Mapping Proton Wires in Proteins: Carbonic Anhydrase and GFP Chromophore Biosynthesis[†]

Ai Shinobu and Noam Agmon*

Fritz Haber Research Center, Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

We have developed an algorithm for mapping proton wires in proteins and applied it to the X-ray structures of human carbonic anhydrase II (CA-II), the green fluorescent protein (GFP), and some of their mutants. For both proteins, we find more extensive proton wires than typically reported. In CA-II the active site wire exits to the protein surface, and leads to Glu69 and Asp72, located on an electronegative patch on the rim of the active site cavity. One possible interpretation of this observation is that positively charged, protonated buffer molecules dock in that area, from which a proton is delivered to the active site when the enzyme works in the dehydration direction. In GFP we find a new internal proton wire, in addition to the previously reported wire involved in excited state proton transfer. The new wire is located on the other face of the chromophore, and we conjecture that it plays a role in chromophore biosynthesis that occurs following protein folding. In the last step of this process, transient carbanion formation was suggested to occur on the bridge carbon [Pouwels et al. *Biochemistry* **2008**, *47*, 10111]. Residues on the new wire (Thr62, His181, Arg96) may participate in proton abstraction from this bridge carbon atom. A possible mechanism involves a rotation of the Thr62 side chain and completion of a short wire by which the proton is transported to His181, while the negative charge is transferred to the imidazolone carbonyl, producing a homoenolate intermediate that is stabilized by Arg96. Finally, comparison of the proton wires in the two proteins reveals common motifs, such as short internalized Ser/Thr-Glu hydrogen-bonded pairs for ultrafast proton abstraction, and threonine side chain rotation functioning as a proton wire switch.

I. Introduction

Proton mobility is central to chemical and biochemical reactivity,[1–5] determining the rate of acid–base catalysis in atmospheric chemistry, solution phase reactions, and enzyme kinetics. Rates of biochemical reactions are influenced by proton mobility in solution, on the surface, and within biomolecules.

A plethora of proton mobility mechanisms have been suggested over the years for liquid water.[6] For example, Hückel suggested[7] that the hydronium, H_3O^+, has one long $O-H^+$ bond which rotates around the C_3 symmetry axis through the oxygen atom, then dissociates to deliver the proton to a neighboring water molecule. Bernal and Fowler[8] proposed that water molecule rotation in the first solvation shell of the hydronium is rate limiting for proton mobility. These elegant suggestions appear not to be relevant to liquid water, where the three hydrogen bonds (HBs) in the first hydration shell of the hydronium are considerably stronger than those in the bulk,[9,10] and thus do not break on the time scale $(1-2$ ps$)$[11] of proton hopping between two water molecules. Rather, proton mobility in liquid water appears to be made up of rapid isomerization between nonsymmetrically hydrated H_3O^+ and $H_5O_2^+$ moieties,[6,12] driven by fluctuations in the HB connectivity.[6,13,14] Thus in room temperature liquid water there are no predetermined "pathways" along which a proton migrates, as these are created as the proton propagates.

This scenario contrasts with the interior of proteins, where well-defined HB networks are thought to exist.[15] Such networks, consisting predominantly of oxygen and nitrogen atoms of amino acid side chains (and perhaps also on the backbone), were therefore termed "proton wires". Atoms along such wires are the "stepping stones" utilized by protons for moving within a protein. In a fluctuating protein, portions of these wires may also break and form dynamically, either due to side change rotations, more global conformational changes, and water diffusion or exchange with the exterior solvent.

Because almost every biochemical reaction has some proton transfer (PT) steps, mapping of proteineous proton wires may help elucidate the mechanisms of PT within proteins. One might think that such maps are routinely generated as a guide for more detailed investigation, yet this is not the case. In a typical study usually only small portions of the overall proton wires are reported, those which seem relevant to the mechanistic question under investigation. To our knowledge, only one project dealt with the problem of systematic generation of proton wires:[16] Starting from the static X-ray conformation, Taraphder and Hummer have devised a Monte-Carlo search algorithm for locating the "optimal pathway" generated by the set of all possible side chain conformations. This computer code has been applied by Roy and Taraphder to human carbonic anhydrase II (CA-II).[17–19] Two generic conclusions may be inferred from their work: (i) In this protein most of the pertinent proton wires are seen already in the static X-ray structure,[20,21] and (ii) these wires are much more extensive than previously reported, suggesting the existence of alternate pathways for conducting protons in CA-II.

Our interest in proteineous proton wires arose from the observation[22] that in the X-ray structures[23–25] of the wild-type

---

SCHEME 1: The Parallelism between the Mechanism of the Second Proton Abstraction from the CA-II Active Site[41] and That of Proton Abstraction from the GFP Chromophore[24] (Arrows Depict Electronic Rearrangements)
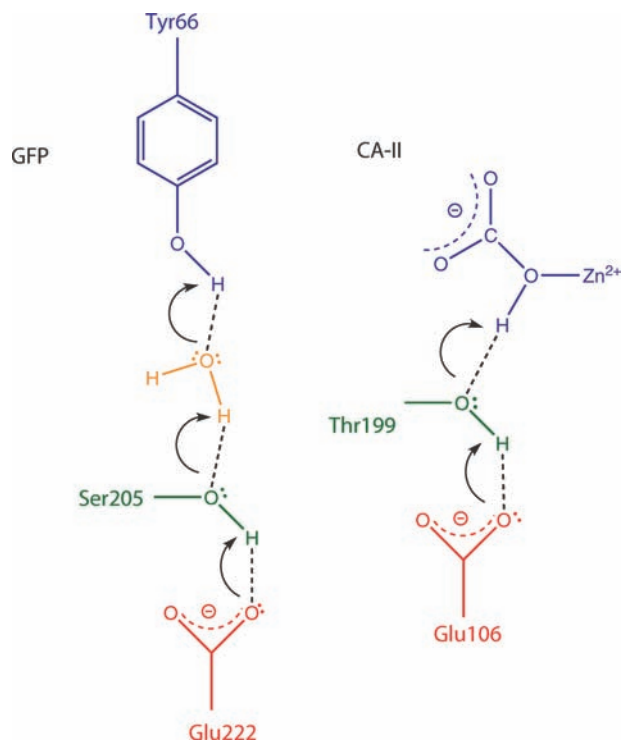
(wt) green fluorescent protein (GFP), the active site proton wire is considerably more extensive than the commonly discussed wire segment that leads from Tyr66 (part of the chromophore), via a water molecule (W) and a serine hydroxyl, to the carboxylate of the buried Glu222:

$$\text{Tyr66-OH}\cdots\text{W}\cdots\text{Ser205-OH}\cdots\text{Glu222-O}^-\cdots \quad (1)$$

see Scheme 1. The wire continues beyond Glu222, to Glu5 and other groups on the protein surface. The proton released from Tyr66-OH by photoexciting the chromophore moves very fast to Glu222,[26,27] but on a longer (ns) time scale may sample longer segments of this wire, with occasional return to the excited Tyr66-O⁻. Reversible recombination, which is coupled to one-dimensional diffusion along the wire, occasionally regenerates the acidic (Tyr66-OH) form, resulting in a $t^{-1/2}$ long-time tail in its time-resolved fluorescence signal.[28,29] The side chain of Thr203 may subsequently rotate, establishing a short pathway for proton exit into solution.[22,29]

GFP is a very rigid protein,[30] so it may not be surprising if most of its proton wires are observed already in its static X-ray structure. Nevertheless, some side chain rotations (such as that of Thr203) may contribute to extending the wire even here. In CA-II a much discussed proton wire segment leads from the active site to His64, which is thought to flip outward and release its proton to solution.[31] Another example of a "proton shuttle" is the rotation of the carboxylate side chain of Asp15 in ferredoxin I, which transfers a proton between solution and a sulfur atom in the enzyme's active site.[32] In contrast, the ascribed role of Thr203 in GFP is not of a shuttle but rather of a "switch", whose rotation can either connect or disconnect additional residues from the main pathway.

Side chain rotations have also been implicated in enhancing the diffusion of small ligands (other than protons) within proteins. For example, leucine side chain rotations (with activation enthalpies of 20 kJ/mol or less) likely facilitate the diffusion of small ligands such as CO, NO, and O_2 within myoglobin (Mb).[33] As the side chain rotates, it applies a small kick to the ligand, propelling it along its diffusive trajectory. Isoleucine side chains are thought to provide "gates" for ligand diffusion in proteins. For example, Ile107 in Mb gates the motion of small ligands from the distal pocket to the opposite side of the porphyrin ring.[34]

Given these examples, it appears useful to develop a simple and fast algorithm for comprehensive mapping of proton wires within proteins. The input would be an atomic coordinate list, such as a Protein Data Bank (PDB) file from X-ray studies. The output would be a tree structure containing all the (oxygen, nitrogen, and sulfur) atoms connected by HBs. From it one could identify all proton wire clusters. As a first step, we do not couple this to Monte-Carlo or molecular dynamics routines for generating additional conformations that may lead to extensions of the proton wire.[16] This limits the current application to rigid proteins, such as GFP or CA-II. To a certain degree, the lack of dynamics can be compensated for by utilizing several X-ray structures for a given protein (which may sample some of the relevant conformations), as well as high resolution ones (that may locate also the more mobile water molecules), and manual analysis of the rotation of select side chains. It is also useful to include a measure for the solvent accessibility of atoms along the wire, from which one could determine whether the wire is buried inside a protein, has exits to external solution, or resides on the protein surface.

Clearly, a newly identified proton wire need not necessarily have a functional role in conducting protons, but it could. Additional knowledge on the function of the protein may help to assess the situation. With time, one could develop intuition on the functional role of specific residues along the wires, such as glutamate shuttles vs threonine switches discussed above. Whenever a pathway is suspect of having a role in the protein activity, its preliminary study could motivate further experiments (e.g., mutations of residues along the pathway) or new computational projects for testing the potential of the pathway in actually transferring protons.

Here we report on two applications of our proton wire mapping routine. First, we test it on CA-II for which extensive data are available.[31] Here the wires located agree with those reported in previous studies. In particular, we corroborate and extend the results of Roy and Taraphder,[18] suggesting a role for their alternate proton wire in funneling protons into the active site. Then we apply the algorithm to GFP, where we find, in addition to the previously reported active site wire,[22] a new internalized wire that is disconnected from the active site. We consider the possibility that this wire has a role in the last step of chromophore biosynthesis, which occurs immediately after the nascent protein folds. We compare the two examples in order to highlight common motifs that may be characteristic of proton wires in proteins.

II. Methods

We have developed a general algorithm for identifying and mapping proton wires within proteins. The input consists of X-ray structures from the Protein Data Bank (PDB). The output consists of all HBed clusters within a protein, where a cluster is defined as a maximal set of atoms connected to each other by HBs. By definition, in a given cluster there is a pathway

connecting every atom to every other atom, whereas atoms outside the cluster are not connected to any atom in the cluster. Atoms typically participating in these clusters are the following: (a) All of the oxygen atoms within a protein. These include oxygens of water molecules, amino acid side chains (specifically of Glu, Asp, Ser, Thr, Tyr, Asn, and Gln), and also backbone carbonyls or oxygens on prosthetic groups (if exist). The motivation for including backbone carbonyls is the possibility for their resonance stabilization across the peptide bond:

$$-C_\alpha-C(=OH^+)-NH- \leftrightarrow -C_\alpha-C(-OH)=NH^+-$$
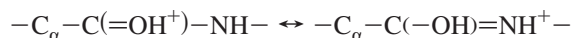
(b) Nitrogen atoms of amino acid side chains (specifically of Asn, Gln, His, Arg, and Lys), and also backbone amide nitrogens. Assuming that nitrogens are more basic than oxygens, they are considered here as "proton traps", which can only reside at chain ends. Thus the present routine does not continue a cluster beyond a nitrogen atom. (c) Sulfur atoms, such as within cysteine residues (not implemented here).

The first step in the algorithm is to determine whether any pair of the specified atoms is HBed. We define a HB using a conventional distance-angle criterion.[35] This geometric definition of HBs utilizes cutoff distances and angles as follows. The cutoff distance for oxygens was assumed to be 3.3 Å, unless one of the atoms is a water oxygen, in which case the cutoff was increased to 3.5 Å. The rationale for this is that water molecules may be slightly more mobile within the protein.

Because we utilize X-ray structures in which hydrogens are not resolved, the HB angle is deduced from the $C-O\cdots O$ angle (where an unresolved hydrogen is presumably in-between the two oxygen atoms). Assuming a linear hydrogen bond, $O-H\cdots O$, the optimal $C-O\cdots O$ angle depends on the carbon atom hybridization: for sp^2 hybridization (e.g., backbone carbonyls) it was taken as 120° whereas for sp^3 hybridization (tetrahedral carbon centers) it was taken as 109.5°. Allowed deviations from these "ideal" values were ±35°.

In comparison to the algorithm of Taraphder and Hummer,[16] our program does not perform a side chain conformational search, and focuses on static structures. On the other hand, we include more atom types, such as backbone carbonyls, and employ the angular criteria discussed above, rather than just a distance cutoff. In practice we find only a few cases where a would-be HB is rejected due to an unfavorable angle, some of these involving backbone carbonyls.

With these HB definitions, the program first scans all oxygen atoms in the structure, generating a "bonding matrix" that enlists the HB partners of each oxygen. This matrix consists of 1's (for bonded atoms) and 0's (for nonbonded atom pairs). The second step is utilizing the bonding matrix to search for chains of HBs, consisting of sequences of more than two HBs. Shorter chains are considered as "structural HBs", which stabilize the protein structure but do not necessarily participate in proton conduction, and are thus not included. Cluster construction is performed by a recursive tree algorithm, which will be described in detail elsewhere. Finally nitrogens are added (but only at chain ends).

The algorithm was written in the C++ programming language using Dev-C++ compiler from Bloodshed, version 4.9.8.0 (http://www.bloodshed.net/). For graphic visualization we used Matlab version 7.4.0.287 (The MathWorks, http://www.mathworks.com/). For each atom on a cluster, we also ran NAccess version 2.1 (http://www.bioinf.manchester.ac.uk/naccess/) to assess its solvent accessibility.[36] In this algorithm a water molecule (radius $R = 1.4$ Å) is rolled over the protein surface. Assuming the same radius for all oxygen atoms, a fully exposed oxygen has an accessible surface area (ASA) of $4\pi(2R)^2 = 99$ Å^2. The generated ASA values help determine whether the cluster found is on the surface or internal, and if so identify exit points from the protein. Typically, one finds more surface clusters than internalized ones.

To obtain the ASA of a proteineous atom, all water molecules are first stripped. Otherwise hydration water molecules may block surface accessibility when they are, in fact, part of the solvent. Likewise, to obtain the ASA of a water molecule, all other water molecules are eliminated. This gives a residual ASA for internalized water molecules (when close to a large cavity that hosted a water molecule), but prevents misleadingly low ASA readings for surface water molecules which are covered by a second layer of water. For example, the active site cavity in CA-II contains several layers of water. If unremoved, one could obtain a false impression that this water body is disconnected from the solvent.

Figure 1a shows the output of our program for the active site cluster of CA-II from an X-ray structure of 1.54 Å resolution.[20] This structure presents a good test for our methodology because it can be directly compared with literature results (see Figure 9 in ref 18). In Figure 1a, atoms are depicted by blue circles, with color intensity proportional to their ASA value. Table 1 gives the ASA values for water molecules in this cluster with and without water stripping, demonstrating why water stripping may be necessary to identify exit points. Lines in the figure indicate HBs, their thickness depicting HB length: A short and strong bond is seen as a thick line while a long and weak bond is shown as a thin line. More specifically, the thinnest line is for $r = 3.5$ Å, and the line thickness increases linearly with $3.5 - r$ ($r < 3.5$ Å). The atom type is indicated by distinct colored labels. Shown is one projection of the three-dimensional Matlab output. An alternative output format involves "truncated PDB files", in which we retain only the coordinates of water molecules and amino acid residues with atoms belonging to the cluster. We have deposited these files as Supporting Information.

The static approach utilized herein is supplemented by the investigation of the reorientation of select side chains, predominantly of threonines. To check whether a given side chain rotation is feasible, a molecular mechanics (MM) force field is utilized, with Allinger's MM2 parameters.[37] The rotational profile is then generated by the dihedral driver of Chem3D (version 9).[38] This program utilizes the MM2 force field to

TABLE 1: The Effect of Partial or Full Water Stripping on the calculated NAccess[36] Accessible Solvent Areas (ASA, measured in Å^2) for Water Oxygen Atoms in the Active Site Cluster of CA-II (PDB code 2CBA[20])

water no.[a]	ASA[b]	ASA[c]	ASA[d]
263	0	0	10.65
264	0	0	0.832
265	0	0	0
292	0	0	5.49
318	0	0	20.16
319	10.98	16.40	30.56
338	0	0	6.93
359	0	0	29.16
369	0	0	8.68
381	13.28	13.28	30.66
389	0.006	0.006	21.04
393	12.86	18.72	44.09
461	3.40	3.40	27.16

[a] Figure 9 in ref 18 omitted several water molecules for brevity. [b] No water stripping. [c] Noncluster water stripping. [d] Full water stripping.
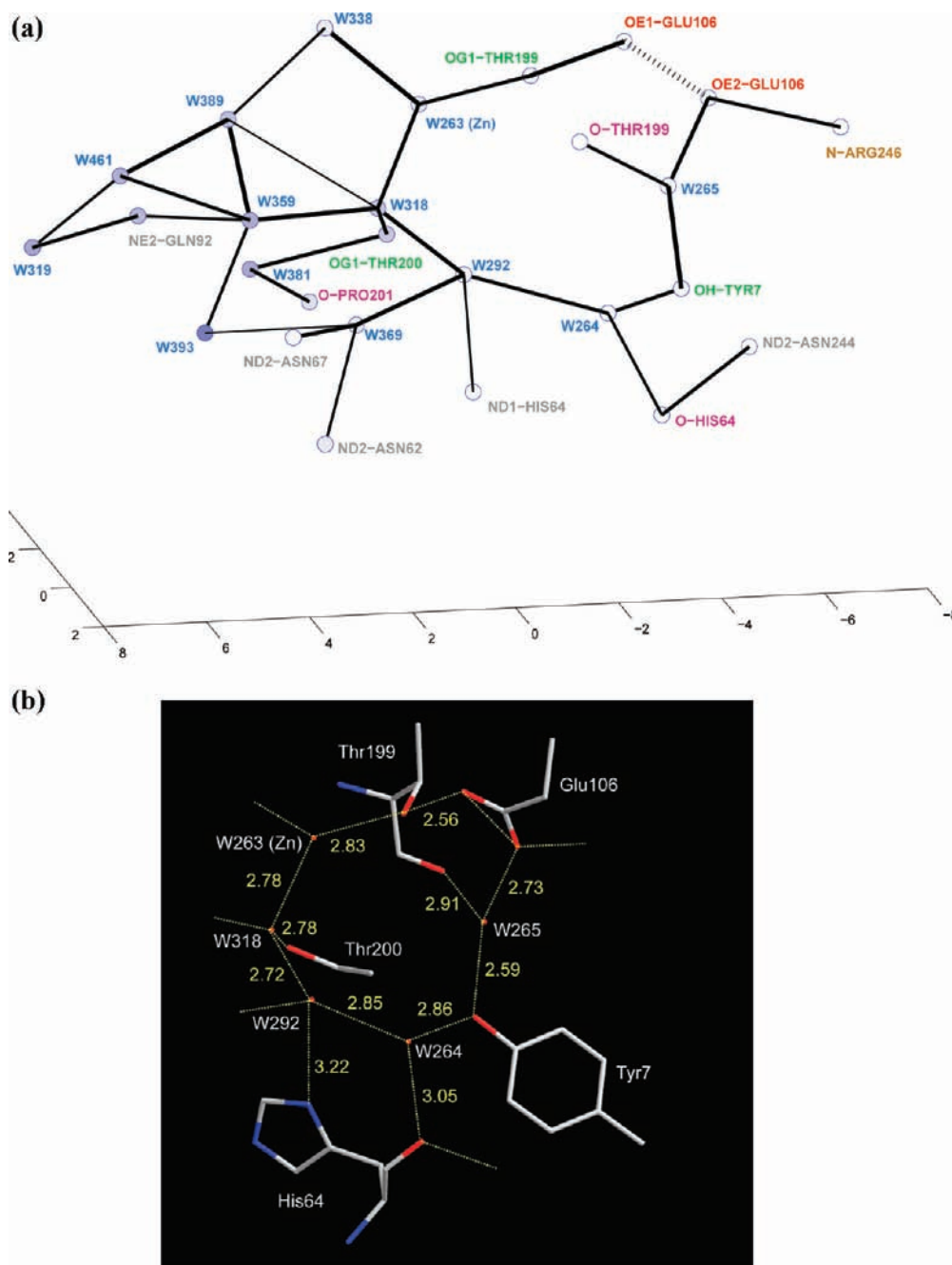
Figure 1. The active site proton wire of CA-II (PDB file 2CBA,[20] 1.54 Å resolution): (a) Output from our program as displayed by Matlab. HBed atoms (circles) are in blue, with intensity proportional to their surface accessibility. Lines denote HBs, with their width inversely proportional to the HB length. (b) A portion of the wire in stick representation with use of Chem3D.[38] Atom colors: gray, carbon; red, oxygen; and blue, nitrogen. Dashed yellow lines represent HBs of the indicated lengths (in Å).

calculate the interaction of the side chain with the rest of the protein (held in its X-ray conformation), as a function of the appropriate dihedral angle.
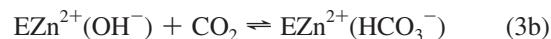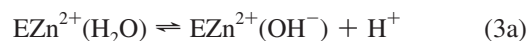
III. Human Carbonic Anhydrase II

Carbonic anhydrase, one of the fastest enzymes (turnover rate 10^6/s), catalyzes the reversible interconversion of carbon dioxide and bicarbonate:[31]

$$CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+ \qquad (2)$$
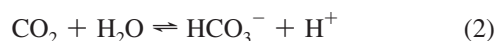
It is made up predominantly of a β-sheet structure, with a cylindrical hydrophilic/hydrophobic cavity in which the zinc

catalytic site is located. The tetrahedral Zn^{2+} cation is coordinated to three histidines (His94, His96, and His119) on the hydrophilic side of the cavity, whereas the fourth ligand is either water or a hydroxide ion. At this enzymatic zinc site (EZn^{2+}) two reactions are catalyzed,

$$EZn^{2+}(H_2O) \rightleftharpoons EZn^{2+}(OH^-) + H^+ \qquad (3a)$$

$$EZn^{2+}(OH^-) + CO_2 \rightleftharpoons EZn^{2+}(HCO_3^-) \qquad (3b)$$

In the forward (hydration) direction, the first step is ionization of water, with intramolecular and subsequently intermolecular

Mapping Proton Wires in Proteins

J. Phys. Chem. A, Vol. 113, No. 26, 2009 **7257**

transfer of the proton to a buffer group (B) that conducts it into solution. Thus, coupled to the first step is a reaction of buffer protonation

$$B + H^+ \rightleftharpoons BH \qquad (4)$$

The second step is nucleophilic attack on the CO_2 carbon by the activated hydroxide. Its pK_a is around 7, and it forms a strong HB to Thr199 thus orienting its loan pair in the direction of the region of the cavity into which the CO_2 enters.[39] The final step is an exchange of bicarbonate for water at the active site.

A. Intramolecular Proton Transfer. It is widely believed that following the nucleophilic attack in eq 3b, and preceding the bicarbonate product release, intramolecular PT moves the second proton from the zinc-bound water to one of the other two oxygen atoms in the nascent HCO_3^-:

$$EZn^{2+}(OH)COO^- \rightarrow EZn^{2+}OC(OH)O^-$$
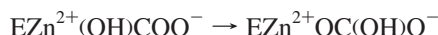
This was suggested to occur either by rotation of the enzyme-bound HCO_3^- (Lindskog mechanism) or direct PT between two of its oxygen atoms (Lipscomb mechanism). If insight can be gained from the discussion of proton mobility mechanisms in water,[6] it may be that none of these two mechanisms is strictly operative. The Lindskog mechanism is analogous to the rotational Hückel or Bernal−Fowler mechanisms, which require cleavage of one or more strong HBs in the first solvation shell of an ion. The Lipscomb mechanism is analogous to proton tunneling between well-separated water molecules, whereas the preferred route is to first bring them close together to form a protonated dimer, $H_5O_2^+$. Thus both these routes may be energetically costly.

Consulting the active site proton wire in Figure 1, it appears that the relevant HB chain to consider is

$$Zn^{2+}\text{-}OH^- \cdots Thr199\text{-}OH \cdots Glu106\text{-}O^- \cdots \qquad (5)$$

cf. Figure 1 in ref 39. This chain is analogous to the fragment of the active pathway in GFP that leads from the Tyr66-OH via a serine to an internalized glutamate (Scheme 1, left side). That connection is known to transfer a proton on the ultrafast time scale of several picoseconds,[26] or even several hundred femtoseconds (in which case it is a synchronous triple-PT).[40] Thus we identify here a motif involving either a serine or threonine hydroxyl, HBed to an internalized glutamate as a potential ultrafast, low-barrier pathway for proton abstraction. This suggests that instead of the direct PT between carbonate oxygens envisioned by Lipscomb, proton wires are utilized to reduce the activation energy. The proton is first transferred to Glu106 (Scheme 1) and subsequently either it or another proton is transferred by another wire to one of the two other oxygen atoms.

This mechanism has recently been demonstrated in a detailed DFT calculation by Bottoni et al.[41] As in GFP, the PT to Glu106 was seen to occur in a concerted fashion, involving double PT from the active site to Thr199 and from Thr199 to Glu106. Interestingly, this work identified a low-energy pathway in which the Thr199 side chain rotates subsequent to PT to Glu106, delivering the proton to one of the two other oxygen atoms.[41] This reinforces the concept of threonine side chains acting as

switches (cf. ref 22): Here Thr199 operates to connect alternate carbonate oxygens to Glu106 and thus mediate PT between them.

B. Intermolecular Proton Transfer. The proton released in the water cleavage step, eq 3a, is believed to be transferred to the $N_{\delta 1}$ imidazole nitrogen of His64, which then flips to an "out" conformation from which it delivers the proton to solution (or to a solvated buffer molecule).[31] Indeed, replacement of this histidine by alanine (a H64A mutation) slows the enzymatic reaction by more than an order of magnitude. A problem with such a mechanism is that the distance from the nearest water oxygen to the $N_{\delta 1}$ position of His64 in the "in" conformation is nearly 3.3 Å,[21] larger by 0.9 Å than the optimal distance for efficient PT. Now, when a soluble proton donor/acceptor (4-methylimidazole, 4-MI) is added to the solution, it enhances the activity of the H64A mutant to near wt level. One might expect this "chemical rescue" molecule to bind near the location of His64. Indeed, its π-stacking to the adjacent Trp5 was observed in the X-ray structure of this mutant,[42] but also docking near Asp72. However, the first docking site seems nonoperational, as judged by the lack of effect of Trp5 mutations on the enzymatic kinetics,[43] so it must be the further docking site near Asp72 (or additional docking sites not seen in this X-ray structure) that contributes the most to the rescue effect.

Another question on the exclusiveness of the His64-shuttle is posed by the observation that the CA-II active site is actually open to solution.[18,19] To identify also the more mobile water molecules, we utilize here the highest resolution (1.05 Å) X-ray structure for wt CA-II from Fisher et al.[21] Figure 2a shows that the active site HB cluster now extends much further than in the lower resolution structure of Figure 1a, likely because some of the more mobile water molecules on the surface could now be detected. As Figure 2b shows, this cluster includes two pathways ascending from the Zn center toward Glu69 and further on to Asp72. One pathway passes via the $N_{\delta 1}$ atoms of Asn62 and Asn67, whereas the other, which meanders past Gln92, consists of only water molecules. In comparison, the H64A mutant structure (complexed with 4-MI)[42] reveals only the pathway through Asn62 and Asn67 (Figure 3). This pathway is apparently responsible for the "chemical rescue" mechanism of 4-MI, when it is found to bind between Asp72 and Glu69 (see Figure 10 in ref 42).

The surface electrostatic potential density of CA-II can be seen in Figure 2B (top) of ref 44. Asp72 and Glu69 are located within a negative region on the rim of the "caldera" surrounding the active site. In contrast, the His64 surface region is about neutral. This surface potential may be relevant in "steering" the proton and the bicarbonate to/from the active site, and in offering docking sites for charged/neutral buffer molecules which carry the proton to/from the enzyme surface.

One clear conclusion from this analysis is that, in agreement with Roy and Taraphder,[18,19] the His64 pathway is not the only one connecting the active site with the outside world. Which pathway does the proton utilize? The current consensus concerning the His64 pathway is based, experimentally, on the effect of the H64A mutation on slowing down the PT rate. Corroborating this requires ruling out the alternative pathways which thus far has not been done.

The opposite view might be that the proton is transported only along the surface pathways in Figure 2b. In this scenario the diminished reactivity of the H64A mutant may be ascribed to the elimination of one surface pathway, forcing the proton to migrate via the $N_{\delta 1}$ atom of Asn62 or Asn67 (Figure 3b). Assuming that PT through nitrogen is slower than through
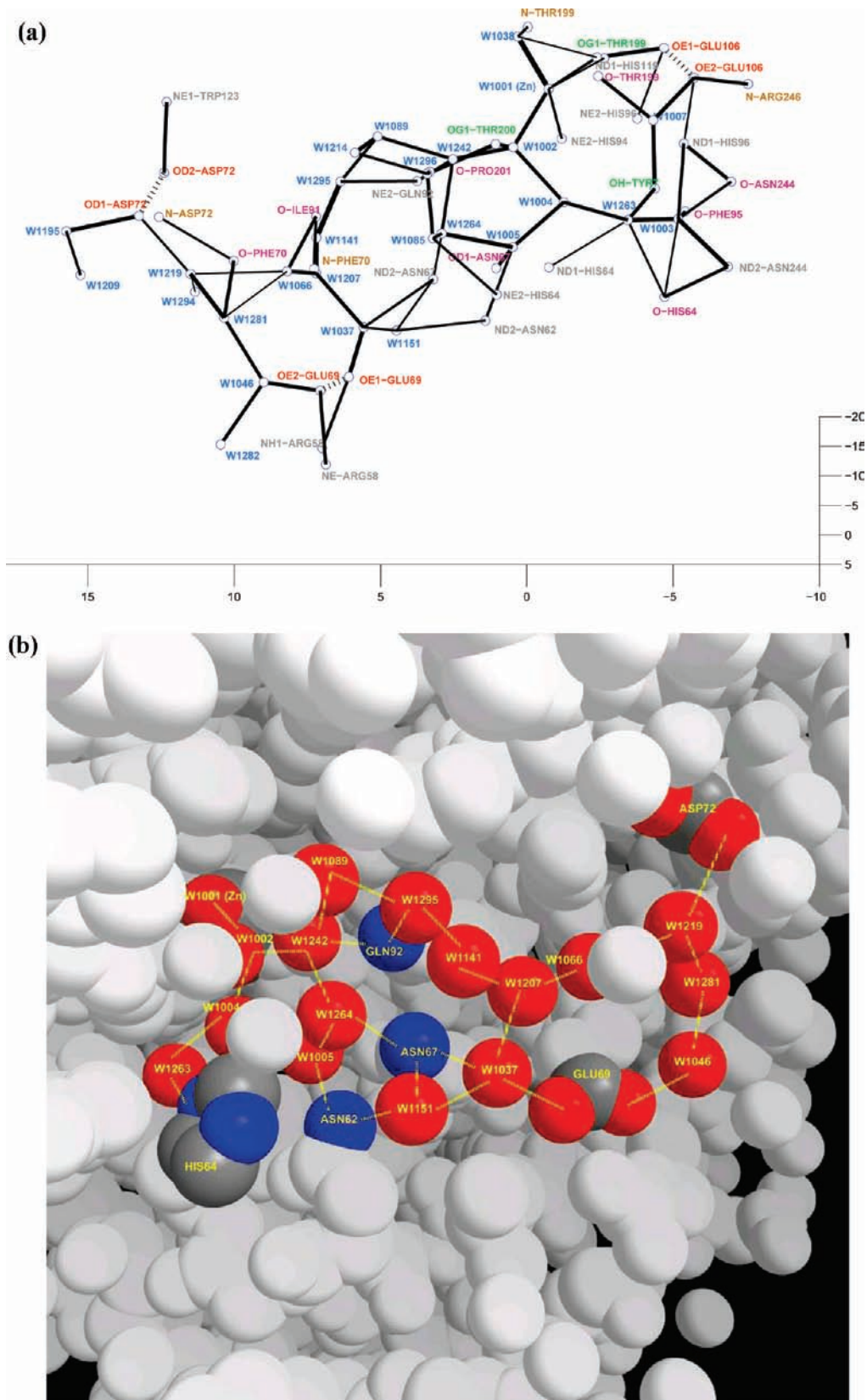
Figure 2. The active site proton wire for high-resolution CA-II structure (PDB file 2ILI,[21] 1.54 Å resolution). (a) Matlab display of the output from our program. (b) Depiction of the portion of the wire that lies on the protein surface with use of Chem3D version 9.[38]

oxygen, this could delay the proton transport process. Such an interpretation may be questioned, because when the protein is immersed in water it is likely that additional layers of water will open up more pathways, so that these specific surface pathways may be less dominant. In other words, surface pathways observed in the X-ray structure might rearrange in solution. On the other hand, they may be significant if water HBs to the protein surface are stronger than bulk HBs, and if
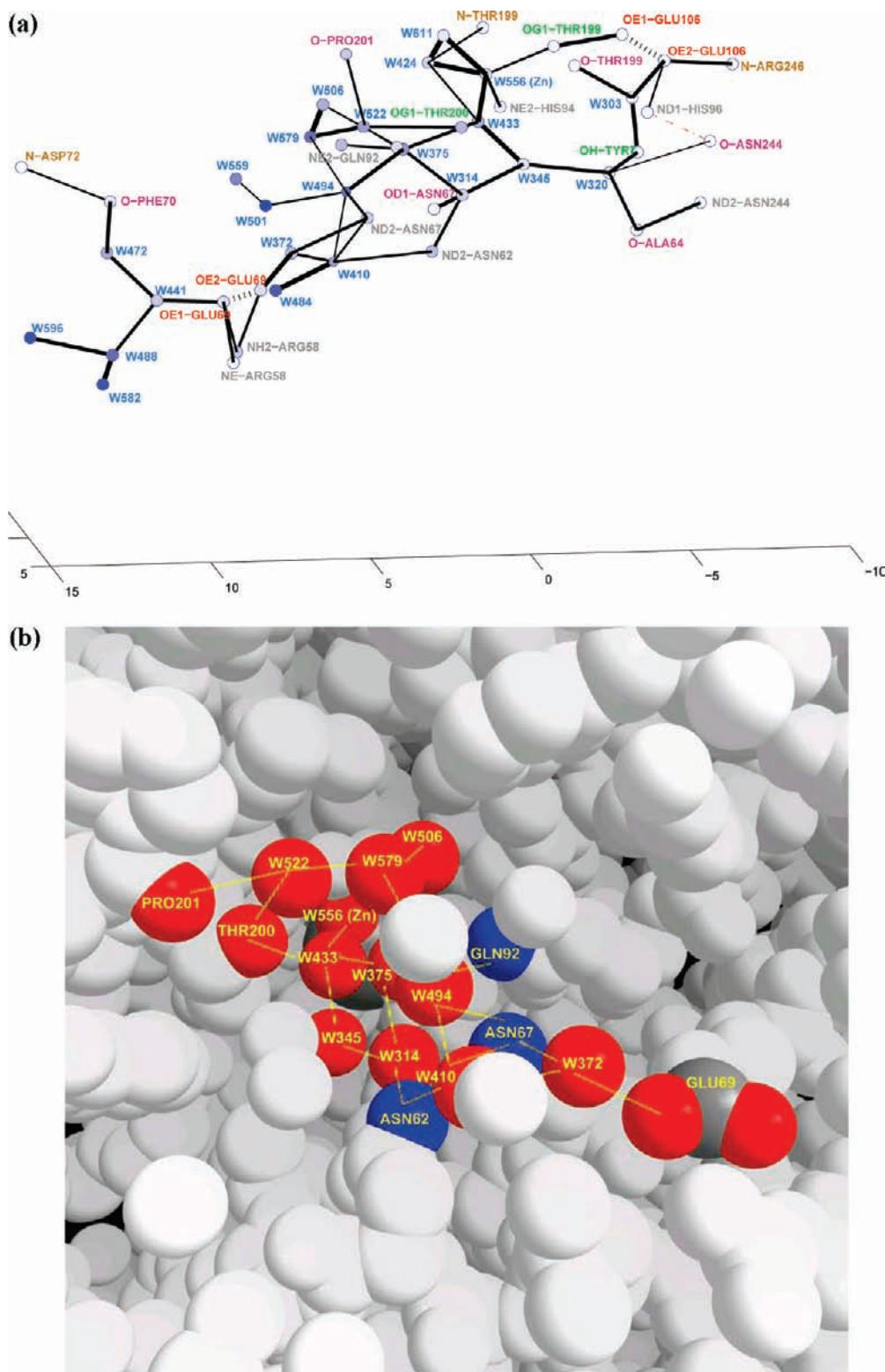
Mapping Proton Wires in Proteins

J. Phys. Chem. A, Vol. 113, No. 26, 2009 **7259**

Figure 3. The active site proton wire for a high-resolution structure (PDB file 1MOO,[42] 1.05 Å resolution) of the H64A mutant of CA-II, complexed with two molecules of 4-MI. (a) Matlab display of the output from our program. (b) Depiction of the portion of the wire that lies on the protein surface with use of Chem3D version 9.[38] Comparison with Figure 2 is meaningful because both structures have the same X-ray resolution.

the proton prefers to migrate on the surface, rather than in the bulk. This may be the case, given the demonstrated interfacial affinity of the hydronium.[45,46]

A third scenario is that both His64 and surface pathways are operative, and their relative importance depends on conditions. For example, one pathway may be important for the hydration

and the other for the dehydration direction. The reason for having the two kinds of pathways may involve the docking of buffer molecules from solution. A neutral buffer molecule may dock near His64 and collect an outgoing proton when the enzyme works in the hydration direction. A positively charged, protonated buffer molecule will preferentially dock on the

SCHEME 2: The Product of the Cyclization Step of GFP Chromophore Synthesis, with Key Atoms Marked and Numbered
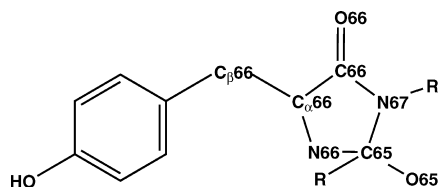
negative surface near Asp72 and deliver a proton (via the surface pathways of Figure 2b) to the enzyme working in the dehydration direction.

IV. Chromophore Biosynthesis in GFP

A. Background. "The remarkable brightly glowing green fluorescent protein, GFP, was first observed in the beautiful jellyfish, *Aequorea victoria* in 1962. Since then, this protein has become one of the most important tools used in contemporary bioscience. With the aid of GFP, researchers have developed ways to watch processes that were previously invisible, such as the development of nerve cells in the brain or how cancer cells spread."[47] The GFP rigid eleven-stranded β-barrel structure is traversed by a distorted α-helical segment: Three of its amino acids, Ser65, Tyr66, and Gly67, undergo a post-translational condensation reaction to form a five-membered imidazolone ring that couples to the tyrosine phenol ring.[48] This constitutes the aromatic chromophore, whose anion fluoresces in the green following an excited state PT reaction.[26,28]

Cyclization occurs only after protein folding, within a tight α-helical turn that places the Gly67 amide in close proximity to the Ser65 carbonyl, resulting in nucleophilic attack and ring formation (Scheme 2). Tsien and co-workers suggested that chromophore biosynthesis in GFP consists of three steps: cyclization, dehydration, and oxidation (see Scheme 3).[49] Kinetic experiments indeed demonstrated that cyclization precedes oxidation.[50] Subsequently, a S65G/Y66G mutant was found to cyclize but not dehydrate, suggesting that cyclization precedes dehydration.[51] Thus both oxidation and dehydration are later steps in GFP maturation.

In contrast to the agreement concerning the cyclization step, the order and detail of the subsequent steps in the mechanism are constantly under revision.[52] Most notably, Wachter and co-workers have found that a Y66L mutant undergoes cyclization and oxidation, but not dehydration, concluding that oxidation is actually the second step in chromophore biogenesis (Scheme 4).[53] This oxidation reaction is centered on the 5-membered ring rather than on the Tyr66 $C_\alpha - C_\beta$ bond, as previously assumed. Subsequently, participation of a pre-oxidation enolate intermediate (not shown here) was demonstrated.[54] It is stabilized by the positive charge on Arg96, which is HBed to the enolate oxygen atom (formerly the Tyr66 carbonyl).[55] In the mechanism of Scheme 4, this enolate formation is coupled to the elimination of a proton from $C_\alpha 66$ (see Scheme 2 for notation).

Finally it was shown[56] that dideuteration at the Tyr66 C_β position does not affect the oxidation rate, but slows the subsequent dehydration step by nearly a factor of 6. This is consistent with a primary isotope effect on the tetrahedral bridge carbon. Its hydrogen atom is therefore not abstracted by molecular oxygen, but rather during a subsequent PT reaction that produces a carbanion in the $C_\beta 66$ position.[56] This carbanion is stabilized by charge delocalization on the two flanking aromatic rings, which explains why chromophore maturation is observed only for aromatic substitutions in position 66. Nevertheless, carbanion formation still requires a rather basic residue ($pK_a = 9.4$) within the protein for abstracting this proton.[56]

The problem is that no such basic moiety is observed within HBing proximity to $C_\beta 66$. Wachter and co-workers suggested two solutions to this enigma (Scheme 5): (a) The proton is translocated, through a chain of two water molecules, to O65 (see Scheme 2 for notations).[53] This then releases a water molecule, completing the dehydration step. Apparently, water is not sufficiently basic to supply the driving force for this PT step, so that this option was abandoned. (b) The $C_\beta 66$ proton is abstracted by the guanidinium group of Arg96, while O65 is subsequently protonated by Glu222 to release a water molecule.[56]

There are evident problems with Arg96 as the deprotonating base, most of which are discussed in the original publication:[56]

(i) Arginines are very basic, their solution $pK_a = 12.5$, so that Arg96 is quite certainly protonated within GFP. Indeed,

SCHEME 3: A Three-Step GFP Chromophore Biosynthesis Involving Cyclization, Dehydration, and Oxidation[49]

Mapping Proton Wires in Proteins

J. Phys. Chem. A, Vol. 113, No. 26, 2009 **7261**
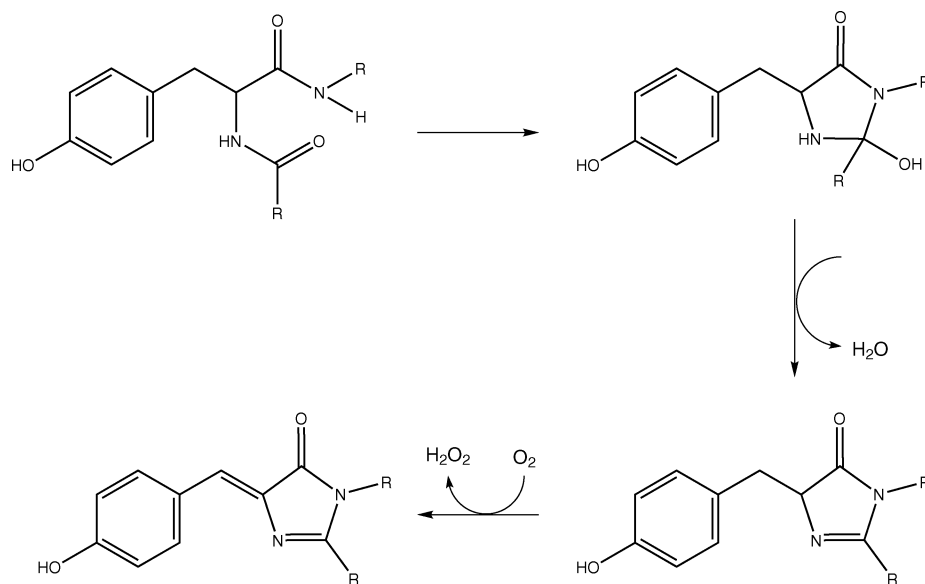
SCHEME 4: A Three-Step GFP Chromophore Biosynthesis Involving Cyclization, Oxidation, and Dehydration[53]
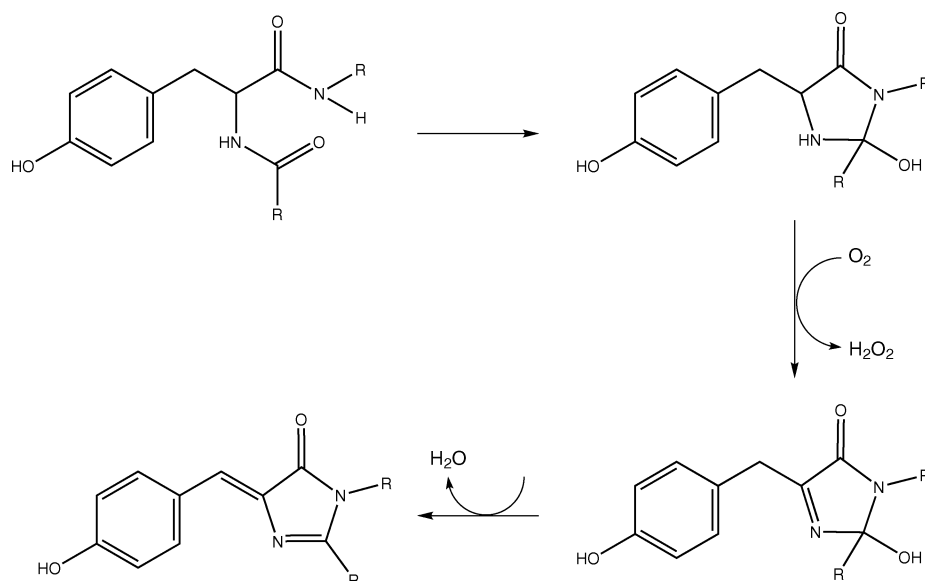
SCHEME 5: Suggested Mechanisms for Carbanion Formation by Proton Extraction from C$_\beta$66: (a) Direct Transfer through a Water−Molecule Chain According to Figure 7b in Reference 53 (arrows depict electronic motion) and (b) PT to a Neutral Arg96 through a Water Molecule (Red) That Does Not Exist in the X-ray Structure, According to Scheme 3 of Reference 56
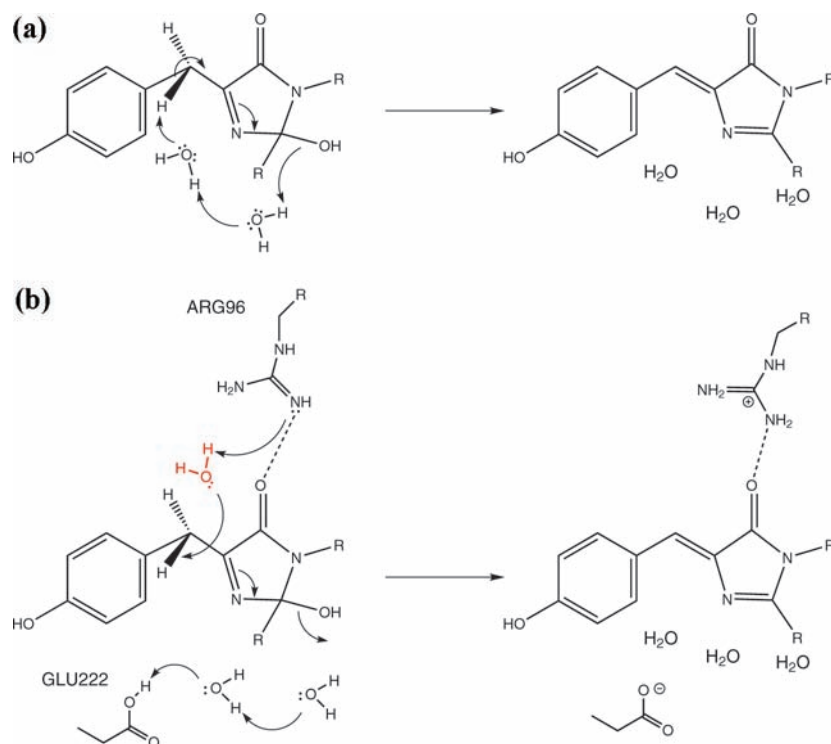
its positive charge is postulated to catalyze enolate formation as well as other steps in the chromophore maturation mechanism.[55]

(ii) Subsequently, one has to assume that "the pK_a of Arg96 may drop below 8 while the biosynthesis reaction is in progress",[56] but it is unclear how such a large pK_a change comes about.

(iii) The closest guanidinium nitrogen to C$_\beta$66 is at a distance of 3.9 Å, too far for direct PT (typical cutoff values for PT in water are around 3.3 Å). Consequently, Wachter and co-workers assume the participation of a mediating water molecule (Scheme

5b, in red), although such a water molecule is not seen in any of the GFP X-ray structures.

B. A New Internal Proton Wire in GFP. We have used our computer code to map out proton wires within GFP. The program located the extensive proton wire that connects Tyr66-OH via Glu222 to the surface of the β-barrel, which was previously found manually.[22] This cluster (not shown here) is thought to participate in the excited state PT reaction.[28,29] In addition, it found the new internal cluster depicted in Figure 4. It is located on the other face of the chromophore, apparently disconnected from the first cluster and from the protein exterior.
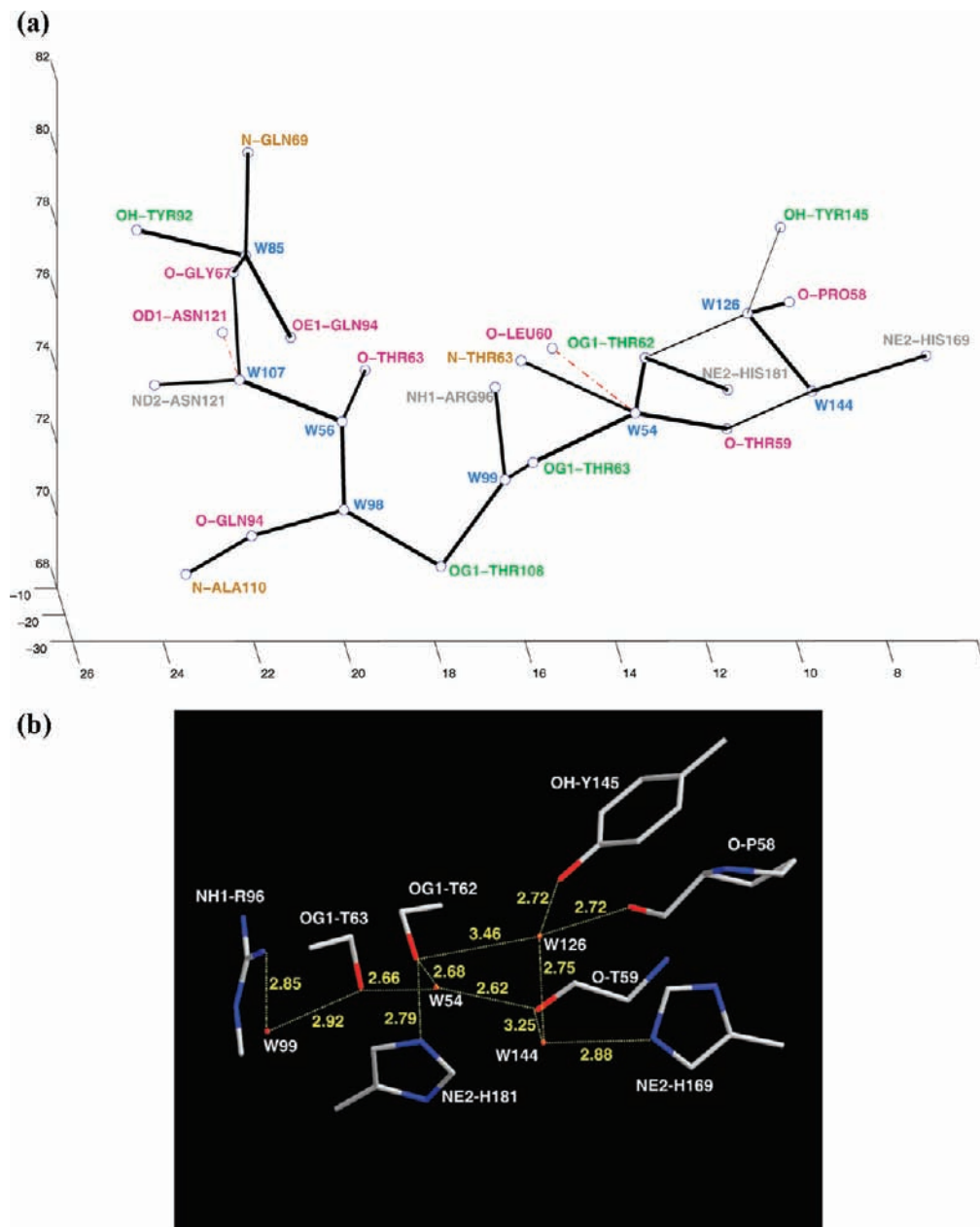
Figure 4. The new internal proton wire in GFP (PDB file 1W7S,[25] 1.85 Å resolution): (a) Output from our program as displayed by Matlab. HBed atoms (circles) are colored blue, with intensity proportional to their surface accessibility. (b) A portion of the wire in stick representation with use of Chem3D.[38] See Figure 1 for notations.

Thus this new wire does not seem to be involved in excited state PT. It connects some key residues near the chromophore, such as His181-$N_{\epsilon 2}$, Thr62-OH, Thr63-OH, Thr108-OH, Tyr145-OH, and Arg96-$N_{\eta 1}$, some of which are thought to be involved in chromophore biosynthesis.[51,56] Hence the new proton wire may be relevant for some of the biosynthesis steps. For example, if indeed Arg96 gets temporarily deprotonated during the dehydration step,[56] it may occur via this wire.

Checking more closely for contact points between the new wire and the chromophore, we found the three rotamers of Thr62 shown in Figure 5. In the X-ray structure (0° rotation) its side chain forms a HB with His181. When it is rotated by about 150°, its hydroxyl oxygen reaches a distance of 3.0 Å from the bridge carbon, C_{β}66. At about 260°, a HB is formed instead with the OH moiety of Tyr145.

The energetics of these rotamers are depicted in Figure 6, which shows a rotational profile calculated from two wt-GFP

and two S65T mutant GFP crystal structures, using the MM2 force field[37] as implemented in the dihedral driver of Chem3D (version 9).[38] These profiles depict the potential as a function of the Thr62 dihedral angle (rotation around its $C_{\alpha}-C_{\beta}$ bond), resulting from the interaction of the side chain with the rest of the protein, which is maintained in its X-ray conformation. Hence these potentials are expected to converge with improvement in X-ray resolution, as indeed appears to be the case. The lowest barrier separating the native (0°) and active (150°) rotamers is in the range 13−18 kJ/mol. These values are consistently smaller than those for the other two threonines on this cluster (Thr63 and Thr108), as well as for Thr62 in precyclized mutants[51] (30−40 kJ/mol, see Figure 7). Thus Thr62 in the cyclized structure may be "engineered" for facile rotation. Compared with rotational time scales for other small side chains with barriers on the order of 20 kJ/mol or less (see the
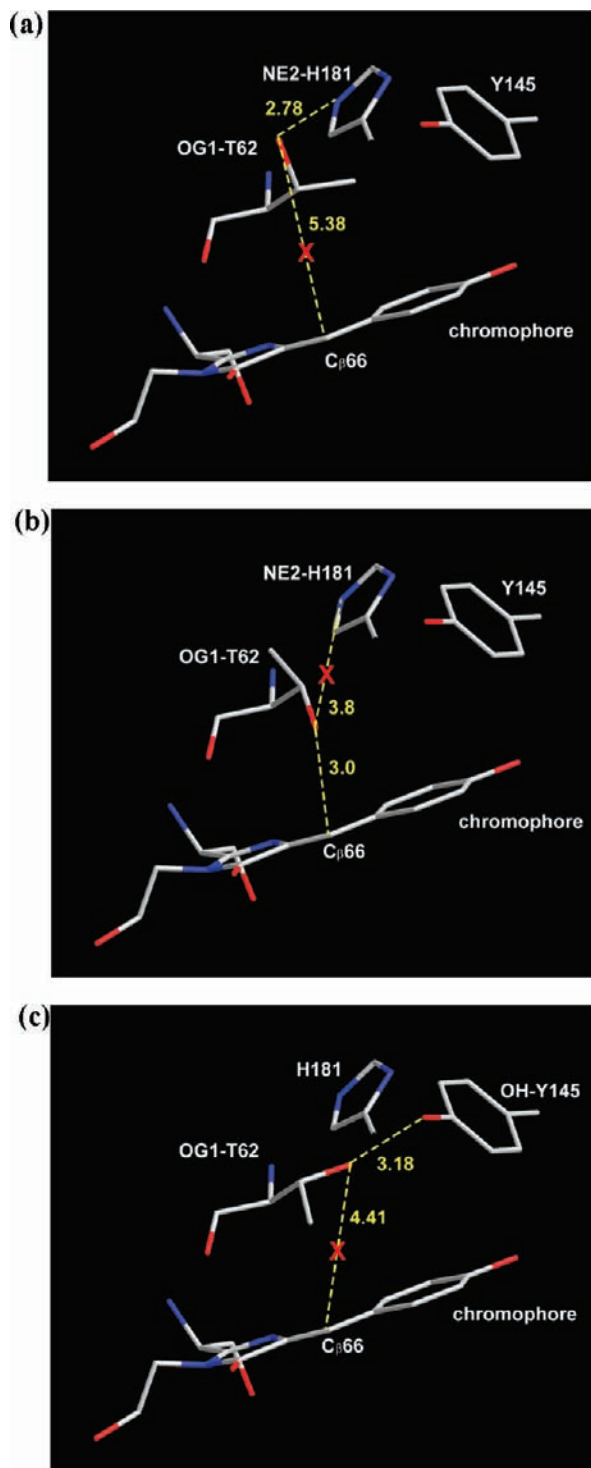
Figure 5. The three conformers of the Thr62 side chain in wt-GFP, PDB file 1W7S:[25] (a) The native conformer, in which the Thr62 side chain is HBed to His181; (b) the active conformer (145° rotation) forms a HB to $C_\beta 66$; (c) a third conformer (255° rotation) forms a HB to Tyr145. Atoms rendered as in Figure 4b.
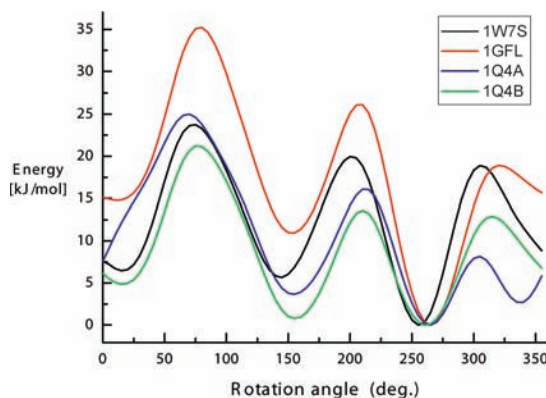

Figure 6. Dihedral chart for the side chain rotation of Thr62 in the folded structures of wt-GFP and some of its mutants. Conformational energy as a function of Thr62 side chain dihedral angle was calculated by using the MM2 dihedral driver of Chem3D.[38] The native rotamer is at 0°, whereas contact with $C_\beta 66$ is formed around 150°. Around 260° a HB is formed with the hydroxyl of Tyr145. Coordinates were taken from the X-ray structures of resolution better than 2 Å in order to reduce the energetic spread. The following PDB files were utilized: 1GFL (1.90 Å resolution)[23] and 1W7S (1.85 Å resolution),[25] both wild-type structures; 1Q4A (1.45 Å resolution) and 1Q4B (1.48 Å resolution), both S65T mutants at pH 8.5 and 5.5, respectively.[64]
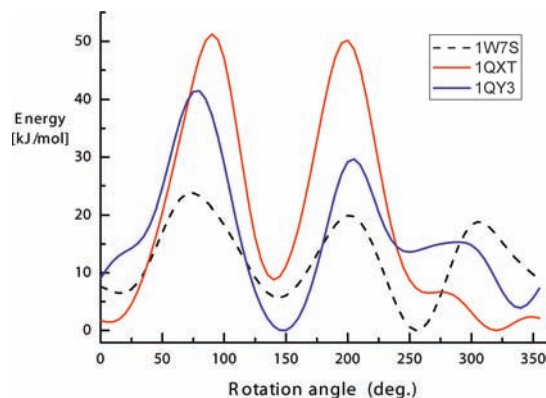

Figure 7. Dihedral chart for the side chain rotation of Thr62 in precyclized GFP, X-ray structures from PDB files 1QXT and 1QY3 (mutations R96A, F99S, M153T, V163A, F64L, S65T).[51] The minimum around 140−150° does not involve a HB to $C_\beta 66$. This minimum for structure 1QY3 is lower probably due to new HBs to water molecules which do not exist in structure 1QXT. See the legend of Figure 6 for other details.

Introduction), one may estimate that this rotation occurs on the submicrosecond time scale,[33] hence it is not rate limiting.

To abstract the proton from the bridge CH_2 moiety, the rotated Thr62 side chain must be able to (a) form a rather strong HB with it and (b) participate in a short HB network leading to a nearby basic moiety. Consider first the possibility of forming a C−H···O HB. In model calculations,[57] HBs between water and methane were shown to behave very much like ordinary HBs, only that they are weaker. When the methane was substituted with fluorines (electron attracting substituents) the HB strength increased to about half its magnitude in water.[57] For $C_\beta 66$, the role of the electron attracting groups is played by the two flanking aromatic rings. Moreover, it was shown that when water was replaced by methanol, the HB strength to the substituted methane further increased.[57] Thus the $C_\beta 66$−H···OH−Thr62 HB may be comparable in strength to the HB between water molecules.

Searching next for an appropriate base that may abstract this proton, we note that the $N_{\epsilon 2}$ atom of His181 is only 3.8 Å away from the hydroxyl oxygen of the rotated Thr62 side chain. Histidines in solution typically have $pK_a = 6$ but, contrary to popular opinion, they tend to become more basic when buried inside a protein.[58] However, the variability in their pK_a values also increases with burial, spanning the range 4−10. Thus it is possible that His181 is sufficiently basic to abstract the proton from $C_\beta 66$. As discussed in Section III, a histidine (His64) has been implicated as the proton acceptor in water ionization in the CA-II active site (eq 3a). More specifically, there are several

SCHEME 6: Proton Abstraction from C$_\beta$66 to His181 and the New Proton Wire via the Thr62 Switch, with Carbanion Stabilization as a Homoenolate[63] by Arg96
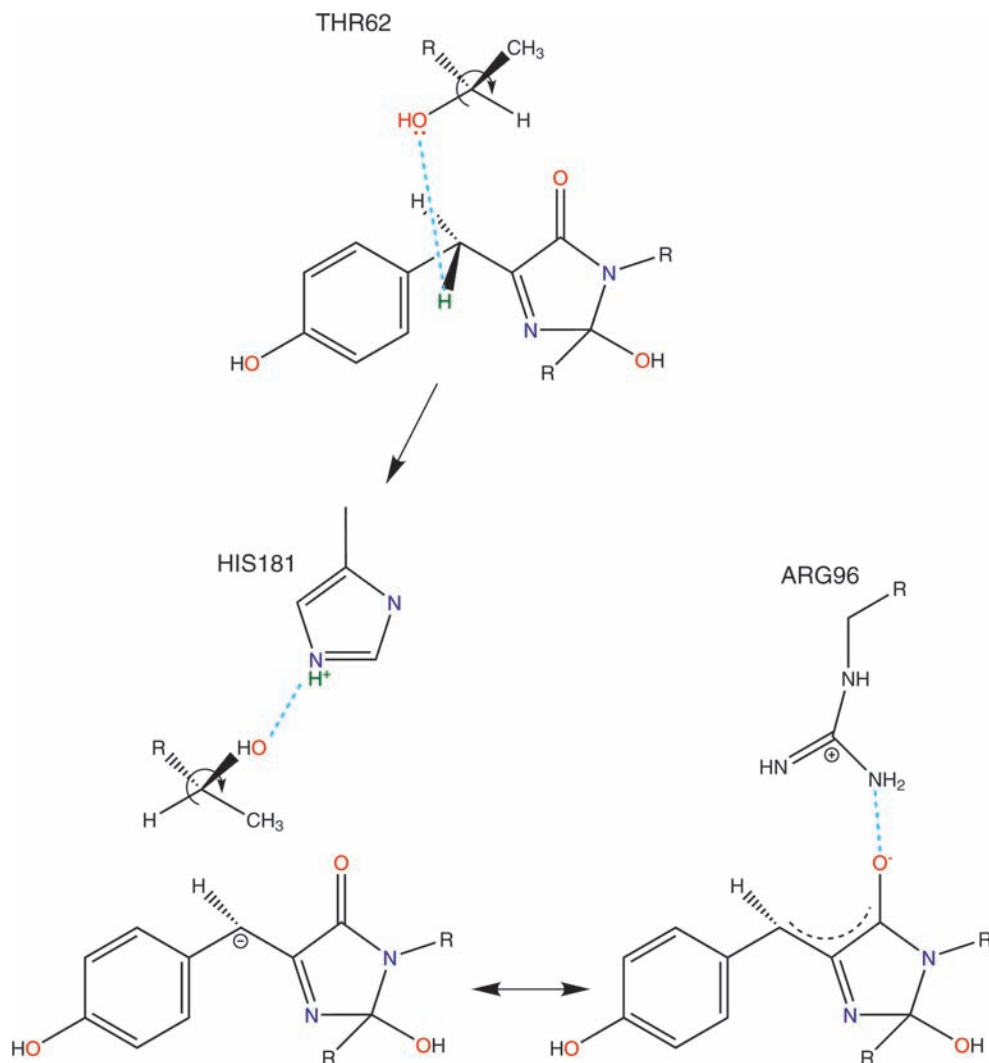
enzymes in which a histidine serves as a base in α-proton abstraction, leading to formation of a carbanion intermediate.[59–62] In these examples, the carbanion is often stabilized by resonance with an adjacent carbonyl, forming an enolate which can be stabilized by the positive charge of an arginine.[60] Such proton abstractions are analogous to the elimination of the proton from C$_\alpha$66 in a preceding step of the GFP chromophore synthesis. The subsequent carbanion formation on C$_\beta$66 can still gain stabilization by delocalizing the negative charge onto the carbonyl, forming a so-called homoenolate intermediate.[63]

With the above observations in mind, we suggest the tentative mechanism in Scheme 6. Thr62 rotation enables proton abstraction to His181, a step that becomes essentially irreversible as the threonine rotates back to its original position, where it forms a HB to His181 and stabilizes the positive charge there. As the original proton wire is reformed, the proton on His181 may delocalize on it, further reducing the probability for back-PT. In the chromophore, the negative charge on the bridge carbon is partly transferred to the carbonyl, through the 3-center bond shown in Scheme 6. The ensuing homoenolate, in turn, is stabilized by a salt bridge to Arg96. Thus Arg96 catalyzes this reaction step without ever becoming deprotonated.

It remains to identify the transient proton wire that may connect C$_\beta$66-H···OH−Thr62 with His181. We note that if the side chains of Thr62 and His181 are tilted, a direct contact can

be formed, C$_\beta$66-H···Thr62-OH···His181-N$_{\epsilon 2}$ in which the two HB lengths are around 3.3 Å. These distances are a bit long for an efficient PT to take place, which may nevertheless proceed by tunneling on a sufficiently long time scale. Alternately, we note that a nearby water molecule may be "squeezed" between Thr62, His181, and Phe165 (see Figure 8). This fit requires a rotation of 30° of the Ile167 side chain, at a cost of about 16 kJ/mol (determined by using the dihedral driver of Chem3D). This suggestion is in line with other cases where an isoleucine side chain was seen to gate the motion of small ligands within proteins.[34] The fit of the water molecule is still a bit tight (distance of only 3.0 Å to the Phe165 side chain), but could be further optimized by either molecular dynamics or Monte-Carlo routine.[16] Such calculations, which are outside the scope of the present work, may require the (unknown) conformation from before the carbanion formation step, which likely differs somewhat from the X-ray structure of the mature GFP.

Finally we note that Tyr145-OH is only 0.7 Å further away than His181. With a pK_a around 10, we do not expect it to be ionized. However, when irradiated (peak absorbance near 275 nm) it could undergo excited state proton transfer, ejecting the hydroxylic proton into the proton wire shown in Figure 4. In a few nanoseconds its anion will decay back to the ground state, producing a basic moiety that could participate in abstracting the proton from the bridge carbon. If this mechanism holds,
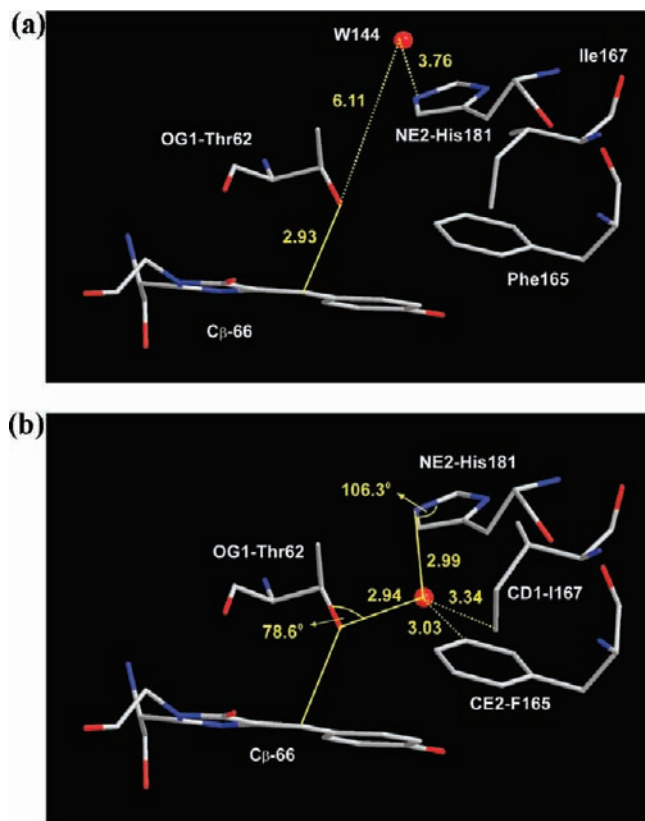
Figure 8. Suggested steps in the formation of a transient proton wire connecting $C_\beta 66$ with the His181 base. (a) Rotation of the Thr62 side chain. (b) Rotation of the Ile167 side chain with translation of a water molecule into the ensuing cavity. From PDB file 1W7S, subunit B.[25] Compare with Figure 5. Full yellow lines denote HBs.

one may expect GFP chromophore maturation to be catalyzed by UV light, an experiment yet to be done.

V. Conclusion

PT in proteins depends crucially on HBed networks ("proton wires") formed by oxygen atoms from internal water molecules and side chains, and (depending on the time scale) possibly also backbone carbonyls, sulfur, and nitrogen atoms. These may form large HBed clusters, of which only select portions were discussed in the literature. Examples are the proton wire fragment connecting the zinc atom to His64 in CA-II, or the Tyr66 to Glu222 fragment in GFP. It may well be that the initial intuition was correct, and these are indeed the only functionally important wire segments, but this can be decided only after all relevant segments are investigated. It is therefore desirable to have a simple to use computer code for systematic mapping of proton wires in proteins.

Here we have developed a simplified version of such an algorithm, which scans the atomic coordinates systematically and reports on all the clusters of interconnected oxygen atoms (see Methods). Application of this algorithm consistently reports larger HB clusters than encountered in the literature. Moreover, with increase in X-ray structural resolution these clusters tend to increase even further. In the present study we have exemplified the potential of the methodology on two proton wires: The rather extensively investigated active site proton wire in CA-II and a previously unknown, nonactive site internal proton wire in GFP.

The active site cluster of CA-II[21] is surprising because it exhibits an unreported direct exit to solution. From the buried

SCHEME 7: The Four Protons, H1–H4, Eliminated during Chromophore Biogenesis in GFP
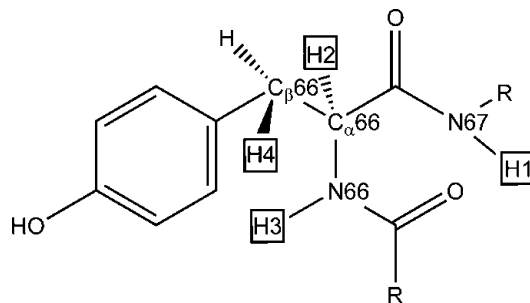
"active site caldera", where the zinc center resides, two proton wires climb on the enzyme surface up to Glu69, which is located within a negatively charged patch on the rim of the caldera (Figure 2b). Do protons utilize these wires rather than the His64 flip for either exit or entry? Are the different paths utilized under different conditions? Without detailed experimentation and calculations it is hard to assess the significance of this observation. Nevertheless, the fact that these surface wires lead to a region of negative charge density hints to the possibility of positively charged buffer molecules docking there,[42] from which protons can be fed into the active site via these surface pathways when the enzyme works in the dehydration direction.

The finding of the new proton wire in GFP, unconnected to the Tyr66 active site, is also intriguing. Does it have a functional role? Here we have argued that residues on this wire may be involved in GFP chromophore biosynthesis. Overall this reaction requires the abstraction of four protons/hydrogens:[56] from N67, $C_\alpha 66$, N66, and $C_\beta 66$ (Scheme 7). Each proton entails a different mechanism for its abstraction, thus there are *four* (not three) steps in the overall mechanism: (i) cyclization requires the transfer of the proton attached to N67; (ii) enolate forms following abstraction of the proton attached to $C_\alpha 66$; (iii) elimination of the hydrogen from N66 occurs by molecular oxygen; and (iv) the rate limiting step for dehydration is abstraction of a proton from $C_\beta 66$. Here we focused on the carbanion formation in step iv.

The newly observed internal HB cluster could be involved in several ways in this transient carbanion formation (Scheme 6). A rotation of the Thr62 side chain brings it into HB contact with one of the hydrogens of $C_\beta 66$. Completion of a short proton wire from Thr62 to His181 (e.g., by diffusion of a water molecule facilitated by Ile167 side chain rotation as in Figure 8) can allow PT from $C_\beta 66$ to $N_{\epsilon 2}$ of His181, with stabilization of the negative charge by its delocalization to the imidazolone carbonyl which forms a strong HB with Arg96. Rotation of Thr62 to its original orientation prevents back PT. Furthermore, it reestablishes the proton wire to His181, allowing the abstracted proton to delocalize on the wire, until the elimination of the water molecule from C65 is completed.

Again, a mechanism deduced by inspection can be only tentative. Yet the comparison of proton wire motifs across proteins suggests that threonines may indeed function as proton wire switches. The comparison of the CA-II and GFP proton abstraction mechanisms also suggests that a buried Ser/Thr-Glu pair near the active site functions as a rapid reversible proton abstractor. With further comparisons of proton wires in different proteins, one may hope to identify additional proton wire motifs, and these could help decipher the function of new proton wires in other proteins.

Supporting Information Available: Truncated PDB files with the coordinates of the amino acid residues, water oxygen, and zinc atoms participating in the HB clusters shown in Figures 1−4 (to view a cluster, the appropriate PDB file should be loaded into a visualization software with HB drawing capabilities; files are labeled by their original PDB names). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Agmon, N. *J. Phys. Chem. A* **2005**, *109*, 13–35.

(2) Voth, G. A. *Acc. Chem. Res.* **2006**, *39*, 143–150.

(3) Marx, D. *ChemPhysChem* **2006**, *7*, 1848–1870; addendum: *ChemPhysChem* **2007**, *8*, 209−210.

(4) Wraight, C. A. *Biochim. Biophys. Acta* **2006**, *1757*, 886–912.

(5) Swanson, J. M. J.; Maupin, C. M.; Chen, H.; Petersen, M. K.; Xu, J.; Wu, Y.; Voth, G. A. *J. Phys. Chem. B* **2007**, *111*, 4300–4314.

(6) Agmon, N. *Chem. Phys. Lett.* **1995**, *244*, 456–462.

(7) Hückel, E. *Z. Electrochem.* **1928**, *34*, 546.

(8) Bernal, J. D.; Fowler, R. H. *J. Chem. Phys.* **1933**, *1*, 515–548.

(9) Agmon, N. *J. Chim. Phys. Phys.-Chim. Biol.* **1996**, *93*, 1714–1736.

(10) Markovitch, O.; Agmon, N. *J. Phys. Chem. A* **2007**, *111*, 2253–2256.

(11) Luz, Z.; Meiboom, S. *J. Am. Chem. Soc.* **1964**, *86*, 4768–4769.

(12) Markovitch, O.; Chen, H.; Izvekov, S.; Paesani, F.; Voth, G. A.; Agmon, N. *J. Phys. Chem. B* **2008**, *112*, 9456–9466.

(13) Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Phys. Chem.* **1995**, *99*, 5749–5752.

(14) Lapid, H.; Agmon, N.; Petersen, M. K.; Voth, G. A. *J. Chem. Phys.* **2005**, *122*, 014506.

(15) Nagle, J. F.; Morowitz, H. J. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 298–302.

(16) Taraphder, S.; Hummer, G. *J. Am. Chem. Soc.* **2003**, *125*, 3931–3940.

(17) Roy, A.; Taraphder, S. *Biopolymers* **2006**, *82*, 623–630.

(18) Roy, A.; Taraphder, S. *J. Phys. Chem. B* **2007**, *111*, 10563–10576.

(19) Roy, A.; Taraphder, S. *J. Phys. Chem. B* **2008**, *112*, 13597–13607.

(20) Håkansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. *J. Mol. Biol.* **1992**, *227*, 1192–1204.

(21) Fisher, S. Z.; Maupin, C. M.; Budayova-Spano, M.; Govindasamy, L.; Tu, C.; Agbandje-McKenna, M.; Silverman, D. N.; Voth, G. A.; McKenna, R. *Biochemistry* **2007**, *46*, 2930–2937.

(22) Agmon, N. *Biophys. J.* **2005**, *88*, 2452–2461.

(23) Yang, F.; Moss, L. G.; Phillips, G. N., Jr. *Nat. Biotechnol.* **1996**, *14*, 1246–1251.

(24) Örmo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392–1395.

(25) van Thor, J. J.; Georgiev, G. Y.; Towrie, M.; Sage, J. T. *J. Biol. Chem.* **2005**, *280*, 33652–33659.

(26) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362–8367.

(27) van Thor, J. J.; Zanetti, G.; Ronayne, K. L.; Towrie, M. *J. Phys. Chem. B* **2005**, *109*, 16099–16108.

(28) Leiderman, P.; Huppert, D.; Agmon, N. *Biophys. J.* **2006**, *90*, 1009–1018.

(29) Agmon, N. *J. Phys. Chem. B* **2007**, *111*, 7870–7878.

(30) Helms, V.; Straatsma, T. P.; McCammon, J. A. *J. Phys. Chem. B* **1999**, *103*, 3263–3269.

(31) Silverman, D. N.; McKenna, R. *Acc. Chem. Res.* **2007**, *40*, 669–675.

(32) Chen, K.; Hirst, J.; Camba, R.; Bonagura, C. A.; Stout, C. D.; Burgess, B. K.; Armstrong, F. A. *Nature* **2000**, *405*, 814–817.

(33) Dantsker, D.; Samuni, U.; Friedman, J. M.; Agmon, N. *Biochim. Biophys. Acta: Proteins Proteomics* **2005**, *1749*, 234–251.

(34) Ostermann, A.; Waschipky, R.; Parak, F. G.; Nienhaus, G. U. *Nature* **2000**, *404*, 205–208.

(35) Luzar, A. *J. Chem. Phys.* **2000**, *113*, 10663–10675.

(36) Hubbard, S. J.; Thornton, J. M. *'NACCESS' Computer Program*, Department of Biochemistry and Molecular Biology, University College London, 1993.

(37) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.

(38) CambridgeSoft Corp., 2005, http://www.cambridgesoft.com/software/ChemOffice/.

(39) Liang, Z.; Xue, Y.; Behravan, G.; Jonsson, B.-H.; Lindskog, S. *Eur. J. Biochem.* **1993**, *211*, 821–827.

(40) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. B* **2008**, *112*, 5500–5511.

(41) Bottoni, A.; Lanza, C. Z.; Miscione, G. P.; Spinelli, D. *J. Am. Chem. Soc.* **2004**, *126*, 1542–1550.

(42) Duda, D.; Govindasamy, L.; Agbandje-McKenna, M.; Tu, C.; Silverman, D. N.; McKenna, R. *Acta Crystallogr. D* **2003**, *59*, 93–104.

(43) An, H.; Tu, C.; Duda, D.; Montanez-Clemente, I.; Math, K.; Laipis, P. J.; McKenna, R.; Silverman, D. N. *Biochemistry* **2002**, *41*, 3235–3242.

(44) Marino, S.; Hayakawa, K.; Hatada, K.; Benfatto, M.; Rizzello, A.; Maffia, M.; Bubacco, L. *Biophys. J.* **2007**, *93*, 2781–2790.

(45) Petersen, M. K.; Iyengar, S. S.; Day, T. J. F.; Voth, G. A. *J. Phys. Chem. B* **2004**, *108*, 14804–14806.

(46) Petersen, M. K.; Voth, G. A. *J. Phys. Chem. B* **2006**, *110*, 7085–7089.

(47) Royal Swedish Academy of Science Press Release, 2008; http://nobelprize.org/nobel_prizes/chemistry/laureates/2008/press.html.

(48) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509–544.

(49) Cubitt, A. G.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *Trends Biochem. Sci.* **1995**, *20*, 448–455.

(50) Reid, B. G.; Flynn, G. C. *Biochemistry* **1997**, *36*, 6786–6791.

(51) Barondeau, D. P.; Putnam, C. D.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12111–12116.

(52) Wachter, R. M. *Acc. Chem. Res.* **2007**, *40*, 120–127.

(53) Rosenow, M. A.; Huffman, H. A.; Phail, M. E.; Wachter, R. M. *Biochemistry* **2004**, *43*, 4464–4472.

(54) Barondeau, D. P.; Tainer, J. A.; Getzoff, E. D. *J. Am. Chem. Soc.* **2006**, *128*, 3166–3168.

(55) Wood, T. I.; Barondeau, D. P.; Hitomi, C.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. *Biochemistry* **2005**, *44*, 16211–16220.

(56) Pouwels, L. J.; Zhang, L.; Chan, N. H.; Dorrestein, P. C.; Wachter, R. M. *Biochemistry* **2008**, *47*, 10111–10122.

(57) Gu, Y.; Kar, T.; Scheiner, S. *J. Am. Chem. Soc.* **1999**, *121*, 9411–9422.

(58) Edgcomb, S. P.; Murphy, K. P. *Proteins: Struct. Funct. Genet.* **2002**, *49*, 1–6.

(59) Medda, R.; Padiglia, A.; Pedersen, J. J.; Floris, G. *Biochem. Biophys. Res. Commun.* **1993**, *196*, 1349–1355.

(60) Yorita, K.; Janko, K.; Aki, K.; Ghisla, S.; Palfey, B. A.; Massey, V. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9590–9595.

(61) Wu, C.-Y.; Lee, H.-J.; Wu, S.-H.; Chen, S.-T.; Chiou, S.-H.; Chang, G.-G. *Biochem. J.* **1998**, *333*, 327–334.

(62) Lehoux, I. E.; Mitra, B. *Biochemistry* **1999**, *38*, 9948–9955.

(63) Buncel, E. *Carbanions: Mechanistic and Isotopic Aspects*; Elsevier: Amsterdam, 1975; Chapter 4, Section 4.

(64) Jain, R. K.; Ranganathan, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 111–116.