Published in final edited form as:

J Am Chem Soc. 2018 January 17; 140(2): 675–682. doi:10.1021/jacs.7b09974.

Slow dynamics of tryptophan-water networks in proteins

R. Bryn Fenwick, David Oyen, H. Jane Dyson, and Peter E. Wright*

The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla CA 92037

Abstract

Water has a profound effect on the dynamics of biomolecules and governs many biological processes, leading to the concept that function is slaved to solvent dynamics within and surrounding the biomolecule. Protein conformational changes on the μ s-ms timescales are frequently associated with protein function, but little is known about the behavior of protein-bound water on these timescales. Here we have used NMR relaxation dispersion measurements to probe the tryptophan indoles in the enzyme dihydrofolate reductase (DHFR). We find that during structural changes on the μ s-ms timescale, large chemical shift changes are often observed for the NH proton on the indole ring, while relatively smaller chemical shift changes are observed for the ring nitrogen atom. Comparison with experimental chemical shifts and density functional theory-based chemical shift predictions show that during the structural change the tryptophan indole NHs remain bound to water, but the geometry of the protein-bound water networks change. These results establish that relaxation dispersion measurements can indirectly probe water dynamics and indicate that water can influence, or be influenced by, protein conformational changes on the μ s-ms timescale. Our data show that structurally conserved bound water molecules can play a critical role in transmitting information between functionally important regions of the protein and provide evidence that internal protein motions can be coupled through the mediation of hydrogen-bonded water bound in the protein structure.

Table of Contents

*To whom correspondence should be addressed. wright@scripps.edu.

Supporting Information

A table containing literature chemical shifts of non-hydrogen bonded tryptophan indoles.

Keywords

CPMG; buried water; dynamic coupling; chemical shifts; NMR

Introduction

The dynamics of water in and around biomolecules defines much of their structure and function, playing a key role in their folding and equilibrium fluctuations.^{1–4} The structure and dynamics of the water molecules that hydrate proteins, as well as waters buried in protein structures, have been extensively characterized by NMR spectroscopy and X-ray crystallography.^{5–9} X-ray crystallography allows the detailed structural characterization of strongly bound water molecules across multiple structures and multiple solvents.^{10–11} Bound water is a general feature of proteins¹² and often the number of water molecules and amino acid residues in protein structures are approximately equal.¹³ Water is frequently found in the core of proteins, especially in enzyme active sites, and defined structural patterns have been observed for both buried and surface exposed residues.^{4, 14} Moreover, the reorganization and/or elimination of water from active sites often accompanies binding of proteins to ligands, cofactors, and other small molecules, and is envisaged to be a crucial component of molecular design.⁴ The high degree of hydrogen bonding between proteins and solvent creates intimate coupling that results in the thermal motion of the solvent being coupled to biomolecules. Although solvent slaving has been proposed to drive the internal motions of biomolecules through strong coupling between the solvent and the protein, there are relatively few studies that have directly probed these interactions.^{15–16} The popular unified model of protein dynamics proposes that large collective motions are controlled by motion of the bulk solvent, while the internal motions of proteins are slaved to the dynamics of tightly bound water molecules in the hydration shell.¹⁷ Recently, Qin et al. have used the time-dependent Stokes shift of tryptophan residues to characterize ultrafast dynamics on the timescale of 300 fs to 120 ps.¹⁸ This work demonstrated the ability of the hydration shell to modulate side chain motions to link water and protein dynamics on these ultrafast timescales. Using NMR spectroscopy it is possible to measure the exchange rates of some bound waters on fast timescales (ps-ns) and this has allowed the lifetimes of waters within the core of proteins to be determined,^{19–20} although this interpretation has recently been questioned.²¹ Complementary information from molecular dynamics simulations of proteins in explicit solvent provides insight into the movement and flow of water across the surface of proteins and allows determination of the residency times of bound water molecules.^{22–24} Through these combined efforts, detailed insights into the water dynamics and the structural coupling between proteins and buried water on timescales on the order of ps- μ s have been obtained. Relaxation dispersion studies undertaken by Koenig and Schillinger have indicated that longer-lived water molecules are present with bound lifetimes of 0.1–10 μ s.²⁵ Kuntz and Kauzmann highlight that many of these measurements of bound water rely on sufficient exchange with the bulk water, essentially limiting these methods to detecting waters with lifetimes less than $\sim$ 30 μ s.²⁶ Thus, the dynamics of waterprotein interactions on longer timescales (μ s-ms) are less well characterized.

Important questions remain: how are the dynamics of bound water on the μ s-ms timescale related to function, and are these waters coupled to the biologically relevant dynamics of proteins? NMR spectroscopy is a powerful technique to study the dynamics of proteins and has been used extensively to demonstrate the coupling between motions and enzyme function.^{27–29} Structured water has previously been directly studied by NMR relaxation dispersion experiments on the short μ s timescale.²⁵ Here we indirectly investigate the slow time scale dynamics of water molecules in *E. coli* dihydrofolate reductase (DHFR), an enzyme in which μ s-ms motions are closely coupled to enzyme function.^{29–30} By studying the transverse ^{15}N and ^1H relaxation of the tryptophan indoles we observe dispersive behavior indicative of coupling between dynamic structural changes in the protein and rearrangement of the hydration networks.

Materials and Methods

NMR samples of the wild type and N23PP/S148A mutant of *E. coli* DHFR were prepared as previously described in amber tubes.³¹ Samples were approximately 1.5 mM in DHFR in NMR buffer (10% D_2O , 50 mM phosphate, pH 6.8, 100 mM KCl, 1 mM EDTA, 1 mM DTT, 0.02% NaN_3) and contained 20 mM NADP^+ and 12 mM folic acid. Resonance assignments have been previously reported for the folate: NADP^+ complex of wild type and N23PP/S148A DHFRs.^{32–33} Relaxation compensated, constant time Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion data were recorded at 301K at two static magnetic fields at proton frequencies of 500 and 800 MHz.³¹ Relaxation dispersion profiles were measured for ^1H and ^{15}N using previously published pulse sequences with modifications to allow scan-interleaved acquisition for the dispersion series.^{34–35} The relaxation data were recorded using optimized non-uniform Poisson-gap sampling schemes in the ^{15}N dimension^{36–37} and processed with a combination of MDDNMR and NMRPipe using IRLS reconstruction in the non-uniformly sampled dimension.^{30, 38–39} Cross-peak intensities were extracted using CcpNmr Analysis.⁴⁰ The data for the N23PP/S148A mutant were grouped and fitted as defined previously for the C-terminal associated region, the CD-loop and the Active-site.³¹ The data fitted to a two-site exchange model (p_A , p_B , k_{ex} , ω_H , ω_N , where ω_H and ω_N are ^1H and ^{15}N chemical shift differences between the states, in ppm) described by the Bloch-McConnell equations using the program Glove.⁴¹ For the WT data individual fits were made for each indole. During the fitting, the minor population p_B was restrained to be less than 10%, while an upper limit of 12000 s^{-1} was imposed on the exchange rate k_{ex} . These limits were used to reduce the conformation search space for the optimum solutions to the Bloch-McConnell equations. Errors were set to 2% and 4% for the 800 and 500 MHz data, respectively unless the estimated error was larger based on duplicate measurements.^{30, 42} The sign of ω was determined by comparing ^1H - ^{15}N HSQC and HMQC spectra.⁴³ Indole hydrogen-deuterium exchange rates were measured for the N23PP/S148A mutant using CLEANEX-PM experiments recorded at 700 MHz at a temperature of 301 K.⁴⁴ Mixing periods of 22, 44, and 66 ms were used with an inter-scan delay of 2 seconds. The collected data were fitted to the previously described equation and errors were determined using the covariance method.⁴⁴

Indole ^{15}N and ^1H chemical shift calculations were prepared using the program SHIFTS,⁴⁵ with the QM part of the calculations being prepared with AFNMR and run with Gaussian.⁴⁶

Aggregate and individual indole chemical shifts were extracted from the RefDB protein chemical shift database and used without re-referencing.⁴⁷ Proteins containing tryptophan residues with indole ¹H chemical shifts < 8 ppm and ¹⁵N shifts < 123 ppm were selected for structural analysis. Ring-current contributions, determined using SHIFTS, were removed from the indole proton and nitrogen chemical shifts.⁴⁵ For the density functional theory (DFT) calculations, the tripeptide GWG was built with Pymol⁴⁸ and a water molecule was placed within hydrogen bonding distance of the indole NH (H----O distance 1.9 Å, N—H----O angle 172.5 degrees). We report the differences between the calculated (DFT) chemical shifts in the presence and absence of the water because the changes in chemical shift are independent of any referencing errors that might arise.

Results

For the current study, we used the *E. coli* DHFR N23PP/S148A mutant, which has a reduced amplitude of motion in the loops that surround the active site compared to the wild-type (WT) protein.³² However, throughout the manuscript, we use residue numbering corresponding to the WT sequence (the N23PP mutation introduces an extra residue) to aid comparison with previously published studies of *E. coli* DHFR structure and dynamics and the conserved water molecule numbering has been taken from the WT structure 4NX6.

Relaxation dispersion for many of the tryptophan indoles (¹⁵N and ¹H nuclei) in the folate:NADP⁺ complex of *E. coli* N23PP/S148A DHFR was noted previously.³¹ However, the nature of the associated conformational changes was not clear. Here we describe the indole relaxation dispersion data in detail and identify the structural perturbations that are responsible for the observed chemical shift changes. The ¹⁵N and ¹H relaxation dispersion profiles for all observed tryptophan indoles are shown in Figure 1 for the folate:NADP⁺ complexes of N23PP/S148A and WT proteins. We observe strong dispersion for the indole NeH resonances of W30 and W47, as well as a smaller degree of dispersion for those of W133. The NeH resonances of the remaining two indoles, W22 and W74, do not exhibit appreciable R₂ dispersion for the N23PP/S148A mutant but do display dispersive behavior in WT DHFR that is associated with the exchange between the closed ground state and occluded excited state.⁴⁹ As a consequence, fitting of the WT dispersion data is complicated by three-state exchange behavior for many residues; the N23PP/S148A mutations simplify data fitting by abrogating the closed/occluded exchange process³² and the dispersion profiles for N23PP/S148A shown in Figure 1 fit well to two-site exchange models with the exchange parameters previously determined by cluster fits of the full amide ¹⁵N and ¹H relaxation dispersion dataset (see Table 1).³¹ The dispersion data for all active-site backbone amides and W30 NeH, were fitted to a rate of 4200 ± 100 s⁻¹ and an excited-state population of 4.2 ± 0.7 %. The dispersion profiles for the CD-loop backbone amides and W47 NeH, were fitted to a rate of 3100 ± 200 s⁻¹ and a population of 2.13 ± 0.06 %. Finally, W133 NeH and the C-terminal associated region backbone amides were fitted to a rate of 630 ± 20 s⁻¹ and a population of 3.32 ± 0.07 %.³¹ The chemical shift changes (Δν) in the minor excited-state conformations of these indoles were also obtained from the fitted relaxation dispersion profiles and are summarized in Table 1.

To determine whether the relaxation dispersion observed for the tryptophan indoles might be caused by hydrogen exchange, we measured the physical proton exchange rates for both backbone and side chain NH groups in the N23PP/S148A mutant under the same conditions as the CPMG experiments using the hydrogen-deuterium exchange CLEANEX-PM experiment.⁴⁴ No exchange was observed for the tryptophan indole protons within the error of the experiment, showing that the hydrogen-deuterium exchange rate is slower than 0.1 s^{-1} (lifetime $>10 \text{ s}$). This rate is very much slower than the process probed by the CPMG experiments (rates $> 500 \text{ s}^{-1}$ and lifetimes $< 2 \text{ ms}$).

The crystallographic water molecules located near tryptophan indoles were analyzed for conservation across WT and mutant DHFR structures. In the crystal structure of the folate:NADP⁺ complex of N23PP/S148A (3QL0), water molecules are observed in proximity to four of the tryptophan indoles (Figure 2). To characterize the bound waters we analyzed the available crystal structures of DHFR using PyWATER,¹⁰ which identified 86 conserved water molecules in 76 high-resolution structures of *E. coli* DHFR. PyWATER determines the percentage conservation for each water molecule, where a percentage of 97% conservation indicates that a water molecule was present in 97% of the superimposed structures. In structures where water molecules were not observed, the water could be missing due to a difference in resolution of the structure or in the method used to place the water. In addition, some structures have waters in very similar places but are outside the default distance cutoff for inclusion as a conserved water. The analysis of the structures shows strong conservation of water molecules in hydrogen-bonding positions adjacent to the tryptophan indole nitrogen atoms. The only exception is W74 Ne, which forms a direct hydrogen bond to the backbone carbonyl oxygen of residue D69. The room temperature structure of the wild-type DHFR:folate:NADP⁺ complex (PDB code 4NX6 and from which the water numbering is used in this work), contains 65 of the 86 conserved water molecules.⁵⁰ W22 Ne is hydrogen bonded to water HOH308, W30 Ne is hydrogen bonded to water HOH301, W47 Ne is hydrogen bonded to water HOH326, and W133 Ne is hydrogen bonded to water HOH327. Full details of the conserved waters associated with the DHFR tryptophan indoles are given in Table 2, including the extent of conservation in the set of *E. coli* DHFR structures and the hydrogen bonding angles and distances determined from the room temperature X-ray structures of the WT (PDB code 4NX6) and N23PP/S148A mutant (PDB code 3QL0) DHFR:folate:NADP⁺ complexes. In both structures, the hydrogen bond distances and angles are very similar.

Analysis of the relaxation dispersion data showed that the tryptophan indole NH chemical shift changes ($|\Delta\omega_{\text{H}}|$ and $|\Delta\omega_{\text{N}}|$) associated with the structural fluctuations are in the range of 0.2–0.9 ppm for proton and 1.4–2.5 ppm for nitrogen. To determine if these chemical shift changes are caused by complete loss of the hydrogen bonding water we examined published data for tryptophan indole resonances (¹H and ¹⁵N) in proteins for which 3D structures are available (see Table S1). Indole ¹H and ¹⁵N resonances are predicted to be shifted upfield by ~2 ppm in nonpolar solvents^{51–52} and if the hydrogen-bonded water was completely removed from the indole of a buried tryptophan, we would expect its environment to become more nonpolar. The published data show that the most upfield-shifted indoles have a range of proton chemical shifts between 6.5 and 8.0 ppm, while the nitrogen chemical shifts are between 121.1 and 123.3 ppm, after the ring current contributions have been factored out.

Inspection of the structures from Table S1 reveals that none of the indole NH groups with upfield-shifted resonances undergo hydrogen bonding interactions and that no water or other hydrogen bond donor is present in the tightly packed core of these structures. Since the normal chemical shifts observed for hydrogen-bonded indoles are 10.1 ppm for ^1H and 129.3 ppm for ^{15}N (Table 3), the mean upfield shifts observed experimentally for non-hydrogen bonded indoles are 2.7 ppm for ^1H (δH) and 7.2 ppm for ^{15}N (δN). Further validation was provided by a density functional theory (DFT) calculation of the indole ^1H and ^{15}N chemical shifts for the tripeptide Gly-Trp-Gly in the presence and absence of a single water molecule hydrogen bonded to the tryptophan indole NH. The results of the DFT chemical shift calculations (Table 3) predict that upon complete loss of the hydrogen bonded water, the ^1H resonance would shift upfield by 2.7 ppm while the indole ^{15}N resonance would be shifted upfield by 8.2 ppm, in excellent agreement with the values obtained from the database analysis. The $|\rho|$ values measured by relaxation dispersion for the indole ^1H and ^{15}N resonances of the N23PP/S148A DHFR:folate:NADP⁺ complex (Table 1) are very much smaller than the chemical shift changes that would accompany complete loss of hydrogen bonded water. We thus conclude that the conformational exchange processes that give rise to the observed tryptophan relaxation dispersion involve subtle fluctuations in the indole hydrogen bonding network rather than loss of bound water.

Discussion

Water-Indole Hydrogen Bonds

While the observation of internal water molecules using X-ray crystallography illustrates their important role in protein structure,^{3, 53–54} the direct involvement of water in conformational changes occurring on the μs -ms timescale has been difficult to observe. NMR spectroscopy has played a critical role in identifying the specific locations of protein-bound water molecules,^{6, 55} as has the combination of molecular dynamics simulations and NMR spectroscopy.⁵⁶ The chemical shifts of hydrated side chains can also give clues about the nature of protein-water interactions.⁵⁷ The relaxation dispersion data for the N23PP/S148A DHFR:folate:NADP⁺ complex can be fitted by two-site exchange models and the indole ^1H and ^{15}N CPMG dispersion profiles for each of the tryptophan residues fit well to the kinetic and thermodynamic parameters that describe the fluctuations of neighboring backbone amides.³¹ This indicates that the backbone and Trp side chain probes are reporting on the same conformational exchange processes. The physical exchange of the indole proton is much too slow to contribute to the observed relaxation dispersion. This does not imply that the lifetime of the bound waters is very long; exchange of bound water with bulk solvent on a time scale much faster than that of the relaxation dispersion experiments would be invisible to our measurements. Indeed, bound water molecules typically have lifetimes shorter than 1 μs , although a longer-lived water has been observed in BPTI.²⁰ We note that the fluctuations on the μs -ms timescale observed here involve transitions between stable ground state and excited state conformations and are accompanied by small changes in the average distances/angles of the water-indole hydrogen bonds. These changes in the conformational ensemble are distinct from the harmonic oscillations of the hydrogen bonds on shorter ps timescales that do not perturb the average configuration.

Tryptophan indole nitrogen atoms are frequently observed to be hydrogen bonded to water molecules in protein structures.^{7, 14} In the folate:NADP⁺ complex of N23PP/S148A DHFR, the indole NH groups of the tryptophan residues that exhibit dispersion (W30, W47, and W133) are hydrogen bonded to structurally conserved water molecules (Table 2). The observed values of ωH are much smaller (< 1 ppm) than the δH of 2.7 ppm associated with the loss of a water hydrogen bond to the tryptophan indole (see Table 3). The ^{15}N chemical shift also reflects hydrogen bonding to the indole NH. The indole nitrogen ωN is observed to change by < 3 ppm while complete loss of the hydrogen bonds to the water would cause an upfield shift of 7.2 ppm (δN in Table 3).

Water-Indole Structural Changes in the Excited State

Based on the relationship between the indole ^{15}N shift and the hydrogen bonding distance established using solid state NMR measurements on tryptophan model compounds,⁵⁸ we estimate that the indole ^{15}N chemical shift changes observed by relaxation dispersion (ω , Table 1) correspond to changes in the ensemble average N-O distances in the range of 0.05 to 0.1 Å. These distances should be regarded as an upper limit on the magnitude of the conformational change since it is possible that changes in the N-H-O bond angle could also contribute to the changes in the ^{15}N chemical shift. The relaxation dispersion experiments reported here reveal μs -ms timescale conformational fluctuations that lead to modest changes in the tryptophan indole chemical shifts. Thus, the observed chemical shift changes (ω) most likely arise from subtle changes in the local water-mediated hydrogen bonding network and do not arise from the loss of hydrogen bonded water molecules.

Relaxation Processes in the C-terminal Region

The C-terminal associated region is a cluster of residues in the FG loop (Y128 – E134) and near the C-terminus (L156 – R158) that exhibit strong backbone ^{15}N and ^1H relaxation dispersion arising from conformational fluctuations (with $k_{\text{ex}} = 630 \text{ s}^{-1}$) between the ground state and a 3.3% population of an excited state.³¹ While the excited state conformation of this region of the protein remains unknown, it is clear that the backbone conformation is altered. The indole ^{15}N and ^1H resonances of W133 also exhibit pronounced dispersion profiles that fit the same exchange parameters as the backbone amides in this region. There is no obvious interaction between the W133 indole NH and other parts of the protein, except through a water molecule (HOH327) (Figure 3) that acts as a bridge between the indole and the backbone amide of E129 as well as the side chain carboxyl of E134. This water molecule is conserved in 62% of the X-ray structures of *E. coli* DHFR. Another less well-conserved water molecule (HOH359, 49%) is also within hydrogen bonding distance of HOH327. These observations suggest that the HOH327 water acts as a bridge that couples conformational changes in the protein backbone to the indole ring of the tryptophan. The observed changes in the indole proton and nitrogen chemical shifts suggest that this water molecule is an integral part of the protein structure that remains bound in the excited state, albeit with subtle changes in its hydrogen bonding interactions with the W133 indole NH. The magnitude and sign of ωN (+1.62 ppm) show that the ^{15}N resonance is shifted downfield in the excited state, consistent with a stronger hydrogen bond to the bound water.

Active-Site Relaxation Processes

Previously we have observed a fast global process (4200 s^{-1}) in the active site of the DHFR mutant N23PP/S148A.³¹ This process was observed in the β -sheet and neighboring V10–V13 loop and modulates the chemical shifts of backbone amides extending from the binding pocket of the cofactor to the substrate binding pocket. The extensive hydrogen bonding network (Figure 4) includes the water molecule HOH301, which is hydrogen bonded to the folate NH₂, W30 indole NH, Y111 CO, and Thr113 O γ . The large ω N for the W30 indole suggests a change in the average hydrogen bond distance to HOH301 of up to 0.1 Å in the excited state. The sign of ω N (–2.46 ppm) shows that the ¹⁵N resonance is shifted upfield in the excited state, indicating weaker hydrogen bonding to the bound water.

The major factor propagating structural change from the active site to the B helix is the hydrogen bonding network within the β -sheet. Chemical shift changes are observed for both ¹H and ¹⁵N from residues in strand A (residues 6–9) to strand F (residues 111–115) as well as on strand H (residue 152).³¹ The conserved water HOH301 constitutes the only physical connection between the β -sheet and the C-terminal region of the B helix, forming a hydrogen bonded bridge that connects the indole side chain of W30 to β -strand F and couples the B helix to the fast conformational fluctuations in the β -sheet and the distant (~20 Å) V10–V13 loop. In a majority of the X-ray structures of *E. coli* DHFR, the HOH301 is positioned where it also hydrogen bonds to the pyrimidine amino group of bound folate. HOH301 thus plays a crucial role in the positioning of the pterin ring in the active site and couples motions at the active site all the way to W30 and F31. Although collective structural changes in networks of hydrogen bonds have been described previously,⁵⁹ our observations additionally show that hydrogen-bonded water molecules can also be instrumental in transducing local motions through proteins.

CD-loop Relaxation Processes

In the CD-loop, the water molecule that is hydrogen bonded to the indole NH of W47 is well conserved (78%) in the X-ray structures of *E. coli* DHFR. The largest chemical shift change in the CD-loop is observed on the indole proton suggesting that this probe is close to the site of structural change. The W47 indole is hydrogen bonded to HOH326, which forms part of an extensive network of water mediated hydrogen bonds in proximity to the CD-loop (see Figure 5). Conformational change of the CD-loop itself is unlikely because no backbone ¹⁵N dispersion was observed for residues in this loop.

Relaxation Dispersion in Indoles of Wild Type DHFR

The relaxation dispersion profiles in Figure 1 show that the conformational exchange in the C-terminal associated, active-site, and CD-loop regions (probed by indoles of residues W133, W30, and W47) occurs in both the N23PP/S148A mutant and wild type DHFR complexes. In addition, the wild type protein, shows large dispersion for the indole ¹⁵N and ¹H resonances of W22. The W22 dispersion profiles fit to a two-site model with a rate of ~1400 s^{–1}; this process is associated with the exchange between the closed ground state and an occluded excited state, a process that is suppressed by the N23PP mutation.³² The conserved water molecule hydrogen bonded to W22 N ϵ is HOH308 and is present in both the WT occluded ground state structure (4NX6, folate:NADP⁺) and the occluded ground

state structure (5CCC, ddTHF:NADP⁺) and appears to play a critical role in the chemical step and product release.^{60–63} In Figure 6 we show the conserved active site waters in the closed ground state conformation of WT *E. coli* DHFR bound to folate:NADP⁺, a model of the Michaelis complex, illustrating the complex hydrogen bonding network where HOH308 is hydrogen bonded to the W22 indole as well as the conserved water HOH305, the keto oxygen of the pterin, and the side chain of D27. HOH301 associated with W30 also appears to be part of this hydrogen bonding network, further emphasizing the role of the water molecules in binding and coupling of the substrate to the protein.

Our observations that water dynamics play an integral role in the protein structural changes accompanying μ s-ms motions are consistent with a recent molecular dynamics study on glutamate dehydrogenase, where changes in hydration appeared to be necessary for collective domain motions.⁶⁴ Water molecules were observed to gate structural transitions by wetting and drying protein cavities, leading to changes in the local hydrogen bonding networks. Results presented here indicate that many of the indoles in proteins are hydrated even though the individual lifetimes of waters may be very short. This has potential implications for the development and bench marking of molecular mechanics force fields, which should be able to accurately capture the hydrated states of these indoles. It has been noted previously that water molecules may function as extensions of protein side chains.⁷ Our present data are consistent with this and show, furthermore, that bound water molecules are an integral part of the protein that are able to transmit μ s-ms timescale conformational fluctuations through the protein structure. This is exemplified by our observation that the indole NH moieties of Trp30 and Trp133 are coupled to conformational changes in distant regions of the protein.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Gerard Kroon for support with NMR instrumentation and expertise and David Case for helpful discussions and advice with chemical shift calculations using AFNMR. This work was supported by the U.S. National Institutes of Health (NIH) grant GM75995 (P.E.W.) and the Skaggs Institute of Chemical Biology (P.E.W.). D.O. was supported by a Collen-Francqui fellowship from the Belgian American Educational Foundation (B.A.E.F.).

References

1. Laage D, Elsaesser T, Hynes JT. Chem. Rev. 2017; 117:10694. [PubMed: 28248491]
2. Bellissent-Funel M-C, Hassanali A, Havenith M, Henchman R, Pohl P, Sterpone F, van der Spoel D, Xu Y, Garcia AE. Chem. Rev. 2016; 116:7673. [PubMed: 27186992]
3. Saenger W. Annu. Rev. Biophys. Biophys. Chem. 1987; 16:93. [PubMed: 3297093]
4. Meyer E. Protein Sci. 1992; 1:1543. [PubMed: 1304887]
5. Nakasako M. Phil. Trans. Roy. Soc. B. 2004; 359:1191. [PubMed: 15306376]
6. Otting G, Liepinsh E, Wuthrich K. Science. 1991; 254:974. [PubMed: 1948083]
7. Jiang L, Kuhlman B, Kortemme T, Baker D. Prot. Struct. Funct. Genet. 2005; 58:893.
8. Nucci NV, Pometun MS, Wand AJ. Nat. Struct. Mol. Biol. 2011; 18:245. [PubMed: 21196937]
9. Halle B. Philos Trans R Soc Lond B Biol Sci. 2004; 359:1207. [PubMed: 15306377]

10. Patel H, Gruning BA, Gunther S, Merfort I. *Bioinformatics*. 2014; 30:2978. [PubMed: 24990608]
11. Mattos C, Bellamacina CR, Peisach E, Pereira A, Vitkup D, Petsko GA, Ringe D. *J. Mol. Biol.* 2006; 357:1471. [PubMed: 16488429]
12. Cooke R, Kuntz ID. *Annu. Rev. Biophys. Bioeng.* 1974; 3:95. [PubMed: 4371481]
13. Brändén C-I, Alwyn Jones T. *Nature*. 1990; 343:687.
14. Thanki N, Thornton JM, Goodfellow JM. *J. Mol. Biol.* 1988; 202:637. [PubMed: 3172231]
15. Frauenfelder H, Fenimore PW, Chen G, McMahon BH. *Proc. Natl. Acad. Sci. U.S.A.* 2006; 103:15469. [PubMed: 17030792]
16. Fenimore PW, Frauenfelder H, McMahon BH, Young RD. *Proc. Natl. Acad. Sci. U.S.A.* 2004; 101:14408. [PubMed: 15448207]
17. Frauenfelder H, Chen G, Berendzen J, Fenimore PW, Jansson H, McMahon BH, Stroe IR, Swenson J, Young RD. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:5129. [PubMed: 19251640]
18. Qin Y, Wang L, Zhong D. *Proc. Natl. Acad. Sci. U.S.A.* 2016; 113:8424. [PubMed: 27339138]
19. Kaieda S, Halle B. *J. Phys. Chem. B*. 2013; 117:14676. [PubMed: 24195787]
20. Persson E, Halle B. *J. Am. Chem. Soc.* 2008; 130:1774. [PubMed: 18183977]
21. Huang Y, Nam K, Westlund PO. *Phys. Chem. Chem. Phys.* 2013; 15:14089. [PubMed: 23868443]
22. Morón MC. *Phys. Chem. Chem. Phys.* 2012; 14:15393. [PubMed: 23059844]
23. Marklund EG, Larsson DS, van der Spoel D, Patriksson A, Caleman C. *Phys. Chem. Chem. Phys.* 2009; 11:8069. [PubMed: 19727514]
24. Salari R, Chong LT. *J. Phys. Chem. Lett.* 2010; 1:2844. [PubMed: 24920993]
25. Koenig SH, Schillinger WE. *J. Biol. Chem.* 1969; 244:3283. [PubMed: 5792660]
26. Kuntz ID Jr, Kauzmann W. *Adv. Prot. Chem.* 1974; 28:239.
27. Kay LE. *J. Mol. Biol.* 2016; 428:323. [PubMed: 26707200]
28. Henzler-Wildman KA, et al. *Nature*. 2007; 450:838. [PubMed: 18026086]
29. Boehr DD, McElheny D, Dyson HJ, Wright PE. *Science*. 2006; 313:1638. [PubMed: 16973882]
30. Oyen D, Fenwick RB, Stanfield RL, Dyson HJ, Wright PE. *J. Am. Chem. Soc.* 2015; 137:9459. [PubMed: 26147643]
31. Fenwick RB, Oyen D, Wright PE. *Phys. Chem. Chem. Phys.* 2016; 18:5789. [PubMed: 26426424]
32. Bhabha G, Lee J, Ekiert DC, Gam J, Wilson IA, Dyson HJ, Benkovic SJ, Wright PE. *Science*. 2011; 332:234. [PubMed: 21474759]
33. Osborne MJ, Venkitakrishnan RP, Dyson HJ, Wright PE. *Protein Sci.* 2003; 12:2230. [PubMed: 14500880]
34. Loria JP, Rance M, Palmer AG III. *J. Am. Chem. Soc.* 1999; 121:2331.
35. Ishima R, Torchia DA. *J. Biomol. NMR.* 2003; 25:243. [PubMed: 12652136]
36. Aoto PC, Fenwick RB, Kroon GJA, Wright PE. *J. Magn. Reson.* 2014; 246:31. [PubMed: 25063954]
37. Hyberts SG, Takeuchi K, Wagner G. *J. Am. Chem. Soc.* 2010; 132:2145. [PubMed: 20121194]
38. Delaglio F, Grzesiek S, Vuister GW, Guang Z, Pfeifer J, Bax A. *J. Biomol. NMR.* 1995; 6:277. [PubMed: 8520220]
39. Kazimierczuk K, Orekhov VY. *Angew. Chem. Int. Ed. Engl.* 2011; 50:5556. [PubMed: 21538743]
40. Vranken WF, et al. *Prot. Struct. Funct. Genet.* 2005; 59:687.
41. Sugase K, Konuma T, Lansing JC, Wright PE. *J. Biomol. NMR.* 2013; 56:275. [PubMed: 23754491]
42. Korzhnev DM, Neudecker P, Mittermaier A, Orekhov VY, Kay LE. *J. Am. Chem. Soc.* 2005; 127:15602. [PubMed: 16262426]
43. Skrynnikov NR, Dahlquist FW, Kay LE. *J. Am. Chem. Soc.* 2002; 124:12352. [PubMed: 12371879]
44. Hwang T-L, van Zijl PCM, Mori S. *J. Biomol. NMR.* 1998; 11:221. [PubMed: 9679296]
45. Xu XP, Case DA. *J. Biomol. NMR.* 2001; 21:321. [PubMed: 11824752]
46. Swails J, Zhu T, He X, Case D. *J. Biomol. NMR.* 2015; 63:125. [PubMed: 26232926]
47. Zhang H, Neal S, Wishart DS. *J. Biomol. NMR.* 2003; 25:173. [PubMed: 12652131]

48. Schrödinger LLC. The PyMOL molecular graphics system, version 1.8.2.3. 2017
49. McElheny D, Schnell JR, Lansing JC, Dyson HJ, Wright PE. Proc. Natl. Acad. Sci. U.S.A. 2005; 102:5032. [PubMed: 15795383]
50. Fenwick RB, van den Bedem H, Fraser JS, Wright PE. Proc. Natl. Acad. Sci. U.S.A. 2014; 111:E445. [PubMed: 24474795]
51. Bean JW, Peishoff CE, Kopple KD. Int. J. Pept. Prot. Res. 1994; 44:223.
52. Mahalakshmi R, Sengupta A, Raghothama S, Shamala N, Balaram P. J Pept Res. 2005; 66:277. [PubMed: 16218995]
53. Watenpaugh KD, Margulis TN, Sieker LC, Jensen LH. J. Mol. Biol. 1978; 122:175. [PubMed: 682189]
54. Blake CC, Pulford WC, Artymiuk PJ. J. Mol. Biol. 1983; 167:693. [PubMed: 6876162]
55. Kovacs H, Agback T, Isaksson J. J. Biomol. NMR. 2012; 53:85. [PubMed: 22528292]
56. Brunne RM, Liepinsh E, Otting G, Wuthrich K, van Gunsteren WF. J. Mol. Biol. 1993; 231:1040. [PubMed: 7685828]
57. Berjanskii MV, Wishart DS. J. Am. Chem. Soc. 2013; 135:14536. [PubMed: 24032347]
58. Petkova AT, Hatanaka M, Jaroniec CP, Hu JG, Belenky M, Verhoeven M, Lugtenburg J, Griffin RG, Herzfeld J. Biochemistry. 2002; 41:2429. [PubMed: 11841237]
59. Ishima R, Baber J, Louis JM, Torchia DA. J. Biomol. NMR. 2004; 29:187. [PubMed: 15014232]
60. Reyes VM, Sawaya MR, Brown KA, Kraut J. Biochemistry. 1995; 34:2710. [PubMed: 7873554]
61. Lee H, Reyes VM, Kraut J. Biochemistry. 1996; 35:7012. [PubMed: 8679526]
62. Warren MS, Brown KA, Farnum MF, Howell EE, Kraut J. Biochemistry. 1991; 30:11092. [PubMed: 1932031]
63. Boehr DD, McElheny D, Dyson HJ, Wright PE. Proc. Natl. Acad. Sci. U.S.A. 2010; 107:1373. [PubMed: 20080605]
64. Oroguchi T, Nakasako M. Sci. Rep. 2016; 6:26302. [PubMed: 27193111]

Figure 1.

Relaxation dispersion profiles for the five tryptophan indole groups in *E. coli* DHFR mutant N23PP/S148A and for the wild-type protein. Nitrogen dispersion profiles measured at 800 and 500 MHz, black and red respectively. Proton dispersion profiles measured at 800 and 500 MHz, green and blue respectively. Data for the wild-type protein have been individually fitted and do not represent global fits to the data.

Figure 2.
Structural environments for the five tryptophan indole groups in *E. coli* DHFR mutant N23PP/S148A. The highest occupancy conformation is shown in the case that more conformations exist for the tryptophan side chain. Water molecules are shown as red spheres.

Figure 3.
The water-mediated dynamic coupling in the C-terminal associated region of N23PP/S148A *E. coli* DHFR. Hydrogen bonds are shown as orange dashed lines. Blue and green spheres indicate the excited-state chemical shift change $\Delta\omega$ for nitrogen and proton respectively.

Figure 4.
The water-mediated dynamic coupling in the active site of N23PP/S148A *E. coli* DHFR.
Hydrogen bonds are shown as orange dashed lines and blue and green spheres indicate the chemical shift change for nitrogen and proton respectively as in Figure 3.

Figure 5.
The water-mediated dynamic coupling in the and the CD-loop of N23PP/S148A *E. coli* DHFR. Hydrogen bonds are shown as orange dashed lines and blue and green spheres indicate the chemical shift change for nitrogen and proton respectively as in Figure 3.

Figure 6. Stereo image of the water-mediated hydrogen bond coupling between the ligands and the protein mediated by water molecules in the active site of WT *E. coli* DHFR. Hydrogen bonds are shown as orange dashed lines.

Chemical shift changes ($\Delta\omega$) and exchange parameter values determined from fits of relaxation dispersion data for the N23PP/S148A mutant as well as the ground state chemical shifts (δN and δH).

Table 1

Indole NeH	k_{ex} (s^{-1})*	p_b (%)*	ωH (ppm)	ωN (ppm)	δH (ppm)	δN (ppm)
W22		no conformational exchange observed			10.23	131.10
W30	4200	4.2	0.45 (0.04)	(-)-2.46 (0.2)	10.35	130.27
W47	3100	2.13	0.86 (0.04)	1.36 (0.08)	10.24	128.16
W74		no conformational exchange observed			10.12	130.23
W133	630	3.32	0.20 (0.01)	(+)-1.62 (0.04)	10.16	127.77

* Values for exchange rate (k_{ex}) and the minor population (p_b) are those previously determined for fits of relaxation dispersion data for DHFR.³¹ The sign of ωN determined from HSQC-HMQC spectra is given in parentheses.

Table 2

Conserved water molecules located in hydrogen bonding positions around the tryptophan indoles of DHFR identified using PyWATER.¹⁰ Water numbering is given for 4NX6 (WT folate:NADP⁺) and 3QL0 (N23PP/S148A folate:NADP⁺).

indole	Acceptor (4NX6)	Dist (Å) (4NX6)	θ (deg.) ⁺ (4NX6)	Acceptor (3QL0)	Dist (Å) (3QL0)	θ (deg.) ⁺ (3QL0)	Cons (%) ⁺⁺
W22 Ne	HOH308	2.1	154.5	HOH216	1.9	167.0	71
W30 Ne	HOH301	2.1	142.4	HOH183	2.0	137.2	97
W47 Ne	HOH326	2.2	167.2	HOH201	2.2	154.8	78
W74 Ne	D69	1.9	158.3	D69	1.9	169.2	NA
W133 Ne	HOH327	2.2	159.7	HOH301	1.9	164.7	62

^{*} the distance between the indole NH proton and the water oxygen

⁺ the angle between the indole NH nitrogen and proton and the water/carbonyl oxygen

⁺⁺ conservation is defined as the percentage of structures for which a similar water is found within 2.4 Å

Table 3

Calculated and observed chemical shifts for tryptophan indoles.

system	δ ^1H (ppm)	δ ^{15}N (ppm)
RefDB mean	10.1	129.3
RefDB non H-bonded mean ⁺	7.4	122.1
RefDB obs δ 	2.7	7.2
DFT pred δ 	2.7	8.2

⁺ see Table S1 for details