# **Functional Genomics and NMR Spectroscopy**

Robert Powers<sup>\*</sup>

Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68522, USA

**Abstract:** The recent success of the human genome project and the continued accomplishment in obtaining DNA sequences for a vast array of organisms is providing an unprecedented wealth of information. Nevertheless, an abundance of the proteome contains hypothetical proteins or proteins of unknown function, where high throughput approaches for genome-wide functional annotation (functional genomics) has evolved as the necessary next step. Nuclear magnetic resonance spectroscopy is playing an important role in functional genomics by providing information on the structure of protein and protein-ligand complexes, from metabolite fingerprinting and profiling, from the analysis of the metabolome, and from ligand affinity screens to identify chemical probes.

**Keywords:** NMR, functional genomics, NMR high throughput screens, protein-ligand binding, protein-ligand co-structures, structural biology, structural genomics, NMR metabolomics, chemical proteomics, protein structure initiative, hypothetical proteins.

## **INTRODUCTION**

The availability of the human genome sequence has just begun to provide a wealth of information on cell biology, development, evolution and physiology [1, 2]. As a result of the success of the Human Genome Project, there has been an explosion of effort in obtaining the complete genome sequence of other organisms, which has included the chimpanzee [3], domestic dog [4], mouse [5] and woolly mammoth [6]. To date, 496 genomes have been completely sequenced with 939 ongoing genome projects [7-11]. The sequencing of the human genome and other genomes provides powerful protein data sets for making beneficial contributions to medicine and human health issues [12], but fully ~50% of these proteomes comprise hypothetical proteins or proteins of unknown function [13-15]. In effect, the acquisition of genomic information has clearly highlighted the extent of information that is currently lacking in our understanding of cell biology.

Contributing to this expanding knowledge base is the extensive amount of protein structures emerging from the Protein Structure Initiative (PSI), which includes four largescale centers and six specialized centers involving collaborations among a total of 45 universities, research institutes and companies [16, 17]. PSI has the ambitious aim of determining the three-dimensional structure for most of the proteins in the proteome. A major expected benefit of this effort is obtaining the necessary functional information for the vast number of hypothetical proteins emerging from ongoing sequencing efforts [16, 17]. The accepted paradigm of structural genomics is to focus on solving structures of proteins where 3D structures cannot be easily predicted from the Protein Database (PDB) [13-15]. By obtaining a representative structure for each of the 20,000-30,000 sequence families, it is expected that a significant structural coverage will be achieved with respect to obtaining an experimental or model structure for all proteins in nature. Unfortunately, numerous structures emerging from structural genomics correspond to folds that provide little insight into function [18-23]. The impact has been the addition of ~2,166 proteins of unknown function in the PDB [24, 25], where this number will continue to expand with the growing success of PSI [26-28].

Readily obtaining functional information to augment the available sequence and structural data is essential for the eventual success of the Human Genome and PSI projects. The prospect of obtaining molecular functional information for an extensive collection of hypothetical proteins by traditional biochemical approaches presents an extremely overwhelming and daunting task [29]. Historically, many years of research are required to identify the function of a single protein. Alternatively, functional genomics applies high throughput technologies to obtain genome-wide functional annotation [30-34]. Common approaches include (i) monitoring gene expression in response to external stress factors (temperature, pH, drugs) using DNA arrays [35], (ii) determining protein-protein interaction maps [36, 37] using yeast two-hybrids [38], (iii) monitoring protein post-translational modifications using mass spectrometry [39], (iv) monitoring protein-DNA interactions [40] using chromatin immunoprecipitation (ChIP) [41], (v) monitoring phenotypes of gene knockouts using traditional genetic engineering methods [42-44] or, more recently, RNA interference [45], (vi) monitoring changes in the metabolome [46, 47] and (vii) using protein arrays [48, 49] to monitor protein-ligand interactions or biochemical activity. Protein arrays still face significant technical challenges that need to be resolved before wholeproteome chips become routinely available.

Bioinformatics approaches are also being applied for the functional annotations of whole genomes and protein structures to augment these experimental methods. DNA and protein sequence similarities using BLAST [50], ClustalW [51], FASTA [52] are well-established methodologies to infer function based on the accepted paradigm that a similarity in sequence implies a similarity in function [53]. More advanced approaches utilize hidden Markov models (HMM) to generate profiles for functional or structural protein families.

<sup>\*</sup>Address correspondence to this author at the Department of Chemistry, University of Nebraska-Lincoln, 722 Hamilton Hall, Lincoln, NE 68588-0304, USA; Tel: (402) 472-3039; Fax: (402) 472-2044; E-mail: rpowers3@unl.edu

These profiles are used in combination with BLAST sequence alignments to identify new members to the functional family. Unfortunately, recent analysis indicate a high (10-30%) error rate in genome annotations [54, 55] and functional conservation decreases significantly when sequence identity falls below 50% [56].

The CATH [57], SCOP [58] and TIGRFAMs [59] databases contain comprehensive structural and HMM profile alignments to create hierarchical classifications of all known protein structures. These protein family classifications are useful for identifying functional and evolutionary relationships between protein structures [60]. Programs like Dali [61], SSAP [62], CE [63], TM-align [64] and Vast [65] are used to identify structural homologs in the absence of sequence similarity to infer function [66]. This occurs because tertiary structures are significantly more evolutionarily stable than protein sequences [67]. There are numerous examples of proteins that share similar structures in the absence of high-sequence identity (<30%) [68-73].

Amino-acid residues associated with the active-sites and biological activities of proteins are evolutionary stable relative to the remainder of the protein's sequence [74, 75]. As a result, significant global sequence and structural divergence may occur while biological activity remains constant. For example, the peptidyl-tRNA hydrolases (Pth & Pth2) have no sequence or structure similarity, but exhibit identical activity [76, 78]. A number of computational methods have been developed to predict the location of active sites or ligand binding sites in hypothetical protein structures to assign function in the absence of global sequence and structural similarity. These methods include: matching 3D templates [79-81], aligning structures to match a few consensus or enzymatic catalytic residues [82-93], identification of binding pockets or clefts [94-98], identification of cavities consistent with shapes of known ligands [99], a sequence independent force field to extract common active site features [100], theoretical prediction of titration curves [101] or energetics of ligand interactions [102-104], using chemical properties and electrostatic potentials of amino-acid residues consistent with active site characteristics [105, 106], neural network analysis of spatial clustering of residues [107], and conserved residues from multiple sequence alignments (phylogenetic motifs) [89, 108-111]. Distant evolutionary relationships between enzymes with no sequence similarity have been identified based on the comparison of conserved active-site structures [112, 113], but these methods also tend to suffer from the inherent uncertainty and ambiguity in correctly predicting a ligand-binding site [60] with error rates as high as 40% [54]. These problems may arise from numerous issues that include: the impact of protein conformational changes upon ligand binding, protein and ligand dynamics, shallow binding sites, surface inaccessible binding pockets in apo-structures, coarse definition or limited number of activesite templates.

Whole-genome functional annotation is clearly a challenging endeavor and will require the application of multiple experimental and computational protocols for its success. Verification of a functional assignment for a hypothetical protein will inevitably involve consensus among a number of high throughput analysis. Nuclear magnetic resonance spectroscopy (NMR) is becoming an important addition to functional genomics with applications in: (i) the structure determination of hypothetical proteins as part of structural genomics [26, 114-117], (ii) the rapid structure determination of protein-ligand complexes [118-121], (iii) the application of NMR ligand affinity screens to identify protein-ligand complexes for functional annotation [121-124] and chemical probes [125, 126], and (iv) analyzing changes in the metabolome in response to changes in protein activity [127-129]. This review will describe some recent examples of the application of NMR spectroscopy in the functional annotation of hypothetical proteins or proteins of unknown function.

## NMR STRUCTURES OF HYPOTHETICAL PRO-TEINS

Obtaining a structure of a protein or protein-ligand complex greatly contributes to our understanding of its biological function [130]. As summarized above, identifying a structural homolog for a hypothetical protein is a routine approach for assigning a function. Simply, the function of the structural homolog is also assigned to the hypothetical protein based on the accepted premise that structure implies function [60]. Multidimensional heteronuclear NMR experiments are routinely used to determine the structure of proteins and their interactions with ligands [131-135]. These methods have been extensively reviewed and will only be briefly summarized here.

There are three basic steps to solving a protein structure by NMR: (i) obtaining the NMR backbone and side-chain sequential assignments, (ii) identification of secondary structure elements, and (iii) solving the three-dimensional fold from structural constraints. Through the application of isotope labeling (<sup>13</sup>C, <sup>15</sup>N) of the protein and 3D tripleresonance NMR experiments, it is routinely achievable to assign each observed <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N peak in the NMR spectra with a specific amino acid in the protein sequence [136-138]. Briefly, the protein assignment protocol utilizes a series of triple-resonance experiments where each experiment correlates a subset of the protein backbone atoms through Jcoupling. By combining overlapping information between the various experiments, it is possible to "walk" down the protein backbone and complete the resonance assignments.

Regions of  $\alpha$ -helical and  $\beta$ -sheet secondary structures are then identified from  $^{13}\text{C}\alpha/^{13}\text{C}\beta$  secondary structure chemical shifts [139], NH exchange rates,  ${}^{3}J_{HNa}$  coupling constants and sequential distance patterns (NH(i)-NH(i+1)  $C^{\alpha}H(i)$ -NH(i+2,3,4), NH(i)-NH(i+2), and  $C^{\alpha}H(i)-C^{\beta}H(i+3))$  or short inter-strand distances between NH and  $C^{\alpha}H$  protons. The NMR assignments then enable the determination of a solution structure using distance constraints derived from NOEs (Nuclear Overhauser Effect - the enhancement of NMR signals by a through space dipole-dipole interaction that is dependent on the distance  $(1/r^6)$  between the nuclei.) present in 3D <sup>15</sup>N- [140, 141] and <sup>13</sup>C-edited NOESY [142, 143] experiments (correlates <sup>1</sup>H nuclei bonded to <sup>13</sup>C or <sup>15</sup>N nuclei to other <sup>1</sup>H nuclei that are  $\leq 5$ Å away), dihedral angles derived from coupling constants and carbon chemical shifts [144, 145] and hydrogen bond constraints from slowly exchanging amides. The structures may also be directly refined against the  ${}^{3}J_{HNa}$  coupling constants [146], secondary

 $^{13}$ C $\alpha'^{13}$ C $\beta$  chemical shift restraints [147], radius of gyration [148], and a conformational database potential [149-151].

To date, a total of 1099 NMR structures have been deposited into the PDB by all the structural genomics consortiums where 758 of these structures are listed with a functional annotation [152]. The following recent example highlights the inherent value of obtaining an NMR structure of hypothetical proteins for functional annotation.

<u>Hypothetical protein AF2095</u>. Thermophilic archaea *Archaeglobus fulgidis* AF2095 is a protein of unknown function that was targeted for structural analysis by the Northeast Structural Genomics Consortium (NESG; http://www.nesg. org). The NMR structure of protein AF2095 was shown to exhibit a similar structural topology to peptidyl-tRNA hydrolase (Pth2) [76, 153]. Dali analysis resulted in a Z-score of 8.2 with a 3.7Å RMSD for the backbone structure alignments. Further analysis of the AF2095 structure indicated a

large cavity with a positive electrostatic potential suggestive of a potential tRNA binding site that was proximal to a catalytic triad (Fig. 1). These structural characteristics are also consistent with Pth2 enzymatic activity. Pth cleaves the peptide from peptidyl-tRNA molecules allowing the freed tRNA to be recycled in the protein synthesis process [154, 155]. The accumulation of peptidyl-tRNA molecules that have prematurely dissociated from the ribosome during protein translation have been associated with cell death [156, 157].

The AF2095 NMR structure was leveraged to assign a structure and function for 55 other proteins. Additionally, the functional assignment of AF2095 provided further insight in the evolution of Pth and Pth2 enzymes and the formation of the mitochondria in eukaryotes. The Pth and Pth2 proteins exhibit similar enzymatic activity but lack any sequence or structure homology. Pth enzymes are essential in bacteria [158] and Pth2 enzymes are only found in archaea. Eukaryotes contain both Pth and Pth2 enzymes, where Pth2 are mito-



Fig. (1). Functional Annotation from NMR Structures. Ribbon diagrams of the aligned views of (a) human Pth2 (PDB code: 1q/s), and (b) A. *fulgidus* AF2095 (GR4; PDB code: 1rzw). (c) Electrostatic surface of AF2095. The electrostatic surface was calculated using a salt concentration of 0.1 M and the color scale is -5 kt (negative, red) to 5 kt (positive, blue). Residues proposed to form the catalytic triad (Lys 19, Asp 80, Thr 90) are labeled. (d) Structure is rotated -45 degrees about the y-axis, relative (c). (Reprinted with permission from reference [76], Copyright 2005 by *Protein Science Online* by The Protein Society).

chondrial enzymes in humans. This analysis suggests that eukaryotes inherited Pth2 enzymes from archaea during the formation of the mitochondria from the endosymbiosis of two prokaryotes.

### NMR LIGAND AFFINITY SCREENS

Despite the inherent value in obtaining a protein structure for functional analysis, there are numerous situations when a structure alone is insufficient. This generally occurs when a hypothetical protein exhibits a novel fold that lacks any similarity to proteins of known function. A fundamental component to our understanding of the biological activity of a protein is through the identification of its functional ligand(s), co-factors, substrates or metabolites [159]. Essential to this understanding is the knowledge that a protein's active-site has been optimized by nature to interact with a unique and specific set of targets. Promiscuity of binding is inherently detrimental to the overall biological process, which is evident by the high specificity of interactions that have been well-documented in numerous metabolic and signaling pathways [160-162]. This understanding is also an essential aspect of drug discovery and supports the observed rational that high-affinity and selective compounds targeting a specific protein can be developed and used therapeutically [163-166].

NMR is an important component of the drug discovery process and is routinely used in the pharmaceutical industry for the screening and validation of chemical leads [122, 167-171]. NMR's sensitivity to weak binders ( $K_D \le 10 \text{ mM}$ ) is an important advantage since initial chemical leads generally have a weak binding affinity, which is improved upon during the iterative design process [172]. Also, the versatility of NMR provides extensive flexibility in the methodology used to monitor a protein-ligand interaction. Thus, the experiment can be optimized for the particulars of the system under study. Protein-ligand binding interactions can be monitored by changes in line-width and/or peak intensity ( $T_1$  and  $T_2$ ) relaxation changes) [173, 174], by changes in ligand diffusion coefficients [175-177], from ligand or protein chemical shift perturbations [178-180], induced transferred NOE (trNOE) for the ligand [181-183], a saturation transfer difference (STD) between either the protein or bulk solvent to the ligand [184-186], appearance of new NOEs and/or intermolecular NOEs between the ligand and protein [183, 187]. A number of high throughput NMR screens for drug discovery have been implemented based on these observables that are equally applicable for identifying ligands that bind hypothetical proteins for functional annotations. A few of these methods will be briefly summarized.

SAR by NMR. The influential manuscript by Shuker *et al.* described the first general application of NMR to screen a library of small molecules for their ability to bind proteins from observed chemical shift perturbations [188]. In addition to initiating subsequent efforts in designing alternative NMR based screens, the "SAR by NMR" method also stimulated the current interest in fragment-based drug discovery screens [189, 190].

In the SAR by NMR approach, chemical shift perturbations are observed from 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectra [191-194] or 2D <sup>1</sup>H-<sup>13</sup>C HSQC spectra [195]. The chemical shift changes are used to verify a positive binding event and to identify the ligand's binding site on the protein's surface. Each amino-acid residue in the protein generally exhibits a unique peak in the 2D  $^{1}H^{-15}N$  HSQC NMR spectrum, where the peak position is dependent on the local environment for each particular amino-acid. Each peak in the 2D  $^{1}H^{-15}N$  HSQC NMR spectrum has been sequentially assigned to a *specific* amino-acid using standard protocols [196]. The protein residues involved in binding the ligand will experience a change in its local environment, which results in a change in chemical shifts. Mapping the chemical shift changes for these residues onto the protein surface identifies the ligand binding site.

A structure of the protein-ligand complex is obtained, where a key component of the protocol is to link two or more fragments that bind in distinct, but proximal locations in the protein's active site. Properly linking the fragments is predicted to result in a dramatic increase in the binding affinity of the new compound. A recent illustration of the SAR by NMR methodology is the design of a potent inhibitor of anti-apoptotic protein Bcl- $x_L$  [197] (Fig. 2).

The overexpression of the Bcl-x<sub>L</sub> protein suppresses the apoptotic process initiated by DNA damage or hypoxia. The antiapoptotic activity of Bcl-x<sub>L</sub> may lead to the development of cancer, implicating Bcl-x<sub>L</sub> as a potential therapeutic target. A fragment library containing 10,000 compounds were screened using NMR chemical shift perturbations to identify chemical leads. A fluoro biaryl acid (K<sub>D</sub> ~ 300µM) was identified to bind Bcl-x<sub>L</sub> in a hydrophobic groove near residues G94 and G138. The X-ray structure of the Bcl-x<sub>L</sub>-Bak peptide complex indicated a second potential ligand binding site proximal to the fluoro biaryl acid binding site. A second screen using 3,500 compounds in the presence of excess fluoro biaryl acid identified naphthol analogues (K<sub>D</sub> ~ 2-13mM) that bound in this second binding site. The proper linking of these fragments with an acylsulfonamide combined with further refinements lead to an inhibitor with a K<sub>i</sub> of 36nM.

RAMPED-UP NMR is a unique modification of SAR by NMR that uses mixtures of proteins to screen compounds for specificity and selectivity against a particular protein [198]. The proteins in the mixture are each uniquely labeled with a specific amino acid (<sup>15</sup>N-Trp, <sup>15</sup>N-Ile, or <sup>15</sup>N-Ala) associated with the active site of each protein. The protein mixture yields a 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectrum with distinct NMR resonances uniquely identified to each protein. The RAMPED-UP NMR approach is a specific application of the general modification of the SAR by NMR technique to use specific labeling of residue types [199, 200], methyl groups [201] or sequential fragments [202, 203] or spin-labeling [204] of the protein to simplify the NMR spectra.

SHAPES. The design and composition of the chemical library is a major component of the SHAPES approach to screening by NMR [205-207]. The SHAPES library is a small, structurally diverse library composed of water soluble compounds that correspond to fragments or molecular frameworks of known drugs. NMR screening for binders in the SHAPES approach is typically accomplished by saturation transfer difference (STD), 1D line-broadening or 2D-trNOE experiments.



**Fig. (2).** SAR by NMR (**a**) Selected region of <sup>15</sup>N HSQC spectra recorded on uniformly <sup>15</sup>N-labeled Bcl-xL alone (black), in the presence of 2 mM biaryl acid (red), and in the presence of 2 mM biaryl acid and 5 mM naphthol derivative (green). (**b**) Superposition of seven lowenergy structures calculated for Bcl-xL complexed to biaryl acid and naphthol derivative. For clarity, the average-minimized structure of the protein is shown as a solvent-accessible surface, color coded as follows: oxygen and oxygen-bound protons are red, and nitrogen and nitrogen-bound protons are blue, while all other atoms are gray. The positions of biaryl acid and naphthol derivative in the average minimized structure are shown in orange. (**c**) NMR-derived model of *trans*-olefin linked compound bound to Bcl-xL. The side chain of F97 divides the first site from the second site. (Reprinted with permission from reference [197], Copyright 2006 by *American Chemical Society*).

The SHAPES method was successfully applied in the development of inhibitors for p38 MAP kinase, which is associated with a number of cancers including breast, stomach, liver and prostrate cancers [208]. 1D <sup>1</sup>H NMR linebroadening experiments was initially used to identify weak binders ( $K_D = 1-7$  mM) to p38 MAP kinase using the SHAPES library [209]. Fragments with common scaffolds were fused to generate compounds with an initial increase in affinity to  $K_D$ 's of 200-300  $\mu$ M. Further refinements of the chemical structure resulted in a chemical lead with a  $K_D$  of ~ 200 nM.

MS/NMR assay. A fundamental limitation to NMR ligand affinity screens is the high utilization of resources required to complete even modest sized screens. It is not uncommon for an NMR screen to require 100s of milligrams to gram quantities of a protein and days to weeks of dedicated NMR instrument time. Conversely, mass spectrometry (MS) has a higher sensitivity relative to NMR, with limits of detection in the femtomolar range [210]. Similarly, MS data collection times may be an order of magnitude or faster compared to NMR techniques, but MS ligand affinity screens also have limitations. MS does not provide any detailed structural information on the protein-ligand complex and is unable to differentiate between multiple, non-specific binding interactions and productive stoichiometric binding in the protein's active site. The MS/NMR assay takes advantage of the inherent strengths of MS and NMR ligand affinity screens and combines the two approaches. A flow-diagram of the MS/NMR screening assay is depicted in Fig. 3.

Compound mixtures are incubated with the protein and passed through a size-exclusion chromatography (SEC) column, where only compounds that bind the protein will be present in the eluant. ESI/MS is used to identify the compounds in the eluent by the observation of a molecular ion peak consistent with the molecular-weight of compound in the original mixture. To eliminate false positives, the compound mixtures are also passed through the SEC column in the absence of the protein. The MS/NMR assay does not require a deconvolution step like SAR by NMR, since each compound in a mixture has a unique molecular weight, which acts as a molecular tag to identify each individual compound. A 2D <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectrum of the protein-ligand complex is obtained for each positive "hit" from the SEC-MS step. The observation of chemical shift perturbations clustered in the vicinity of the protein's active site verifies a biologically relevant binding interaction. Conversely, the absence of chemical shift perturbations or a random distribution of chemical shift changes on the protein surface would imply a lack of an interaction of the compound with the protein or potentially the existence of nonspecific binding. The NMR data effectively filters out nonspecific and non-productive binders that may be present in the SEC-MS screen.

The MS/NMR assay was successfully applied using a 32,000 compound library that was screened for RGS4 ligands [211]. Regulators of G-protein signaling (RGS) act as attenuators of the G-protein signal cascade by binding to the Ga subunit of G-proteins and inducing a 30-fold increase in the intrinsic Ga GTPase activity (for reviews see [212-215]). RGS activity has been associated with a variety of cellular functions including proliferation, differentiation, membrane trafficking, and embryonic development. The SEC-MS screening component of the MS/NMR assay identified 50 potential hits. The NMR chemical shift perturbation analysis of these hits identified a compound that bound specifically to RGS4 in a unique allosteric binding site on RGS4 [216]. This binding site comprises helix  $\alpha_1$  and  $\alpha_2$  and the intervening loop in the RGS4 structure and suggested a potential mechanism for the inhibition of the formation of the RGS4-Ga complex by preventing the observed conforma-



Fig. (3). Pictorial flow diagram of MS/NMR assay using data from the MMP-1 binding assay. (Reprinted with permission from reference [211], Copyright 2001 by American Chemical Society).

tional change in RGS4 that occurs upon binding G $\alpha$  [216, 217]. The compound was then shown to inhibit the interaction of RGS4 with G $\alpha$ . These results represented the identification of the first known small molecule inhibitor of RGS4.

### **RAPID PROTEIN-LIGAND CO-STRUCTURE**

NMR affinity screens have been successful in identifying numerous drug leads against a variety of therapeutic targets. Obtaining a protein-ligand co-structure as part of a structurebased drug discovery program or as an aid in functional annotation is the necessary next step of the process. Solving a high-resolution NMR structure for the protein-ligand complex would still require the same expenditure of time and effort that was required for determining the original protein structure. This may require upwards of months to a year of NMR data collection and analysis. A number of NMR-based approaches have been described to shorten this time-frame and obtain rapid protein-ligand co-structures.

<u>NOE-Guided Protein-Ligand Docked Models</u>. The NMR structure and dataset for the free protein is used as the initial model for determining the structure of the complex [118].



The ligand is docked into the free protein structure primarily based on intermolecular NOEs observed in a 3D <sup>13</sup>C-edited/<sup>12</sup>C-filtered NOESY spectrum [218]. The NMR assignments for the ligand in the complex are determined by a combination of <sup>12</sup>C-filtered COSY, TOCSY and NOESY experiments [219-221]. The basic premise of this approach is to use a minimal set of the standard 3D NMR experiments to re-assign the protein NMR resonances that are perturbed in the complex to permit a quick assignment of the protein-ligand intermolecular NOEs.

The reliability of protein-ligand co-structures determined from a minimal number of constraints was illustrated by the comparison of an MMP-1:CGS-27023A structure calculated with a complete NMR data set and an idealized minimal set of constraints [119] (Fig. **4a**). Further validation of the utility of NOE-directed protein-ligand co-structures was demonstrated by the structure-based design of a selective and potent inhibitor of MMP-13 from a MMP-13:CL-82198 NMR model (Fig. **4b**).

SOS-NMR. The SOS-NMR approach also uses the existing structure of the protein to dock the ligand of interest



**Fig. (4).** NOE-directed Protein Ligand Co-structures. (a) Comparison of the active site for the MMP-1-CGS-27023A structure calculated with minimal constraints (red) and the MMP-1-CGS-27023A complex calculated with a complete set of constraints (green). The active-site residues that are displayed are 80-82, 112-115, and 138-140. (b) *top* Expanded 2D plane from the 3D <sup>13</sup>C-edited/<sup>12</sup>C-filtered NOESY experiment corresponding to NOEs from L82  $\delta$  and L115  $\delta$  to the labeled resonances from CL-82198. *bottom* Expanded view of the NMR MMP-13:CL-82198 complex where the MMP-13 active site is shown as a transparent surface with CL-82198 shown as liquorice bonds. (Reprinted with permission from references [119, 120], Copyright 2000 by American Chemical Society).

[186]. Instead of using NOEs, the SOS-NMR method uses <sup>1</sup>H STD experiments with selectively labeled protein samples to position the ligand in the protein's active-site. In a general STD experiment, all the NMR resonances of the protein are saturated (no-signal), where this saturation is transferred to the ligand during binding (reduced signal). The STD experiment is conducted by alternatively subtracting a 1D spectrum with on-resonance and off-resonance irradiation of protein resonances. If the ligand binds the protein, there will be a reduction in the intensity of the ligand spectrum during the on-resonance irradiation compared to the off-resonance irradiation cycle. Since the spectrum is being alternatively subtracted, the difference spectra should yield the 1D NMR spectrum of the ligand. Conversely, if the ligand does not bind the protein, no change in the ligand's intensity occurs for either the on- or off-resonance irradiation cycle so the end result is a null spectrum.

In the case of the SOS-NMR experiment, only one residue type is selectively protonated and the remainder of the protein is deuterated. A saturation transfer will only occur if the ligand is within 5Å of this labeled amino-acid (Fig. **5a**). The experiment is repeated with different labeled aminoacids. A ligand binding surface is determined by the overlap of protein regions that contain all the amino-acid types that exhibit an STD to the ligand while excluding the amino-acid types that did not exhibit an STD to the ligand (Fig. **5b**). The ligand is then docked into the binding surface using molecular modeling software to generate multiple conformers. A list of ambiguous NOEs [222] to each amino-acid residue in the binding surface that exhibited an STD is used to generate an energy profile for each protein-ligand complex. The lowest energy structure is then chosen as the best model. The SOS-NMR protocol was demonstrated using FKBP complexed to FK506. The SOS-NMR complex compared to the X-ray structure yielded a 1.1Å rmsd difference for the heavy atoms of FK506.

<u>NMR-DOC</u>. The NMR-DOC protocol is closely related to the SOS-NMR process for determining a rapid proteinligand co-structure [223]. Both methods use selective labeling of a protein and STD experiments to monitor a binding event. NMR-DOC uses uniformly deuterium labeled proteins combined with <sup>13</sup>C methyl labeling of the Met, Ile and Thr residues. These specific <sup>13</sup>C methyl labels are sequentially assigned by using a reference protein-ligand X-ray structure. This existing structure identifies which protein methyl resonance are within the 5Å NOE distance to the ligand. These NOEs are simply measured using a 2D <sup>1</sup>H NOESY spectrum.

A new compound is shown to bind the protein from a <sup>1</sup>H STD experiment where the <sup>13</sup>C methyl labels are selectively saturated. A docked structure is then simply generated using NOEs between the ligand and the <sup>13</sup>C labeled Met, Ile, Thr methyls in a 2D <sup>1</sup>H NOESY spectrum. The intermolecular NOEs are assigned based on the assignments obtained for the reference protein-ligand structure.

Chemical shift differences between two chemically related ligands bound to the same protein provide an alternative approach to assign the <sup>13</sup>C labeled Met, Ile, Thr methyls. Chemical shift changes are measured using a 2D <sup>1</sup>H-<sup>13</sup>C HSQC spectrum. Again, the reference protein-ligand structure and the chemical perturbations in the two ligand struc-



**Fig. (5).** SOS-NMR (**a**) STD-NMR spectra of 2-(3'-pyridyl)-benzimidazole (**1**) in the presence of (A) unlabeled FKBP, (B) perdeuterated, Ile-protonated FKBP, (C) perdeuterated, Val-protonated FKBP, (D) perdeuterated, Leu-protonated FKBP, (E) perdeuterated, Met-protonated FKBP, and (F) perdeuterated FKBP. The resonances corresponding to the ligand are indicated. (**b**) 2-(3'-pyridyl)-benzimidazole complexed to the FK506 binding site of FKBP (gray ribbons, orange surface) generated using the program DOCK after filtering with the SOS-NMR data. (Reprinted with permission from reference [186], Copyright 2004 by American Chemical Society).

tures would indicate which methyl NMR resonance peaks are expected to change. This occurs because a chemical change in one of the ligands removes an interaction with a specific methyl in the protein resulting in a large chemical shift change for that methyl NMR resonance. Differential chemical shifts have also been shown to be sufficient to rapidly determine a protein-ligand co-structure [224]. Chemical shift differences are measured between a series of compounds that have subtle structural changes relative to the compound of interest. Each differential chemical shift change for a pair of compounds is mapped onto the protein's active site. The compound of interest is then docked into the active-site by correlating the chemicals shift differences with the structural differences in the compound.

Chemical Shift Differences. Chemical shift differences observed by comparing the 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectra of the free protein relative to a protein-ligand complex can also be used to generate a rapid protein-ligand co-structure model [225, 226]. This approach is analogous to NOE-guided docked structures. The NHs that exhibit an above average chemical shift change are converted to ambiguous interaction restraints that are used to dock the ligand in a molecular dynamics simulation (Fig. 6a). Simply, a distant constraint is used between each NH with an above average chemical shift change with the ligand. Since the constraints are ambiguous, the energy potential for the distance constraints is minimized by maximizing the number of satisfied ambiguous constraints. A simultaneous goal is to minimize the van der Waals interaction energy in the docked complex (Fig. 6b). The approach was validated with three existing apo-protein (PTP1b, PKA, p38) structures and their corresponding protein-ligand X-ray structures. The ligands were re-docked into the apo-protein structure based on 2D <sup>1</sup>H-<sup>15</sup>N HSOC chemical shift perturbation data and compared to the original protein-ligand X-ray structure. The resulting rmsds ranged from 1 to 2.9 Å (Fig. 6c).

#### NMR CHEMICAL PROTEOMICS

Drug discovery projects effectively use the protein as a probe to identify inhibitors and potential drug candidates. Chemical proteomics simply reverses this process and uses a defined chemical entity (co-factor, inhibitor) as a probe to profile many proteins [125, 126]. These chemical probes can be used in multiple *in vivo* and *in vitro* assays to monitor and analyze the biological activity of a protein for functional annotation (Fig. **7**).

One such approach is the utility of a chemical probe in affinity chromatography to "fish-out" functionally related proteins from a proteome. Godl *et al.* (2003) illustrates this approach using a selective p38 protein kinase inhibitor (SB203580) along with the protein-inhibitor X-ray structure to identify related protein kinases from HeLa cells or transfected COS-7 cells [227]. Alternatively, the chemical probe can be simply used to label the protein by the addition of a fluorescent tag for gel analysis [228] or a molecular-weight tag for MS analysis [229, 230]. Similarly, the chemical probes can be designed to be detected by NMR. Spence *et al.* (2001) demonstrated the application of hyperpolarized xenon [231, 232] to detected nanomoles of avidin binding biotin [233]. A cryptophane-A cage molecule is used to bind a



**Fig. (6).** Protein-Ligand Co-structures from Chemical Shift Perturbations. (a) Residues with the largest chemical shift pertubations (red) and residues for which flexibility was allowed during the simulation (blue), mapped onto the X-ray structure of apo-PTP1b. (b) Selection plot of the resulting complexes. The accuracy of each structure relative to the reference X-ray complex is indicated in gray scale. This represents the positional rmsd in the coordinates of the ligand. The selection plot shows that in this case one of the two energies would be sufficient for the selection of a structure. (c) A selection of accepted structures from NMR data (green/yellow) corresponds well with the X-ray reference complex (blue), with an rmsd of ca. 1 Å. (Reprinted with permission from reference [226], Copyright 2002 by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim).



Fig. (7). Chemical proteomic analysis of the p38 kinase inhibitor SB203580. Based on structural information from the known co-crystal structure with p38 kinase, a suitable derivative of SB203580 was synthesised and used for the selective isolation of cellular protein targets by affinity chromatography. After gel electrophoresis and MS analysis, previously unknown kinase targets of SB203580 were further characterized in *in vitro* and cellular kinase assays. (Reprinted with permission from reference [227], Copyright 2005 by Elesvier).

xenon molecule, where the cage is attached to biotin through a peptide tether to increase solubility [129]. Xe chemical shift perturbations are observed when avidin binds biotin (Fig. 8).

These tags can also be used to specifically follow the relative expression level of a protein in response to environmental stimuli (addition of a drug, hypoxia, temperature change, etc) or a protein knockout.

Chemical probes can also be used to simulate genetic knockouts or deletion mutants by inhibiting the protein of interest. A particular advantage of chemical knockouts is the transient nature of the protein inhibition. The protein is only inactivated as long as the drug is present allowing for a reversible system. Jaeschke *et al.* (2006) demonstrated this application by monitoring the cellular impact of inactivating JNK2, a key mediator of cell responses to environmental stimuli, with the general protein kinase inhibitor PP1 [234]. Inhibiting JNK2 was shown to affect the phosphorylation and expression of cJun and cell proliferation (Fig. 9).

The same types of NMR affinity screens described above for identifying drug leads can be similarly applied to develop probes for chemical proteomics. Additionally, the NMR methods for solving protein structures and for rapidly determining protein-ligand co-structures will also assist in these endeavors. In effect, NMR affinity screens are used for identifying the tools required by chemical proteomics for functional annotation.

# NMR LIGAND AFFINITY SCREENS FOR FUNC-TIONAL ANNOTATIONS

The NMR ligand affinity screens described above for drug discovery can also have applications in the functional annotations of hypothetical proteins [122, 123]. The primary difference is the use of a compound library composed of cofactors, substrates and inhibitors instead of a random compound library or a fragment-based library of known drugs [235]. The basic premise is to use the identity of the ligand that binds the hypothetical protein to leverage a functional assignment. Hajduk et al. (2002) demonstrated the general protocol with hypothetical protein HI-0033 from Haemophilus influenza. The goal was to identify a chemical probe to develop a high throughput screening (HTS) assay for a protein of unknown function. 160 compounds were screened using the SAR by NMR methodology, where deoxyadenosine monophosphate (dAMP) and S-adenosylhomocysteine (SAH) were shown to bind HI-0033 in the same binding site. A fluorescent SAH analog was then used in an HTS displacement assay.

Other similar NMR ligand affinity screening approaches have been used to aid in the functional assignments of hypothetical proteins. Cort *et al.* (2000) evaluated the function of hypothetical protein MTH538 from *Methanobacterium thermoautotrophicum* from limited ligand binding [236]. The NMR structure of MTH538 was similar to known structures



**Fig. (8).** Chemical Proteomics using NMR. (**a**) Structure of a biosensor molecule designed to bind xenon to a protein with high affinity and specificity. (**b**) Xenon-129 NMR spectra monitoring the binding of biotin-functionalized xenon to avidin. *Top* shows the functionalized xenon before the addition of avidin, with the more intense peak corresponding to functionalized xenon and the smaller peak corresponding to xenon in the cage without linker and ligand, serving as both a chemical shift and signal intensity reference. *Bottom* shows the spectrum on the addition of avidin monomer. A third peak, corresponding to functionalized xenon bound to avidin, has appeared, and the unbound functionalized xenon peak has decreased in intensity. All chemical shifts are referenced to that of xenon gas. (Reprinted with permission from reference [233], Copyright 2001 by The National Academy of Sciences of the United States of America).

of receiver domains from two component response system like CheY and to the structures of flavodoxins and GTPbinding proteins. NMR binding studies were conducted with FMN, F420 coenzyme,  $Mg^{+2}$ , and acetyl phosphate. MTH538 exhibited a possible binding interaction only to  $Mg^{+2}$ , implying that MTH539 may be a phosphorylationindependent two-component response regulator system. Similarly, Yao & Sem (2005) used <sup>1</sup>H STD experiments and a mixture of six co-factors and cyclic nucleotides to analyze the function of radial spoke protein-2 (RSP2) from *Chlamydomonas reinhardtii*. RSP2 was tentatively assigned as a new class of cGMP receptors based on its sequence homology to a GAF domain. The protein was expected to bind either cAMP or cGMP, but was shown to preferentially bind cCMP (Fig. **10a,b**).



**Fig. (9).** Gene Knock-out using a Chemical Probe. (a) JNK1 and JNK2 protein kinase activity contributes to cJun expression Primary MEFs prepared from wild-type embryos (J1+/+ J2+/+) or mutated embryos (J1+/+ J2MG/MG or J12/2 J2MG/MG) were incubated in medium supplemented with solvent (DMSO) or 10 mM 1NM-PP1. The cells were harvested, and the expression of cJun and tubulin was examined by immunoblot analysis. (b) JNK1 and JNK2 protein kinase activities are required for cell proliferation. Primary MEFs prepared from wild-type embryos (J1+/+ J2HG/MG or J12/2 J2MG/MG) were incubated in medium supplemented with solvent (DMSO) or 10 mM 1NM-PP1. The cells were harvested, and the expression of cJun and tubulin was examined by immunoblot analysis. (b) JNK1 and JNK2 protein kinase activities are required for cell proliferation. Primary MEFs prepared from wild-type embryos (J1+/+ J2+/+) or mutated embryos (J1+/+ J2MG/MG or J12/2 J2MG/MG) were incubated in medium supplemented with solvent (DMSO) or 10 mM 1NM-PP1. Relative cell numbers were measured by staining with crystal violet. (Reprinted with permission from reference [234], Copyright 2006 by Elesvier).



**Fig. (10).** Cofactor fingerprinting with STD NMR applied to RSP2. (**a**) 1D <sup>1</sup>H NMR spectrum of the mixture of cAMP (\*), cGMP ( $\circ$ ) and cCMP ( $\diamond$ ). (**b**) STD NMR spectrum of the mixture of RSP2 and the three cyclic nucleotides. (Reprinted with permission from reference [274], Copyright 2005 by Elesvier). Ribbon diagrams showing the view of the (**c**) intersubunit cavity with the invariant residues mapped onto the structure. (**d**) showing the location of amide groups (red spheres) in HI0719 that have perturbed chemical shifts when 2-ketobutyrate is added. (Reprinted with permission from reference [237], Copyright 2003 by American Chemical Society).



Fig. (11). Pictorial flow diagram of FAST-NMR assay. (Reprinted with permission from reference [121], Copyright 2006 by American Chemical Society).

Parsons *et al.* (2003) used a directed library of 93 compounds based on prior biological studies and the structure of HI0719 that suggested a potential role in either isoleucine biosynthetic pathways, translation inhibition, purine regulation, or 2-aminomuconate deaminase activity [237]. 2-ketobutyrate and three other isoleucine analogs were shown to bind HI0719 based on chemical shift perturbations in a 2D <sup>1</sup>H-<sup>15</sup>N HSQC experiment. The residues that experienced chemical shift changes in the ligand are all located in the putative ligandbinding cavity (Fig. **10c,d**). These results suggest that HI0719 is involved in the isoleucine biosynthetic pathway [237].

## **FAST-NMR SCREEN**

The Functional Annotation Screening Technology by NMR (FAST-NMR) combines structural biology, NMR ligand affinity screens and bioinformatics in a high throughput mode to provide a functional assignment to hypothetical proteins [121]. The FAST-NMR assay contains four major steps: (i) identify functional ligands that bind the protein, (ii) use the ligands to determine protein-ligand co-structures, (iii) use the co-structure with bioinformatics to infer function and (iv) use the ligand-binding profile to infer function [238, 239]. The FAST-NMR protocol is outlined in Fig. 11. FAST-NMR uses a screening library composed of aminoacids, carbohydrates, co-factors, fatty-acids, hormones, metabolites, neurotransmitters, nucleic acids and vitamins that bind specific proteins or a functional class of proteins [235]. The functional library is screened against a hypothetical protein using a tiered set of NMR experiments to minimize resources and increase throughput.

In a tiered approach, the first NMR experiment is more suitable for efficiently filtering the large compound library and providing preliminary binding information. 1D <sup>1</sup>H linebroadening (LB) NMR experiments are applied in the FAST-NMR assay, where an increase in line-width for the ligand NMR signals in the presence of the protein will indicate a positive binding event [173, 174]. The second NMR experiment is more informative, but also more resource intensive, so it is only conducted on positive results from the first NMR experiment. The second NMR experiment further filters the "hits" for determining a protein-ligand structure by identifying compounds that interact with the protein in a defined binding-site determined from chemical shift perturbations while eliminating non-specific binders that are not functionally related to the activity of the protein. In the FAST-NMR assay, the positive hits from the 1D LB experiments are further evaluated by obtaining a 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectrum [179] using an <sup>15</sup>N labeled protein [240]. A binding interaction is determined from the observation of chemical shift perturbations for the protein in the presence of the ligand. Mapping the protein residues that incurred a chemical shift change onto the protein surface identifies the ligand binding site. The 2D <sup>1</sup>H-<sup>15</sup>N HSQC NMR experiment also eliminates non-informative, non-specific binders by the lack of chemical shift changes that cluster together on the protein surface. A protein-ligand co-structure is then rapidly determined by using AutoDock [102], where a "grid" based on the NMR determined binding site directs the docking of the ligand.

The protein-ligand co-structure combined with the identity of the functional ligands is an important tool for the detailed bioinformatic analysis of the hypothetical protein. The <u>Comparison of Protein Active-Site Structures (CPASS)</u> database and software is used as part of the FAST-NMR assay to aid in functional annotations [241]. CPASS determines a sequence and structural alignment of the *experimentally*  identified ligand-defined active-site from our FAST-NMR assay with corresponding *experimental* ligand-defined active-sites of proteins of known function from the PDB [242]. A functional annotation can be made by using CPASS to identify proteins of known function that share similar active-site sequence and structural characteristics with the hypothetical protein.

The CPASS database contains ~21,000 ligand-defined protein active-sites identified from the ~55,000 X-ray and NMR structures that are currently available in the PDB [242]. The CPASS database also includes protein active-sites defined by the presence of small peptides or oligonucleotides ( $\leq$  13 residues). Conversely, the CPASS database excludes 121 common and abundant buffer reagents (2-mercaptanol, glycerol), salts (Na<sup>+</sup>, Cl<sup>-</sup>, SO<sub>4</sub><sup>-</sup>), solvents (water, MES, DMSO) and chemical fragments or clusters (acetyl, methyl) that generally exhibit promiscuous or non-specific binding irrelevant to functional activity.

The ligands identified from protein-ligand complexes in the PDB are then used to determine ligand-defined activesites within the protein structure. The amino-acid residues that comprise an active-site are identified by having at least one atom that is  $\leq 6\text{\AA}$  from any ligand atom. The CPASS program determines the optimal sequence and structural alignment between two compared active-sites *without* maintaining sequence connectivity by maximizing a scoring function [241] based on a C $\alpha$  distance-weighted BLOSUM62 score [243]. The CPASS alignment is also independent of the bound ligands.

The FAST-NMR assay combined with CPASS was used to assign a function to hypothetical protein SAV1430 from *Staphylococcus aureus*. SAV1430 was determined to be part of a multi-protein complex within the [Fe-S] cluster assembly network that may exhibit activity comparable to NifU or may regulate NifU activity [111, 244-247]. *S. aureus* hypothetical protein SAV0936 is postulated as being a binding partner of SAV1430, where the complex formation may be regulated by phosphorylation of SAV0936.

## NMR METABOLOMICS

NMR analysis of the cellular metabolome provides another potential avenue for the functional analysis of hypothetical proteins. NMR-based metabolomics complements ligand affinity assays, such as FAST-NMR, by providing in vivo information on the biological activity of a protein. NMR metabolomics is primarily being applied in the analysis of biofluids to identify disease markers and monitor drug efficacy and toxicity [248-256]. NMR has also been used to analyze cell lysis to classify clinical bacterial strains [257] and to infer a function for a silent mutation in yeast [258]. Raamsdonk et al. (2001) demonstrates a methodology for functional annotation that does not require a detailed analysis of the metabolome. Instead, the approach requires the identification of a functionally related gene for a comparative analysis. The approach was demonstrated using six deletion yeast mutants, where two genes (PFK26 and PFK27) encode the same enzyme (6phosphofructo-2-kinase). Multiple NMR spectra of yeast cell extracts for all six mutants were analyzed using a combined principal component analysis (PCA) and discriminant function analysis (DFA) (Fig. 12a). The DFA plot of the six yeast mutants clearly indicates five distinct clusters, where data for

genes *PFK26* and *PFK27* form a single cluster as expected. If the function for gene *PFK27*was unknown, then it would be attributed a function similar to gene *PFK26* from the overlapped clustering in the DFA plot (Fig. **12b**).

A similar differential approach for the analysis of NMRbased metabolomics was demonstrated with Aspergillus nidulans deletion mutants [129]. In this method, the metabolome of wild-type and mutant cells in the presence and absence of a protein inhibitor are compared. This provides information on the impact on the metabolome from the inactivated enzyme using a chemical probe. The approach identifies the affected metabolites and the corresponding metabolic pathways, which may be used to assist in the functional annotation of a hypothetical protein. Urate oxidase is an enzyme in the purine degradation pathway that oxidizes urate to 5-hydroxyisourate (Fig. 13a). The A. nidulans uaZ14 mutant eliminates the normal function of urate oxidase. Also, 8-azaxanthine (AZA) is a known inhibitor of urate oxidase. PCA analysis of NMR spectra from lysed wild-type and mutant A. nidulans cells in both the presence and absence of AZA identified two distinct clusters (Fig. 13b). As expected, the wild-type A. nidulans cells formed a separate cluster from the uaZ14 mutant cells and the wild-type cells in the presence of AZA because of the different activity of urate oxidase. The PCA scores plot can also be used to identify changes in metabolite concentration that primarily contributed to these differences. Comparison of the <sup>1</sup>H NMR spectra of the wild-type A. nidulans cells with the uaZ14 mutant cells and the wild-type cells in the presence of AZA identifies a number of metabolites involved in the purine degradation pathway from an increase in concentration due to the inactivation of urate oxidase (Fig. 13c). In principal, this approach can be similarly applied to identify the function of a hypothetical protein.

# PRACTICAL CONSIDERATIONS AND NMR LIMI-TATIONS

The type of NMR experiments that is applicable to a particular research project is dependent on a number of variables and the specifics of the system. A particular protein may not be amenable for analysis by NMR if obtaining the necessary quantity of protein is prohibitive or if the protein solubility is too low for analysis. The amount of protein required is highly dependent on the type of NMR experiment. 1D <sup>1</sup>H STD NMR experiments are the preferred choice to minimize protein usage and avoid isotope labeled protein in a ligand affinity screen. Typically the ligand concentration is at a >30 fold excess relative to the protein in an STD experiment. Conversely, NMR screens based on 2D HSQC experiments may require 100s of milligrams of labeled protein. Similarly, determining a protein structure by NMR will routinely require 10s of milligrams of  ${}^{13}C$  and  ${}^{15}N$  labeled proteins. Structure-based NMR experiments also require relatively high protein concentrations of > 0.5-1 mM that are stable for weeks or longer. Conversely, ligand screening samples require low sample concentrations  $< 10-100 \,\mu\text{M}$  that only require a few hours of stability.

The length of time required to collect a specific NMR experiment also varies greatly. The 1D and 2D NMR experiments used for ligand affinity screens can be acquired in < 15 minutes per sample. Nevertheless, this may still require days to weeks of NMR instrument time for relatively large



**Fig. (12).** Functional Analysis of Gene Paralogs using NMR Metabolomics. Cluster analysis of NMR spectra from cell extracts. (**a**) Flowchart to show chemometric approach used to cluster the NMR data. Step 1: The region around the most prominent NMR peak due to water (4.4 - 5.5 p.p.m.) is removed and the internal standard (0 p.p.m.) used to normalize each ordinate (thus allowing quantitative comparison of spectra) before it, too, is removed from the spectrum. The region beyond 5.5 p.p.m. (the aromatic region) is also removed. The resulting reduced data set describes the subspectral region between 0 p.p.m. and 4.4 p.p.m. (i.e., 1,300 variables). Step 2: PCA transforms the original set of variables to a new set of orthogonal variables called principal components (PCs). Step 3: DFA has "a priori" information based on spectral replicates and uses this to minimize within-group variance and maximize between-group variance. (**b**) DFA plot based on the first eight PC projections from the NMR spectral data. The numbers represent the NMR spectra of extracts of the following strains: (1) FY23.cox5a $\Delta$ ; (2) FY23.ho $\Delta$ ; (3) FY23.p<sup>0</sup>; (4) FY23.pet191 $\Delta$ ; (5) FY23.pfk26 $\Delta$ ; (6) FY23.pfk27 $\Delta$ . (Reprinted with permission from reference [258], Copyright 2001 by Nature Publishing Group).

compound libraries. The experiment time can be significantly reduced by the application of compound mixtures, where a recent study determined the optimal mixture size to minimize data collection time [259]. Determining a protein structure by NMR may require weeks to months of dedicated instrument time. The availability of cryoprobes [260] combined with recent developments in NMR pulse sequences [261] has drastically reduced the experiment time needed to determine a protein structure to a few days at most.

Obtaining a protein structure and complete NMR resonance assignments is also limited by the protein's size. In general, high-resolution NMR structures and assignments can be routinely obtained for proteins <25 kDa using standard <sup>13</sup>C and <sup>15</sup>N protein labeling techniques [136, 240]. The

molecular-weight upper-limit can be extended by the application of deuterium labeling, specific methyl labeling and TROSY-based experiments [262-264]. Some large MW complexes have been determined using these methods: 900kDa GroEL-GroES complex, 300-kDa cylindrical protease ClpP [265], 95 kDa homotrimeric complex of the acyltransferase protein [266], 82.4 kDa of malate synthase [267, 268], 69 kDa  $\alpha$ 1-proteinase inhibitor Pittsburgh-trypsin covalent complex [269], the 45.3 kDa catalytic domain of human BACE1 [270] and the 44 kDa nucleotide-binding domain [271] among others. Protein molecular-weight can also affect the choice of NMR experiment used in a ligandbinding screen. 1D NMR experiments generally perform better for larger molecular-weight proteins where 2D experiments follow the same limitations that affect complete



**Fig. (13).** Functional Analysis of Gene Knockouts (genetic or chemical) using NMR metabolomics. (a) uaZ14 mutant in *A. nidulans* disrupts the oxidation of urate in the purine degradation pathway. (b) The PCA scores plot comparing *A. nidulans* uaZ14 mutant (x), wild-type with 8-azaxanthine (AZA) (**u**), uaZ14 mutant with AZA (•), and wild-type cells (•). (c) Aromatic regions of NMR spectra in order from top to bottom are: uaZ14 mutant, wild-type with drug, and wild-type untreated. Arrows identify the adenine, hypoxanthine, inosine and xanthine peaks which are absent or significantly decreased in AZA untreated wild-type mycelia. The peaks are all normalized to the most intense peak in the spectra. (Reprinted with permission from reference [129], Copyright 2006 by American Chemical Society).

structure determination. Additionally, membrane bound proteins are problematic targets due to the challenges of expressing significant quantities of properly folded and stabile proteins that require membrane or membrane mimics for solubility [272, 273].

# CONCLUSION

The beneficial impacts of the Human Genome Project and the Protein Structure Initiative necessitate a conclusion at functional genomics. While there is an inherent value in knowing the sequence and structure of a protein, it is still imperative that we understand its biological activity to further our understandings of cell biology, development, evolution and physiology. Furthermore, this functional information is essential for the identification of new therapeutic targets and the development of novel drugs. The complexity of analyzing the proteome for functional information requires making a variety of unique measurements of biological activity. A confirmed functional assignment will then come from overlapping annotations from these multiple functional screens. NMR spectroscopy is playing an integral role in functional genomics by providing multiple measurements of the biological activity of a hypothetical protein. This includes roles in: (i) structural genomics (ii) ligand affinity screens, (iii) chemical proteomics and (iv) metabolomics.

In essence, NMR spectroscopy is important for determining the structure of hypothetical proteins, for monitoring the in vivo activity of the protein through the analysis of the metabolome, and for identifying functional ligands that bind the hypothetical protein. Specifically, the efficient identification of ligands that bind a hypothetical protein is an extremely valuable contribution to functional genomics. First, the identity of the ligand, the location of the ligand-binding site and the determination of a protein-ligand co-structure all provide important clues toward understanding the function of a hypothetical protein. Second, the ligand can be used as an important tool in the design and implementation of other functional screens. The compound can be used to find other proteins with similar activity; it can be used in a chemical approach to generate a knockout mutant and it can be used as part of a traditional biological activity assay.

Functional genomics is still in the early stages of development and new technologies for evaluating biological activity will inevitably emerge from this effort. Nevertheless, the versatility of NMR has already established itself as a critical component of functional genomics.

## ABBREVIATIONS

FAST-NMR = Functional Annotation Screening Technology using NMR

CPASS	= Comparison of Protein Active-Site Structures
NMR	= Nuclear Magnetic Resonance
HTS	= High Throughput Screening
1D	= One-dimensional
2D	= Two-dimensional
PCA	= Principal component analysis
PSI	= Protein Structure Initiative (PSI)
DFA	= Discriminant function analysis
PDB	= Protein Data Bank
HMM	= Hidden Markov models
Pth & Pth2	= Peptidyl-tRNA hydrolases
STD	= Saturation transfer difference
LB	= Line-broadening
NOE	= Nuclear Overhauser effect
SAR	= Structure-Activity Relationship
MS	= Mass spectrometry
RGS	= Regulators of G-protein signaling
dAMP	= Deoxyadenosine monophosphate
AZA	= 8-Azaxanthine
RSP2	= Radial spoke protein-2
dAMP	= Deoxyadenosine monophosphate
SAH	= S-adenosylhomocysteine
ChIP	= Chromatin immunoprecipitation
SEC	= Size-exclusion chromatography

#### REFERENCES

[1] Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; Funke, R.; Gage, D.; Harris, K.; Heaford, A.; Howland, J.; Kann, L.; Lehoczky, J.; LeVine, R.; McEwan, P.; McKernan, K.; Meldrim, J.; Mesirov, J.P.; Miranda, C.; Morris, W.; Navlor, J.; Raymond, C.; Rosetti, M.; Santos, R.; Sheridan, A.; Sougnez, C.; Stange-Thomann, N.; Stojanovic, N.; Subramanian, A.; Wyman, D.; Rogers, J.; Sulston, J.; Ainscough, R.; Beck, S.; Bentley, D.; Burton, J.; Clee, C.; Carter, N.; Coulson, A.; Deadman, R.; Deloukas, P.; Dunham, A.; Dunham, I.; Durbin, R.; French, L.; Grafham, D.; Gregory, S.; Hubbard, T.; Humphray, S.; Hunt, A.; Jones, M.; Lloyd, C.; McMurray, A.; Matthews, L.; Mercer, S.; Milne, S.; Mullikin, J.C.; Mungall, A.; Plumb, R.; Ross, M.; Shownkeen, R.; Sims, S.; Waterston, R.H.; Wilson, R.K.; Hillier, L.W.; McPherson, J.D.; Marra, M.A.; Mardis, E.R.; Fulton, L.A.; Chinwalla, A.T.; Pepin, K.H.; Gish, W.R.; Chissoe, S.L.; Wendl, M.C.; Delehaunty, K.D.; Miner, T.L.; Delehaunty, A.; Kramer, J.B.; Cook, L.L.; Fulton, R.S.; Johnson, D.L.; Minx, P.J.; Clifton, S.W.; Hawkins, T.; Branscomb, E.; Predki, P.; Richardson, P.; Wenning, S.; Slezak, T.; Doggett, N.; Cheng, J.F.; Olsen, A.; Lucas, S.; Elkin, C.; Uberbacher, E.; Frazier, M.; Gibbs, R.A.; Muzny, D.M.; Scherer, S.E.; Bouck, J.B.; Sodergren, E.J.; Worley, K.C.; Rives, C.M.; Gorrell, J.H.; Metzker, M.L.; Naylor, S.L.; Kucherlapati, R.S.; Nelson, D.L.; Weinstock, G.M.; Sakaki, Y.; Fujiyama, A.; Hattori, M.; Yada, T.; Toyoda, A.; Itoh, T.; Kawagoe, C.; Watanabe, H.; Totoki, Y.; Taylor, T.; Weissenbach, J.; Heilig, R.; Saurin, W.; Artiguenave, F.; Brottier, P.; Bruls, T.; Pelletier, E.; Robert, C.; Wincker, P.; Smith, D.R.; Doucette-Stamm, L.; Rubenfield, M.; Weinstock, K.; Lee, H.M.; Dubois, J.; Rosenthal, A.; Platzer, M.; Nyakatura, G.; Taudien, S.; Rump, A.; Yang, H.; Yu, J.; Wang, J.; Huang, G.; Gu, J.; Hood, L.; Rowen, L.; Madan, A.; Qin, S.; Davis, R.W.; Federspiel, N.A.; Abola, A.P.; Proctor, M.J.; Myers, R.M.; Schmutz, J.; Dickson, M.; Grimwood, J.; Cox, D.R.; Olson, M.V.; Kaul, R.; Raymond, C.; Shimizu, N.; Kawasaki, K.; Minoshima, S.; Evans, G.A.; Athanasiou, M.; Schultz, R.; Roe, B.A.; Chen, F.;

Evans, G.A.; Athanasiou, M.; Schultz, R.; Roe, B.A.; Chen, F.; Pan, H.; Ramser, J.; Lehrach, H.; Reinhardt, R.; McCombie, W.R.; de la Bastide, M.; Dedhia, N.; Blocker, H.; Hornischer, K.; Nordsiek, G.; Agarwala, R.; Aravind, L.; Bailey, J.A.; Bateman, A.; Batzoglou, S.; Birney, E.; Bork, P.; Brown, D.G.; Burge, C.B.; Cerutti, L.; Chen, H.C.; Church, D.; Clamp, M.; Copley, R.R.; Doerks, T.; Eddy, S.R.; Eichler, E.E.; Furey, T.S.; Galagan, J.; Gilbert, J.G.; Harmon, C.; Hayashizaki, Y.; Haussler, D.; Hermjakob, H.; Hokamp, K.; Jang, W.; Johnson, L.S.; Jones, T.A.; Kasif, S.; Kaspryzk, A.; Kennedy, S.; Kent, W.J.; Kitts, P.; Koonin, E.V.; Korf, I.; Kulp, D.; Lancet, D.; Lowe, T.M.; McLysaght, A.; Mikkelsen, T.; Moran, J.V.; Mulder, N.; Pollara, V.J.; Ponting, C.P.; Schuler, G.; Schultz, J.; Slater, G.; Smit, A.F.; Stupka, E.; Szustakowski, J.; Thierry-Mieg, D.; Thierry-Mieg, J.; Wagner, L.; Wallis, J.; Wheeler, R.; Williams, A.; Wolf, Y.I.; Wolfe, K.H.; Yang, S.P.; Yeh, R.F.; Collins, F.; Guyer, M.S.; Peterson, J.; Felsenfeld, A.; Wetterstrand, K.A.; Patrinos, A.; Morgan, M.J.; Szustakowki, J.; de Jong, P.; Catanese, J.J.; Osoegawa, K.; Shizuya, H.; Choi, S.; Chen, Y. J.; Consortium, I.H.G.S. Nature, 2001, 409, 860.

- [2] Venter, C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A.; Gocayne, J.D.; Amanatides, P.; Ballew, R.M.; Huson, D.H.; Wortman, J.R.; Zhang, Q.; Kodira, C.D.; Zheng, X.H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P.D.; Zhang, J.; Miklos, G.L.G.; Nelson, C.; Broder, S.; Clark, A.G.; Nadeau, J.; McKusick, V.A.; Zinder, N.; Levine, A.J.; Roberts, R.J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T.J.; Higgins, M.E.; Ji, R.-R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G.V.; Milshina, N.; Moore, H.M.; Naik, A.K.; Narayan, V.A.; Neelam, B.; Nusskern, D.; Rusch, D.B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z. Y.; Wang, A.; Wang, X.; Wang, J.; Wei, M.-H.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S.C.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M.L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y.-H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N.N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigó, R.; Campbell, M. J.; Sjolander, K.V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y.-H.; Coyne, M.; Dahlke, C.; Mays, A.D.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson, M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X. Science, 2001, 291, 1304.
- [3] Mikkelsen, T.S.; Hillier, L.W.; Eichler, E.E.; Zody, M.C.; Jaffe, D.B.; Yang, S.-P.; Enard, W.; Hellnann, I.; Lindbladd-Toh, K.; Altheide, T.K.; Archidiacono, N.; Bork, P.; Butler, J.; Chang, J.L.;

Cheng, Z.; Chinwalla, A.T.; de Jong, P.; Delehaunty, K.D.; Fronick, C.C.; Fulton, L.L.; Gilad, Y.; Glusman, G.; Gneere, S.; Graves, T.A.; Hayakawa, T.; Hayden, K.E.; Huang, X.; Ji, H.; Kent, W.J.; King, M.-C.; Kulbokaslll, E.J.; Lee, M.K.; Liu, G.; Lopez-Otin, C.; Makova, K.D.; Man, O.; Mardis, E.R.; Mauceli, E.; Miner, T.L.; Nash, W.E.; Nelson, J.O.; Paeaebo, S.; Patterson, N.J.; Pohl, C.S.; Pollard, K.S.; Pruefer, K.; Puente, X.S.; Reich, D.; Rocchi, M.; Rosenbloom, K.; Ruvolo, M.; Richter, D.J.; Schaffner, S.F.; Smit, A.F. A.; Smith, S.M.; Suyama, M.; Taylor, J.; Torrents, D.; Tuzun, E.; Varki, A.; Velasco, G.; Ventura, M.; Wallis, J.W.; Wendl, M.C.; Wilson, R.K.; Lander, E.S.; Waterston, R. H. *Nature*, **2005**, *437*, 69.

- [4] Lindblad-Toh, K.; Wade, C.M.; Mikkelsen, T.S.; Karlsson, E.K.; Jaffe, D.B.; Kamal, M.; Clamp, M.; Chang, J.L.; Kulbokas, E.J.; Zody, M.C.; Mauceli, E.; Xie, X.; Breen, M.; Wayne, R.K.; Ostrander, E.A.; Ponting, C.P.; Galibert, F.; Smith, D.R.; de Jong, P.J.; Kirkness, E.; Alvarez, P.; Biagi, T.; Brockman, W.; Butler, J.; Chin, C.-W.; Cook, A.; Cuff, J.; Daly, M.J.; DeCaprio, D.; Gnerre, S.; Grabherr, M.; Kellis, M.; Kleber, M.; Bardeleben, C.; Goodstadt, L.; Heger, A.; Hitte, C.; Kim, L.; Koepfli, K.-P.; Parker, H.G.; Pollinger, J.P.; Searle, S.M.J.; Sutter, N.B.; Thomas, R.; Webber, C.; Baldwin, J.; Abebe, A.; Abouelleil, A.; Aftuck, L.; Ait-zahra, M.; Aldredge, T.; Allen, N.; An, P.; Anderson, S.; Antoine, C.; Arachchi, H.; Aslam, A.; Ayotte, L.; Bachantsang, P.; Barry, A.; Bayul, T.; Benamara, M.; Berlin, A.; Bessette, D.; Blitshteyn, B.; Bloom, T.; Blye, J.; Boguslavskiy, L.; Bonnet, C.; Boukhgalter, B.; Brown, A.; Cahill, P.; Calixte, N.; Camarata, J.; Cheshatsang, Y.; Chu, J.; Citroen, M.; Collymore, A.; Cooke, P.; Dawoe, T.; Daza, R.; Decktor, K.; DeGray, S.; Dhargay, N.; Dooley, K.; Dooley, K.; Dorje, P.; Dorjee, K.; Dorris, L.; Duffey, N.; Dupes, A.; Egbiremolen, O.; Elong, R.; Falk, J.; Farina, A.; Faro, S.; Ferguson, D.; Ferreira, P.; Fisher, S.; FitzGerald, M.; Foley, K.; Foley, C.; Franke, A.; Friedrich, D.; Gage, D.; Garber, M.; Gearin, G.; Giannoukos, G.; Goode, T.; Goyette, A.; Graham, J.; Grandbois, E.; Gyaltsen, K.; Hafez, N.; Hagopian, D.; Hagos, B.; Hall, J.; Healy, C.; Hegarty, R.; Honan, T.; Horn, A.; Houde, N.; Hughes, L.; Hunnicutt, L.; Husby, M.; Jester, B.; Jones, C.; Kamat, A.; Kanga, B.; Kells, C.; Khazanovich, D.; Kieu, A.C.; Kisner, P.; Kumar, M.; Lance, K.; Landers, T.; Lara, M.; Lee, W.; Leger, J.-P.; Lennon, N.; Leuper, L.; LeVine, S.; Liu, J.; Liu, X.; Lokyitsang, Y.; Lokyitsang, T.; Lui, A.; Macdonald, J.; Major, J.; Marabella, R.; Maru, K.; Matthews, C.; McDonough, S.; Mehta, T.; Meldrim, J.; Melnikov, A.; Meneus, L.; Mihalev, A.; Mihova, T.; Miller, K.; Mittelman, R.; Mlenga, V.; Mulrain, L.; Munson, G.; Navidi, A.; Naylor, J.; Nguyen, T.; Nguyen, N.; Nguyen, C.; Nguyen, T.; Nicol, R.; Norbu, N.; Norbu, C.; Novod, N.; Nyima, T.; Olandt, P.; O'Neill, B.; O'Neill, K.; Osman, S.; Oyono, L.; Patti, C.; Perrin, D.; Phunkhang, P.; Pierre, F.; Priest, M.; Rachupka, A.; Raghuraman, S.; Rameau, R.; Ray, V.; Raymond, C.; Rege, F.; Rise, C.; Rogers, J.; Rogov, P.; Sahalie, J.; Settipalli, S.; Sharpe, T.; Shea, T.; Sheehan, M.; Sherpa, N.; Shi, J.; Shih, D.; Sloan, J.; Smith, C.; Sparrow, T.; Stalker, J.; Stange-Thomann, N.; Stavropoulos, S.; Stone, C.; Stone, S.; Sykes, S.; Tchuinga, P.; Tenzing, P.; Tesfaye, S.; Thoulutsang, D.; Thoulutsang, Y.; Topham, K.; Topping, I.; Tsamla, T.; Vassiliev, H.; Venkataraman, V.; Vo, A.; Wangchuk, T.; Wangdi, T.; Weiand, M.; Wilkinson, J.; Wilson, A.; Yadav, S.; Yang, S.; Yang, X.; Young, G.; Yu, Q.; Zainoun, J.; Zembek, L.; Zimmer, A.; Lander, E. S. Nature, 2005, 438, 803.
- Waterston, R.H.; Lindblad-Toh, K.; Birney, E.; Rogers, J.; Abril, [5] J.F.; Agarwal, P.; Agarwala, R.; Ainscough, R.; Alexandersson, M.; An, P.; Antonarakis, S.E.; Attwood, J.; Baertsch, R.; Bailey, J.; Barlow, K.; Beck, S.; Berry, E.; Birren, B.; Bloom, T.; Bork, P.; Botcherby, M.; Bray, N.; Brent, M.R.; Brown, D.G.; Brown, S.D.; Bult, C.; Burton, J.; Butler, J.; Campbell, R.D.; Carninci, P.; Cawley, S.; Chiaromonte, F.; Chinwalla, A.T.; Church, D.M.; Clamp, M.; Clee, C.; Collins, F.S.; Cook, L.L.; Copley, R.R.; Coulson, A.; Couronne, O.; Cuff, J.; Curwen, V.; Cutts, T.; Daly, M.; David, R.; Davies, J.; Delehaunty, K.D.; Deri, J.; Dermitzakis, E.T.; Dewey, C.; Dickens, N.J.; Diekhans, M.; Dodge, S.; Dubchak, I.; Dunn, D.M.; Eddy, S.R.; Elnitski, L.; Emes, R.D.; Eswara, P.; Eyras, E.; Felsenfeld, A.; Fewell, G.A.; Flicek, P.; Foley, K.; Frankel, W.N.; Fulton, L.A.; Fulton, R.S.; Furey, T.S.; Gage, D.; Gibbs, R.A.; Glusman, G.; Gnerre, S.; Goldman, N.; Goodstadt, L.; Grafham, D.; Graves, T.A.; Green, E.D.; Gregory, S.; Guigo, R.; Guyer, M.; Hardison, R.C.; Haussler, D.; Hayashizaki, Y.; Hillier, L.W.; Hinrichs, A.; Hlavina, W.; Holzer, T.; Hsu, F.; Hua, A.;

Hubbard, T.; Hunt, A.; Jackson, I.; Jaffe, D.B.; Johnson, L. S.; Jones, M.; Jones, T.A.; Joy, A.; Kamal, M.; Karlsson, E.K.; Karolchik, D.; Kasprzyk, A.; Kawai, J.; Keibler, E.; Kells, C.; Kent, W.J.; Kirby, A.; Kolbe, D.L.; Korf, I.; Kucherlapati, R.S.; Kulbokas, E.J.; Kulp, D.; Landers, T.; Leger, J.P.; Leonard, S.; Letunic, I.; Levine, R.; Li, J.; Li, M.; Lloyd, C.; Lucas, S.; Ma, B.; Maglott, D.R.; Mardis, E.R.; Matthews, L.; Mauceli, E.; Mayer, J.H.; McCarthy, M.; McCombie, W.R.; McLaren, S.; McLay, K.; McPherson, J.D.; Meldrim, J.; Meredith, B.; Mesirov, J.P.; Miller, W.; Miner, T.L.; Mongin, E.; Montgomery, K.T.; Morgan, M.; Mott, R.; Mullikin, J.C.; Muzny, D.M.; Nash, W.E.; Nelson, J.O.; Nhan, M.N.; Nicol, R.; Ning, Z.; Nusbaum, C.; O'Connor, M.J.; Okazaki, Y.; Oliver, K.; Overton-Larty, E.; Pachter, L.; Parra, G.; Pepin, K.H.; Peterson, J.; Pevzner, P.; Plumb, R.; Pohl, C.S.; Poliakov, A.; Ponce, T.C.; Ponting, C.P.; Potter, S.; Quail, M.; Reymond, A.; Roe, B.A.; Roskin, K.M.; Rubin, E.M.; Rust, A.G.; Santos, R.; Sapojnikov, V.; Schultz, B.; Schultz, J.; Schwartz, M.S.; Schwartz, S.; Scott, C.; Seaman, S.; Searle, S.; Sharpe, T.; Sheridan, A.; Shownkeen, R.; Sims, S.; Singer, J.B.; Slater, G.; Smit, A.; Smith, D.R.; Spencer, B.; Stabenau, A.; Stange-Thomann, N.; Sugnet, C.; Suyama, M.; Tesler, G.; Thompson, J.; Torrents, D.; Trevaskis, E.; Tromp, J.; Ucla, C.; Ureta-Vidal, A.; Vinson, J.P.; von Niederhausern, A.C.; Wade, C.M.; Wall, M.; Weber, R.J.; Weiss, R.B.; Wendl, M.C.; West, A.P.; Wetterstrand, K.; Wheeler, R.; Whelan, S.; Wierzbowski, J.; Willey, D.; Williams, S.; Wilson, R.K.; Winter, E.; Worley, K.C.; Wyman, D.; Yang, S.; Yang, S.-P.; Zdobnov, E.M.; Zody, M.C.; Lander, E.S. Nature, 2002, 420, 520.

- [6] Poinar, H.N.; Schwarz, C.; Qi, J.; Shapiro, B.; MacPhee, R.D.E.; Buigues, B.; Tikhonov, A.; Huson, D.H.; Tomsho, L.P.; Auch, A.; Rampp, M.; Miller, W.; Schuster, S.C. Science, 2006, 311, 392.
- [7] Frishman, D.; Mokrejs, M.; Kosykh, D.; Kastenmuller, G.; Kolesov, G.; Zubrzycki, I.; Gruber, C.; Geier, B.; Kaps, A.; Albermann, K.; Volz, A.; Wagner, C.; Fellenberg, M.; Heumann, K.; Mewes, H.-W. *Nucleic Acids Res.*, **2003**, *31*, 207.
- [8] Mewes, H.W.; Frishman, D.; Gueldener, U.; Mannhaupt, G.; Mayer, K.; Mokrejs, M.; Morgenstern, B.; Muensterkoetter, M.; Rudd, S.; Weil, B. *Nucleic Acids Res.*, 2002, 30, 31.
- [9] Thomas, P.D.; Kejariwal, A.; Campbell, M.J.; Mi, H.; Diemer, K.; Guo, N.; Ladunga, I.; Ulitsky-Lazareva, B.; Muruganujan, A.; Rabkin, S.; Vandergriff, J.A.; Doremieux, O. *Nucleic Acids Res.*, 2003, 31, 334.
- [10] Bernal, A.; Ear, U.; Kyrpides, N. Nucleic Acids Res., 2001, 29, 126.
- [11] Wheeler, D.L.; Barrett, T.; Benson, D.A.; Bryant, S.H.; Canese, K.; Chetvernin, V.; Church, D.M.; DiCuccio, M.; Edgar, R.; Federhen, S.; Geer, L.Y.; Kapustin, Y.; Khovayko, O.; Landsman, D.; Lipman, D.J.; Madden, T.L.; Maglott, D.R.; Ostell, J.; Miller, V.; Pruitt, K.D.; Schuler, G.D.; Sequeira, E.; Sherry, S.T.; Sirotkin, K.; Souvorov, A.; Starchenko, G.; Tatusov, R.L.; Tatusova, T.A.; Wagner, L.; Yaschenko, E. *Nucleic Acids Res.*, **2007**, *35*, D5.
- [12] Caskey, C.T.; Rossiter, B.J.F. J. Pharm. Pharmacol., 1992, 44, 198.
- [13] Jensen, L.J.; Gupta, R.; Staerfeldt, H.H.; Brunak, S. *Bioinformatics*, 2003, 19, 635.
- [14] Pouliot, Y.; Gao, J.; Su, Q.J.; Liu, G.G.; Ling, X.B. Genome Res., 2001, 11, 1766.
- [15] Brenner, S.E. Nat. Struct. Biol., 2000, 7, 967.
- [16] Burley, S.K.; Almo, S.C.; Bonanno, J.B.; Capel, M.; Chance, M.R.; Gaasterland, T.; Lin, D.; Sali, A.; Studier, F.W.; Swaminathan, S. *Nat. Genet.*, **1999**, 23, 151.
- [17] Burley, S.K. Nat. Struct. Biol., 2000, 7, 932.
- [18] Shin, D.H.; Yokota, H.; Kim, R.; Kim, S.-H. J. Struct. Func. Genomics, 2002, 2, 53.
- [19] Shin, D.H.; Yokota, H.; Kim, R.; Kim, S.-H. Proc. Nat. Acad. Sci. USA, 2002, 99, 7980.
- [20] Pineda-Lucena, A.; Liao, J.; Wu, B.; Yee, A.; Cort, J.R.; Kennedy, M.A.; Edwards, A.M.; Arrowsmith, C.H. Proteins: Struct. Funct. Genet., 2002, 47, 572.
- [21] Schulze-Gahmen, U.; Pelaschier, J.; Yokota, H.; Kim, R.; Kim, S.-H. Proteins: Struct. Funct. Genet., 2003, 50, 526.
- [22] Kumaran, D.; Eswaramoorthy, S.; Gerchman, S.E.; Kycia, H.; Studier, F.W.; Swaminathan, S. Proteins: Struct. Funct. Genet., 2003, 52, 283.
- [23] Das, K.; Xiao, R.; Wahlberg, E.; Hsu, F.; Arrowsmith, C.H.; Montelione, G.T.; Arnold, E. Proteins: Struct. Funct. Genet., 2001, 45, 486.

- [24] von Grotthuss, M.; Plewczynski, D.; Ginalski, K.; Rychlewski, L.; Shakhnovich, E. I. BMC Bioinformatics, 2006, 7, 53.
- [25] Friedberg, I.; Jambon, M.; Godzik, A. Protein Sci., 2006, 15, 1527.
- [26] Bonanno, J.B.; Almo, S.C.; Bresnick, A.; Chance, M.R.; Fiser, A.; Swaminathan, S.; Jiang, J.; Studier, F.W.; Shapiro, L.; Lima, C.D.; Gaasterland, T.M.; Sali, A.; Bain, K.; Feil, I.; Gao, X.; Lorimer, D.; Ramos, A.; Sauder, J.M.; Wasserman, S.R.; Emtage, S.; D'Amico, K.L.; Burley, S.K. J. Struct. Funct. Genomics, 2005, 6, 225.
- [27] Anon Proteins: Struct. Funct. Bioinformatics, 2004, 56, 181.
- [28] Lattman, E. Proteins: Struct. Funct. Bioinformatics, 2004, 54, 611.
  [29] White, R.H. J. Bacteriol., 2006, 188, 3431.
- [30] Pan, Y. Frontiers of Biotechnol. Pharmac., 2001, 2, 1.
- [31] Auffray, C.; Imbeaud, S.; Roux-Rouquie, M.; Hood, L. Comptes Rendus Biologies 2003, 326, 879.
- [32] Friedman, A.; Perrimon, N. Curr. Opin. Genet. Devel., 2004, 14, 470.
- [33] Hughes, T.R.; Robinson, M.D.; Mitsakakis, N.; Johnston, M. Curr. Opin. Microbiol., 2004, 7, 546.
- [34] Sauer, S.; Konthur, Z.; Lehrach, H. Comb. Chem. High Throughput Screen., 2005, 8, 659.
- [35] Brown, P.O.; Botstein, D. Nat. Genet., 1999, 21, 33.
- [36] Giot, L.; Bader, J.S.; Brouwer, C.; Chaudhuri, A.; Kuang, B.; Li, Y.; Hao, Y.L.; Ooi, C.E.; Godwin, B.; Vitols, E.; Vijayadamodar, G.; Pochart, P.; Machineni, H.; Welsh, M.; Kong, Y.; Zerhusen, B.; Malcolm, R.; Varrone, Z.; Collis, A.; Minto, M.; Burgess, S.; McDaniel, L.; Stimpson, E.; Spriggs, F.; Williams, J.; Neurath, K.; Ioime, N.; Agee, M.; Voss, E.; Furtak, K.; Renzulli, R.; Aanensen, N.; Carrolla, S.; Bickelhaupt, E.; Lazovatsky, Y.; DaSilva, A.; Zhong, J.; Stanyon, C.A.; Finley, R.L., Jr.; White, K.P.; Braverman, M.; Jarvie, T.; Gold, S.; Leach, M.; Knight, J.; Shimkets, R.A.; McKenna, M.P.; Chant, J.; Rothberg, J.M. Science, 2003, 302, 1727.
- [37] Kang, S.-H.; Lim, W.-S.; Kim, K.-H. Mol. Cells, 2004, 18, 122.
- [38] Phizicky, E.; Bastiaens, P.I.H.; Zhu, H.; Snyder, M.; Fields, S. *Nature*, 2003, 422, 208.
- [39] Jensen, O.N. Nat. Rev. Mol. Cell Biol., 2006, 7, 391.
- [40] Kim, T.H.; Ren, B. Annu. Rev. Genom. Hum. Genet., 2006, 7, 81.
- [41] Solomon, M.J.; Varshavsky, A. Proc. Nat. Acad. Sci. USA, 1985, 82, 6470.
- [42] Durick, K.; Mendlein, J.; Xanthopoulos, K.G. Genome Res., 1999, 9, 1019.
- [43] Parinov, S.; Sundaresan, V. Curr. Opin. Biotechnol., 2000, 11, (2), 157.
- [44] Vidan, S.; Snyder, M. Curr. Opin. Biotechnol., 2001, 12, 28.
- [45] Vanhecke, D.; Janitz, M. Drug Discov. Today, 2005, 10, 205.
- [46] Forster, J.; Gombert, A.K.; Nielsen, J. Biotechnol. Bioengin., 2002, 79, 703.
- [47] Mendes, P. Brief Bioinformatics, 2002, 3, 134.
- [48] Maercker, C. Biosci. Rep., 2005, 25, 57.
- [49] Schweitzer, B.; Predki, P.; Snyder, M. Proteomics, 2003, 3, 2190.
- [50] Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Nucleic Acid Res., 1997, 25, 3389.
- [51] Thompson, J.D.; Higgins, D.G.; Gibson, T.J. Nucleic. Acid Res., 1994, 22, 4673.
- [52] Pearson, W.R. Methods Mol. Biol., 2000, 132, 185.
- [53] He, J.; Dai, X.; Zhao, X. BMC Bioinformatics, 2007, 8, 53.
- [54] Devos, D.; Valencia, A. Trends Genet., 2001, 17, 429.
- [55] Brenner, S.E. *Trends Genet.*, **1999**, *15*, 132.
- [56] Rost, B. J. Mol. Biol., 2002, 318, 595.
- [57] Orengo, C.A.; Michie, A.D.; Jones, S.; Jones, D.T.; Swindells, M.B.; Thornton, J.M. *Structure*, **1997**, *5*, 1093.
- [58] Andreeva, A.; Howorth, D.; Brenner, S.E.; Hubbard, T.J.P.; Chothia, C.; Murzin, A.G. Nucleic Acid Res., 2004, 32, D226.
- [59] Selengut, J.D.; Haft, D.H.; Davidsen, T.; Ganapathy, A.; Gwinn-Giglio, M.; Nelson, W.C.; Richter, A.R.; White, O. *Nucleic Acid Res.*, 2007, 35, D260.
- [60] Watson, J.D.; Laskowski, R.A.; Thornton, J.M. Curr. Opin. Struct. Biol., 2005, 15, 275.
- [61] Dietmann, S.; Park, J.; Notredame, C.; Heger, A.; Lappe, M.; Holm, L. Nucleic Acid Res., 2001, 29, 55.
- [62] Orengo, C.A.; Taylor, W.R. Methods Enzymol., 1996, 266, 617.
- [63] Shindyalov, I.N.; Bourne, P.E. Protein Engin., 1998, 11, 739.
- [64] Zhang, Y.; Skolnick, J. *Nucleic Acid Res.*, **2005**, *33*, 2302.
- [65] Gibrat, J.-F.; Madej, T.; Bryant, S.H. Curr. Opin. Struct. Biol., 1996, 6, 377.
- [66] Koehl, P. Curr. Opin. Struct. Biol., 2001, 11, (3), 348.

- [67] Mirny, L.A.; Shakhnovich, E.I. J. Mol. Biol., 1999, 291, 177.
- [68] Bauer, F.; Schweimer, K.; Kluver, E.; Conejo-Garcia, J.-R.; Forssmann, W.-G.; Rosch, P.; Adermann, K.; Sticht, H. *Protein Sci.*, 2001, 10, 2470.
- [69] Moy, F.J.; Diblasio, E.; Wilhelm, J.; Powers, R. J. Mol. Biol., 2001, 310, (1), 219.
- [70] Powers, R.; Garrett, D.S.; March, C.J.; Frieden, E.A.; Gronenborn, A.M.; Clore, G.M. *Biochemistry*, **1993**, *32*, 6744.
- [71] Powers, R.; Garrett, D.S.; March, C.J.; Frieden, E.A.; Gronenborn, A.M.; Clore, G.M. Science (Washington, DC, United States), 1992, 256, 1673.
- [72] Gajiwala, K.S.; Chen, H.; Cornille, F.; Roques, B.P.; Reith, W.; Mach, B.; Burley, S.K. *Nature*, **2000**, 403, 916.
- [73] Wu, R.; Skaar, E.P.; Zhang, R.; Joachimiak, G.; Gornicki, P.; Schneewind, O.; Joachimiak, A. J. Biol. Chem., 2005, 280, 2840.
- [74] Zvelebil, M.J.; Barton, G.J.; Taylor, W.R.; Sternberg, M.J.E. J. Mol. Biol., 1987, 195, 957.
- [75] Livingstone, C.D.; Barton, G.J. CABIOS, 1993, 9, 745.
- [76] Powers, R.; Mirkovic, N.; Goldsmith-Fischman, S.; Acton, T.B.; Chiang, Y.; Huang, Y.J.; Ma, L.; Rajan, P.K.; Cort, J.R.; Kennedy, M.A.; Liu, J.; Rost, B.; Honig, B.; Murray, D.; Montelione, G.T. *Protein Sci.*, 2005, 14, 2849.
- [77] de Pereda, J.M.; Waas, W.F.; Jan, Y.; Ruoslahti, E.; Schimmel, P.; Pascual, J. J. Biol. Chem., 2004, 279, 8111.
- [78] Schmitt, E.; Mechulam, Y.; Fromant, M.; Plateau, P.; Blanquet, S. EMBO J., 1997, 16, 4760.
- [79] Kleywegt, G.J. J. Mol. Biol., **1999**, 285, (4), 1887.
- [80] Barker, J.A.; Thornton, J. M. Bioinformatics, 2003, 19, 1644.
- [81] Wallace, A.C.; Borkakoti, N.; Thornton, J.M. Protein Sci., 1997, 6, (11), 2308.
- [82] Cammer, S.A.; Hoffman, B.T.; Speir, J.A.; Canady, M.A.; Nelson, M.R.; Knutson, S.; Gallina, M.; Baxter, S.M.; Fetrow, J.S. J. Mol. Biol., 2003, 334, 387.
- [83] Goldman, B.B.; Wipke, W.T. J. Chem. Info. Comp. Sci., 2000, 40, 644.
- [84] Kinoshita, K.; Nakamura, H. Protein Sci., 2003, 12, 1589.
- [85] Fetrow, J.S.; Skolnick, J. J. Mol. Biol., 1998, 281, 949.
- [86] Guo, T.; Shi, Y.; Sun, Z. Protein Engin. Des. Sel., 2005, 18, 65.
- [87] Kitson, D.H.; Badretdinov, A.; Zhu, Z.-Y.; Velikanov, M.; Edwards, D.J.; Olszewski, K.; Szalma, S.; Yan, L. *Briefings in Bioinformatics*, 2002, *3*, 32.
- [88] Fetrow, J.S.; Godzik, A.; Skolnick, J. J. Mol. Biol., 1998, 282, 703.
- [89] Johnson, J.M.; Church, G.M. Proc. Nat. Acad. Sci. USA, 2000, 97, 3965.
- [90] Gutteridge, A.; Thornton, J.M. Trends Biochemi. Sci., 2005, 30, 622.
- [91] Laskowski, R.A.; Watson, J.D.; Thornton, J.M. J. Mol. Biol., 2005, 351, 614.
- [92] Tian, W.; Arakaki, A.K.; Skolnick, J. Nucleic Acid Res., 2004, 32, 6226.
- [93] Panchenko, A.R.; Kondrashov, F.; Bryant, S. Protein Sci., 2004, 13, 884.
- [94] Burgoyne, N.J.; Jackson, R.M. Bioinformatics, 2006, 22, 1335.
- [95] Laurie, A.T.R.; Jackson, R.M. Bioinformatics, 2005, 21, 1908.
- [96] Pazos, F.; Sternberg, M. J. E. Proc. Nat. Acad. Sci. USA, 2004, 101, 14754.
- [97] Huang, B.; Schroeder, M. BMC Struct. Biol., 2006, 6, No pp given.
- [98] Brady, G.P., Jr.; Stouten, P.F.W. J. Comp.-Aided Mol. Des., 2000, 14, 383.
- [99] Schmitt, S.; Kuhn, D.; Klebe, G. J. Mol. Biol., 2002, 323, 387.
- [100] Naumann, T.; Matter, H. J. Med. Chem., 2002, 45, 2366.
- [101] Ko, J.; Murga, L.F.; Wei, Y.; Ondrechen, M.J. *Bioinformatics*, 2005, 21, i258.
- [102] Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. J. Comp. Chem., **1998**, *19*, 1639.
- [103] Chakrabarti, R.; Klibanov, A.M.; Friesner, R.A. Proc. Nat. Acad. Sci. USA, 2005, 102, 10153.
- [104] Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Proteins: Struct. Funct. Genetics, 2002, 47, 409.
- [105] Greaves, R.; Warwicker, J. J. Mol. Biol., 2005, 349, 547.
- [106] Ringe, D.; Wei, Y.; Boino, K.R.; Ondrechen, M.J. Cell. Mol. Life Sci., 2004, 61, 387.
- [107] Gutteridge, A.; Bartlett, G.J.; Thornton, J.M. J. Mol. Biol., 2003, 330, 719.
- [108] La, D.; Sutch, B.; Livesay, D. R. Proteins: Struct. Funct. Bioinformatics, 2004, 58,309.

- [109] Lopez-Romero, P.; Gomez, M.J.; Gomez-Puertas, P.; Valencia, A. *Methods in Proteome and Protein Anal.*, **2004**, 319.
- [110] Pupko, T.; Bell Rachel, E.; Mayrose, I.; Glaser, F.; Ben-Tal, N. Bioinformatics, 2002, 18, S71.
- [111] Glaser, F.; Pupko, T.; Paz, I.; Bell, R.E.; Bechor-Shental, D.; Martz, E.; Ben-Tal, N. *Bioinformatics*, 2003, 19, 163.
- [112] Hasson, M.S.; Schlichting, I.; Moulai, J.; Taylor, K.; Barrett, W.; Kenyon, G.L.; Babbitt, P.C.; Gerlt, J.A.; Petsko, G.A.; Ringe, D. *Proc. Nat. Acad. Sci. USA*, **1998**, 95, 10396.
- [113] Holm, L.; Sander, C. Proteins, 1997, 28, 72.
- [114] Montelione, G.T.; Zheng, D.; Huang, Y.J.; Gunsalus, K.C.; Szyperski, T. Nat. Struct. Biol., 2000, 7, 982.
- [115] Staunton, D.; Owen, J.; Campbell, I.D. Acc. Chem. Res., 2003, 36, 207.
- [116] Kennedy, M.A.; Montelione, G.T.; Arrowsmith, C.H.; Markley, J. L. J. Str. Funct. Genomics, 2002, 2, 155.
- [117] Yee, A.; Gutmanas, A.; Arrowsmith, C.H. Curr. Opin. Struct. Biol., 2006, 16, 611.
- [118] Moy, F.J.; Chanda, P.K.; Chen, J.M.; Cosmi, S.; Edris, W.; Skotnicki, J.S.; Wilhelm, J.; Powers, R. *Biochemistry*, **1999**, *38*, 7085.
- [119] Huang, X.; Moy, F.; Powers, R. Biochemistry, 2000, 39, 13365.
- [120] Chen, J.M.; Nelson, F.C.; Levin, J.I.; Mobilio, D.; Moy, F.J.; Nilakantan, R.; Zask, A.; Powers, R. J. Amer. Chem. Soc., 2000, 122, 9648.
- [121] Mercier, K.A.; Baran, M.; Ramanathan, V.; Revesz, P.; Xiao, R.; Montelione, G.T.; Powers, R. J. Amer. Chem. Soc., 2006, 128, 15292.
- [122] Stockman, B.J.; Dalvit, C. Prog. Nucleic. Magn. Reson. Spectr., 2002, 41, 187.
- [123] Hajduk, P.J.; Betz, S.F.; Mack, J.; Ruan, X.; Towne, D.L.; Lerner, C.G.; Beutel, B.A.; Fesik, S.W. J. Biomol. Screen., 2002, 7, 429.
- [124] Modi, S.; Paine, M.J.; Sutcliffe, M.J.; Lian, L.Y.; Primrose, W.U.; Wolf, C.R.; Roberts, G.C.K. *Biochemistry*, **1996**, *35*, 4540.
- [125] Sem, D.S. Exp. Rev. Proteomics, 2004, 1, 165.
  [126] Pullela, P.K.; Sem, D.S. Separation Methods in Proteomics, 2006,
- [127] Griffin, J.L.; Bonney, S.A.; Mann, C.; Hebbachi, A.M.; Gibbons,
- G. F.; Nicholson, J.K.; Shoulders, C.C.; Scott, J. Physiol. Genom., 2004, 17, 140.
- [128] Moing, A.; Maucourt, M.; Renaud, C.; Gaudillere, M.; Brouquisse, R.; Lebouteiller, B.; Gousset-Dupont, A.; Vidal, J.; Granot, D.; Denoyes-Rothan, B.; Lerceteau-Koehler, E.; Rolin, D. Funct. Plant Biol., 2004, 31, 889.
- [129] Forgue, P.; Halouska, S.; Werth, M.; Xu, K.; Harris, S.; Powers, R. J. Proteome Res., 2006, 5, 1916.
- [130] Bartlett, G.J.; Todd, A.E.; Thornton, J.M. Methods Biochem. Anal., 2003, 44, 387.
- [131] Wishart, D. Curr. Pharmaceut. Biotechnol. 2005, 6, 105.
- [132] Nietlispach, D.; Mott, H.R.; Stott, K.M.; Nielsen, P.R.; Thiru, A.; Laue, E.D. *Methods Mol. Biol.*, **2004**, *278*, 255.
- [133] Huang, Y.J.; Moseley, H.N.B.; Baran, M.C.; Arrowsmith, C.; Powers, R.; Tejero, R.; Szyperski, T.; Montelione, G.T. *Methods Enzy*mol., 2005, 394, 111.
- [134] Kanelis, V.; Forman-Kay, J.D.; Kay, L.E. *IUBMB Life*, **2001**, *52*, 291.
- [135] Clore, G.M.; Gronenborn, A.M. Biol. Magn. Reson., 1998, 16, 3.
- [136] Ferentz, A.E.; Wagner, G. Quart. Rev. Biophys., 2000, 33, 29.
- [137] Clore, G.M.; Gronenborn, A.M. Trends Biotechnol., 1998, 16, 22.
- [138] Clore, G.M.; Gronenborn, A.M. Annu. Rev. Biophys. Biophys. Chem., **1991**, 20, 29.
- [139] Spera, S.; Bax, A. J. Am. Chem. Soc., **1991**, 113, 5490.
- [140] Marion, D.; Driscoll, P.C.; Kay, L.E.; Wingfield, P.T.; Bax, A.; Gronenborn, A.M.; Clore, G.M. *Biochemistry*, **1989**, 28, 6150.
- [141] Zuiderweg, E.R.P.; Fesik, S. W. *Biochemistry*, **1989**, *28*, 2387.
- [142] Zuiderweg, E.R.P.; McIntosh, L.P.; Dahlquist, F.W.; Fesik, S.W. J. Magn. Reson., 1990, 86, 210.
- [143] Ikura, M.; Kay, L.E.; Tschudin, R.; Bax, A. J. Magn. Reson., 1990, 86, 204.
- [144] Cornilescu, G.; Delaglio, F.; Bax, A. J. Biomol. NMR, 1999, 13, 289.
- [145] Vuister, G.W.; Bax, A. J. Am. Chem. Soc., 1993, 115, 7772.
- [146] Garrett, D.S.; Kuszewski, J.; Hancock, T.J.; Lodi, P.J.; Vuister, G.W.; Gronenborn, A.M.; Clore, G.M. J. Magn. Reson., Ser. B, 1994, 104, 99.
- [147] Kuszewski, J.; Gronenborn, A.M.; Clore, G.M. J. Magn. Reson. Ser. B, 1995, 107, 293.

- [148] Kuszewski, J.; Gronenborn, A.M.; Clore, G.M. J. Am. Chem. Soc., 1999, 121, 2337.
- [149] Kuszewski, J.; Gronenborn, A.M.; Clore, G.M. Protein Sci., 1996, 5, 1067.
- [150] Kuszewski, J.; Gronenborn, A.M.; Clore, G.M. J. Magn. Reson., 1997, 125, 171.
- [151] Kuszewski, J.; Clore, G.M. J. Magn. Reson., 2000, 146, 249.
- [152] Chen, L.; Oughtred, R.; Berman, H.M.; Westbrook, J. Bioinformatics, 2004, 20, 2860.
- [153] Powers, R.; Acton Thomas, B.; Chiang, Y.; Rajan, P. K.; Cort John, R.; Kennedy Michael, A.; Liu, J.; Ma, L.; Rost, B.; Montelione Gaetano, T. J. Biomol. NMR, 2004, 30, 107.
- [154] Meinnel, T.; Mechulam, Y.; Blanquet, S. *Biochimie*, **1993**, *75*, 1061.
- [155] Schmitt, E.; Guillon, J.M.; Meinnel, T.; Mechulam, Y.; Dardel, F.; Blanquet, S. *Biochimie*, **1996**, 78, 543.
- [156] Menninger, J.R. J. Bacteriol., 1979, 137, 694.
- [157] Heurgue-Hamard, V.; Mora, L.; Guarneros, G.; Buckingham, R.H. EMBO J., 1996, 15, 2826.
- [158] Menez, J.; Buckingham Richard, H.; de Zamaroczy, M.; Campelli Celine, K. Mol. Microbiol., 2002, 45, 123.
- [159] Thornton, J.M.; Todd, A.E.; Milburn, D.; Borkakoti, N.; Orengo, C.A. Nat. Struct. Biol., 2000, 7, 991.
- [160] Takai-Igarashi, T.; Kaminuma, T. In Silico Biol., 1999, 1, 129.
- [161] Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. Nucleic Acid Res., 2004, 32, D277.
- [162] Krieger, C.J.; Zhang, P.; Mueller, L.A.; Wang, A.; Paley, S.; Arnaud, M.; Pick, J.; Rhee, S.Y.; Karp, P.D. *Nucleic Acid Res.*, 2004, 32, D438.
- [163] Gubernator, K.; Boehm, H.J. Methods Princ. Med. Chem., 1998, 6, 15.
- [164] Kubinyi, H. Curr. Opin. Drug Discov. Dev., 1998, 1, 4.
- [165] Sintchak, M.D.; Nimmesgern, E. Immunopharmacol., 2000, 47, 163.
- [166] Mihelich, E.D.; Schevitz, R.W. Biochimica et Biophysica Acta, 1999, 1441, 223.
- [167] Dalvit, C.; Caronni, D.; Mongelli, N.; Veronesi, M.; Vulpetti, A. Curr. Drug Discov. Technol., 2006, 3, 115.
- [168] Pellecchia, M. Chem. Biol., 2005, 12, 961.
- [169] Peng, J.W.; Moore, J.; Abdul-Manan, N. Progr. Nucleic. Magn. Reson. Spectr., **2004**, 44, 225.
- [170] Lepre, C.A. Methods Princip. Med. Chem., 2003, 16, 391.
- [171] Powers, R. J. Struct. Funct. Genom., 2002, 2, 113.
- [172] Mestres, J.; Veeneman, G.H. J. Med. Chem., 2003, 46, 3441.
- [173] Rossi, C.; Donati, A.; Sansoni, M.R. Chem. Phys. Lett., 1992, 189, 278.
- [174] Hajduk, P.J.; Olejniczak, E.T.; Fesik, S.W. J. Am. Chem. Soc., 1997, 119, 12257.
- [175] Lin, M.; Shapiro, M.J.; Wareing, J.R. J. Am. Chem. Soc., 1997, 119, 5249.
- [176] Lin, M.; Shapiro, M.J.; Wareing, J.R. J. Org. Chem., 1997, 62, 8930.
- [177] Waldeck, A. R.; Kuchel, P. W.; Lennon, A. J.; Chapman, B. E. Prog. Nucleic Magn. Reson. Spectrosc., 1997, 30, 39.
- [178] Wüthrich, K., NMR of proteins and nucleic acids. ed.; John Wiley & Sons, Inc.: New York, 1986; pp. 292.
- [179] Otting, G. Curr. Opin. Struct. Biol., 1993, 3, 760.
- [180] Shirakawa, M.; Lee, S.J.; Takimoto, M.; Matsuo, H.; Akutsu, H.; Kyogoku, Y. J. Mol. Struct., 1991, 242, 355.
- [181] Ni, F. Prog. Nucleic Magn. Reson. Spectrosc., 1994, 26, 517.
- [182] Vogtherr, M.; Peters, T. J. Am. Chem. Soc., 2000, 122, 6093.
- [183] Leone, M.; Freeze, H.H.; Chan, C.S.; Pellecchia, M. Curr. Drug Discov. Technol., 2006, 3, (2), 91.
- [184] Mayer, M.; Meyer, B. Angewandte Chemie, Int. Ed., **1999**, 38, 1784.
- [185] Dalvit, C.; Pevarello, P.; Tato, M.; Veronesi, M.; Vulpetti, A.; Sundstrom, M. J. Biomol. NMR, 2000, 18, 65.
- [186] Hajduk, P.J.; Mack, J.C.; Olejniczak, E.T.; Park, C.; Dandliker, P.J.; Beutel, B.A. J. Am. Chem. Soc., 2004, 126, 2390.
- [187] Chen, A.; Shapiro, M.J. J. Am. Chem. Soc., 1998, 120, 10258.
- [188] Shuker, S.B.; Hajduk, P.J.; Meadows, R.P.; Fesik, S.W. Science (Washington, DC), **1996**, 274, 1531.
- [189] Erlanson, D.A. *Curr. Opin. Biotechnol.*, **2006**, *17*, 643.
- [190] Huth, J.R.; Sun, C.; Sauer, D.R.; Hajduk, P.J. *Methods Enzymol.*,
- **2005**, *394*, 549.

- [191] Shuker, S.B.; Hajduk, P.J.; Meadows, R.P.; Fesik, S.W. Science, 1996, 274, 1531.
- [192] 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.
- [193] 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.
- [194] Hajduk, P.J.; Gerfin, T.; Boehlen, J.-M.; Haeberli, M.; Marek, D.; Fesik, S.W. J. Med. Chem., 1999, 42, 2315.
- [195] 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.
- [196] Moseley, H.N.B.; Monleon, D.; Montelione, G.T. *Methods Enzy*mol., 2001, 339, 91.
- [197] Petros, A.M.; Dinges, J.; Augeri, D.J.; Baumeister, S.A.; Betebenner, D.A.; Bures, M.G.; Elmore, S.W.; Hajduk, P.J.; Joseph, M.K.; Landis, S.K.; Nettesheim, D.G.; Rosenberg, S.H.; Shen, W.; Thomas, S.; Wang, X.; Zanze, I.; Zhang, H.; Fesik, S.W. J. Med. Chem., 2006, 49, 656.
- [198] Zartler, E.R.; Hanson, J.; Jones, B.E.; Kline, A.D.; Martin, G.; Mo, H.; Shapiro, M.J.; Wang, R.; Wu, H.; Yan, J. J. Am. Chem. Soc., 2003, 125, 10941.
- [199] Liu, G.; Xin, Z.; Pei, Z.; Hajduk, P. J.; Abad-Zapatero, C.; Hutchins, C.W.; Zhao, H.; Lubben, T.H.; Ballaron, S.J.; Haasch, D.L.; Kaszubska, W.; Rondinone, C.M.; Trevillyan, J.M.; Jirousek, M.R. J. Med. Chem., 2003, 46, 4232.
- [200] Liu, G.; Xin, Z.; Liang, H.; Abad-Zapatero, C.; Hajduk, P.J.; Janowick, D.A.; Szczepankiewicz, B.G.; Pei, Z.; Hutchins, C.W.; Ballaron, S.J.; Stashko, M.A.; Lubben, T.H.; Berg, C.E.; Rondinone, C.M.; Trevillyan, J.M.; Jirousek, M.R. J. Med. Chem., 2003, 46, 3437.
- [201] 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.
- [202] Yamazaki, T.; Otomo, T.; Oda, N.; Kyogoku, Y.; Uegaki, K.; Ito, N.; Ishino, Y.; Nakamura, H. J. Am. Chem. Soc., 1998, 120, 5591.
- [203] Xu, R.; Ayers, B.; Cowburn, D.; Muir, T.W. Proc. Nat. Acad. Sci. USA, 1999, 96, 388.
- [204] Jahnke, W.; Ruedisser, S.; Zurini, M. J. Am. Chem. Soc., 2001, 123, 3149.
- [205] Moore, J.; Abdul-Manan, N.; Fejzo, J.; Jacobs, M.; Lepre, C.; Peng, J.; Xie, X. J. Synchr. Rad., 2004, 11, 97.
- [206] Lepre, C.A.; Peng, J.; Fejzo, J.; Abdul-Manan, N.; Pocas, J.; Jacobs, M.; Xie, X.; Moore, J.M. Comb. Chem. High Throughput Screen., 2002, 5, 583.
- [207] Fejzo, J.; Lepre, C.A.; Peng, J.W.; Bemis, G.W.; Ajay; Murcko, M.A.; Moore, J.M. Chem. Biol., 1999, 6, 755.
- [208] Matsuzawa, A.; Ichijo, H. J. Biochem., 2001, 130, 1.
- [209] Fejzo, J.; Lepre, C.A.; Peng, J.W.; Bemis, G.W.; Ajay; Murcko, M.A.; Moore, J.M. Chem. Biol., 1999, 6, 755.
- [210] Siegel, M.M. Curr. Top. Med. Chem., 2002, 2, 13.
- [211] Moy, F.J.; Haraki, K.; Mobilio, D.; Walker, G.; Tabei, K.; Tong, H.; Siegel, M.M.; Powers, R. Anal. Chem., 2001, 73, 571.
- [212] Dohlman, H.G.; Thorner, J. J. Biol. Chem., 1997, 272, 3871.
- [213] Kehrl, J.H. Immunity, **1998**, 8, 1.
- [214] Arshavsky, V.Y.; Pugh, E.N., Jr. Neuron, 1998, 20, 11.
- [215] De Vries, L.; Farquhar, M.G. *Trends Cell Biol.*, **1999**, *9*, 138.
- [216] Powers, R.; Moy, F.; Chanda, P.K.; Cockett, M.I.; Jones, P.; Mason, K.; Semus, S.; Young, K.H.; Shuey, D. Structure of free regulator of G-protein signaling (RGS4) and methods of identifying agonists and antagonists using same. WO 2001085769 A2 20011115 WO 2001-US15100 20010509 US 2000-569836 A 20000511, 2001.
- [217] Moy, F.J.; Chanda, P.K.; Cosmi, S.; Edris, W.; Levin, J.I.; Powers, R. J. Biomol. NMR, 2000, 17, 269.
- [218] Lee, W.; Revington, M.J.; Arrowsmith, C.; Kay, L.E. FEBS Lett., 1994, 350, 87.
- [219] Petros, A.M.; Kawai, M.; Luly, J.R.; Fesik, S.W. FEBS Lett., 1992, 308, 309.
- [220] Gemmecker, G.; Olejniczak, E.T.; Fesik, S.W. J. Magn. Reson., 1992, 96, 199.
- [221] Ikura, M.; Bax, A. J. Am. Chem. Soc., 1992, 114, 2433.
- [222] Nilges, M.; O'Donoghue, S.I. Prog. Nucleic. Magn. Reson. Spectrosc., 1998, 32, 107.

- [223] Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D.S. J. Biomol. NMR, 2002, 22, 165.
- [224] Medek, A.; Hajduk, P.J.; Mack, J.; Fesik, S.W. J. Am. Chem. Soc., 2000, 122, 1241.
- [225] Dominguez, C.; Boelens, R.; Bonvin, A.M.J.J. J. Am. Chem. Soc., 2003, 125, 1731.
- [226] Schieborr, U.; Vogtherr, M.; Elshorst, B.; Betz, M.; Grimme, S.; Pescatore, B.; Langer, T.; Saxena, K.; Schwalbe, H. ChemBio-Chem., 2005, 6, 1891.
- [227] Daub, H. Biochimica et Biophysica Acta, Proteins and Proteomics, 2005, 1754, 183.
- [228] Hu, Y.; Huang, X.; Chen, G.Y.J.; Yao, S.Q. Mol. Biotechnol., 2004, 28, 63.
- [229] Gygi, S.P.; Rist, B.; Gerber, S.A.; Turecek, F.; Gelb, M.H.; Aebersold, R. Nat. Biotechnol., 1999, 17, 994.
- [230] Griffin, T.J.; Lock, C.M.; Li, X.-j.; Patel, A.; Chervetsova, I.; Lee, H.; Wright, M.E.; Ranish, J.A.; Chen, S.S.; Aebersold, R. Anal. *Chem.*, **2003**, 75, 867.
- [231] Cherubini, A.; Bifone, A. Prog. Nucleic. Magn. Reson. Spectrosc., 2003, 42, 1.
- [232] Oros, A.-M.; Shah, N.J. Phys. Med. Biol., 2004, 49, R105.
- [233] Spence, M.M.; Rubin, S.M.; Dimitrov, I. E.; Ruiz, E. J.; Wemmer, D. E.; Pines, A.; Yao, S.Q.; Tian, F.; Schultz, P.G. *Proc. Nat. Acad. Sci. USA*, **2001**, *98*, 10654.
- [234] Jaeschke, A.; Karasarides, M.; Ventura, J.-J.; Ehrhardt, A.; Zhang, C.; Flavell, R.A.; Shokat, K.M.; Davis, R.J. *Mol. Cell*, **2006**, *23*, 899.
- [235] Mercier, K.A.; Germer, K.; Powers, R. Comb. Chem. High Throughput Screen., 2006, 9, 515.
- [236] Cort, J.R.; Yee, A.; Edwards, A.M.; Arrowsmith, C.H.; Kennedy, M.A. J. Mol. Biol., 2000, 302, (1), 189.
- [237] Parsons, L.; Bonander, N.; Eisenstein, E.; Gilson, M.; Kairys, V.; Orban, J. *Biochemistry*, **2003**, *42*, (1), 80.
- [238] Greenbaum, D.C.; Arnold, W.D.; Lu, F.; Hayrapetian, L.; Baruch, A.; Krumrine, J.; Toba, S.; Chehade, K.; Bromme, D.; Kuntz, I. D.; Bogyo, M. Chem. Biol., 2002, 9, 1085.
- [239] Kauvar, L.M.; Higgins, D.L.; Villar, H.O.; Sportsman, J.R.; Engqvist-Goldstein, A.; Bukar, R.; Bauer, K.E.; Dilley, H.; Rocke, D.M. *Chem. Biol.*, **1995**, *2*, 107.
- [240] Lian, L.-Y.; Middleton, D. A. Prog. Nucleic. Magn. Reson. Spectrosc., 2001, 39, 171.
- [241] Powers, R.; Copeland, J.C.; Germer, K.; Mercier, K.A.; Ramanathan, V.; Revesz, P. PROTEINS: Struct. Funct. Bioinformatics, 2006, 65, 124.
- [242] Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. Nucleic Acid Res., 2000, 28, 235.
- [243] Henikoff, S.; Henikoff, J.G. Proc. Nat. Acad. Sci. USA, 1992, 89, 10915.
- [244] Mercier, K.A.; Baran, M.; Xiao, R.; Powers, R.; Montelione, G.T. Abstracts, 39th Midwest Regional Meeting of the American Chemical Society, Manhattan, KS, United States, October 20-22 2004, MID04.
- [245] Baran, M.C.; Huang, J.; Acton, T.; Xiao, R.; Montelione, G.T. In NESG Rosetta Stones Unturned, 2004 International Conference on Structural Genomics, Washington, D.C., November 17-21.
- [246] Frazzon, J.; Dean, D.R. Curr. Opin. Chem. Biol., 2003, 7, 166.
- [247] Dos Santos, P.C.; Smith, A.D.; Frazzon, J.; Cash, V.L.; Johnson, M.K.; Dean, D.R. J. Biol. Chem., 2004, 279, 19705.
- [248] Shockcor, J.P.; Holmes, E. *Curr. Top. Med. Chem.*, **2002**, *2*, 35.
   [249] Robertson, D.G.; Reily, M.D.; Sigler, R.E.; Wells, D.F.; Paterson,
- D.A.; Braden, T.K. *Toxicol. Sci.*, **2000**, *57*, 326. [250] Harrigan, G.G., Goodacre, R. (Eds.), *Metabolic Profiling: Its Role*
- *in Biomarker Discovery and Gene Function Analysis*, Kluwer Academic Publishers, Norwell, Mass, **2003**..
- [251] Nicholson, J.K.; Connelly, J.; Lindon, J.C.; Holmes, E. Nat. Rev. Drug Discov., 2002, 1, 153.
- [252] Navon, G.; Burrows, H.; Cohen, J.S. FEBS Lett., 1983, 162, 320.
- [253] Pfeuffer, J.; Tkac, I.; Provencher, S. W.; Gruetter, R. J. Magn. Reson., 1999, 141, 104.
- [254] Keun, H.C.; Ebbels, T.M.D.; Antti, H.; Bollard, M.E.; Beckonert, O.; Holmes, E.; Lindon, J.C.; Nicholson, J.K. Analytica Chimica Acta, 2003, 490, 265.
- [255] Lamers, R.-J. A.N.; DeGroot, J.; Spies-Faber, E.J.; Jellema, R.H.; Kraus, V.B.; Verzijl, N.; TeKoppele, J.M.; Spijksma, G.K.; Vogels,

J.T.W.E.; van der Greef, J.; van Nesselrooij, J.H.J. J. Nutr., 2003, 133, 1776.

- [256] Nicholson, J.K.; Lindon, J.C.; Holmes, E. Xenobiotica, 1999, 29, 1181.
- [257] Bundy, J.G.; Willey, T.L.; Castell, R.S.; Ellar, D.J.; Brindle, K.M. FEMS Microbiol. Lett., 2005, 242, 127.
- [258] Raamsdonk, L.M.; Teusink, B.; Broadhurst, D.; Zhang, N.; Hayes, A.; Walsh, M.C.; Berden, J.A.; Brindle, K.M.; Kell, D.B.; Rowland, J.J.; Westerhoff, H.V.; Van Dam, K.; Oliver, S.G. Nat. Biotechnol., 2001, 19, 45.
- Mercier, K.A.; Powers, R. J. Biomol. NMR, 2005, 31, 243. [259]
- [260] Moskau, D.; Ritcher, C.; Kovaks, H.; Salzmann, M.; Baselgia, L.; Haeberli, M.; Marek, D.; Schett, O. Spectra Analyse, 2003, 32, 39.
- [261] Atreya, H.S.; Szyperski, T. Methods Enzymol., 2005, 394, 78.
- Riek, R.; Pervushin, K.; Wuthrich, K. Trends Biochem. Sci., 2000, [262] 25, 462.
- [263] Gardner, K.H.; Kay, L.E. Annu. Rev. Biophys. Biomol. Struct., 1998, 27, 357.

Received: April 12, 2007

Revised: July 19, 2007

[264]

[265]

Accepted: July 23, 2007

Jain, N.U.; Wyckoff, T.J.O.; Raetz, C.R.H.; Prestegard, J.H. J. Mol. [266] Biol., 2004, 343, 1379.

Proc. Nat. Acad. Sci. USA, 2005, 102, 16678.

[267] Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L.E. J. Am. Chem. Soc., 2002, 124, 10025.

Sprangers, R.; Gribun, A.; Hwang, P.M.; Houry, W.A.; Kay, L.E.

- Tugarinov, V.; Kay, L.E. Biochemistry, 2005, 44, 15970. [268]
- Peterson, F.C.; Gettins, P.G.W. Biochemistry, 2001, 40, 6284. [269]
- [270] Liu, D.; Wang, Y.-S.; Gesell, J.J.; Wilson, E.; Beyer, B.M.; Wyss, D.F. J. Biomol. NMR, 2004, 29, 425.
- [271] Revington, M.; Zuiderweg, E.R.P. J. Biomol. NMR, 2004, 30, 113. Mokhonova, E.I.; Mokhonov, V.V.; Akama, H.; Nakae, T. Bio-[272]
- chem. Biophys. Res. Commun., 2005, 327, 650.
- [273] Kiefer, H. Biochimica et biophysica Acta, 2003, 1610, 57.
- Yao, H.; Sem, D.S. FEBS Lett., 2005, 579, 661. [274]