nature COMMUNICATIONS

ARTICLE

Received 23 May 2013 | Accepted 13 Sep 2013 | Published 1 Nov 2013

DOI: 10.1038/ncomms3611

Probing water micro-solvation in proteins by water catalysed proton-transfer tautomerism

Jiun-Yi Shen[1,*], Wei-Chih Chao[1,*], Chun Liu[1], Hsiao-An Pan[1], Hsiao-Ching Yang[2], Chi-Lin Chen[1,2], Yi-Kang Lan[2], Li-Ju Lin[3], Jinn-Shyan Wang[3], Jyh-Feng Lu[3], Steven Chun-Wei Chou[1], Kuo-Chun Tang[1] & Pi-Tai Chou[1]

Scientists have made tremendous efforts to gain understanding of the water molecules in proteins via indirect measurements such as molecular dynamic simulation and/or probing the polarity of the local environment. Here we present a tryptophan analogue that exhibits remarkable water catalysed proton-transfer properties. The resulting multiple emissions provide unique fingerprints that can be exploited for direct sensing of a site-specific water environment in a protein without disrupting its native structure. Replacing tryptophan with the newly developed tryptophan analogue we sense different water environments surrounding the five tryptophans in human thromboxane A_2 synthase. This development may lead to future research to probe how water molecules affect the folding, structures and activities of proteins.

[1] Department of Chemistry, Center for Emerging Material and Advanced Devices, National Taiwan University, Taipei 10617, Taiwan. [2] Department of Chemistry, Fu-Jen Catholic University, New Taipei City 24205, Taiwan. [3] School of Medicine, Fu-Jen Catholic University, New Taipei City 24205, Taiwan. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to H.-C.Y. (email: hcyang_chem@mail.fju.edu.tw) or to P.-T.C. (email: chop@ntu.edu.tw).

In proteins, water ubiquitously participates in dictating structure and functionality, including protein secondary and tertiary structures, the spontaneous formation of membranes and the recognition of signaling molecules in signal transduction[1–3]. Recent advances have provided more convincing evidence that the water molecules in proteins are key elements in activating bio-functionality such as enzymatic reactions[4–7] or *vice versa*, the protein (for example, an antifreeze protein) may affect the organization of water molecules[8]. Probing the water environment of a specifically interesting site in proteins thus may pave a way to understanding the underlying mechanisms and functionality. Unfortunately, although enormous efforts have been made in the characterization of fundamental aqueous hydration phenomena on protein surfaces[9–11], little insight has been gained into water micro-solvation in protein[12,13]. Most of the relevant approaches have focused instead on probing the polarity of the local environment in proteins, particular around the active sites, by means of fluorescence solvatochromism[14,15].

To date, **Trp** in bacterial proteins has been popularly substituted by **(7-aza)Trp** or in part with **(2-aza)Trp**, in which the **indole** moiety in **Trp** is replaced by **7-azaindole** or **2-azaindole** (Fig. 1), for either investigating the antimetabolic properties of the **Trp** isosteres[16] or accessing the local polarity of **Trp** in proteins with unknown tertiary structures[17–21]. Catalysed by protic solvent molecules such as alcohols (for example, methanol), **7-azaindole** undergoes solvent-assisted excited-state proton transfer (ESPT), resulting in an N_7-H tautomer form (Fig. 2a) that exhibits green emission ($\sim$510 nm)[22–24]. In water, however, the N_7-H tautomer green emission is virtually non-observable[25]; this lack has been attributed to the much slower proton-transfer rate constant ($\sim 10^9\ s^{-1}$) together with dominant radiation-less deactivation pathways[22]. The photophysical properties of **(7-aza)Trp** are very similar to those of **7-azaindole**, exhibiting essentially no N_7-H tautomer form emission[16,26–28] and thus cannot be used for probing any water-associated photophysical phenomena in proteins.

To surpass the limitation of the traditional **Trp** analogues, we strategically design a new type of aza-Trp by overlapping two azaindole parent moieties, that is, **2-azaindole** and **7-azaindole**, yielding a new core moiety **2,7-diazaindole** (Fig. 1) that is capable of sensing a protein water environment. The underlying concept originates from the molecule **3-cyano-7-azaindole**[29] (Fig. 1), in which the cyano electron-withdrawing ability increases hydrogen acidity of the N_1-H form and hence facilitates the overall ESPT rate in water, rendering an intensive N_7-H proton-transfer tautomer emission in the green (Fig. 2a). For **2,7-diazaindole**, the replacement of pyrrole by the pyrazole moiety, in which the N(2) atom acts as an efficient electron-withdrawing group[30], is expected to increase the N_1-H proton acidity without perturbing the parent geometry. Using **2,7-diazaindole** as the core moiety, we further develop a new **Trp** analogue, **(2,7-aza)Trp** (Fig. 1), that exhibits remarkably different properties from classical aza-Trp analogues in that water-assisted proton-transfer isomerization takes place in both the excited and the ground states (Fig. 2a,b). The proposed schematic diagram of the proton-transfer cycle for **2,7-diazaindole** and **(2,7-aza)Trp** is depicted in Fig. 2c. Figure 2c unveils the water-catalysed isomerization between the N_1-H form and the N_2-H form in the ground state and ESPT for the N_1-H form in the excited state, as supported by the following firm spectroscopy and dynamics evidence. The associated multiple emissions thus serve as a unique fingerprint

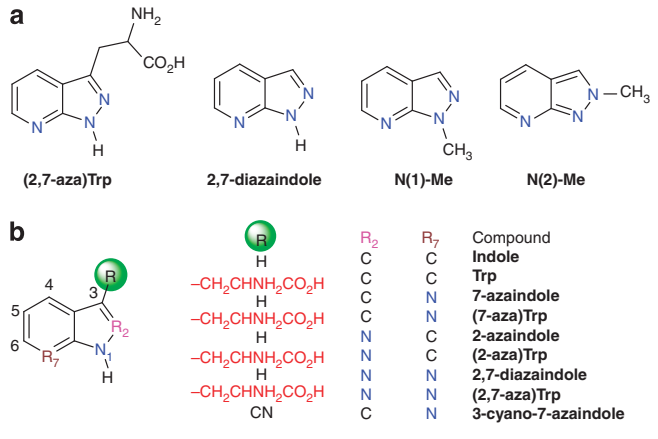

Figure 1 | The structural formulae of relevant indole and tryptophan derivatives. (**a**) The chemical structures of **(2,7-aza)Trp**, **2,7-diazaindole**, **N(1)-Me** and **N(2)-Me**. (**b**) The chemical structures of various indole and tryptophan analogues.

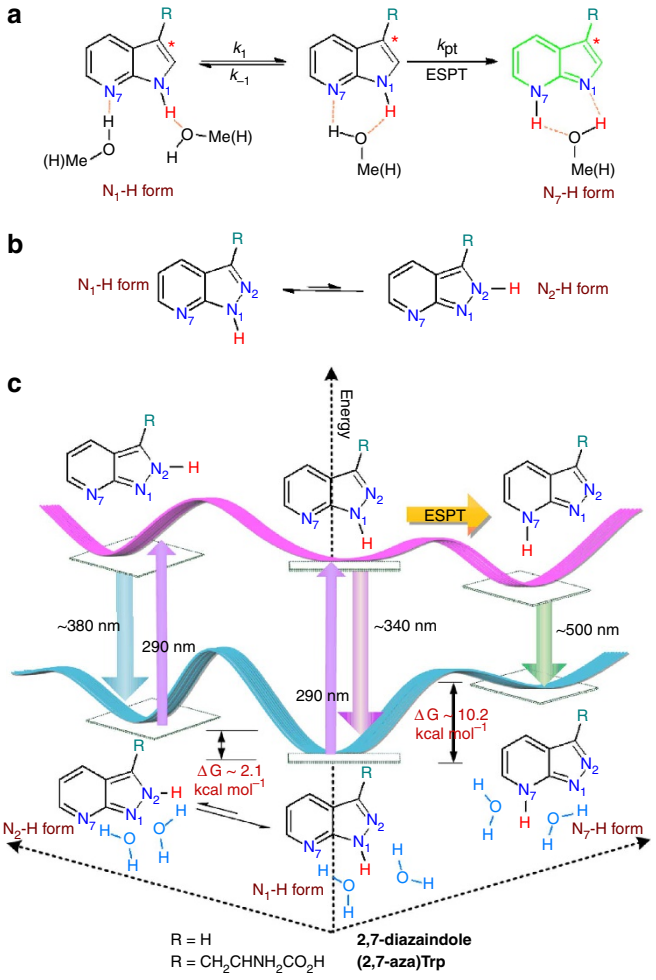

Figure 2 | Proposed ground and electronically excited states mechanism. (**a**) The proposed protic solvent catalysed ESPT mechanism for **7-azaindole** and its derivatives. The asterisk * indicates the electronically excited state. This mechanism incorporates a fast excited-state equilibrium between polysolvated and 1:1 cyclic hydrogen-bonded N_1-H/H_2O (or methanol) followed by proton tunnelling k_{pt}[22], which has been unambiguously proved by the kinetic deuterium isotope effect[29]. Under the assumption of $k_{-1} > k_1$ and k_{pt}, the overall ESPT rate constant k_{rxn} can be expressed as $k_{rxn} = (k_1/k_{-1})k_{pt}$. (**b**) The ground-state equilibrium for **2,7-diazaindole** and its derivatives between N_1-H form and N_2-H form isomers in water. (**c**) The proposed schematic diagram of a proton-transfer cycle for **2,7-diazaindole** and **(2,7-aza)Trp** in water, in which the barrier and hence potential energy surface are arbitrarily chosen.

for sensing the water environment. The proof of concept is made by the site-specific substitution **(2,7-aza)Trp** for tryptophan in human thromboxane A_2 synthase (TXAS) to resolve a distinct water environment in various **Trp** sites of TXAS.

Results

Proton-transfer tautomerism. Manifested in neutral water (pH = 7.0), either **(2,7-aza)Trp** or **2,7-diazaindole** reveals remarkable triply fluorescent bands (Fig. 3a,b and Table 1), consisting of an apparent shoulder at ~330 nm, a peak wavelength at ~370 nm and a much red-shifted emission band maximized at ~500 nm, corresponding to the existence of N_1-H, N_2-H and N_7-H isomers, respectively, in the excited state. The excitation spectra at 335 nm and at the very red edge (for example, 530 nm) of the emission are identical and they also resemble the absorption spectrum (Supplementary Fig. S1). As **2,7-diazaindole** and **(2,7-aza)Trp** were synthesized independently from different methods (Supplementary Fig. S2 and Supplementary Methods), the possibility of trace impurities contributing to the triple emission bands can be eliminated. Rather, the triple emission can be well rationalized by the proton-transfer tautomerism in both ground and excited states. Support is provided chemically by the methyl derivatives of **2,7-diazaindole**, **N(1)-Me** and **N(2)-Me** (Fig. 1), which exhibit prominent emissions at 342 and 367 nm, respectively, in neutral water (Fig. 3c and Supplementary Fig. S3). The emission of **N(2)-Me**, in view of the peak position (367 nm), is nearly identical to that (370 nm) of **2,7-diazaindole**, so the assignment of **2,7-diazaindole** 370 nm emission to the N_2-H isomer appears unambiguous. In addition, given the similarity of its emission peak wavelength (342 nm) to that of **N(1)-Me**, the 335 nm band in **2,7-diazaindole** can reasonably be ascribed to the emission of the N_1-H isomer. The difference in excitation origins for 335 and 370 nm emission bands (Supplementary Fig. S1) clearly indicates the existence of ground-state equilibrium between N_1-H and N_2-H isomers for **2,7-diazaindole** (Fig. 2b). In accordance with this assignment, we also propose the occurrence of water-catalysed ESPT in the N_1-H form, forming the N_7-H isomer, to account for the ~500 nm green emission. Taking account of the structural and spectral similarity, the same assignment of the triple emission should hold true for **(2,7-aza)Trp**.

Firm support for the tautomerism is given by the correlation of relaxation dynamics among emission bands (Table 1). At the short wavelength shoulder of 320 nm, **(2,7-aza)Trp** exhibits a single decay component of 260 ps in water. At the red edge of the green emission (540 nm), the relaxation dynamics consist of a finite rise component (263 ps) and a decay component with a lifetime of 0.58 ns (Fig. 3d and Table 1). The 263 ps rise dynamic, within experimental uncertainty of ± 30 ps, correlates well with the decay of the 340 nm band, supporting the precursor (340 nm band) and successor (500 nm band) type of reaction kinetics and hence the water-assisted ESPT for the N_1-H isomer. On the other hand, the relaxation dynamics at the peak wavelength of 380 nm consists of an instant, system-response-limited rise time (<60 ps), accompanied by two decay components, one with a small amplitude (~12%), fast decay (265 ps), and a major, long decay component (88%, 10.07 ns, see Fig. 3e). The former 265 ps

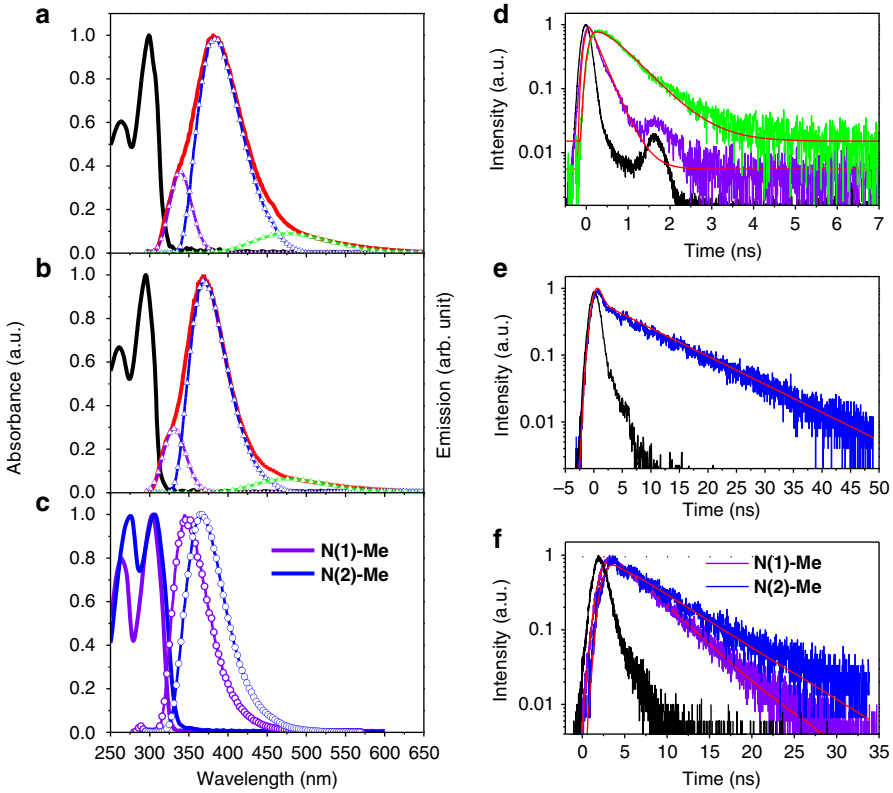

Figure 3 | Identification of the triple emission bands. (**a**,**b**) Absorption (black) and emission spectra (red) of **(2,7-aza)Trp** are shown in **a** and **2,7-diazaindole** are shown in (**b**) in neutral water. Further, the decomposed PL spectra are shown (violet, blue and green triangles). λ_{ex} = 290 nm. (**c**) Absorption and emission spectra of **N(1)-Me** and **N(2)-Me** in neutral water. (**d**) The emission decay dynamics of **(2,7-aza)Trp** monitored at 320 nm (violet) and 540 nm (green). (**e**) The emission decay dynamics of **(2,7-aza)Trp** monitored at 380 nm. (**f**) The emission decay dynamics of **N(1)-Me** and **N(2)-Me**. The instrument response and the fitting curves are marked in black and red lines, respectively. The residuals of the fit of emission decay dynamics as shown in Supplementary Fig. S8.

Table 1 | Photophysical properties.

	λ_{abs} (nm)	λ_{em} (nm)	λ_{mon} (nm)†	τ (ns)‡
(7-aza)indole§	288	386		0.91
(7-aza)Trp	289	400		0.65
2,7-diazaindole	295	335	320	0.22
		370	400	10.10
		495	550	0.21(rise), 1.30
(2,7-aza)Trp	300	340	320	0.26
		380	380	0.27(12%), 10.07(88%)
		500	540	0.26(rise), 0.58
N(1)-Me	304	342	360	3.92
N(2)-Me	306	367	370	5.20
(2,7-aza)Trp^{31,65,133,203,446}-TXAS	300	338	350	0.31(83%), 1.60(17%)
		507	520	0.27(rise), 1.18
(2,7-aza)Trp^{31,133,203,446}Phe^{65}-TXAS	300	340	350	0.28
		503	520	0.31(rise), 1.12
(2,7-aza)Trp^{31}Phe^{65,133,203,446}-TXAS	300	341	350	0.29
		507	520	0.30(rise), 1.23
(2,7-aza)Trp^{65}Phe^{31,133,203,446}-TXAS	300	332	350	1.55
(2,7-aza)Trp^{133}Phe^{31,65,203,446}-TXAS	300	341	350	0.21
		504	520	0.22(rise), 1.32
(2,7-aza)Trp^{203}Phe^{31,65,133,446}-TXAS	300	341	350	0.28
		503	520	0.32(rise), 1.47
(2,7-aza)Trp^{446}Phe^{31,65,133,203}-TXAS	300	341	350	0.30
		503	520	0.30(rise), 1.41

Photophysical properties of various azaindole and azatryptophan analogues in neutral water and mutated TXAS in buffer (20 mM sodium phosphate (pH 7.5), containing 10% glycerol, 0.2% cholic acid and 0.05% lubrol).
†The wavelength at which the measurement of relaxation dynamics was monitored.
‡Values are ± 0.03 ns in uncertainty. Values in parentheses indicate percentages and in brackets denote the rise component.
§Data from the study by Chen *et al.*[28]

component in the 380 nm emission decay, within the experimental error, is identical to 260 ps monitored at 320 nm and thus is assigned to the residue of the N_1-H emission band. The latter is the population decay of the N_2-H form emission. The same correlation among the triple emission bands is also resolved for **2,7-diazaindole** (Table 1), supporting firmly that both **2,7-diazaindole** and **(2,7-aza)Trp** in water undergo proton-transfer tautomerism in the ground state (N_1-H/N_2-H equilibrium) and the excited state (N_1-H $\rightarrow N_7$-H proton transfer).

One fundamental issue lies in the thermodynamics between N_1-H and N_2-H isomers in water (Fig. 2b). Experimentally, the molar absorption coefficient of **N(2)-Me**, $5.3 \times 10^3\ M^{-1}\ cm^{-1}$ (at 320 nm), was adopted for the N_2-H isomer of **2,7-diazaindole** at 320 nm, a wavelength that is free from N_1-H form interference. A known concentration of **2,7-diazaindole** was prepared, and the N_2-H/N_1-H equilibrium constant was measured to be 0.024 (detail in Supplementary Methods), corresponding to a difference of free energy of 2.2 kcal mol^{-1} at 298 K. This value is in good agreement with the computational calculations, estimating that the N_2-H form is 2.1 kcal mol^{-1} higher in energy than the N_1-H form for **2,7-diazaindole** in water. The same trend is observed but with a rising difference if we consider that less polar solvation by benzene, N_2-H form, has a higher energy by 6.3 kcal mol^{-1} for **2,7-diazaindole**. We find that this is true with no detectable N_2-H isomer form emission resolved in organic solvents for **2,7-diazaindole** (Supplementary Figs S4 and S5). This proton transfer induced multiple emission in water is remarkable and proves to provide fingerprints for probing a protein water environment.

TXAS homology structure and substrate/water gating channel. The focus of probing water in proteins is on the cytochrome P450 superfamily because recent advances have generated more convincing data indicating that water molecules have pivotal roles in the organism metabolic pathways in P450s[31–33]. We herein selected the structurally pending TXAS[34–36], an endoplasmic reticulum membrane protein containing five tryptophans, as a prototype to demonstrate the feasibility of **(2,7-aza)Trp** in probing water environments. The atypical catalysis without reductase or oxygen and the pathogenesis of major diseases make TXAS of prime interest to elucidate the corresponding structure–function relationship[37]. One recent advance indicates that W65 of TXAS is likely to be in a hydrophobic pocket surrounded by the Phe-cluster, whereas the environment of the other four tryptophans is yet to be determined[38]. A prior molecular dynamic (MD) simulation in this study unveils the spatial arrangement of these five tryptophan residues relative to the haem group and gating channels of TXAS (Fig. 4a,b), where W65 residue is on the substrate channel, W133 is located on the aqueduct water channel, W31 lies in the NH_2-terminal loop and W203 and W446 are close to the protein surface (detail in Supplementary Fig. S6).

Expression and analytics of (2,7-aza)Trp-TXAS. To examine the relevant tryptophans positions, TXAS mutants were synthesized in Trp-auxotrophic *Escherichia coli* cells[39], and they proved to be capable of co-translational incorporation by site-specifically replacing **Trp** residues with **(2,7-aza)Trp** (see Supplementary Methods). Fig. 4d and e depict the fluorescence spectra of various mutants. As for single **(2,7-aza)Trp** substitution at W31, W133, W203 and W446, the resulting **(2,7-aza)Trp^{31}Phe^{65,133,203,446}-TXAS**, **(2,7-aza)Trp^{133}Phe^{31,65,203,446}-TXAS**, **(2,7-aza)Trp^{203}Phe^{31,65,133,446}-TXAS** and **(2,7-aza)Trp^{446}Phe^{31,65,133,203}-TXAS** exhibit prominent, well-resolved dual emissions maximized at ~340 nm (N_1-H) and ~505 nm (N_7-H), indicating the occurrence of N_1-H to N_7-H proton-transfer in the excited state. By taking advantage of the MD simulations, we were able to examine the proximal environment around various Trp sites of TXAS. We found no relevant amino acids forming direct H-bonds with the mutant **(2,7-aza)Trps**, also eliminated the possibility of the ESPT catalysed from the protic residues. Therefore, the results ensure the water-catalysed ESPT for the single **(2,7-aza)Trp** substitution at each of four Trp sites in TXAS. For **(2,7-aza)Trp** substitution

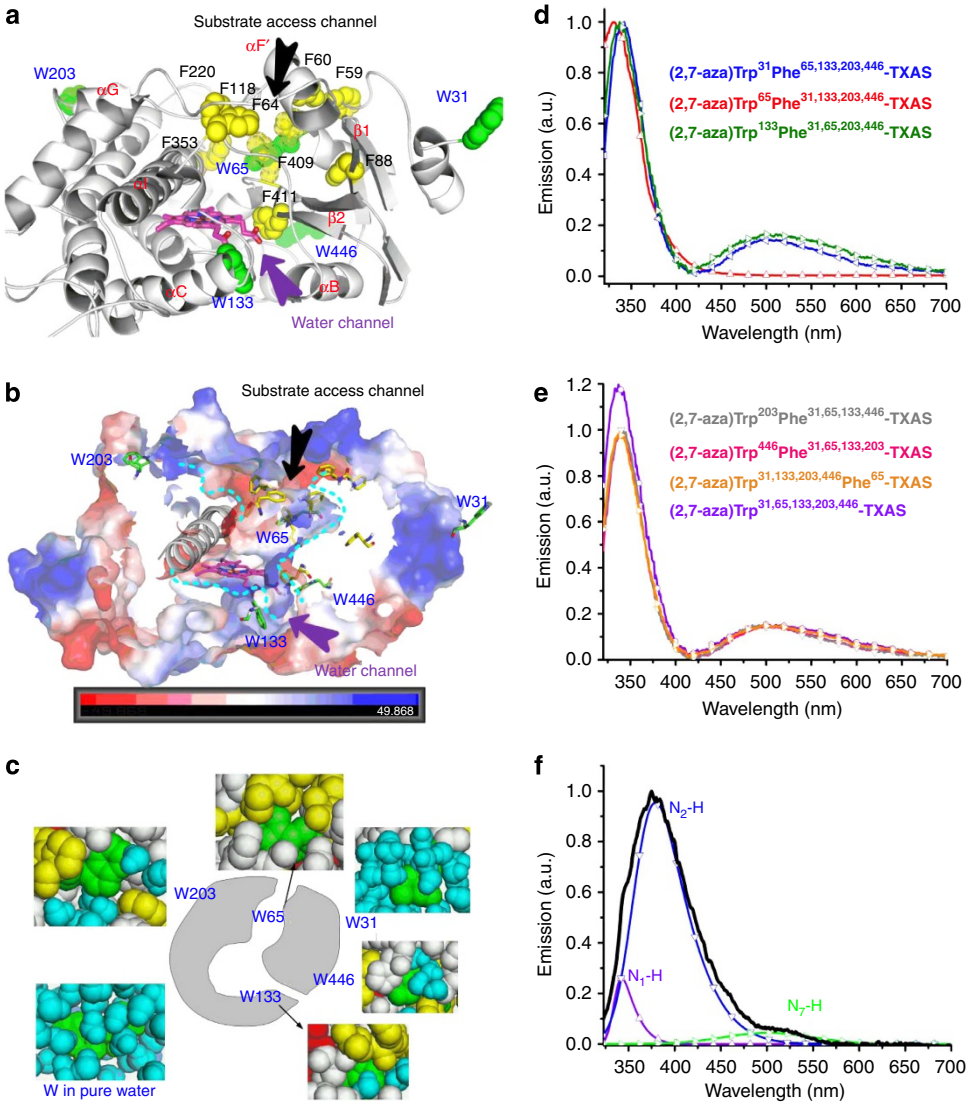

Figure 4 | MD simulation of TXAS protein and expression of TXAS mutants. (**a**) TXAS homology structure shows the characteristic P450 fold, illustrating the phenylalanine residues (yellow), five tryptophan positions (green) and haem (stick figures in violet). (**b**) The solvent-accessible surfaces of the active site cavities calculated using an H probe under an electrostatic potential (ESP) surface with electrophile/nucleophile (blue/red). Proposed gating channels are indicated. (**c**) The local environment of tryptophan (W) in hydrated TXAS (or water), in which W is shown in green, water is blue, acidic residues are red, hydrophobic residues are yellow and other residues are white. (**d**,**e**) Emission spectra of various TXAS mutants excited at 310 nm. The use of two separated figures is to avoid spectral congestion. (**f**) Emission spectra of denatured (**2,7-aza**)**Trp**^{31,65,133,203,446}**-TXAS** and the decomposed PL spectra.

at W31, W203 and W446, time-resolved measurement monitored at 350 nm clearly reveals a major 0.28–0.30 ns (±0.03) decay component, which correlates well with the rise time of the ~505 nm emission band of 0.30–0.32 ns (±0.03) (Table 1). Both spectroscopy and dynamics data are reminiscent of the water-catalysed ESPT of the **(2,7-aza)Trp** N_1-H isomer in water. As shown in Fig. 4d, the observed dual emission also guarantees the occurrence of water-catalysed ESPT for single **(2,7-aza)Trp**-substituted TXAS mutants at W133. However, compared with that of W31, the slight increase of green tautomer emission intensity and the associated faster rate of ESPT (0.21 ± 0.03 ns, Table 1) may imply its difference in a water environment. Supported by water titration experiments for **7-azaindole**[25,40], the breakdown of bulk water clusters may facilitate the rate of ESPT. Accordingly, we tentatively propose that there is a reduced water density region at the W133 site (*cf.* W31). The results of MD simulation elaborated in a later section seem to support this viewpoint.

In sharp contrast, for the single **(2,7-aza)Trp** substitution at W65, the resulting **(2,7-aza)Trp**^{65}**Phe**^{31,133,203,446}**-TXAS** exhibits solely an emission band maximized at 332 nm with a long lifetime of ~1.55 ns (Fig. 4d and Table 1). The apparent lack of a 500 nm emission band speaks to the nature of the prohibition of ESPT at the W65 position in TXAS, concluding the deficiency of water molecules surrounding the proximity of the W65 site. Support is further given by the full **(2,7-aza)Trp** substitution of all five tryptophan residues in the expression of **(2,7-aza)Trp**^{31,65,133,203,446}**-TXAS** mutant, which exhibits well-resolved dual emissions maximized at ~340 and 507 nm as well (Fig. 4e). Notice that the intensity ratio for the 340 nm emission band versus the 500 nm emission band is slightly higher

than that of quadruple **(2,7-aza)Trp** substitution of **(2,7-aza)Trp^{31,133,203,446}Phe^{65}-TXAS** mutant, indicating the increase of 340 nm emission due to the prohibition of ESPT in (2,7-aza)Trp65. Statistically, however, the mixture of incomplete substitution of **(2,7-aza)Trp** at any Trp sites is likely to occur in **(2,7-aza)Trp^{31,65,133,203,446}-TXAS**. Therefore, the corresponding quantitative spectral analysis is not possible. Nevertheless, the selected excitation wavelength at 310 nm should be virtually free of absorption by trace Trp residues. Unambiguously, the results present a unique feature, that the position at W65 is subject to a water-deficient microenvironment wherein the lack of surrounding water molecules prohibits ESPT, resulting in a solely normal N_1-H form 340 nm emission. The other four tryptophans (W31, W133, W203 and W446) in TXAS are in a water-accessible microenvironment that catalyses ESPT, resulting in an N_7-H tautomer green emission.

Importantly, unlike the triple emission of dissolved **(2,7-aza)Trp** in water, all native **(2,7-aza)Trp-TXAS** mutants reveal an absence of N_2-H form 380 nm emission (*cf.* Figs 3a and 4d,e). This seems to describe the vital folding structure of a solvated protein. In a native folding, the associated framework in protein (TXAS) is different from the unfolding residues exposed in the bulk water solution, in which the N_2-H isomer is solvated and stabilized at the sufficient water hydration level. Moreover, the TXAS conformer polarity may not be high enough to stabilize the N_2-H isomer, being similar to the lack of population of N_2-H isomer in organic solvents (*vide supra*, see Supplementary Fig. S4, in methanol). Firm support of this viewpoint is given by the reappearance of the N_2-H form 380 nm emission upon denaturing of **(2,7-aza)Trp^{31,65,133,203,446}-TXAS** variants (Fig. 4f). The lack of N_2-H isomer emission thus indicates that none of the **(2,7-aza)Trp** in **(2,7-aza)Trp-TXAS** mutants are bulk water-solvated, and so the Trps in wild-type TXAS, due to the structural similarity.

Explicit solvent MD simulations on both wild-type TXAS and **(2,7-aza)Trp**-replaced TXAS have been performed and showed no detectable differences in terms of the solvated homology structure. The results indicate that the TXAS matrices are dynamically coupled to the hydration water by a relay-type of hydrogen-bonding configuration (Fig. 4c). This, together with the close-up inspections of the hydrogen-bonding configurations for each **(2,7-aza)Trp**-replaced tryptophans in TXAS (Supplementary Fig. S6), draws two important remarks. First, there are clearly no water molecules surrounding the W65 site. Second, except for W65, water micro-solvation does exist for all other **(2,7-aza)Trp**. When compared with explicit water MD simulations, a locally fluctuating density profile is revealed (detail in Supplementary Methods and Supplementary Fig. S7), so it seems to be apparent that fewer trapped water molecules associate with the W133 site. Both results are consistent with the experimental observation (*vide supra*). The combination of experimental and MD approaches thus proves the feasibility of **(2,7-aza)Trp** in probing the micro-solvation structures in protein.

As manifested by explicit solvent MD simulations, two distinct channels are identified as the substrate access channel (W65 locus) and the aqueduct water channel (W133 locus) in protein TXAS (Fig. 4b). The associated faster rate of ESPT in (2,7-aza)Trp133 seems to specify the destination of hydration water behaviours between the protein surface and aqueduct water channel in protein. Also worthy of note is the hydrophobic region located on the substrate channel; this region, with phenylalanine residues (Phe60, Phe64, Phe127, Phe220, Phe409 and Phe411) and W65, forms a 'Phe-cluster'. This distinctive Phe-cluster and hydrophobic residues on the TXAS substrate channel[38] result in a less polar and water-deficient microenvironment at the active site, which is in stark contrast to its counter-enzyme prostacyclin synthase (PGIS), which has greater water accessibility to the active haem site[39,41]. This prediction is firmly supported by the prohibition of ESPT in (2,7-aza)Trp65. The location of the Phe-cluster results in a gating conformation of the active site of TXAS. MD simulations also reveal that conformational movement of the Phe-cluster residues could open/close the channel and consequently increase/decrease the volume of the active site of TXAS. We hypothesize that this relevant locus of the 'Phe-cluster' is crucial in gating different ligand/water access in an early step of allowance to the active haem site.

Discussion

In summary, we have developed a new tryptophan analogue, **(2,7-aza)Trp**, which undergoes water-catalysed proton-transfer tautomerism in both ground and excited states, giving a multiple emission that is sensitive to the water environment. Incorporating **(2,7-aza)Trp** into human TXAS, we then demonstrate for the first time the feasibility to probe water molecules surrounding in proteins. Manifested by explicit solvent MD simulation, **(2,7-aza)Trp-TXAS** mutant maintains its parent wild-type TXAS structure. Exploiting co-translational **(2,7-aza)Trp** as a probe, correlation between these biocompatible probing spectra and a water micro-solvation environment has been established, thereby identifying specific tryptophan locations on TXAS. The incorporation of **(2,7-aza)Trp** into TXAS proves **(2,7-aza)Trp** to be a novel optical probe that allows transmission of its surrounding water environment into unique fluorescence spectral features. In combination with in-depth MD simulation, future quantitative correlation for the ESPT dynamics versus number/orientation of water molecules is feasible, showing different hydration environments between the substrate access channel, aqueduct water channel and protein surface in, for example, TXAS. The superb water-sensing capability in terms of ESPT and ground-state N_1-H/N_2-H equilibrium thus provides an unprecedented tool for probing the water environment in bio-systems on a structural basis.

Methods

General synthesis and characterization. Full experimental details, synthesis and characterization for all compounds, computational methodology, growth and induction of bacteria are included in the Supplementary Methods.

Steady-state and time-resolved fluorescence spectroscopy. Steady-state absorption and emission spectra were recorded on a Hitachi (U-3310) spectrophotometer and an Edinburgh (FS920) fluorimeter, respectively. Both wavelength-dependent excitation and emission response of the fluorimeter had been calibrated. Time-resolved spectroscopic measurements were carried out by means previously reported elsewhere in detail[42,43]. Briefly, nanosecond (ns) lifetime studies were performed with a time-correlated single-photon counting (TCSPC, Edinburgh FL 900), using a hydrogen-filled lamp as the excitation source. The emission decays were analysed by the sum of exponential functions as below, which allows partial elimination of instrument time broadening and thus renders a temporal resolution of $\sim$300 ps. To achieve a faster time resolution, studies were also performed using a TCSPC system (Edinburgh OB-900L) coupled with an excitation light from the third harmonic generation (THG, at 306 nm) of pulse-selected femtosecond laser pulses at 920 nm (90 fs, Tsunami and Model 3980 pulse picker, Spectra-Physics). The temporal resolution is $\sim$60 ps. Data were analysed by using the nonlinear least-squares procedure in combination with an iterative convolution method.

Tautomer structure calculations. The theoretical approach used the Second-Order Approximate Coupled-Cluster method (CC2)[44] together with a resolution-of-the-identity technique to enhance performance. All atoms were employed using correlation-consistent basis sets of double-ζ quality with the diffuse and polarization functions (aug-cc-pVDZ)[45]. All calculations incorporated the solvation (COSMO model)[46] for ground-state geometry optimization. All the theoretical calculations were performed with the TURBOMOLE programme package (version 6.3)[47–50]. Thermodynamics of various isomers in different solvents were calculated using the DFT approach (B3LYP 6–311 + g(d,p)) incorporating a continuum solvation model (Gaussian 09)[51].

References

1. Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **108,** 74–108 (2008).
2. Chaplin, M. Do we underestimate the importance of water in cell biology? *Nat. Rev. Mol. Cell Biol.* **7,** 861–866 (2006).
3. Levy, Y. & Onuchic, J. N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* **35,** 389–415 (2006).
4. Abel, R. *et al.* Contribution of explicit solvent effects to the binding affinity of small-molecule inhibitors in blood coagulation factor serine proteases. *Chem. Med. Chem.* **6,** 1049–1066 (2011).
5. Robinson, D. D., Sherman, W. & Farid, R. Understanding kinase selectivity through energetic analysis of binding site waters. *Chem. Med. Chem.* **5,** 618–627 (2010).
6. Beuming, T. *et al.* Thermodynamic analysis of water molecules at the surface of proteins and applications to binding site prediction and characterization. *Proteins* **80,** 871–883 (2012).
7. Wang, L., Berne, B. J. & Friesner, R. A. Ligand binding to protein-binding pockets with wet and dry regions. *Proc. Natl Acad. Sci. USA* **108,** 1326–1330 (2011).
8. Meister, K. *et al.* Long-range protein-water dynamics in hyperactive insect antifreeze proteins. *Proc. Natl Acad. Sci. USA* **110,** 1617–1622 (2013).
9. Pal, S. K. & Zewail, A. H. Dynamics of water in biological recognition. *Chem. Rev.* **104,** 2099–2123 (2004).
10. Zhang, L. *et al.* Mapping hydration dynamics around a protein surface. *Proc. Natl Acad. Sci. USA* **104,** 18461–18466 (2007).
11. Lakshmikanth, G. S. & Krishnamoorthy, G. Solvent-exposed tryptophans probe the dynamics at protein surfaces. *Biophys. J.* **77,** 1100–1106 (1999).
12. Ebbinghaus, S. *et al.* An extended dynamical hydration shell around proteins. *Proc. Natl Acad. Sci. USA* **104,** 20749–20752 (2007).
13. Okada, T. *et al.* Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc. Natl Acad. Sci. USA* **99,** 5982–5987 (2002).
14. Brand, L. & Gohlke, J. R. Fluorescence probes for structure. *Annu. Rev. Biochem.* **41,** 843–868 (1972).
15. Sackett, D. L. & Wolff, J. Nile red as a polarity-sensitive fluorescent probe of hydrophobic protein surfaces. *Anal. Biochem.* **167,** 228–234 (1987).
16. Ross, J. B., Szabo, A. G. & Hogue, C. W. Enhancement of protein spectra with tryptophan analogs: fluorescence spectroscopy of protein-protein and protein-nucleic acid interactions. *Methods Enzymol.* **278,** 151–190 (1997).
17. Cornish, V. W. *et al.* Site-specific incorporation of biophysical probes into proteins. *Proc. Natl Acad. Sci. USA* **91,** 2910–2914 (1994).
18. Scott, D. J. *et al.* Quaternary re-arrangement analysed by spectral enhancement: the interaction of a sporulation repressor with its antagonist. *J. Mol. Biol.* **293,** 997–1004 (1999).
19. Blouse, G. E. *et al.* A concerted structural transition in the plasminogen activator inhibitor-1 mechanism of inhibition. *Biochemistry* **41,** 11997–12009 (2002).
20. De Filippis, V. *et al.* Incorporation of the fluorescent amino acid 7-azatryptophan into the core domain 1-47 of hirudin as a probe of hirudin folding and thrombin recognition. *Protein Sci.* **13,** 1489–1502 (2004).
21. Richmond, M. H. The effect of amino acid analogues on growth and protein synthesis in microorganisms. *Bacteriol. Rev.* **26,** 398–420 (1962).
22. Mente, S. & Maroncelli, M. Solvation and the excited-state tautomerization of 7-azaindole and 1-azacarbazole: Computer simulations in water and alcohol solvents. *J. Phys. Chem. A* **102,** 3860–3876 (1998).
23. Smirnov, A. V. *et al.* Photophysics and biological applications of 7-azaindole and its analogs. *J. Phys. Chem. B* **101,** 2758–2769 (1997).
24. Waluk, J. Hydrogen-bonding-induced phenomena in bifunctional heteroazaaromatics. *Acc. Chem. Res.* **36,** 832–838 (2003).
25. Chapman, C. F. & Maroncelli, M. Excited-state tautomerization of 7-azaindole in water. *J. Phys. Chem.* **96,** 8430–8441 (1992).
26. Rich, R. L., Smirnov, A. V., Schwabacher, A. W. & Petrich, J. W. Synthesis and photophysics of the optical probe N_1-methyl-7-azatryptophan. *J. Am. Chem. Soc.* **117,** 11850–11853 (1995).
27. Hoesl, M. G., Larregola, M., Cui, H. & Budisa, N. Azatryptophans as tools to study polarity requirements for folding of green fluorescent protein. *J. Pept. Sci.* **16,** 589–595 (2010).
28. Chen, Y., Gai, F. & Petrich, J. W. Single-exponential fluorescence decay of the nonnatural amino acid 7-azatryptophan and the nonexponential fluorescence decay of tryptophan in water. *J. Phys. Chem.* **98,** 2203–2209 (1994).
29. Chou, P. T., Yu, W. S., Wei, C. Y., Cheng, Y. M. & Yang, C. Y. Water-catalysed excited-state double proton transfer in 3-cyano-7-azaindole: the resolution of the proton-transfer mechanism for 7-azaindoles in pure water. *J. Am. Chem. Soc.* **123,** 3599–3600 (2001).
30. Chou, P. T. & Chi, Y. Phosphorescent dyes for organic light-emitting diodes. *Chem. Eur. J.* **13,** 380–395 (2007).
31. Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chem. Rev.* **105,** 2253–2277 (2005).
32. Zhao, B., Guengerich, F. P., Voehler, M. & Waterman, M. R. Role of active site water molecules and substrate hydroxyl groups in oxygen activation by cytochrome P450 158A2: a new mechanism of proton transfer. *J. Biol. Chem.* **280,** 42188–42197 (2005).
33. Yano, J. K. *et al.* The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-A resolution. *J. Biol. Chem.* **279,** 38091–38094 (2004).
34. Samuelsson, B. *et al.* Prostaglandins and thromboxanes. *Annu. Rev. Biochem.* **47,** 997–1029 (1978).
35. Cheng, Y. *et al.* Role of prostacyclin in the cardiovascular response to thromboxane A2. *Science* **296,** 539–541 (2002).
36. Williams, P. A. *et al.* Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **305,** 683–686 (2004).
37. Attwell, D. *et al.* Glial and neuronal control of brain blood flow. *Nature* **468,** 232–243 (2010).
38. Chao, W. C. *et al.* Probing ligand binding to thromboxane synthase. *Biochemistry* **52,** 1113–1121 (2013).
39. Chao, W. C. *et al.* Probing the interaction between prostacyclin synthase and prostaglandin H2 analogues or inhibitors via a combination of resonance Raman spectroscopy and molecular dynamics simulation approaches. *J. Am. Chem. Soc.* **133,** 18870–18879 (2011).
40. Chou, P. T. *et al.* Monohydrate catalysis of excited-state double-proton transfer in 7-azaindole. *J. Phys. Chem.* **96,** 5203–5205 (1992).
41. Li, Y. C. *et al.* Structures of prostacyclin synthase and its complexes with substrate analog and inhibitor reveal a ligand-specific heme conformation change. *J. Biol. Chem.* **283,** 2917–2926 (2008).
42. Chou, P. T. *et al.* Excited-state intramolecular proton transfer in 10-hydroxybenzo(h)quinoline. *J. Phys. Chem. A* **105,** 1731–1740 (2001).
43. Chou, P. T. *et al.* Femtosecond dynamics on excited-state proton/ charge-transfer reaction in 4′-N,N-diethylamino-3-hydroxyflavone. The role of dipolar vectors in constructing a rational mechanism. *J. Phys. Chem. A* **109,** 3777–3787 (2005).
44. Christiansen, O., Koch, H. & Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **243,** 409–418 (1995).
45. Dunning, J. T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90,** 1007–1023 (1989).
46. Eckert, F. & Klamt, A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* **48,** 369–385 (2002).
47. Ahlrichs, R., Bär, M., Häser, M., Horn, H. & Kölmel, C. Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* **162,** 165–169 (1989).
48. Hattig, C. & Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **113,** 5154–5161 (2000).
49. Hattig, C. Geometry optimizations with the coupled-cluster model CC2 using the resolution-of-the-identity approximation. *J. Chem. Phys.* **118,** 7751–7761 (2003).
50. Kohn, A. & Hattig, C. Analytic gradients for excited states in the coupled-cluster model CC2 employing the resolution-of-the-identity approximation. *J. Chem. Phys.* **119,** 5021–5036 (2003).
51. Frisch, M. J. *et al. Gaussian 09*. revision A.02 (Gaussian Inc., 2009).

Acknowledgements

This work was supported by Grants NSC 101-2811-M-002-134 and 101-2113-M-030-009-MY3 from the National Science Council of the Republic of China, a Cutting-Edge research grant from the National Taiwan University and a Medical Research Grant from the Medical School, Fu-Jen Catholic University. We thank Dr Ah-Lim Tsai and Lee-Ho Wang for their valuable comments. We also thank the National Center for High-Performance Computing for computational resources.

Author contributions

J.-Y.S. and S.C.-W.C. performed the chemical syntheses. C.L. and K.-C.T. measured and analysed the data. H.-A.P. performed the theoretical calculations. W.-C.C., L.-J.L., J.-S.W. and J.-F.L. performed TXAS mutants, experiments and analysis. C.-L.C. and Y.-K.L. performed the MD simulation. H.-C.Y. and P.-T.C. co-wrote the paper. All the authors discussed the results and commented on the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Shen, J.-Y. *et al.* Probing water micro-solvation in proteins by water catalysed proton-transfer tautomerism. *Nat. Commun.* 4:2611 doi: 10.1038/ncomms3611 (2013).