J|A|C|S
COMMUNICATIONS

Published on Web 02/26/2002

Site-Selective Screening by NMR Spectroscopy with Labeled Amino Acid Pairs

Johan Weigelt,* Maria van Dongen, Jonas Uppenberg, Johan Schultz, and Mats Wikström

Department of Structural Chemistry, Biovitrum, S-112 76 Stockholm, Sweden

Received December 20, 2001

In recent years, NMR spectroscopy has become an important tool in the drug discovery process through the advent of NMR based screening to identify lead templates.[1,2] Perhaps the most well-known method is the "SAR-by-NMR" scheme described by Fesik and coworkers in 1996.[1] The SAR-by-NMR technique relies on detecting chemical shift changes in a 2D $^{1}H-^{15}N$ correlation spectrum to identify compounds that bind to the target protein. Provided the three-dimensional structure of the protein is known, and sequence specific NMR assignments of the protein backbone resonances have been obtained, ligands with nonoverlapping binding sites may be linked based on structural data, resulting in high affinity binders even if the starting compounds only had millimolar to micromolar affinities. However, the SAR-by-NMR method is restricted to relatively small proteins (M_w < 35 kDa) since it requires that backbone resonance assignments of the target protein have been obtained, a task that generally demands months of experimental work and analysis even for a small protein. The method may be extended to slightly larger systems by the introduction of ^{13}C-labeled methyl groups and detecting changes in a 2D $^{1}H-^{13}C$ correlation spectrum, but the need for sequence specific assignments remains.[3]

Here we present a technique that facilitates screening for binding to a selected site. The core of the new method is an indirect sequence specific labeling scheme.[4] Amino acid X is labeled with ^{13}C and amino acid Y is labeled with ^{15}N. Provided only one XY pair occurs in the amino acid sequence, only one signal in the 1D carbonyl ^{13}C spectrum will display a splitting due to the $^{1}J_{C'N}$ coupling,[4] and only one peak will appear in an HNCO[5] type correlation spectrum. The occurrence of a unique pair of labeled amino acid residues confers the sequence specificity. It is possible to identify a relevant unique amino acid pair for several known drug targets.[6] This labeling strategy thus makes it possible to screen for binding to a selected site without the need for sequence specific assignments. An HNCO spectrum can be used either directly as a screening experiment (1D or 2D versions) or indirectly to identify what signals to monitor in a 2D $^{1}H-^{15}N$ correlation spectrum. Chemical shift perturbations upon addition of a potential ligand are easily detected even for large proteins due to the reduced spectral complexity resulting from the use of a selectively labeled sample. A 2D $^{1}H-^{15}N$ correlation experiment is most likely the method of choice because of the inherent lower sensitivity of the HNCO experiment.

The principle of the site-selective screening method is demonstrated for the human adipocyte fatty acid binding protein FABP-4. FABP-4 has been implicated in insulin resistance[7] and is thus a potential drug target. Selectivity of FABP-4 binding compounds is important. In particular, binding to FABP-3, which is expressed in heart and muscle, should be avoided. Analysis of the fatty acid binding sites of FABP-3 and FABP-4 shows that residues valine 114 and valine 115 in FABP-4 are replaced by isoleucine and leucine in FABP-3.[8] By screening for binders to this epitope, potential FABP-4 selective compounds may be identified. Valine 114 and valine 115 comprise a unique amino acid residue pair in the amino acid sequence of FABP-4, and are thus suitable for selective labeling. A $^{13}C/^{15}N$-Val selectively labeled sample of FABP-4 was produced from prototrophic *Escherichia coli* cells harboring a plasmid containing the gene coding for a 146 amino acid protein construct of FABP-4. The construct encompassed a C-terminal *his-tag* included for purification purposes. Cells were grown in a growth medium containing all amino acids and nucleotides according to the protocol of Muchmore et al.[9] Valine was supplied uniformly $^{13}C_5/^{15}N$-labeled (Cambridge Isotope Laboratories). Following purification the protein mass was determined with use of MALDI-TOF mass spectrometry. The extent of incorporation of $^{13}C/^{15}N$ in valine residues was estimated to be at least 95%. The mass spectrum displayed a very homogeneous FABP-4 peak, indicating no or very little isotopic scrambling despite the use of a prototrophic bacterial strain.[10] This finding was further corroborated by the very clean $^{1}H-^{15}N$ correlation spectrum of $^{13}C/^{15}N$-Val FABP-4 (Figure 1a). An attractive alternative approach to produce a selectively labeled sample is cell-free synthesis,[11] which eliminates the risk of isotopic scrambling and has the additional benefit of lower consumption of labeled amino acids per milligram of produced protein.

Figure 1 shows NMR spectra of selectively labeled FABP-4. The signals visible in a $^{1}H-^{15}N$ correlation experiment (Figure 1a) correspond perfectly to the valine resonances of FABP-4 as assigned by Constantine et al.[14] The 2D HNCO experiment yields cross-peaks only for $^{13}C'-^{15}N-^{1}H$ moieties. There is only one such instance in the selectively labeled FABP-4 protein, namely ($^{13}C/^{15}N$)Val114-($^{13}C/^{15}N$)Val115. The HNCO spectrum in Figure 1b contains only one cross-peak, which in agreement with the published resonance assignment[14] identifies the valine 115 resonance. The spectrum displayed in Figure 1c demonstrates the feasibility of the use of a 1D HNCO experiment for detection.

Five test compounds were mixed into a compound cocktail. One compound (compound **5**) had been identified as an FABP-4 binder in an NMR line broadening assay (data not shown).

Figure 2 shows spectral expansions of spectra recorded on selectively labeled FABP-4 in different mixtures. The valine 115 cross-peak experiences a chemical shift change upon addition of the test cocktail (Figure 2b). Mixing FABP-4 with compound **5** alone yields a spectrum similar (Figure 2d) to that observed when FABP-4 is mixed with the cocktail. Omitting compound **5** from the test cocktail leaves the spectrum unchanged (Figure 2c). The crystal structure of FABP-4 in complex with compound **5** shows

* To whom correspondence should be addressed. E-mail: johan.weigelt@biovitrum.com.

10.1021/ja0178261 CCC: $22.00

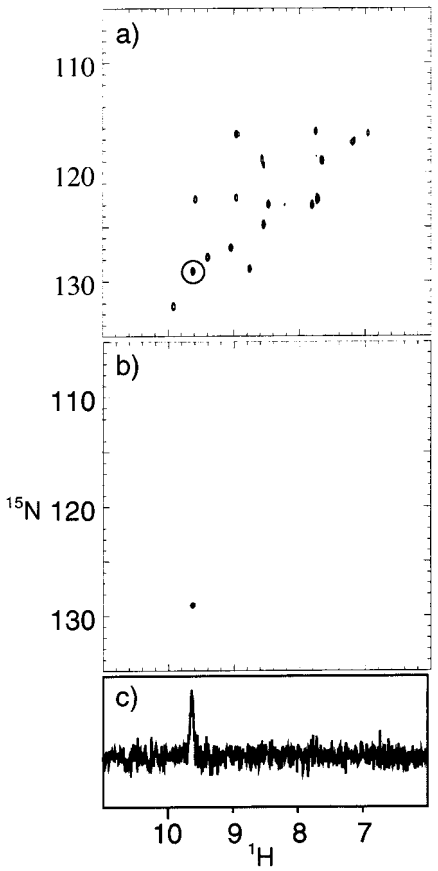

Figure 1. NMR spectra recorded on an 0.2 mM $^{13}C/^{15}N$-Val FABP-4 sample. The spectra were recorded at 20 °C on a 600 MHz Varian Unity INOVA NMR spectrometer with standard techniques. The sample was prepared in 20 mM sodium phosphate buffer (10% D_2O/90% H_2O), pH 7.6, containing 50 mM NaCl, 5 mM DTT, 1 mM EDTA, and 0.005% (w/v) NaN_3. nmrPipe[12] was used for data processing. (a) 2D $^1H-^{15}N$ fast-HSQC,[13] (b) 2D HNCO,[5] and (c) 1D HNCO.[5] The HNCO spectra in panels b and c identify the valine 115 resonance (circled in panel a). Experimental times were 25 min, 1 h 45 min, and 40 min, respectively.

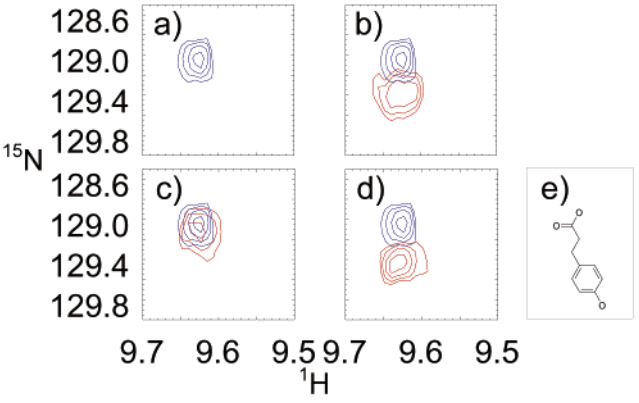

Figure 2. Plots of $^1H-^{15}N$ fast-HSQC[13] spectra recorded at 600 MHz showing the region corresponding to the Val 115 cross-peak of Figure 1a. Experimental conditions were identical with those described in Figure 1 except the protein concentration was 0.1 mM and the samples contained additionally 1% *d*-DMSO. (a−d) Reference spectrum of free FABP-4 (blue). (b) 1:9 mixture of FABP-4 and compound cocktail (red). (c) 1:9 mixture FABP-4 and compound cocktail omitting compound **5** (red). (d) 1:1 mixture of FABP-4 and compound **5** (red). (e) Compound **5**.

that the compound is interacting with the Val114-Val115 epitope. (See Supporting Information.)

The site-selective labeling method promises to be a valuable tool for identifying compounds that bind to a selected site. Two pitfalls should, however, be pointed out. Since only one peak is monitored, false positives might be detected due to a secondary (remote/global) effect. Care should thus be taken when numerous chemical shift changes are observed. Furthermore, false negatives may occur when no chemical shift change is observed despite a compound binding in the desired pocket. This may be caused by multiple counteracting effects on the chemical shift. The risk of false negatives can be reduced by labeling more than one amino acid pair. Potential applications of the method include the following: screening for binders to a selected site that has been identified either through X-ray, NMR, or modeling studies; an assay to confirm binding of ligands identified by other means to a desired site; and screening for ligands that bind to a site that could confer specificity for one target protein within a protein family. Due to the reduced spectral complexity resulting from the use of a selectively labeled sample, the method should be applicable to larger proteins than are conventional methods.

Acknowledgment. Plots were prepared using Azara provided by Wayne Boucher and the Department of Biochemistry, University of Cambridge. The authors thank Annette Elmblad for cloning of the FABP-4 construct and Agneta Tjernberg for mass spectroscopic analysis.

Supporting Information Available: Amino acid sequence of FABP-4, plots corresponding to those of Figure 2 showing larger spectral areas, a table of the compounds in the test cocktail, the X-ray structure of FABP-4 in complex with compound **5**, along with experimental details, and a list of unique amino acid pairs in the active sites/ligand binding sites of known drug targets (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534.

(2) Lin, M.; Shapiro, M. J.; Wareing, J. R. *J. Am. Chem. Soc.* **1997**, *119*, 5249−5250. Hajduk, P. J.; Olejniczak, E. T.; Fesik, S..W. *J. Am. Chem. Soc.* **1997**, *129*, 12257−12261. Chen, A.; Shapiro, M. J. *J. Am. Chem. Soc.* **1998**, *120*, 10258−10259. Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 1784−1788. Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. *J. Biomol NMR* **2000**, *18*, 65−68. Jahnke, W.; Perez, L. B.; Paris, C. G.; Strauss, A.; Fendrich, G.; Nalin, C. M. *J. Am. Chem. Soc.* **2000**, *122*, 7394−7395. Peng, J. W.; Lepre, C. A.; Fejzo, J.; Abdul-Manan, N.; Moore, J. M. *Methods Enzymol.* **2001**, *338*, 202−229. Stockman, B. J.; Farley, K. A.; Angwin, D. T. *Methods Enzymol.* **2001**, *338*, 230−246.

(3) 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.

(4) Kainosho, M.; Tsuji, T. *Biochemistry* **1982**, *21*, 6273−6279. Kato, K.; Matsunaga, C.; Igarishi, T.; Kim, H.; Odaka, A.; Sjimada, I.; Arata, Y. *Biochemistry* **1991**, *30*, 270−278.

(5) Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *96*, 432−440.

(6) Upon examination of the active sites/ligand binding sites of five known drug targets it was possible to identify several unique amino acids pairs in all cases. See Supporting Information.

(7) Hotamisligil, G. S.; Johnson, R. S.; Distel, R. J.; Ellis, R.; Papaioannou, V. E.; Spiegelman, B. M. *Science* **1996**, *274*, 1377−1379.

(8) Young, A. C.; Scapin, G.; Kromminga, A.; Patel, S. B.; Veerkamp, J. H.; Sacchettini, J. C. *Structure* **1994**, *2*, 523−34.

(9) Muchmore, D. C.; McIntosh, L. P.; Russell, C. B.; Anderson, D. E.; Dahlquist, F. W. *Methods Enzymol.* **1989**, 177, 44−73.

(10) Muchmore et al. advocates the use of auxotrophic strains with gene deletions suitable for the particular type of labeling that is desired.

(11) Yabuki, T.; Kigawa, T.; Dohmae, N.; Takio, K.; Terada, T.; Ito, Y.; Laue, E. D.; Cooper, J. A.; Kainosho, M.; Yokoyama, S. *J. Biomol. NMR* **1998**, *11*, 295−306.

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

(13) Mori, S.; Abeygunawardana, C.; Johnson, M. O.; vanZijl, P. C. M. *J. Magn. Reson.* **1995**, *B108*, 94−98.

(14) Constantine, K. L.; Friedrichs, M. S.; Wittekind, M.; Jamil, H.; Chu, C.-H.; Parker, R. A.; Goldfarb, V.; Mueller, L.; Farmer, B. T. *Biochemistry* **1998**, *37*, 7965−7980.

JA0178261