Published in final edited form as:

Chem Biol Drug Des. 2011 July ; 78(1): 14–24. doi:10.1111/j.1747-0285.2011.01096.x.

Transferred NOESY NMR studies of biotin mimetic peptide (FSHPQNT) bound to streptavidin: A structural model for studies of peptide-protein interactions

Dawit Gizachew¹ and Edward Dratz²

¹ Department of Medicine, Texas A&M University Health Science Center College of Medicine and Scott & White Memorial Hospital, Temple, TX 76508

² Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717

Abstract

Protein-protein interactions control signaling, specific adhesion and many other biological functions. The three dimensional structures of the interfaces and bound ligand can be approached with Tr-NOESY NMR, which can be applied to much larger proteins than conventional NMR and requires less concentrated protein. However, it is not clear how accurately the structures of protein-bound peptides can be determined by Tr-NOESY. We studied the structure of a biotin-mimetic peptide (FSHPQNT) bound to streptavidin, since the x-ray structure of the complex is available to 1.74Å resolution and we found that conditions could be adjusted so that the off-rates were fast enough for Tr-NOESY NMR. The off-rate was determined with ¹⁹F NMR, using a para-fluoro-phenylalanine analog of the peptide. A new criterion for a lower limit on kinetic off-rate was found, which allowed accurate structure determination at a slower off-rate. Non-specific binding of the peptide to streptavidin was not significant, since biotin blocked the peptide Tr-NOESY. Protein mediation for the long range peptide Tr-NOESY cross-peaks was corrected by a Tr-NOESY/ROESY averaging procedure. The protein-bound structure of the peptide was determined by Tr-NOESY constrained and simulated annealing. The structure deduced from the NMR was close to the x-ray structure.

Keywords

Peptide-protein interactions; Transferred NOESY; Tr-NOESY accuracy; Ligand off-rates; Fluorine NMR; comparison of x-ray and NMR structures

Protein-protein interactions play important roles in cell signaling and specific adhesion (1). As genomic data proliferates the evolutionary relationships of different signaling cascades and adhesion systems are being increasingly appreciated, and the protein-protein interactions governing signaling and adhesion systems are being revealed (2). In order to understand the detailed mechanisms of cell signaling and specific adhesion as well as to design molecules with pharmacological benefit it is enormously helpful to obtain the structures of the interacting protein-protein or peptide-protein complexes to atomic resolution (3). In most cases of interest it has not yet been possible to determine the 3D structures of the interacting complexes involved in signaling and adhesion, because suitable crystals of the complexes

Address correspondence to: Dawit Gizachew, Ph.D., 702 SW H.K. Dogen Loop, Medical Research Building, Room 111, Temple, TX 76508. dgizachew@tamu.edu and Edward Dratz, Ph.D., Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717. dratz@chemistry.montana.edu.

are not available for diffraction studies and/or the interacting proteins are too large or not soluble enough for determination of full solution structures by NMR.

The interacting interfaces in the protein-protein complexes of interest can often be mapped using peptide competition. Peptides have been found that block protein-protein interactions of interest in a large number of cases, either by testing synthetic sections of the interacting proteins (4) or by selection of effective peptides from random peptide libraries (5,6). The 3D structures of peptides or other ligands, which block protein-protein interactions of interest can be determined, in favorable cases, by transferred nuclear Overhauser effect spectroscopy (Tr-NOESY) NMR when the peptides are bound to their protein binding sites (7–12). Also, it has been shown that Tr-NOESY can provide detailed information on membrane bound receptor-ligand interaction in living cells (13). NMR structure determination is limited due to the size of the protein-protein complex or solubility. However, for ligands (such as bioactive peptides) in exchange with protein binding sites, the Tr-NOESY NMR provides a way to obtain amplified structural information on the bound ligand, which is applicable to much higher molecular weight systems, even if solubility is ten to one hundred times too low for full structure determination by NMR.

When a ligand binds to a protein binding site, the slow tumbling of the protein greatly enhances the efficiency of the proton-proton nuclear Overhauser effect (NOE) cross relaxation between protons of the ligand. The NOE cross relaxation changes the spin populations of the proton energy levels (in a magnetic field), as a function of the spatial proximity of the protons. If the off-rate of the ligand is sufficiently fast, the proton spin populations act as a “magnetic memory” and can be transferred to and read on the sharp NMR spectrum of the free ligand. It is not necessary for the free ligand to retain any vestige of the bound conformation, as the proton proximity information in the bound conformation is carried by the magnetic spin level populations. Small molecules have positive NOE enhancements and large molecules have negative NOE enhancements. The NOE of intermediate sized molecules, such as medium sized peptides (about 5–15 amino acids) the NOE nearly cancels to zero (14), but upon binding to large proteins the ligands acquire large NOE negative enhancements. Thus, NOE enhancements can be used to determine the protein-bound structures of ligands using the Tr-NOESY experiment, under certain conditions. If the ligand off-rate is fast compared to the bound NOE cross relaxation rate, that the observed Tr-NOESY signal intensities (I_{ij}^{obs}) are a weighted sum of the fraction bound (f_b) times the bound intensities (I_{ij}^{bound}) and the fraction free (f_f) times the free intensities (I_{ij}^{free}) for each proton pair ($I_{ij}^{obs} = f_b \cdot I_{ij}^{bound} + f_f \cdot I_{ij}^{free}$) (15). Any Tr-NOESY intensities of the free ligand can be independently measured, and subtracted from the observed signals to obtain the Tr-NOESY intensities of the bound ligand. The NOESY intensities derived for the bound ligand can then be related to inter-proton distances in the bound ligand in the usual ways. Though the fraction of the ligand bound (f_b) may be modest (1–10%), due to the slow tumbling of the ligand bound to the large protein, the NOE intensities of the bound peptides are amplified for structurally useful signals to be obtained from the bound state.

Despite the advantages indicated above, there are a number of potential difficulties in using Tr-NOESY NMR for determining accurate structures of protein-bound ligands (16). Including the need for relatively fast off-rates from the protein binding sites and concerns that Tr-NOESY signals might originate in part from non-specific binding, broad protein NOESY background making it difficult to detect weak, long range Tr-NOESY signals, potential problems of protein proton mediation and distortion of ligand NOE cross relaxation, possible multistage binding mechanisms or multiple bound conformations and possible difficulties with non-uniform local motions in the bound peptides.

In order to explore factors that affect the accuracy of structures of protein-bound peptides obtained from Tr-NOESY, we sought a system where a x-ray crystal structure of a peptide-protein complex was available at high resolution and where the conditions could be adjusted so that the peptide ligand is in sufficiently fast exchange. Most of the peptide-protein complexes with known x-ray structures have much too high an affinity and thus too slow an off-rate to be suitable for Tr-NOESY. A biotin-mimetic peptide that bind to streptavidin, by selection from random peptide libraries (17) appear to provide an excellent model system for evaluation and optimization of Tr-NOESY methodology. The highest affinity peptide (FSHPQNT) had a dissociation constant of about 10^{-4} M at pH=7 (18), much lower affinity for streptavidin than biotin ($K_d \approx 10^{-14}$ M). However, it was found that the enthalpies of binding of the peptides to streptavidin do not differ so greatly from biotin and the energetics of the binding is not so different. The entropy of binding of the linear peptides is unfavorable, since they are flexible in solution and must adopt a conformationally restricted structure when bound.

Sufficiently fast ligand off-rate kinetic is an essential requirement for obtaining accurate protein-bound peptide structures by Tr-NOESY (19). We have found that peptide ligand off-rates can be measured conveniently using ^{19}F substituted ligands and ^{19}F NMR linewidth analysis. With optimized ligand binding kinetics, assessment of non-specific ligand binding, removal of broad protein background and correcting for protein mediated peptide Tr-NOESY, we have obtained the streptavidin-bound structure of the linear biotin-mimetic peptide, using Tr-NOESY, that is nearly as accurate as the high quality 1.74 Å x-ray structure (20). The peptide Tr-NOESY approach validated in the present paper together with other NMR methods such as saturation transfer difference (STD) that complements the Tr-NOESY and provides information on ligand protons in contact with a large protein or receptor by elucidating the binding surface promises to be widely applicable to revealing structural features of the active interfaces of interacting protein systems, on complexes containing large proteins, and on membrane protein complexes involved in signaling and specific adhesion (21,22).

Materials and methods

Peptide synthesis and preparation of streptavidin-peptide samples

FSHPQNT, the N-trifluoroacetyl analogue of FSHPQNT (CF₃-CO-FSHPQNT) and the parafluoro phenylalanine analog (parafluoro-phe-SHPQNT) peptides were synthesized on a continuous flow Milligen 9050 automated peptide synthesizer using Fmoc chemistry. The Fmoc-para-fluoro-L-phenylalanine was from Advanced Chem Tech and used without further purification. The crude peptides were purified by preparative HPLC on a Vydac C18 reverse phase column using acetonitrile/water gradients in 0.1% HCl, freeze dried and stored frozen. The masses of the purified peptides were determined using electrospray and MALDI mass spectrometry and were found to be 830 g/mole for FSHPQNT, 927 g/mole for N-CF₃-CO-FSHPQNT and 848 g/mole for parafluoro-phe-SHPQNT, consistent with their theoretical monoprotonated masses. Recombinant, core streptavidin, expressed in *E. Coli*, which is a tetramer and has a 52 kDa molecular weight, was purchased from Boehringer-Mannheim. The recombinant core streptavidin is desirable for these experiments since the native streptavidin is heterogeneous, due to frayed ends and some of the heterogeneous species have poor solubility. Solutions were made of at pH 4.8, 5.5, 6.5 and 7.5 using deuterioacetate or phosphate buffers.

NMR experiments

^{19}F NMR for exchange rate was measured at 470 MHz on a Bruker DRX500 spectrometer in a 5mm $^1\text{H}/^{19}\text{F}$ selective probe, tuned to ^{19}F at different temperatures and pHs. We

referenced all our ^{19}F spectra using -78.5 ppm of CF_3COO^- chemical shift as external reference. However, in order to easily visualize the relative shifts between the bound and free/fused signals of the predominant trans-proline peptide isomer, the peaks of CF_3COO^- for the $\text{CF}_3\text{CO-FSHPQNT}$ peptide and cis-proline isomer peak for para-fluoro-Phe-SHPQNT peptide were used as an internal references and set to 0 ppm at all temperatures. The exchange rates were assessed by two site exchange spectral simulation (8). Tr-NOESY and Tr-ROESY (transferred nuclear overhauser enhancement spectroscopy in the rotating frame) experiments were carried out on the peptide FSHPQNT and FSHPQNT-streptavidin complexes with 2.2 mM peptide and 231 μM protein concentration, which gives approximately 9.5:1 peptide to protein ratio, in 25 mM phosphate or deuterioacetate buffers and 0.02% sodium azide. Spectra were measured using proton detection triple resonance probes with three axis gradients on a DRX500 at Montana State University, DMX750 at the Nuclear Magnetic Resonance Facility at Madison, and a DMX600 at the NIH. The spectra shown were measured at 300 K and pH 4.8. Tr-NOESY was recorded at 75, 100, 150, 200 and 300 ms mixing times using the 3-9-19 Watergate pulse sequence for water suppression.

A spin-echo filter was used before the t_1 evolution time to reduce the protein NOESY background. We optimized the spin-echo filter time by comparing The Tr-NOESY cross peaks before and after spin-echo filter and found that a 4 msec spin-echo filter removes the protein background without affecting the intensities of the Tr-NOESY peaks of the peptide. Frequency discrimination in the t_1 dimension was achieved using time proportion phase incrementation (TPPI). Spectra were recorded with 512×4096 data matrices and spectral widths of 15 ppm. Spectra were processed using Felix 95 (MSI). Gaussian line broadening window function was used in the t_2 dimension and a shifted sine bell squared window function was applied in the t_1 dimension. Auto baseline (ABL) flattening was used for baseline correction in the t_2 dimension. Conventional 2D TOCSY and COSY methods together with 2D NOESY and ROESY were used for chemical shift assignment of the peptide. In addition, the natural abundance ^{13}C - ^1H HSQC was used to confirm ambiguous assignments.

Tr-ROESY spectra were run at 150 ms spin-lock mixing time and 4 KHz cw spin-lock power with presaturation or watergate for water suppression. The Tr-ROESY intensities were corrected using compensated ROESY pulse sequence for the fall off of spin-lock effectiveness from the irradiation frequency (23). The cross-peak intensities of both the Tr-NOESY and Tr-ROESY were assigned using TOCSY spectra. The volume of the cross-peak intensities and the Isolated Spin Pair Approximation (ISPA) distances were determined using Felix95. The MARDIGRAS (24) program was used to calculate cross-relaxation matrixes to take spin-diffusion within the ligand into account. Structures were refined using simulated annealing (25,26) based on the distance constraints derived from the Tr-NOESY and Tr-ROESY for the protein bound peptides.

Results

The criteria and analysis of kinetic off-rate

Theoretical arguments indicate that if ligand dissociation from the complex is not sufficiently fast compared to the cross relaxation rate of the bound ligand, then distortions occur in Tr-NOESY i.e. $k_{\text{off}} \gg \sigma_{ij}$ (fastest cross relaxation rate of the bound ligand). Fluorine NMR is sensitive to the change in chemical environment and it is small enough to replace hydrogen atoms without causing structural hindrance (27–29), which makes ^{19}F NMR suitable for off-rate analysis. The ^{19}F NMR spectra of the free $\text{CF}_3\text{CO-FSHPQNT}$ and the streptavidin-bound peptide at 278K are shown in figure 1A. The chemical shift reference peak at 0 ppm is the trifluoro acetate (CF_3COO^-), which was present in the synthetic peptide and appeared to be produced by slow hydrolysis of the trifluoroacetate

group from the peptide. The large peak at 0.20 ppm is attributed to the *trans* proline isomer of the CF₃CO-FSHPQNT peptide and the shoulder at 0.22 ppm is attributed to the *cis*-proline isomer of the CF₃CO-FSHPQNT peptide. Minor amount of *cis* proline is expected in proline-containing peptides with the *trans*-proline conformation being the major species. The broader peak at 0.25 ppm, and temperature at 278 K, is attributed to streptavidin-bound *trans*-CF₃CO-FSHPQNT peptide. The intensity of this peak increases with increasing concentration of the streptavidin protein and the *trans* isomer is the predominant species. The presence of separate peaks (0.05 ppm) for the free and the bound peptide indicates slow exchange, on the chemical shift scale at this temperature. At 298 K the peaks of the bound and the free collapse to the single peak, as shown in the upper spectrum in figure 1A. However, it was difficult to obtain accurate values for the peptide off-rate because the separation between the free and the bound signals was only slightly greater than the line width of the free and protein-bound peptide signals. The amino terminal is least well ordered and this may have contributed to the modest shift of the fluorine signal between the bound and the free peptide.

We also used ¹⁹F NMR spectra of para-fluoro-Phe-SHPQNT for the free and the streptavidin bound peptide at different temperatures. As shown in the figure 1B, the peptide is in slow exchange at temperatures 278 K, with the bound peptide peak at 0.4 ppm well separated from the free peptide resonance at -0.15 ppm. The small, somewhat sharper peak that is set to 0 ppm as discussed in the method section in these spectra is attributed to the peptide with a *cis*-proline isomer, which shows no evidence for binding to streptavidin, consistent with the *trans*-proline conformation in the x-ray structure of the bound complex (20). The para-fluoro-Phe-SHPQNT shows over ten times the shift (0.55 ppm) from the free peptide resonance compared to the N-terminal trifluoroacetate derivative of the same peptide and the para-fluoro-Phe-SHPQNT is suitable for kinetic analysis. Increasing the temperature of the peptide-protein mixture from 278 K results in the broadening of the peak attributed to the protein-bound peptide, as shown in figure 1B. In addition, there is a shift of the peak (attributed to the protein-bound peptide) toward the free, a small shift of the free peak and a broadening of the free peak (near -0.2 ppm) as the off rate increases at elevated temperatures. The bound peak disappears and signal area measurements show that it fuses with the free as the temperature is increased. The peptide peak attributed to the *cis*-proline conformation shows no change in linewidth and height indicating that it is not binding to streptavidin. There is a monotonic shift in the ¹⁹F spectra with temperature, which is due to the relatively large difference between the temperature-dependent shift of the D₂O lock signal and the ¹⁹F NMR spectra. The peaks of the *cis*-proline isomer were set to 0 ppm at all temperatures as discussed in the method section. The linewidth of the fused peak broadens as the exchange rate increases with increasing temperature (figure 1B), and this broadening is reflected in a decreasing height of the fused peak with increasing temperature as shown in figure 1C. Simulations of two-site exchange that correlate linewidth and frequency as a function of exchange kinetics (8) shows that at maximum broadening of the fused peak, the off- rate equals the frequency separation between the bound and the free ligands [$k_{\text{off}} = \Delta\nu = (\nu_{\text{bound}} - \nu_{\text{free}})$]. Thus we take $k_{\text{off}} \approx 260 \text{ sec}^{-1}$ at 314 K for FSHPQNT bound to streptavidin. The fastest bound cross relaxation rate possible in peptides would be between protons that are 1.8–2.0 Å (e.g. -CH₂ or NH₂), which are tightly bound. We estimated 22 ns for the tumbling rate of streptavidin using stoke's law (a reasonable approximation for 52 kDa protein in the spin diffusion limit (16)), and the cross relaxation rate for these protons would be approximately 20 sec^{-1} (14). However the maximum rate of cross relaxation rate growth at different mixing time observed was 2 sec^{-1} (data not shown). The fraction bound of 0.27 estimated from the integration of the free and bound ¹⁹F signals implies that the maximum bound relaxation rate was about 8 sec^{-1} . Therefore, an off-rate of 260 sec^{-1} that was estimated from the ¹⁹F linewidth analysis at 314 K was well above ten times the maximum bound cross relaxation rate. However, the amide proton NOESY signals at this

temperature are either lost or weak due to exchange with solvent. The temperature was therefore reduced to below the point where the signal loss from amide protons was a problem. Simulation can be used to estimate the exchange rate at other temperatures near 314 K where there are significant effects of the exchange on the spectra (8). The Tr-NOESY experiments were carried out at 300 K and pH 4.8 where the observed linewidth differences at 314 K and 300 K as shown in figure 1C and the exchange kinetics simulation (8) implied k_{off} is approximately 50 sec^{-1} . We should note that this rate is close to the rate of the fusion of the bound and free $\text{CF}_3\text{CO-FSHPQNT } ^{19}\text{F}$ signals at 298 K, as shown in figure 1A, where $\nu_{\text{bound}} - \nu_{\text{free}} = 23 \text{ Hz}$. $k_{\text{off}} = 50 \text{ sec}^{-1}$ was below the fast exchange limit for the protons in closest proximity to one another. These closest protons are either chemically bonded methylenes or amide side chain protons of Gln and Asn, the distance between these protons are fixed and do not depend on peptide conformation. However, the NOE build up curves indicate that all conformationally significant signals are in the fast exchange limit with a value of $k_{\text{off}} = 50 \text{ sec}^{-1}$. One also could estimate the off-rate using the simulation at other temperatures based on the difference in chemical shifts between the free and the bound peptide and the linewidth of the fused peak. Addition of two fold excess of biotin to the mixture of FSHPQNT-streptavidin completely removes the bound signal from the low temperature samples and sharpens the peptide peak attributed to the trans-proline isomer at higher temperatures (figure 1D). There is a wide range of specific protein-protein interactions that govern cellular pathways. Discovering such interactions that are mostly mediated by peptide recognition motifs is therefore a fundamental step towards unraveling the complexity of varying pathways. Also, identification of preferential binding to a drug-target receptor or protein by a particular peptide or ligand from a mixture of molecules is important. The dissociation constants of these interactions may not tell us accurately about the kinetic off-rates as many factors have influence on rates. ^{19}F NMR approach for direct measurements and optimization of sufficiently fast off-rates for Tr-NOESY studies could be valuable.

Correcting for protein mediation of ligand Tr-NOESY

Protein protons in intimate contact with the ligand can cause important distortions of the apparent distances between ligand protons though there is a suggestion that these effects are minimal and can be ignored (16). However, it is possible to quantitatively remove protein mediation effects on long range Tr-NOESY cross peaks within the ligand using a combination of Tr-NOESY and Tr-ROESY experiments. Cross relaxation due to spin pairs builds twice as fast in ROESY than in NOESY (30). Indirect (three spin) effects of other protons (including protein protons) on the ligand two spin Overhauser effects have the opposite sign and are four times larger in ROESY as in NOESY (30,31), as shown in equations 1 and 2 below. The three spin effects depend on the square of the mixing time and dominate the “spin diffusion” (at all but impractically long mixing times). The three spin effects greatly reduce ROESY at longer mixing times (ROESY is dissipative) but add to NOESY signals. If the Tr-ROESY experiment is carried out at half the mixing time as the Tr-NOESY the two spin effects are equal and the three spin effects are equal in size, but opposite in sign. Therefore, if the ROESY and NOESY cross peaks are phased to have the same sign, and are normalized by diagonal peak intensities to account for differences in T_1 and $T_{1\rho}$, the intensities are averaged and the three spin effects cancel out:

$$I_{ij}^{\text{N}} = - (K\sigma_{ij}\tau_{\text{mix}} + 2K^2(\tau_{\text{mix}})^2 \sum_k \sigma_{ik} \sum_k \sigma_{kj} + \dots) \quad (1)$$

$$I_{ij}^{\text{R}} = 2K\sigma_{ij}\tau_{\text{mix}} - 8K^2(\tau_{\text{mix}})^2 \sum_k \sigma_{ik} \sum_k \sigma_{kj} + \dots \quad (2)$$

$$-I_{ij}^N(\tau_{\text{mix}}) + I_{ij}^R(\tau_{\text{mix}}/2) = 2 K \sigma_{ij} \tau_{\text{mix}}, \text{ a pure two spin term} \quad (3)$$

Where $K = (\gamma^4 h^2 / 40 \pi^2 r_{ij}^6) \tau_c$. Tr-ROESY at 150 ms mixing time is compared to Tr-NOESY at 300 ms mixing time in figure 2. Some of the cross peaks, which appear in the Tr-NOESY spectrum, are weak or absent in the Tr-ROESY spectrum. This is because in the Tr-ROESY the indirect effects subtract from the direct effects and so by comparison of the Tr-ROESY and Tr-NOESY, the indirect effects can be qualitatively and quantitatively assessed and compensated. In addition, some as yet unassigned protein cross peaks and ligand-protein cross peaks (indicated in figure 2) are also missing in the ROESY, since there are three-spin effects within the protein that all contribute to quenching the ROESY. However, for J-coupled spins (limited to spins on the same residues) homonuclear Hartmann-Hahn or TOCSY transfer may have contribution to the ROESY cross peaks. Since the TOCSY has the same sign as the indirect cross-relaxation rate, it becomes difficult to evaluate the contribution of the indirect effect to the direct cross-relaxation rate for J-coupled spins. Fortunately TOCSY does not affect the ROESY of the long range or medium-range cross-peaks. Thus, the Tr-NOESY/Tr-ROESY averaging method can be used to correct the long and medium range Tr-NOESY cross peak intensities for three spin effects from both the protein and other ligand protons.

Assessment of non-specific ligand binding

Since the peptide concentration tends to be rather high in Tr-NOESY experiments it is a concern that nonspecific binding could contribute to the observed Tr-NOESY signals. The specificity of Tr-NOESY signals can be assessed in cases where it is practical to use a high affinity ligand, or a whole protein to block the specific ligand binding site. The present case is ideal from this point of view because the high affinity biotin ligand blocks the binding site when added in stoichiometric amounts. Figure 3A shows a segment of the NOESY spectrum of the free peptide. Figure 3B shows the corresponding region of the Tr-NOESY spectrum when streptavidin is added under conditions of fast exchange. Figure 3C shows the Tr-NOESY spectrum after a two fold excess of biotin is added to the sample, which gave the data in figure 3B. The peptide Tr-NOESY peaks are totally abolished by blocking their protein binding sites with biotin. The residual peaks in figure 3C, in the presence of biotin, are all protein peaks, as shown by the lack of agreement of the positions with any of the peptide resonances and by their occurrence in the protein solution in the absence of ligands, which could be due to the mobile regions of the protein. The protein resonances are indicated with a "p" designation for protein in figure 3B.

It is clear from comparison of the data in figure 3B and 3C (and knowledge that the residual peaks in figure 3C are due to protein) that blocking the specific protein binding site totally inhibits the peptide Tr-NOESY. This conclusion is confirmed quantitatively by integration of signal intensities, which might have been below the plotting threshold in the figure. Therefore, the Tr-NOESY signals of the bound peptide are completely due to binding to the specific biotin binding site. In other experimental systems, where there might be some contributions to Tr-NOESY signals from non-specific binding, any such non-specific signals can be subtracted from the observed Tr-NOESY signals to obtain the signals due to the ligand bound to the specific binding site of interest, if suitable high affinity agents are available to block the specific site. If high affinity binding site blocking agents are not available then more detailed experiments, varying the ligand/protein ratio, are required to assess non-specific binding (32).

Comparison of interresidual proton distances from Tr-NOESY and x-ray

The ligand interresidual proton distances deduced from Tr-NOESY and Tr-ROESY (when far enough apart to be free of HOHAHA contributions) were compared to the distances deduced from the 1.74 Å crystal structure of the FSHPQNT-streptavidin complex (20). In the crystal structure, the crystal symmetry led to two different types of interprotein contacts. The two A sites per tetramer are located at the interfaces between two different tetrameric streptavidin molecules, whereas the B sites are exposed to the solvent.

We reasoned that the B sites were likely to be more like the streptavidin sites in solution (that we studied by Tr-NOESY/Tr-ROESY NMR) and that the A sites were more effected by crystal packing forces. The published analysis of the structure of the FSHPQNT-streptavidin complex primarily focused on the A sites as they showed somewhat more order. We compared the interresidual proton distances found from Tr-NOESY/Tr-ROESY NMR with the corresponding distances in the B site of the crystal structure as shown in figure 4. Long range interresidual proton distances that were corrected using the Tr-ROESY/Tr-NOESY averaging procedure are shown in figure 4 as the open triangle points (Δ). Medium range or adjacent amino acids had much smaller corrections to their interresidual proton distances on the order of 0.1 up to a maximum of 0.3 Å. The longest range Tr-NOESY interresidual proton distances were increased by about 0.5 Å when the Tr-ROESY/Tr-NOESY averaging procedure was applied. As is common in NMR vs. x-ray structure comparisons, the distance discrepancies shown in figure 4 tend to be greatest for the longer distances (greater than 3.5Å). NOESY NMR tends to give short values for long distances, because there is increasing interference from intervening protons at the longer distances. The distances deduced from the Tr-NOESY/Tr-ROESY averaging method shown by the open triangle points in figure 4 are all in better agreement with the interresidual proton distances deduced from the crystal structure than without the corrections. Thus, for long range cross peaks the Tr-NOESY/Tr-ROESY averaging gives substantial improved agreement with the x-ray distances in this system. The error bars in figure 4 pertain to the uncertainties in the interproton distances from x-ray. We derived the error bars on the x-ray distances from the size of the so-called thermal B factors in the refined x-ray structures. The thermal B factor derived from x-ray diffraction analysis is due to dynamic disorder in the crystal, caused by the temperature-dependent vibrations of the structure as well as static disorder, which is caused by parts of molecules occupying somewhat different positions in various unit cells of the crystal. The thermal B factor is related to the root mean square uncertainty of the atom position.

MARDIGRAS was used in order to take into account spin diffusion from neighboring spins within the peptide, using a full relaxation matrix approach (33). In this way, effects of all the protons in the structure are considered on the interproton distances corresponding to the experimentally observed cross peaks. MARDIGRAS starts with a trial structure and iterates until it finds the most consistent fit between interproton distances and the experimental NMR data. The application of MARDIGRAS with a wide range of trial correlation times did not improve the agreement between the NMR-derived distances and the x-ray derived distances. Some of the MARDIGRAS distances moved away from better agreement with the x-ray distances even when using the best-fit correlation times, which suggests that from spin diffusion of the peptide protons to the protein binding site was much more important than spin diffusion within the bound peptide.

Qualitative assessment of local motions of the protein-bound peptide

It is also of great interest to assess the importance of nonuniform correlation times at different positions in the protein-bound peptide. Tr-NOESY is primarily determined by the amplitude of the spectral density function near zero frequency, which produces zero

quantum cross relaxation (34). The spectral density function near zero frequency is proportional to the correlation time, and so binding to larger proteins gives faster bound cross relaxation rates and stronger Tr-NOESY effects. However, fast local motions tend to quench the Tr-NOESY because the amplitude at twice the NMR Larmor frequency (2ν) favors double quantum cross relaxation, which subtracts from the NOESY favored by low frequency motions. Thus, if high frequency local motions are important at sites in the bound peptides, the Tr-NOESY cross peaks are expected to increase in amplitude at higher spectrometer frequencies. Many of the observed FSHPQNT-streptavidin Tr-NOESY cross peak intensities increase considerably at higher fields (500 vs. 600 vs. 750) when measured on the same sample, as shown in figure 5. However, when the cross peaks were normalized by the diagonal peak intensities, most of the frequency dependence of the cross peak intensities in this system were removed. Therefore, the increase in intensities of the cross relaxation at higher field were likely due to the increase in T_1 relaxation instead of local motions in the bound peptide because the cross peaks and the diagonal peaks were reduced in the same proportion by T_1 relaxation at different spectrometer frequencies.

The above does not imply that there are no differences in local motion in different regions of the streptavidin bound peptide. It means that any differences in local motion do not significantly influence the amplitude of the spectral density functions near twice the Larmor frequencies, to cause changes in the Tr-NOESY cross relaxation rates at different Larmor frequencies, in the particular system under study. In addition, those portions of the ends of the peptide that appear to have the fastest motions have little NOESY intensity, even at the highest fields available. In addition, the above studies imply that it may be favorable to use substantially longer mixing times in Tr-NOESY at higher fields (and still remain at or near in the linear portion of the diagonal-normalized build up curve) and facilitate the observation of weaker signals that provide crucial, long range structural information.

Structure determination of the bound peptide

We ran simulated annealing starting from randomized coordinates of the extended chain conformation of FSHPQNT peptide, The N-terminal Phe of the peptide was poorly defined by the NMR data and it was omitted from the simulated annealing. The HPQN main chain and side chain regions in the NMR structures are quite well defined. This is the region of the peptide that contains the consensus biotin-mimetic sequence (HPQ) and has the smallest B factors except for the side chains of the Gln residue in the x-ray structure (20). The least defined regions are the Ser next to the N-terminus and Thr at the C-terminus, which have also very large B factors in the x-ray structure. We observed very few cross peaks, which involve the Phe at the N-terminus or the Ser residues.

Tr-NOESY cross peaks between Phe ring protons and other amino acids were not observed though the Phe side chain is clearly bound to the streptavidin site in solution from the para-fluoro-Phe-SHPQNT ^{19}F NMR experiments shown in figure 1D. The T_1 and the T_2 of the ring are quite long, and the bound Phe evidently has considerable local motion. The His ring, on the other hand, has an effective protein-bound correlation time (about 20 ns), which shows that it is strongly immobilized by binding. The correlation time of the protein as a rigid sphere would be 22 ns, and 20 ns is quite reasonable as a correlation time for immobilized regions of the peptide.

Figure 6 shows ten superimposed structures (red) with lowest energies derived from the Tr-NOESY NMR distance constraints by simulated annealing starting from randomized extended structures using nih-xplor protocol (35). The convergence of the structures is very good, as is demonstrated by the rmsd of the backbone (0.5 Å) and all of the atoms (0.9 Å). The rmsd between the crystal and NMR structures for the backbone and all atoms of the four immobilized amino acids (HPQN) are 0.65 Å and 1.38 Å respectively. The convergence is

much better for the three most immobilized amino acids (HPQ) with rmsd of 0.3 Å for the backbone between the crystal and NMR structures.

Discussion

We sought a peptide test case where there is a high resolution crystal structure of a peptide-protein complex. The biotin-mimetic peptide FSHPQNT bound to streptavidin has excellent properties desired for a Tr-NOESY test bed. Streptavidin is stable over a wide range of pH, temperature and salt concentration, a genetically engineered core streptavidin has higher solubility and is commercially available. The binding constant of FSHPQNT had been measured at pH 7 and was in a range thought to be highly suitable for Tr-NOESY ($K_d = 125 \mu\text{M}$). It may be difficult or even impossible to adjust conditions to make the off-rate of ligands from proteins fast enough to make Tr-NOESY useful for biologically interesting interactions. In this study we set the exchange rate at the edge of being sufficiently fast and found that even though the shortest distances (methylenes and amino acid amide side chain protons) were not in fast exchange the conformationally significant interproton distances were in fast exchange on the NOESY cross relaxation time scale. Our ability to deduce a reasonably accurate bound peptide structure when the off-rate was at the edge of being fast enough is very encouraging. This implies there is no need to go to extremely fast off-rates and very low binding constants to get rather accurate structural data from Tr-NOESY. It is not possible to judge off-rates sufficiently accurately from dissociation constants, as many factors can effect on rates and direct measurements of off-rates seem very important for reliable Tr-NOESY studies. It is also important to understand the specificity of the interfacial peptides. High affinity blockers of the binding sites are important, either small ligands (such as biotin used here) or using random peptide library studies to assess the specificity of the interfacial peptides.

The x-ray crystallographic B factors were found large, particularly for the least immobilized Ser and Thr residues. However, the B factors of the central segment of the peptide are very comparable to the amino acid side chains of the binding pocket and other regions of streptavidin. There is a tendency to think that such a relatively high resolution peptide-protein complex structure has a much better defined peptide than it actually is when the thermal B factors are converted to distance uncertainties. The best fit crystal structures look so well defined in published pictures, whereas the standard way of presentation of NMR structures as a nest of superimposed structures gives a graphic idea of the uncertainties. Tr-NOESY together with other NMR techniques such as STD are playing important role in identification of protein interfaces, drug screening and determination of high resolution flexible peptide structures. However, the accuracy of information from Tr-NOESY studies could also be distorted due to non-specific binding, spin diffusion from the protein. Protein cross relaxation of the ligand can be overcome by deuteration of the target protein, but protein deuteration is difficult or even impossible for many protein systems of interest. If deuteration of the target protein can be accomplished the long range distance constraints that may be obtained from Tr-NOESY can extend much further.

Conclusion

In conclusion, we have obtained the streptavidin-bound structure of the linear biotin-mimetic peptide, using Tr-NOESY, that is nearly as accurate as the high quality 1.74 Å x-ray structure after optimization of ligand binding kinetics, assessment of non-specific ligand binding, removal of broad protein background and correcting for protein mediated peptide Tr-NOESY. This is particularly a fortunate case because we had a high resolution x-ray structure as well as a biotin molecule that binds to the protein with a much higher affinity than the peptide that allowed us to optimize the parameters. ^{19}F NMR of bound peptides and

ligands can be used to determine off-rates that provide optimum conditions for Tr-NOESY experiments. The detailed three dimensional structures of protein-bound peptides that specifically block protein-protein interactions provide information on the structural basis of the protein interactions of interest and useful structural templates for the design of peptide mimetic small molecules with high affinity to the target protein. Most transient protein-protein interactions involve only a small fraction of the protein surfaces. Also, recent studies on the use of Tr-NOESY experiments to elucidate structural details of a ligand bound to a receptor in living cells suggest the potential of Tr-NOESY to design novel molecules in the natural environment that have pharmacological benefit. Thus, Tr-NOESY experiments may be used as a competitive and complementary method to reveal the structures of protein-protein or ligand-protein interfaces that are difficult or impossible to study with other methods when the approaches described in this paper and improvements of the approaches discussed are implemented.

Acknowledgments

This work was supported by NIH grant R01EY06913, NSF grant OSR-9350546, and an NSF NMR upgrade grant to E.D and Scott & White Memorial Hospital research grant to D.G. NMR studies were carried out at the National Magnetic Resonance Facility at Madison, NIH and Montana State University. We thank Dr. Scott Busse for assistance with NMR spectroscopy at Montana State University. We thank Dr. Milo Westler and Dr. Martin Teintze for helpful discussions.

References

1. Jones S, Thornton JM. Principles of protein-protein interactions. *Proc Natl Acad Sci U S A*. 1996; 93:13–20. [PubMed: 8552589]
2. Jones S, Thornton JM. Analysis of protein-protein interaction sites using surface patches. *J Mol Biol*. 1997; 272:121–32. [PubMed: 9299342]
3. Gizachew D, Moffett DB, Busse SC, Westler WM, Dratz EA, Teintze M. NMR studies on the conformation of the CD4 36–59 peptide bound to HIV-1 gp120. *Biochemistry*. 1998; 37:10616–25. [PubMed: 9692951]
4. Cortese R, Monaci P, Luzzago A, Santini C, Bartoli F, Cortese I, et al. Selection of biologically active peptides by phage display of random peptide libraries. *Curr Opin Biotechnol*. 1996; 7:616–21. [PubMed: 8939640]
5. Rarick HM, Artemyev NO, Mills JS, Skiba NP, Hamm HE. Specific peptide probes for G-protein interaction with effectors. *Methods Enzymol*. 1994; 238:13–28. [PubMed: 7799779]
6. Bratkovic T. Progress in phage display: evolution of the technique and its applications. *Cell Mol Life Sci*. 2009
7. Campbell AP, Sykes BD. The two-dimensional transferred nuclear Overhauser effect: theory and practice. *Annu Rev Biophys Biomol Struct*. 1993; 22:99–122. [PubMed: 8348000]
8. Ni F. Recent developments in transferred NOE methods. *Prog NMR Spectrosc*. 1994; 26:517–606.
9. Burritt JB, Busse SC, Gizachew D, Siemsen DW, Quinn MT, Bond CW, et al. Antibody imprint of a membrane protein surface. Phagocyte flavocytochrome b. *J Biol Chem*. 1998; 273:24847–52. [PubMed: 9733789]
10. Carlomagno T. Ligand-target interactions: what can we learn from NMR? *Annu Rev Biophys Biomol Struct*. 2005; 34:245–66. [PubMed: 15869390]
11. Macnaughtan MA, Kamar M, Alvarez-Manilla G, Venot A, Glushka J, Pierce JM, et al. NMR structural characterization of substrates bound to N-acetylglucosaminyltransferase V. *J Mol Biol*. 2007; 366:1266–81. [PubMed: 17204285]
12. Wang Y, Zhang T, Xu J, Du W. Comparison of the binding affinity of chlorogenic acid with two serum albumins. *Int J Biol Macromol*. 2011; 48:81–6. [PubMed: 20923680]
13. Potenza D, Belvisi L. Transferred-NOE NMR experiments on intact human platelets: receptor-bound conformation of RGD-peptide mimics. *Org Biomol Chem*. 2008; 6:258–62. [PubMed: 18174992]

14. Cavanagh, J.; Fairbrother, WJ.; Palmer, AG., III; Skelton, NJ. Protein NMR Spectroscopy: principles and practice. Academic Press; 1996.
15. Landy SBaR BDN. Dynamical NOE in Multiple-spin systems undergoing Chemical exchange. *J Magn Reson.* 1989; 81:371–7.
16. Post CB. Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Curr Opin Struct Biol.* 2003; 13:581–8. [PubMed: 14568612]
17. Devlin JJ, Panganiban LC, Devlin PE. Random peptide libraries: a source of specific protein binding molecules. *Science.* 1990; 249:404–6. [PubMed: 2143033]
18. Weber PC, Pantoliano MW, Thompson LD. Crystal structure and ligand-binding studies of a screened peptide complexed with streptavidin. *Biochemistry.* 1992; 31:9350–4. [PubMed: 1390720]
19. London RE, Pearlman ME, Davis DG. Relaxation Matrix Analysis of the Transferred Nuclear Overhauser Effect for Finite Exchange Rates. *J Magn Reson.* 1992; 97:79–98.
20. Katz BA. Binding to protein targets of peptidic leads discovered by phage display: crystal structures of streptavidin-bound linear and cyclic peptide ligands containing the HPQ sequence. *Biochemistry.* 1995; 34:15421–9. [PubMed: 7492542]
21. Ludwig C, Guenther UL. Ligand based NMR methods for drug discovery. *Front Biosci.* 2009; 14:4565–74. [PubMed: 19273371]
22. Meyer B, Peters T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl.* 2003; 42:864–90. [PubMed: 12596167]
23. Griesinger, CaE; RR. Frequency offset effects and their elimination in NMR rotating-frame cross-relaxation spectroscopy. *J Magn Reson.* 1987; 75:261–71.
24. James TL, Borgias BA, Bianucci AM, Zhou N. Determination of DNA and protein structures in solution via complete relaxation matrix analysis of 2D NOE spectra. *Basic Life Sci.* 1990; 56:135–54. [PubMed: 2078169]
25. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson.* 2003; 160:65–73. [PubMed: 12565051]
26. Nilges M, Clore GM, Gronenborn AM. Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* 1988; 229:317–24. [PubMed: 3345845]
27. Akcay G, Kumar K. A New Paradigm for Protein Design and Biological Self-Assembly. *J Fluor Chem.* 2009; 130:1178–82. [PubMed: 20161419]
28. Kitevski-LeBlanc JL, Evanics F, Scott Prosser R. Optimizing (1)F NMR protein spectroscopy by fractional biosynthetic labeling. *J Biomol NMR.* 2010; 48:113–21. [PubMed: 20734112]
29. Kitevski-Leblanc JL, Evanics F, Scott Prosser R. Approaches to the assignment of (19)F resonances from 3-fluorophenylalanine labeled calmodulin using solution state NMR. *J Biomol NMR.* 2010; 47:113–23. [PubMed: 20401735]
30. Fejzo J, Zolnai Z, Macura S, Markley JL. Analysis of Laboratory-Frame and Rotating Frame Cross Relaxation Buildup Rates from Macromolecules. *J Magn Reson.* 1989; 82:518–28.
31. Perlman ME, Davis DG, Koszalka GW, Tuttle JV, London RE. Studies of inhibitor binding to *Escherichia coli* purine nucleoside phosphorylase using the transferred nuclear Overhauser effect and rotating-frame nuclear Overhauser enhancement. *Biochemistry.* 1994; 33:7547–59. [PubMed: 8011620]
32. Jarori GK, Murali N, Rao BD. Two-dimensional transferred nuclear Overhauser effect spectroscopy study of the confirmation of MgATP bound at the active and ancillary sites of rabbit muscle pyruvate kinase. *Biochemistry.* 1994; 33:6784–91. [PubMed: 8204612]
33. Thomas PD, Basus VJ, James TL. Protein solution structure determination using distances from two-dimensional nuclear Overhauser effect experiments: effect of approximations on the accuracy of derived structures. *Proc Natl Acad Sci U S A.* 1991; 88:1237–41. [PubMed: 1996325]
34. Neuhaus DaW, M. The Nuclear Overhauser Effect in Structural and Conformational analysis. VCH; 1989.
35. Nilges M, Gronenborn AM, Brunger AT, Clore GM. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to

crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Eng.* 1988; 2:27–38. [PubMed: 2855369]

Figure 1A

Figure 1B

Figure 1C

Figure 1D

Figure 1.

^{19}F NMR spectra of fluorine labeled FSHPQNT peptide bound to streptavidin at different temperatures were used for ligand off-rate analysis. (A) ^{19}F NMR spectra of the CF_3CO -FSHPQNT peptide bound to streptavidin as described in materials and methods at 278 K (lower curve) and 298 K (upper curve). The spectra have been calibrated so that the trifluoroacetate reference signal is at 0.0 ppm. (B) Parafluoro-phe-SHPQNT peptide bound to streptavidin complex at pH 4.8 and different temperatures as described in materials and methods. The peak attributed to the cis-proline isomeric peptide was used as an internal standard and was set to 0.0 ppm at each temperature. The peak near 0.4 ppm is due to the peptide bound to streptavidin. The larger peak near -0.15 ppm is due to the free trans-proline peptide at the lowest temperatures and shows broadening and peak height changes as a function of temperature due to faster exchange at higher temperatures. (C) The peak height and width at half height of the major peak in B are plotted as a function of temperature from 298 to 320 K. The maximum width and minimum height occurred at 314 K where the kinetic off-rate equals the chemical shift splitting between the bound and free signals as explained in the text. (D) The effect of adding a two-fold excess of biotin relative to the concentration of streptavidin binding sites to the sample used to record the data in B. The lower pair of spectra show the effect of biotin addition at 278 K, the middle pair of spectra show the effect of biotin addition at 300 K and the upper pair of spectra show the effect of biotin addition at 314 K. In all cases the biotin releases all the peptide from the specific binding sites on streptavidin, removing the bound peaks at 278 and 300 K and sharpening the free peaks at 300 and 314 K.

Figure 2.

Comparison of 2D Tr-NOESY and 2D ROESY NMR for the streptavidin bound FSHPQNT peptide (A) 2D Tr-NOESY NMR at 300 msec mixing time and (B) 2D ROESY NMR spectra at 150 msec mixing time at pH = 4.8 and 300 K. Many of the peaks are weaker or missing in the ROESY spectrum. Some of the missing peaks are due to sharp protein peaks and some of the weaker or missing peaks are due to peptide Tr-NOESY peaks that involve some protein proton mediation, as explained in the text.

Figure 3.

All of the peptide Tr-NOESY signals are removed by the addition of biotin, which blocks the specific peptide binding sites. The remaining peaks after the addition of biotin are due to protein peaks that originate from relatively mobile regions of the protein. (A) spectrum of the free FSHPQNT peptide in the absence of streptavidin. (B) The Tr-NOESY spectrum of the FSHPQNT peptide observed upon the addition of streptavidin, where the fraction of the peptide bound is 0.27. Various peptide Tr-NOESY peaks are labeled as well as a few sharp protein peaks. (C) Tr-NOESY spectrum of the same sample as shown in B, but after the addition of a two fold excess of biotin over the concentration of streptavidin binding sites.

Figure 4.

H---H distances in the FSHPQNT peptide bound to streptavidin derived from Tr-NOESY NMR as described in the text compared to the same distances measured from the 1.74 Å resolution x-ray crystal structure. The error bars for distances with Δ symbol indicate the uncertainties in some of the long range H---H distances derived from the x-ray structure based on the size of the crystallographic thermal B-factors.

Figure 5.

Tr-NOESY spectra of the same sample of FSHPQNT-streptavidin at pH 4.8 and 300 K run for the same length of time at (A) 500 MHz, (B) 600 MHz and (C) 750 MHz. The spectra were acquired under identical conditions, processed identically, plotted at the same contour level.

A

B

Figure 6. Superimposed structures of SHPQNT derived from Tr-NOESY NMR and x-ray crystallography show good agreement. (A) Ten lowest energy structures of SHPQNT

peptide bound to streptavidin calculated by simulated annealing using the Tr-NOESY-derived distance constraints (red) starting from randomized coordinates. The single structure in cyan is from the x-ray crystallography. (B) Backbone structures of SHPQNT derived from Tr-NOESY NMR (red) and x-ray crystallography (cyan) are shown. The rmsd between the crystal and NMR structures for the backbone of the four immobilized amino acids (HPQN) is 0.65 Å.