isotope Labs).

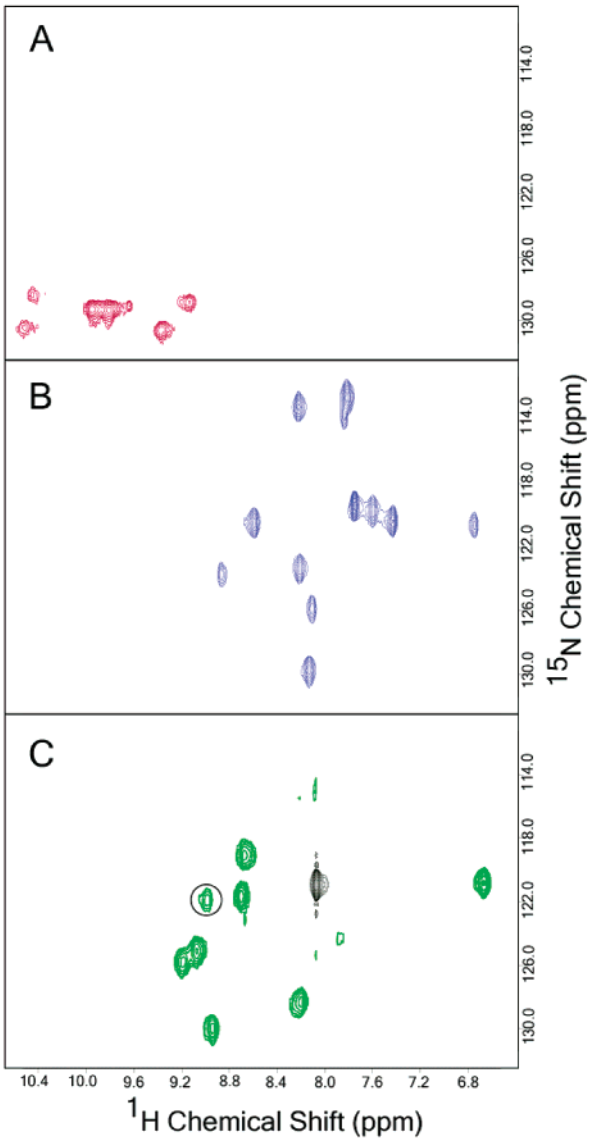


Figure 1. Spectra of uniquely labeled proteins. (A) ^{1}H−^{15}N HMQC spectrum of 300 μM Target A (PTP1b $^{15}N^{\epsilon}$-Trp_6). (B) ^{1}H−^{15}N HMQC spectrum of 250 μM Target B (k-RAS ^{15}N-Ile_{13}). There are only 12 apparent peaks. The peak at 7.91 ppm ^{1}H and 119.4 ppm ^{15}N integrates to two protons; thus, all 13 peaks are accounted for. (C) ^{1}H−^{15}N HMQC spectrum of 200 μM Target C (GFP ^{15}N-Ala_8). The circled resonance (8.99 ppm ^{1}H and 121.8 ppm ^{15}N does not appear in the spectra with all three proteins due to lower Target C concentration in the mixture (discussed in the text). The peak (black) at 8.06 ppm ^{1}H and 120.9 ppm ^{15}N is residual imidazole in a CHAPS micelle which is a product of the purification scheme.

NMR Spectroscopy. Spectra were acquired on a Bruker DRX system operating at 600.13 MHz (Bruker Biospin, Billerica, MA) at 25 °C with a TXI probe with triple-axis gradients. ^{1}H−^{15}N HMQC spectra with watergate solvent supression[28] were run with 128 scans, 2k points in the direct dimension covering 8400 Hz sweep width (113.9744 ms acquisition time) and 128 indirect points covering a 2400 Hz sweep width. A 5 Hz exponential was applied in the direct dimension and cosine function in the indirect dimension. No water deconvolution was performed. Reference was done relative to the carrier (4.78 ppm ^{1}H and 120 ppm ^{15}N). Processing and analysis were performed using Felix (Accelrys, San Diego, CA) and NMRPipe.[29]

(24) Iverson, L. F.; Anderson, H. S.; Moller, K. B.; Olsen, O. H.; Peters, G. H.; Branner, S.; Mortensen, S. B.; Hansen, T. K.; Lau, J.; Ge, Y.; Holsworth, D. D.; Newman, M. J.; Moller, N. P. H. *Biochemistry* **2001**, *40*, 14812−14820.
(25) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(26) Barford, D.; Keller, J. C.; Flint, A. J.; Tonks, N. K. *J. Mol. Biol.* **1994**, *239*, 726−730.
(27) Ormo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392−1395.
(28) Sklenar, V.; Piotto, M.; Leppik, R.; Saudek, V. *J. Magn. Reson., Ser. A* **1993**, *102*, 241−245.
(29) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biol. NMR* **1995**, *6*, 277−293.

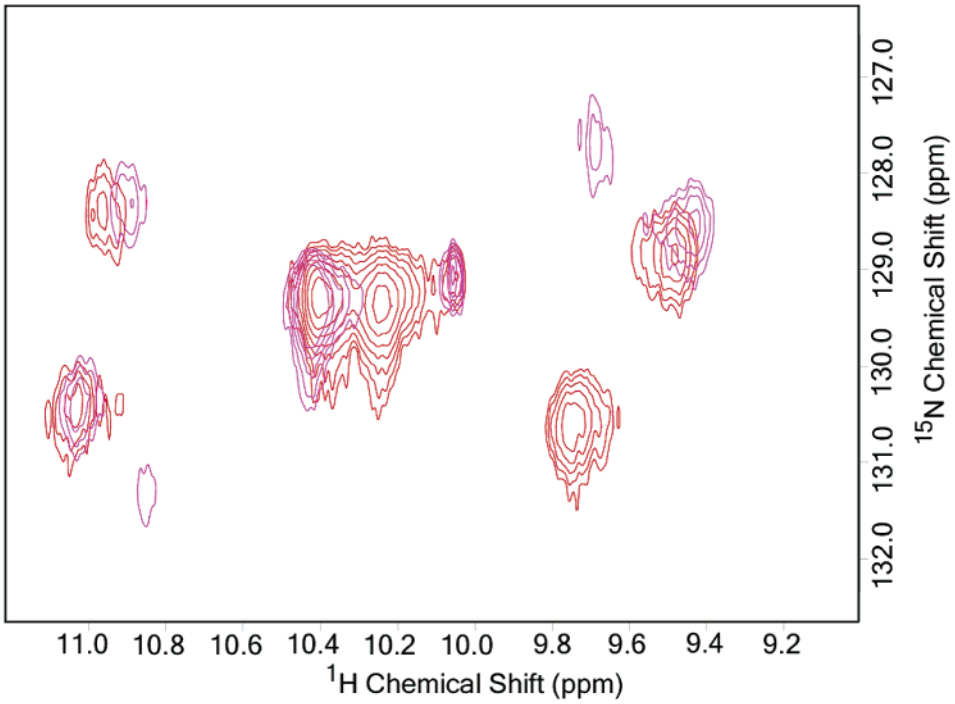


Figure 2. Chemical shift perturbations in Target A ($^{15}N^{\epsilon}$-Trp$_6$) caused by **1**. The spectrum of ~100 μM Target A ($^{15}N^{\epsilon}$-Trp$_6$) plus 0.5 mM **1** was overlaid by the spectrum of ~300 μM Target A ($^{15}N^{\epsilon}$-Trp$_6$) without **1**. The red peaks are Target A in the absence of inhibitor. The purple peaks are Target A in the presence of inhibitor. Experimental conditions were identical for both spectra.

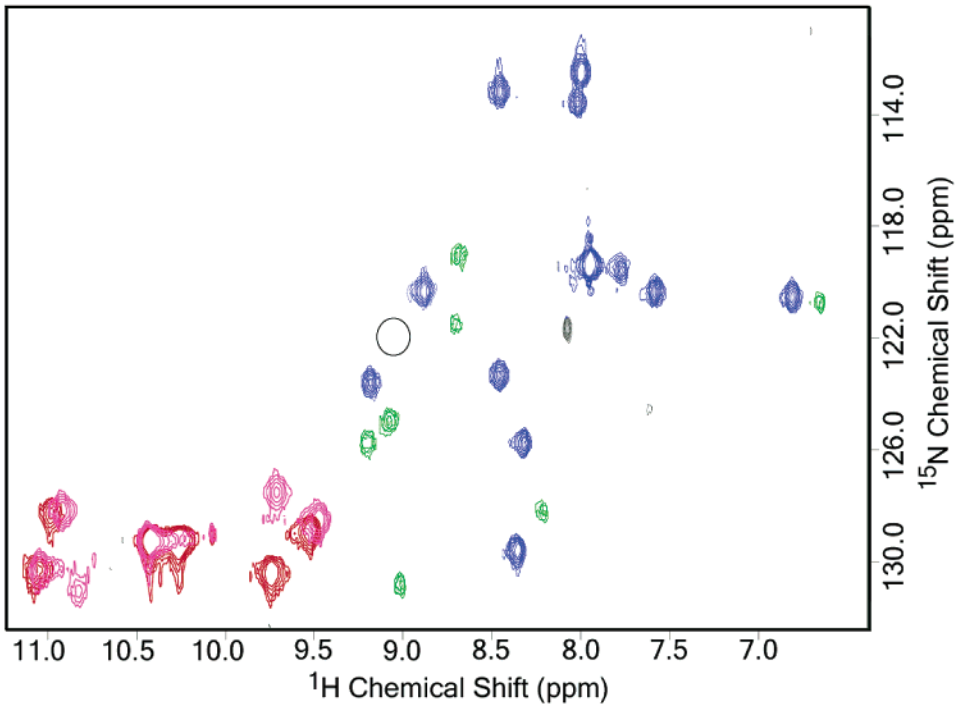


Figure 3. Chemical shift perturbations in Target A ($^{15}N^{\epsilon}$-Trp$_6$) in the presence of Target B and Target C caused by **1**. The color scheme is the same as used in Figures 1 and 2. The circle indicates the position of the Target C resonance that is below the level of noise.

Multiplexed NMR Screening

^{1}H−^{15}N HMQC spectra were taken for each protein separately. (Figure 1A−C). Serial additions of each protein were then performed and spectra acquired. No change in the spectrum of any protein as the other proteins were added was observed. The sequential acquisition of spectra is an important aspect of this method. Analysis of each spectra indicates whether there are protein−protein interactions (which in this case do not exist). One peak (circled in Figure 1C) in Target C is below the threshold in the spectra with all three proteins. Target C is known to aggregate and precipitate at high concentrations.[30] This peak has the lowest signal-to-noise in the Target C sample (Figure 1C). As that peak is monitored through the series of spectra, it is apparent that Target C is precipitating and thus

(30) Seifert, M. H. J.; Ksiazek, D.; Azim, M. K.; Smialowski, P.; Budisa, N.; Holak, T. A. *J. Am. Chem. Soc.* **2002**, *124*, 7932−7942.

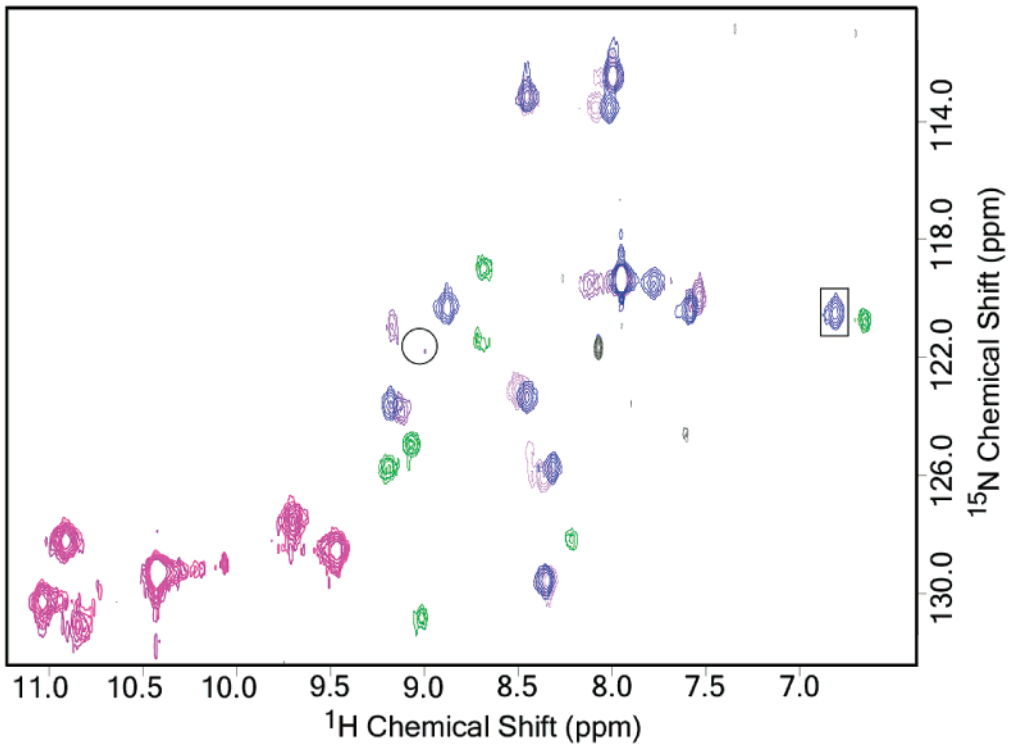


Figure 4. Chemical shift perturbations in Target B in the presence of Target C and **1**-inhibited Target A. The light purple peaks indicate the Target B peaks in the presence of **2**. There are only 11 observable resonances (12 if the two overlapped peaks are accounted for) in Target B. The peak that is lost is marked with a box in Figure 4. The Target A and Target C peaks are unaffected. The boxed resonance at 6.78 ppm ^{1}H and 120.6 ppm ^{15}N is completely eliminated in the presence of the ligand. No attempt at assignments was made.

resulting in a lower Target C concentration. This was confirmed by observation of precipitate in the NMR tube after acquisition. However, the chemical shifts of the visible Target C resonances are identical to the chemical shift of Target C alone, showing that no conformational change has occurred. The final spectrum of all three proteins is identical to a superposition of the spectra of all three alone (data not shown).

Compound **1** inhibits Target A with an IC_{50} of 20 μM.[24] Addition of **1** to Target A results in chemical shift perturbations which give a clearly different pattern than the uninhibited protein (Figure 2). Addition of **1** to the solution of all three proteins results in identical perturbations in the Target A resonances (magenta peaks) compared to Target A alone (red peaks) (Figure 3). There is no perturbation in the resonances corresponding to Target B or Target C. Addition of **2**, a non-hydrolyzable substrate analogue of Target B (γ-S-GTP), to the sample with all three proteins and **1** results in perturbation of Target B peaks only (purple peaks compared to blue peaks) (Figure 4). There are no perturbations in either the Target A peaks or the Target C peaks upon addition of **2**. We had no a priori knowledge whether Target A would respond to **2**, if Target B would respond to **1**, or if Target C would respond to either compound. These data show the ability to quickly and cleanly discriminate binding of weak compounds in solution. The lack of perturbation of the Target C resonances seems to show that neither compound binds to it. It may be that either or both compounds are binding to Target C, but in a nonspecific way.

Our unique labeling while reducing spectral complexity also eliminates signals which may lead to spurious conclusions caused by uniform labeling. In the absence of structure, however, it is still possible to get false results. Here we demonstrate the simplest case, one amino acid type label, but additional labeled amino acids (or the addition of other isotopes) can be applied to affect more of the available "resonance space" without a big increase in spectral complexity. The variety of labels that can be utilized is limited solely by their availability. The selection of amino acids for labeling is very important. If possible, the labeled amino acid should be in the active site of the target protein; if the structure is not known, bioinformatics tools or homology modeling can be used to approximate the active site and determine which amino acids might be in the active site. There is a tradeoff of spectral simplicity for complete coverage of the protein. If one chose to use this method as simply a multiplexed "SAR by NMR", it could be envisioned that different multiple amino acid labeling schemes could be used. Proteins that are highly homologous can be labeled with completely different amino acids and are readily distinguished. Allosteric changes can also be detected by labeling outside the active site, as was seen in the k-Ras spectra. Compounds that merely broaden resonances, as opposed to shifting them, will most likely not be detected either. In RAMPED-UP, an important point in the selection of amino acid to label is to maximize the likelihood of detecting a binding event, not to guarantee that every binding event is detected.

One technical problem that we found very difficult to overcome is the requirement that all proteins be in identical buffers. For three unrelated proteins (as in this paper) this can be limiting. However, for three closely related proteins such as in the HIV case, this requirement should be more easily realized. Protein−protein interactions can be confounding, but the serial addition of proteins and acquisition of spectra help identify these interactions. We believe that if no shifts or broadening is observed in the spectra (albeit we only a have selective view of the proteins by the very nature of our experiment if we label

only a single amino acid type), then this interaction can be ignored. However, the utility of this method in screening for compounds that block protein−protein interactions becomes readily apparent.

Conclusions

By uniquely labeling these proteins, in essence NMR-tagging each one, we have reduced the complexity of 729 $^{1}H-^{15}N$ resonances down to 27 easily resolved resonances, ∼4% of the total. In the case of Target A, we observed 2 of 6 peaks undergo significant shifts. For Target B we observed 3 of 13 peaks shift significantly. While this percentage is much larger than that seen in the original case (∼20−30% vs 7.5%), the total number of shifted peaks is still smaller (5 vs 8) for seven times the total number of amino acids (757 vs 107). Although we only observed less than 1% of the total number of amino acids in solution responding to ligand, it is a very high percentage of the total observed peaks (∼20%). By selecting amino acids more likely to respond to ligand binding we have simplified the spectra and made any chemical shift perturbations more meaningful. The ability to analyze quickly a small set of peaks for chemical shift perturbations makes this method extremely useful. This method should not be confused with "SAR by NMR" times *N* (although it could be used in that manner by extending the labeling to more amino acids per protein); it is really aimed at the question of selectivity and specificity. We can discriminate binding of a weak inhibitor to its target from other potential targets rapidly. Our NMR-tagged proteins give us the powerful ability to study multiple proteins simultaneously and open new vistas for NMR in drug discovery.

Acknowledgment. We thank R. Higgs and M. Bures for fruitful discussions about this technology, A. Kaerner for discussions about this paper, G. Hite for making the PTP1b inhibitor, J. Munroe for helpful comments, and J. Moyers. E.R.Z. thanks R. T. Pickard for the γ-S-GTP and S. Comella and everyone else at RPEG for taking a leap of faith and supporting this research.

Supporting Information Available: The three primary sequences of the proteins used in this study with the presence of labels indicated by a colored letter (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0348593