Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 29, Issue Number 4, (2012)

Fast Mapping of Biomolecular Interfaces by Random Spin Labeling (RSL)

http://www.jbsdonline.com

Maximilian J. Hartl[1,3]
Björn M. Burmann[1]
Stefan J. Prasch[1]
Clemens Schwarzinger[2]
Kristian Schweimer[1,3]
Birgitta M. Wöhrl[1,3]
Paul Rösch[1,3]
Stephan Schwarzinger[1,3]*

[1]Lehrstuhl Biopolymere, Universität Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany
[2]Institut für Chemische Technologie Organischer Stoffe, Johannes Kepler Universität, Altenbergerstraße 69, 4040 Linz, Austria
[3]Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany

Abstract

Random spin labeling (RSL) is a method for rapid mapping of biomolecular interaction surfaces using an interaction partner with SL and an interaction partner enriched in ^{13}C or ^{15}N nuclei for paramagnetic relaxation enhanced NMR-based detection. The SL reaction is conducted in a manner resulting in a heterogeneous reaction product consisting of different populations of the protein carrying a varying number of spin labels at different positions. Preparation of the paramagnetic probe is complete within a few hours and hence much faster than site selective SL. RSL is applicable to tightly interacting systems but shows its particular strength when applied to systems involving weak or transient contacts.

Introduction

Site-specific labeling with reagents that induce paramagnetic relaxation enhancement (PRE) of nuclear spins has become a widely applied technique for studying biological macromolecules and their complexes by NMR in solution. Site-specific spin labeling (SSSL) has proven particularly useful to study protein folding, to map contacts in transient complexes, to determine protein folds, and study large multicomponent assemblies (1-8). Typically, SSSL requires a surface cysteine, either present in the wild-type protein sequence or artificially introduced, whose free sulfhydryl group is modified with, *e.g.*, a methane thiosulfonate spin label (SL). However, labeling of a specifically introduced cysteine is not straightforward, in particular in cases where a protein already contains cysteine residues. In this case, sequence specificity may be lost and undesired disulfide shuffling may reduce expression yields or lead to incorrectly folded proteins.

In an alternative approach a soluble Gd^{31}-chelate has been used to establish molecular contact sites using methyl-group resonances of ^{13}C-labeled proteins (9). In contrast to SSSL, surface cysteines are not required, but the method lacks the structural detail originating from knowledge of the location of the spin probe and in case of very weak, transient complexes results may be less clear than with SSSL. The same type of spin label, however, can be used to edit spectra to aid spectral assignment (10) and the use of soluble spin labels has been reported to be an efficient tool for structure calculation and rigid body docking of bio-macromolecules (11, 12).

Here we present a modification of the SLAPSTIC experiment developed for lead substance discovery (13). In our approach, one binding partner is isotopically enriched with, *e.g.,* ^{15}N, while the other partner is randomly spin labeled (RSL) at its lysine residues (Figure 1). Typically, lysines are abundant on the protein surface and can be conveniently modified by succinimide chemistry, a fact widely used in biochemical studies. Due to the broad range of the paramagnetic enhancement, PRE is expected to be induced in an ^{15}N-enriched binding partner and can be detected by difference

*Corresponding author:
Stephan Schwarzinger
Phone: +49 (0)921 552046
Fax: +49 (0)921 553684
E-mail: stephan.schwarzinger@uni-bayreuth.de

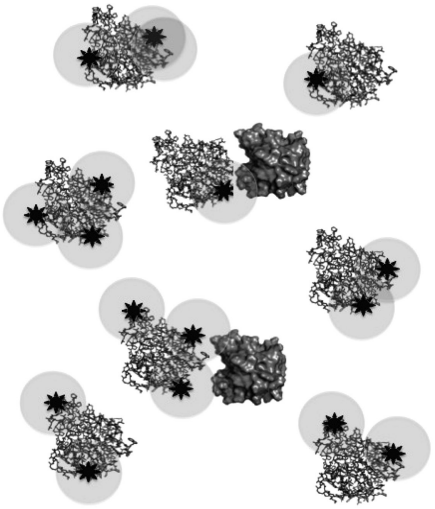

Figure 1: Principle of interaction surface mapping by RSL: RSL results in a sample with an ensemble of protein molecules (wire frames) carrying different numbers of SL (black stars with sphere symbolizing the range of the PRE) attached at different positions on the protein's surface. The binding partner (depicted in surface presentation) is enriched with an NMR-active nucleus, such as ^{13}C or ^{15}N, enabling sensitive NMR-detection. The SL gives rise to PRE, indicated by gray circles. Upon contact of the two protein molecules – one isotope labeled, one spin labeled - a signal decrease at residues in spatial proximity to spin labels can be detected in the ^{13}C- or ^{15}N-labeled binding partner by NMR, which sketches the molecular binding interface.

measurements, comparing spectra with the spin label in its paramagnetic and in its reduced, diamagnetic form. Alternatively, the spin-labeled compound can be displaced from the complex by adding the respective compound carrying no spin label in high excess (14). Quantification of PRE can simply be performed by calculating the ratio of the signal intensities of the spectra with oxidized and reduced spin label or, more robustly, by measuring the paramagnetic relaxation rate in both samples (15).

Materials and Methods

Proteins were expressed in *E. coli* grown in M9 minimal media. In case of ^{15}N labeled proteins, growth temperature, induction and time for harvesting were individually optimized. Proteins were purified following published protocols (16-19).

Spin labeling of the ε-amino groups of lysine residues with 1-oxyl-2,2,5,5-tetramethylpyrolline-3-carboxylate-*N*-hydroxysuccimide ester (OXYL-1-NHS; Toronto Research Chemicals Inc., North York, ON, Canada) was carried out following a published protocol (20) with minor modifications. Proteins were dissolved either directly in labeling buffer (10 mM sodium carbonate, pH 9.2) or the buffer of the protein solution was exchanged with labeling buffer using Vivaspin concentrators (Vivascience, MWCO 5 kDa). Concentrations were adjusted to be in the range of 100-500 µM, allowing a 5- to 10-fold molar access of OXYL-1 -NHS over lysine residues to be added. A stock solution was prepared by dissolving 100 mg of OXYL-1-NHS in 500 µl of DMSO and was stored under an Argon atmosphere at −20°C until used. After addition of OXYL-1-NHS to the protein solution the reaction mixture was incubated in the dark for 1 hour at room temperature and an additional 4 hours at 4° C. To remove excess spin label and to exchange the samples into the desired NMR-buffer, samples were washed with approximately 20 volumes of NMR-buffer using Vivaspin concentrators (Vivascience, MWCO 5 kDa). Spin labels were reduced by the addition of ascorbate from a freshly prepared stock solution (500 mM) to yield an approximately 10-fold molar excess of the reducing agent over the spin label.

Paramagnetic relaxation enhancement ($R_{2,PRE} = R_{2,spinlabel} - R_{2,reduced\_spinlabel}$) was determined by a two point sampling scheme using an HSQC (hetereonuclear single quantum coherence) experiment with an additional spin echo period inserted into the first INEPT (insensitive nuclei enhanced by polarization transfer) transfer (15). Spin echo intervals were set to 0.1 ms and 10.2 ms, respectively. Spectra were recorded on Bruker Avance spectrometers equipped with regular and cryogenic inverse triple resonance probe heads at ^{1}H frequencies of 600.2 MHz and 700.2 MHz, respectively.

Matrix assisted laser desorption ionization time-of-flight (MALDI ToF) mass spectra of SFV-PRshort were recorded on a Bruker Autoflex III Smartbeam in positive linear mode using a double layer method. Lyophilized protein was dissolved in 0.1% trifluoroacetic acid (TFA) in water at a concentration of 0.1% (w/v) and purified by means of a ZipTip C18 pipette tip. After elution the protein solution was mixed with an equal volume of a saturated solution of sinapinic acid in a mixture of 0.1% TFA and acetonitrile (70:30). 0.5 µl of the matrix/analyte solution was then spotted on a ground steel target, which was previously coated with a saturated solution of sinapinic acid in ethanol. At least 1000 spectra were averaged and calibrated against a protein standard mixture.

Results and Discussion

For this study, random spin labeling of the ε-amino groups of lysines was carried out with OXYL-1-NHS in 5- to 10-fold molar excess over the number of lysine residues in 10 mM Na_2CO_3-buffer at pH 9.2 for three different protein complexes. Prior to NMR measurements un-reacted spin label was removed and the randomly spin labeled protein was exchanged for the respective NMR-buffer. Of note, in cases

where only a small number of lysine residues is present in a protein switching to a maleimide based chemistry may be considered. In contrast to succinimide, maleimide based reagents can also be attached to cysteine (21) or even tyrosine residues. Furthermore, spin labeling reagents selective for tyrosine, histidine, arginine, and serine have been described (22-25). Thus, even sequential labeling utilizing chemistries targeted to different amino acid types may be considered to obtain suitable surface coverage. In addition to different labeling chemistries the reactivity can also be influenced by the reaction conditions. Preliminary spin labeling reactions targeting ε-amino groups of lysines that were carried out at pH values of 8 and 7, respectively, and at different reactions times indicate that variation of these reaction parameters is suitable to fine-tune the labeling reaction. Preliminary analysis by mass spectrometry shows that such variations result in a change of the populations of protein carrying a certain number of spin labels (data not shown). In general, a lower pH results in a lower reactivity of the ε-amino groups of lysines and will result in a lower number of spin labels being attached. Labeling reactions have also been carried out with up to 50% DMSO in the reaction mixture.

Randomness of spin labeling is critical for this application. Complete labeling of all lysine residues is likely to produce a protein whose surface topology is severely distorted, which, in turn, may result in decrease or even complete loss of affinity to the interacting partner. Hence, in many cases incomplete labeling is advantageous for complex formation studies. The extent of spin labeling can be controlled by modifying the reaction parameters such as pH or reaction time. Consequently, RSL does not result in a uniform reaction product where all positions are occupied by spin labels. Although this mixture of differently labeled molecules is advantageous for complex formation and subsequent determination of interaction surfaces, it presents an obstacle for quantitative analysis such as determination of dissociation constants.

Using mass spectrometry (MS) and NMR analyses of an ^{15}N-enriched spin labeled protein randomness of spin labeling has been shown here for the separately expressed protease of Simian Foamy Virus (SFV-PRshort). In the absence of an activator the SFV-protease forms transient homodimers that can only be detected by PRE (18, 19, 26, 27). To demonstrate that after the random spin labeling reaction indeed a mixture of the labeled protein carrying different numbers of spin labels at various positions is obtained, the above mentioned labeling protocol was applied to isotopically unlabeled as well as to uniformly ^{15}N enriched SFV-PRshort. The ^{15}N-enriched randomly spin labeled protein was analyzed by NMR spectroscopy. Covalent binding of the spin labels induces small chemical shift changes. Assignments were transferred accordingly from ^{15}N-labeld SFV-PRshort without spin labels. As may be expected, chemical shift changes were most pronounced at positions close to lysine residues, but essentially absent at positions distant from lysines. In fact, spectra with oxidized spin labels show that most resonances close to lysine residues are broadened beyond detection providing evidence for labeling at essentially all lysine residues (Figure 2A). Residues still showing signals cluster in a region of SFV-PRshort distant from lysines (Figure 2B). An exception is Lys15, which apparently is protected from spin labeling, presumably due to the formation of a salt bridge with Glu40. However, NMR proves that in the ensemble of protein molecules in the sample all lysines, except for Lys15, carry spin labels.

Isotopically unlabeled SFV-PRshort, prepared accordingly, was subjected to mass spectrometric analysis, which – in contrast to NMR spectroscopy – is able to characterize the distribution of protein molecules carrying different numbers of spin labels. Figure 2C depicts the corresponding mass spectrum, which clearly shows a distribution of peaks separated by 166 Da, the mass of the spin label. It can unequivocally be seen that proteins with intermediate numbers of spin labels attached correspond to the most abundant populations. The question remains, whether or not the population of proteins carrying, *e.g.*, four spin labels carries these spin labels at identical positions. If this were the case different levels of line

broadening for residues close to different lysine residues would be expected in the NMR experiments. However, this is not the case. Consequently, MS reveals various populations of the protein carrying different numbers of spin labels, while NMR spectroscopy proves that labeling occurs at essentially all possible positions of SFV-PRshort. Hence, labeling is truly random in terms of position and number of spin labels attached. To demonstrate the broad applicability of RSL, the method has been applied to two systems representative for tight binding and transient complex formation, respectively.

The α-C-terminal domain of RNA polymerase (αCTD) and the acidic repeat domain 2 (AR2) of N utilization substance A (NusA) from *E. coli* form a tight complex with a dissociation constant $K_D \approx 10\,\mu M$. Spin-labeling of AR2 leads to signal broadening in the spectrum of ^{15}N-labeled αCTD (Figure 3A) corroborating the complex structure determined by established NMR experiments (28).

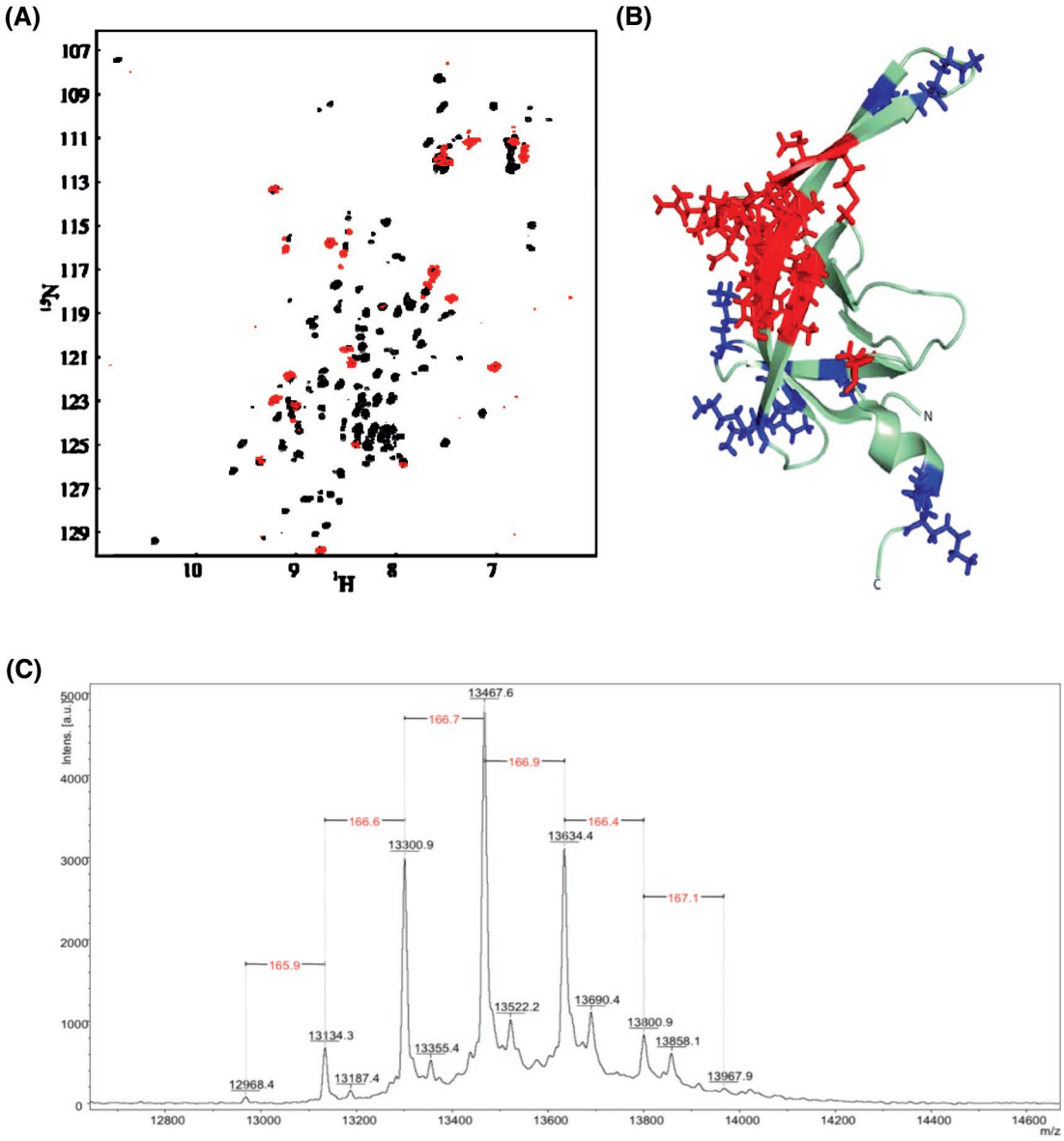

Figure 2: (**A**) HSQC-NMR spectrum of ^{15}N-SFV-PRshort after random spin labeling. The spectrum with the spin label in its paramagnetic form is shown in red, the spectrum with the reduced, diamagnetic form is depicted in black. (**B**) NMR structure of SFV-PRshort (PDB code 2JYS; 19). Lysine residues that can potentially carry a spin label are colored blue. Residues corresponding to signals in the spectrum with paramagnetic spin label are shown in red. (Figure prepared with the program Pymol) (**C**) Mass spectrum of randomly spin labeled SFV-PRshort. Peaks corresponding to populations with different numbers of spin labels are separated by 166 Da (indicated in red). The smaller peaks to the right of these peaks correspond to protein molecules presumably bound to a Ni^{2+}-ion. This is due to the protein purification procedure *via* Ni-affinity chromatography.

Hence, RSL data can be used as additional restraints in structure calculation and docking simulations, similar to the most recently described approach by Madl and colleagues (11).

The *E. coli* two domain protein NusG plays a key role in linking bacterial transcription and translation (29, 17). To determine whether these domains are independent or whether they are involved in formation of a domain-swapped dimer as proposed for the homologous proteins from *Aquifex aeolicus* (30), both domains of *E. coli* NusG were expressed separately. Titration of the NusG-N-terminal domain (NTD) with the NusG-C-terminal domain (CTD) and *vice versa*, one carrying randomly attached spin labels and the other one being ^{15}N-labeled, revealed a transient contact of NusG-NTD and NusG-CTD (Figure 3B) (31).

Hence, RSL quickly reveals the binding interface on the isotopically labeled binding partners in both, tight as well as transient complexes. SFV-PRshort homodimer formation as well as NusG hetero-domain interaction were not evidenced by any other method. By simply reversing isotopic labeling and random spin labeling of the two partner molecules it is possible to map the interaction surface of both binding partners.

In contrast to SSSL, the current method does not reveal the precise relative orientation of the two interacting molecules as identification of the orientation of the two interacting molecules typically requires several different positions to be labeled site-specifically. RSL, however, can provide valuable initial information for further studies aiding identification of sites best suited for SSSL. Moreover, it can help to reduce the number of site-specific mutations required. RSL does not require mutagenesis and specific spin labeling, and purification can be achieved within few hours. Thus, it can significantly reduce time and costs for mapping interaction sites in bio-macromolecules.

In summary, a widely applicable method for mapping intermolecular contact sites is presented. The method's strength includes very fast preparation of the randomly spin-labeled probe and sensitive NMR analysis employing HSQC-type spectra. It is particularly useful when applied to systems involving weak or transient contacts that

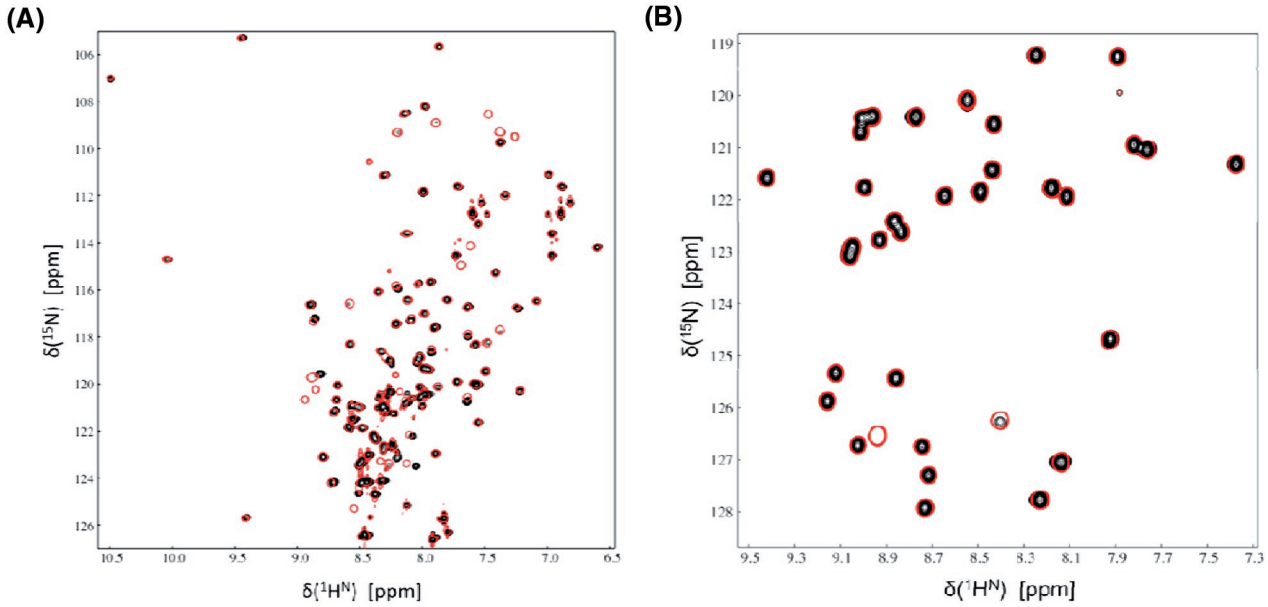

Figure 3: ^{15}N-HSQC spectra of **A**) ^{15}N-labeled αCTD with RSL-AR2 domain of NusA (28). **B**) ^{15}N-labeled NusG-CTD with RSL-NusG-NTD (31). Spectra containing the paramagnetic form of the SL are depicted in black, those with the reduced, diamagnetic SL are shown in red. Residues affected were verified and confirmed by additional NMR-studies in both cases.

cannot be established by any other experimental method. We have demonstrated the technique using ^{15}N-labeled proteins for NMR-detection. RSL can easily be extended to other nuclei. In addition to mapping interaction surfaces, RSL provides a quick tool to check for transient contacts in homo- as well as in hetero-oligomeric complexes and, furthermore, allows determination of contacts as intra- or intermolecular in protein folding studies. As PRE methods are extremely sensitive even at K_D-values in the high μM to mM range (7), it is possible to rule out transient contacts to a certain extent using RSL.

Acknowledgments

StS, PR, and BMW acknowledge support by the Deutsche Forschungsgemeinschaft (Schw901/3-1, Ro617/16-2, Wo630/7-3). MJH, BMB, and SJP contributed equally to this work.

References

1. J. R. Gillespie and D. Shortle. *J Mol Biol 268*, 158-169 (1997).
2. J.R. Gillespie and D. Shortle. *J Mol Biol 268*, 170-184 (1997).
3. M. A. Lietzow, M. Jamin, H. J. Dyson, and P. E. Wright. *J Mol Biol 322*, 655-662 (2002).
4. J. L. Battiste and G. Wagner. *Biochemistry 39*, 5355-5365 (2000).
5. A. N. Volkov, J. A. Worrall, E. Holtzmann and M. Ubbink. *Proc Natl Acad Sci U S A 103*, 18945-18950 (2006).
6. C. Tang, J. Iwahara, and G. M. Clore. *Nature 444*, 383-386 (2006).
7. C. Tang, R. Ghirlando, and G.M. Clore. *J Am Chem Soc 130*, 4048-4056 (2008).
8. K.-P. Wu and J. Baum. *J Am Chem Soc 132*, 5546-5547 (2010).
9. G. Pintacuda and G. Otting. *J Am Chem Soc 124*, 372-373 (2002).
10. R. Kellner, C. Mangels, K. Schweimer, S. J. Prasch, P. R. Weiglmeier, P. Rösch, and S. Schwarzinger. *J Am Chem Soc 131*, 18016-18017 (2009).
11. T. Madl, W. Bermel, and K. Zangger. *Angew Chem Intl Ed 48*, 8259-8262 (2009).
12. T. Madl, T. Güttler, D. Görlich, M. Sattler. *Angew Chem Intl Ed 50*, 3393-3397 (2011).
13. W. Jahnke, S. Rüdisser, M. Zurini. *J Am Chem Soc 123*, 3149-3150 (2001).
14. C. Mangels, R. Kellner, J. Einsiedel, P. R. Weiglmeier, P. Rösch, P. Gmeiner, and S. Schwarzinger. *J Biomol Struct Dyn 28*, 13-22 (2010).
15. J. Iwahara, C. Tang, and G. M. Clore. *J Magn Reson 184*, 185-195 (2007).
16. A. Eisenmann, S. Schwarz, S. J. Prasch, K. Schweimer, and P. Rösch, *Protein Sci 14*, 2018 (2005).
17. B. M. Burmann, K. Schweimer, X. Luo, M. C. Wahl, B. L. Stitt, M. Gottesmann, and P. Rösch. *Science 238*, 501-504 (2010).
18. M. J. Hartl, K. Schweimer, M. H. Reger, S. Schwarzinger, J. Bodem, P. Rösch, and B. M. Wöhrl. *Biochem J 427*, 197-203 (2010).
19. M. J. Hartl, B. M. Wöhrl, P. Rösch, K. Schweimer. *J Mol Biol 381*, 141-149 (2008).
20. J. J. Lawrence, L. Berne, J. L. Ouvrier-Buffet, L. H. Piette. *Eur J Biochem 107*, 263-269 (1980).
21. A. Hata-Tanaka, G. Kawai, K. Yamasaki, Y. Ito, H. Kajiura, J. M. Ha, T. Miyazawa, S. Yokoyama, S. Nishimura. *Biochemistry 28*, 9550-9556 (1989).
22. J.-L. Girardet, J.-J. Lawrence. *Eur J Biochem 132*, 195-200 (1983).
23. R. W. Wien, J. D. Morrisett, H. M. McConnell. *Biochemistry 11*, 3707-3716 (1972).
24. O. H. Hankovszky, K. Hideg, E. v. Goldammer, E. Matuszak, H. Kolkenbrock, H. Tschesche, H. R. Wenzel. *Biochem Biophys Acta 916*, 152-155 (1987).
25. J. D. Morrisett, C. A. Broomfield, and B. E. Hackley, Jr. *J Biol Chem 244*, 5758-5761 (1969).
26. M. J. Hartl, F. Mayr, A. Rethwilm, B. M. Wöhrl. *Retrovirol 7*, 5 (2010).
27. M. J. Hartl, J. Bodem, F. Jochheim, A. Rethwilm, P. Rösch, B. M. Wöhrl. *J Virol 85*, 4462-4469 (2011).
28. K. Schweimer, S. J. Prasch, P. S. Sujatha, M. Bubunenko, M. E. Gottesman, P. Rösch. *Structure 19*, 945-954 (2011).
29. R. A. Mooney, K. Schweimer, P. Rösch, M. Gottesman, and R. Landick. *J Mol Biol 39*, 341-358 (2009).
30. J. R. Knowlton, M. Bubunenko, M. Andrykovitch, W. Guo, K. M. Routzahn, D. S. Waugh, D. L. Court, X. Ji. *Biochemistry 42*, 2275-2281 (2003).
31. B. M. Burmann, U. Scheckenhofer, K. Schweimer, P. Rösch. *Biochem J 435*, 783-789 (2011).

Date Received: September 21, 2011

Communicated by the Editor Ramaswamy H. Sarma