REVIEW

Internal water molecules and H-bonding in biological macromolecules: A review of structural features with functional implications¹

EDGAR MEYER

Biographics Laboratory, Department of Biochemistry and Biophysics,
Texas A&M University, College Station, Texas 77843-2128

(RECEIVED February 13, 1992; REVISED MANUSCRIPT RECEIVED August 26, 1992)

Abstract

Conserved structural patterns of internal water molecules and/or H-bond chains were observed and are here correlated in this review, which then describes two functional properties: equilibration of hydrostatic pressure and proton transport. Available evidence in support of these hypotheses is presented, together with suggested experiments to test them. High-resolution crystal structures of a variety of proteins were studied with interactive computer graphics. Conserved H-bonding linkages may be used as a paradigm for a rationalization of proton transport in membranes. The concept of the "proton wire," which links buried active-site amino acids with the surface of the protein raises the more general question of the functional role of the various molecular components

Keywords: enzymology; H-bonding; internal water; proton transport; three-dimensional computer modeling

As a biological molecule, water is an enigma. In clearly defined cases, a single water molecule can be shown to be catalytically or structurally essential, but in many structure–function studies, its role ranges from uncertain to unknown. In living systems, water is pervasive and ubiquitous, if not expressly excluded. This review documents and integrates known structural roles of water and chains of H-bonds; it points to established functional roles and proposes hypotheses for additional functional roles. A rationale is proposed that provides a plausible mechanism to unify various molecular systems: the observable H-bonding linkage of the buried active site with bulk solution. More than one functional role may be possible.

Survey of structural water in biological systems

Structurally ordered water molecules are best identified by the crystallographic method (X-ray, or better, neutron diffraction) or in special cases by NMR spectroscopy (Otting et al., 1991). Only as a structure analysis approaches higher (ca. beyond 2.5 Å) resolution is it possible to begin to identify the positions of individual water molecules.² Different criteria are used by different authors to determine water positions, so the results reported by investigators must be weighted accordingly. As a general rule, in order for a residual electron density peak to be identified as a water molecule, it must have at least one contact (ca. 3 Å) and have a temperature factor less than some preset maximum; accordingly, it must persist in a difference "omit" Fourier map. The solvent content usually found in protein crystals has been discussed by Matthews (1968).

Reprint requests to: Edgar Meyer, Biographics Laboratory, Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843-2128.

¹ Preliminary presentation at the symposium "Molecular Structure: Chemical Reactivity and Biological Activity," Beijing, People's Republic of China, September 1986.

Abbreviations: HLE, human leucocyte (neutrophil) elastase; MD, molecular dynamics; PEX, PFIGS extension of X-windows; PPE, porcine pancreatic elastase; RMS, root mean square; 3-D, three-dimensional; when available, the four-character-code (e.g., 1TIM) from the Protein Data Bank (Bernstein et al., 1977) is given.

² The scattering powers (i.e., numbers of electrons) of H₂O, OH[−], NH₄⁺, and Na⁺ are virtually indistinguishable in X-ray protein crystallography. The 3-Å "ring" is a characteristic property of diffraction diagrams, primarily representing amorphous or bulk water in the sample.

Without precise neutron diffraction studies, the structural aspects of hydrogen bonds are likewise a matter of deduction or conjecture. Unless the H atom position is determined, the angular component (Baker & Hubbard, 1984; Scheiner, 1988; Geller et al., 1990) cannot be evaluated, so only the heavy atom (O, N, etc.) donor-acceptor positions can be used as a necessary but not always sufficient determinant of an H-bond. Inspection of the interaction geometry makes the assignment of an H-bond sometimes questionable, both on the basis of the distance (<3.2 Å) but especially on the basis of the angle. For example, the catalytic Ser-His ($>\text{N}\epsilon\text{-(H)}\cdots\text{O}\gamma\text{-C}\beta$) angle should be optimally 111° for a linear H-bond but is generally lower (range of 76.6° to 117.8° in trypsin [Marquart et al., 1983]), especially for nonliganded enzymes. Thus, the most crucial H-bond in the serine proteases often appears to be the least substantial.

Careful identification and refinement of water molecules is occasionally neglected as a tedious aspect during the "end game" of structure analysis. Significant structures have been deposited in the Protein Data Bank (Bernstein et al., 1977) with well-characterized water structures (e.g., aprotinin [Wlodawer et al., 1984], crambin [Hendrickson & Teeter, 1981], ribonuclease A [Wlodawer et al., 1986], subtilisin + eglin [Bode et al., 1986a], ribonuclease T1 [Malin et al., 1991], etc.) and are available for comparison.³ Perutz (1992) discussed physiologically crucial H-bonding for the stabilization of the O_2 complex of oxymyoglobin and other relevant H-bonding interactions.

Surface properties of water have been demonstrated in the selective crystallization of materials (Weissbuch et al., 1991). The structural properties of surface water have been recently reviewed (Waks, 1986; Teeter, 1991). A variety of software is available for the prediction of surface water molecules (cf. Vedani & Huhta, 1991).

Bound water

Globular proteins are generally characterized as having a hydrophobic interior and a hydrophilic exterior. Water can be identified as "surface" or "bound," according to the ratio of contacts to protein vs. other water molecules. Bound water molecules can further be divided into three categories (Finney, 1979): bound to (1) metal ions, (2) charged groups, and (3) polar groups. In clearly defined cases, buried polar and especially charged groups frequently have internal water molecules nearby. Conversely, in careful neutron diffraction studies of myoglobin (Cheng & Schoenborn, 1991), distinct water molecules are

never found by diffraction methods in the absence of a polar group. Because of limited experimental information, the functional significance of bound water is often a matter of some speculation. Authors of individual papers have pointed to regions of bound water and suggested possible functional features (polarization, altered dielectric constants in the vicinity of the active site, structural stability of functional groups, screening effects, etc.), but no coherent pattern has emerged until now from surveys of the scientific literature. In some cases, authors indicate that localized water "may play an important role in the catalytic activity of the enzyme" (Suguna et al., 1987, p. 898), or "stabilization" could be achieved through both a hydrogen bond and/or electrostatic interaction (Christianson & Lipscomb, 1986, p. 7571).

The conserved presence of internally bound water channels in the serine proteases provided the initial impetus to study their structure-function relationships. Two functional roles of internal water molecules have been presented (Meyer et al., 1988b), hydrodynamic and ionic linkage. The first is briefly treated here; the second forms the basis for the remainder of the review. By considering in detail specific and structurally conserved water molecules in the serine proteases, the two roles discussed below form a basis for exploring a variety of enzymic and transport systems.

Hydrodynamic molecules

As a hydrated substrate begins to coalesce with a hydrated receptor site, intervening water molecules become trapped (like inserting a sandy key into a sandy lock). This begins to become a serious situation as the nascent complex begins to seal all possible exits along the edges of the contact interface. Before letting the ligand advance into the receptor cavity, trapped water molecules must be displaced by a dynamic mechanism: (1) openings in the "seal" occur and water can slip out, (2) some "rear exit" opens up to permit trapped water to escape, or (3) the receptor and/or ligand must modify the tertiary structure to accommodate the displaced water molecules.

A single intervening water molecule (M. Geller et al., 1990) can keep a ligand from binding to a protein.

The serine proteases have conserved water channels linking the S1 (primary specificity) site (notation of Schechter & Berger, 1976) with a remote site on the surface of the enzyme. In some cases, this channel is unobstructed: A large exit channel in the thrombin structure has been described by Bode et al. (1989a) (Kinemage 1). In other cases, a blocking group (e.g., His 40 imidazole in PPE [Meyer et al., 1988b,c], neighboring Val 190, and/or Val 216 groups in HLE [Bode et al., 1986b, fig. 6, 1989b]) must be dynamically moved, like the locks in a canal, to permit the displacement of individual water molecules. The RMS displacement of bulk water has been estimated to be ca. 6 Å/ps from MD simulations (Ji et al., 1992)

³ However, a few remain deficient, either because the authors did not submit the complete coordinate set, or, in the more general case, the limits of resolution of the experimental data are such as to impede the determination of bound and surface water molecules, whose presence and possible role must be left to conjecture.

with a mean-free path of ca. 0.2 Å; internal water is generally much less mobile. An equality exists between the displacement volume of the binding ligand and the known volume of a water molecule to give a precise count of the number of water molecules that must be displaced, thus perhaps offering some hope for experimental observations that may help illuminate this proposed mechanism.

Subsequent product release is facilitated by water molecules flowing back into the volume left by the vacating ligand, thus further facilitating the final step of enzymatic turnover. Even though the interaction energies cancel out, both these steps need to be factored into any energetic analysis of catalysis. Strictly speaking, hydrodynamics deals with astronomical numbers of water molecules; here the dynamic displacement of a small number (2, 3, etc.) is absolutely crucial for tight, specific binding prior to catalysis or ligand + receptor-induced responses. Water tunnels of various sizes in the serine proteases and other enzymic systems represent a conserved structural property ("rear exit") that can resolve this problem for the functioning enzyme and thus serve as an example of a mechanism for pressure equilibration. Experimental data to support the proposed mechanism might come from studies of enzymes with internal, tritiated water (Rose et al., 1989, 1990). The careful analysis of water and H-bonding in trypsin by means of neutron diffraction (Finer-Moore et al., 1992) independently reaches the same conclusion, which they buttress by a comparison with subtilisin.⁴ A mechanism for displacing trapped water will become an essential aspect in future efforts in molecular design. It is hoped that this review will stimulate the search for experimental tools to evaluate the functional mechanism for displacement of a trapped water molecule.

Ionic (H-bond) linkages

The H-bond

The hydrogen bridge or hydrogen bond was first proposed by Huggins (1922, 1971) in 1919; however, Pauling (1960) credits Moore and Winmill (1912) with the first inference with the structure of trimethylammonium hydroxide. The H-bonding properties of water have been described and reviewed comprehensively by Jeffrey and Saenger (1990), who noted that bound water molecules tend to have Debye-Waller (B) temperature factors commensurate with the bound protein (C, N, O) atoms, thus confirming their structural significance. Specific internal water molecules are located within globular proteins (Rashin et al., 1986), thus confirming their role in protein folding and stability. Conversely, sizable internal cav-

ities have been introduced by site-directed mutagenesis (e.g., Phe → Ala) and shown by high (2 Å or better) resolution studies (Eriksson et al., 1992b) to be void of water or ions, suggesting a deterministic, nonrandom function of bound water. One of these cavities, buried 7 Å, was shown to bind a benzene molecule (Eriksson et al., 1992a). In favorable cases, spectroscopic (IR, NMR) studies have identified and characterized internally bound water molecules (Otting et al., 1991; Maeda et al., 1992). H-bonding in globular proteins was reviewed by Baker and Hubbard (1984), who applied statistical analyses to surface and bound water molecules.

The polarizability of a neutral molecule in the liquid state was first considered by Onsager (1936) and later applied to ice (Runnels, 1966; Onsager, 1973) and to a chain of H-bonds across a nonaqueous medium (Onsager, 1969). H-bond chains were proposed explicitly by Williams (1961) and implied by Mitchell (1961) and form the basis of the current understanding of oxidative phosphorylation in metabolism. Yet, as this review will explore, their existence is either overlooked or not related functionally in most of the structural literature. H-bond chains were treated theoretically by Nagle et al. (1980)⁵ and mechanistically by Goodenough (1986).⁶ Ovchinnikov (1981) discusses various types of ion channels and gives a pictorial view of a water channel in a membrane (Ovchinnikov, 1981, fig. 10) but, as pointed out by Nagle et al. (1980), a transmembrane water channel would not be functional for proton transport because of (1) permeability of other ions (Na⁺, etc.), (2) dielectric breakdown,⁷ and (3) need for a simple switching mechanism without major conformational changes in the protein (too slow and energetically costly).

Serine proteases

The serine proteases, whose catalytic mechanism is (or appears to be) well understood, are one of the most thor-

⁵ "The basic molecular construct in our theories is a chain of hydrogen bonds that traverses a nonaqueous region. . . . Such a chain, originally discussed by Onsager (1969), may be embedded in membrane bound proteins and may connect the intracellular and the extracellular aqueous regions. The hydrogen bonds would form between various hydrogen donor-acceptor side groups . . . and could also involve tightly bound water molecules" (Nagle et al., 1980, p. 3959).

⁶ "A bare proton may also diffuse through a hydrogen bond system to give a dc current. It does so by the Grotthuss mechanism, which consists of a combination of piggyback reorientations and proton displacement" (Goodenough, 1986, p. 266).

⁷ "Along such a well there would be a nearly constant electrical potential V . This would lead to an electric field across the nonaqueous portion of $\Delta V/(L - D)$, where ΔV is the transmembrane potential, L the membrane thickness, and D the well depth. As D approaches L the field strength would lead to dielectric breakdown. For $\Delta V = 100$ mV and $L - D$ equal 5 Å the field strength would be 2×10^6 V/cm, which is in the unstable range. . . . In contrast to the water pore or well with bulk water properties, a hydrogen bonded chain will support a non-zero electrical field along it even though the chain is blocked (open circuit) and no current is flowing" (Nagle et al., 1980, p. 3960).

⁴ "The existence of water channels in two serine proteases, which are not structurally homologous, provides a stronger argument that such channels provide conserved mechanisms for water displacement by substrate" (Finer-Moore et al., 1992, p. 218).

oughly studied enzymatic families in the structural literature. A conserved, familiar structural homology found in chymotrypsin (Tsukada & Blow, 1985), trypsin (Bode & Schwager, 1975), kallikrein (Chen & Bode, 1983), α -lytic protease (Bone et al., 1989), rat mast cell protease (Reynolds et al., 1985), porcine pancreatic elastase (Sawyer et al., 1978; Meyer et al., 1988b; Takahashi et al., 1989), human leucocyte elastase (Bode et al., 1986b), thrombin (Bode et al., 1989a), rat submaxillary gland tonin (Fujinaga & James, 1987), proteinase K (Betzel et al., 1988), *Streptomyces griseus* protease (James et al., 1980), subtilisin BPN' (Wright et al., 1969) and NOVO (Drenth et al., 1971; McPhalen et al., 1985), subtilisin Carlsberg + eglin c (Bode et al., 1986a), *Bacillus amyloliquefaciens* subtilisin (Bott et al., 1988), thermitase + eglin c (Gros et al., 1989), and tissue-kininogenase (Vijayalakshmi et al., in prep.) has been reported in the crystallographic literature to sufficient resolution that water and H-bonding donor-acceptor pairs are well resolved. The structural, chemical, and physiological roles of the proteases have been reviewed (Barrett, 1986; Bieth, 1986; Bode & Huber, 1986).

Based upon a conserved catalytic triad (Ser 195-His 57-Asp 102, chymotrypsinogen numbering system), Blow et al. (1969) proposed a charge relay mechanism, which now serves as the prime text-book example of enzyme catalysis. This original observation, as discussed by Wright (1972) is based on the assumption that short-lived charged species are created and shielded from bulk water. Possible functional roles proposed for observed active-site and nearby water molecules included "polarization" and modification of the dielectric constant in the environment of individual catalytic residues (Blow et al., 1969). The carboxylate group of Asp 102 is clamped in place by H-bonding to Ser 214 O γ on the one side and by an oxyanion hole-like interaction with >NH groups of His 57 and Ala 56 (both absolutely conserved, Fig. 1a). This latter linkage could help fix an intermediate carboxylate ($-\text{COO}^-$) group and thus stabilize a formal charge separation ($-\text{COO}^- \dots \text{His}^+ \dots \text{O}\gamma$) in one or more plausible microreversible steps along the catalytic pathway (Dutler & Bizzozero, 1989). Further mechanistic insights were offered by Carter and Wells (1988). A water-excluded, stabilizing environment could contribute greatly to the stability of such a charge separation (Warshel et al., 1989; Meot-Ner [Mautner], 1988).

The tetrad

To the conventionally recognized catalytic triad (Blow et al., 1969) must be added the Ser 214 residue, which is conserved in 42 available gene sequences (MIPS protein database) and structurally homologous in all serine protease entries in the Protein Data Bank, establishing a catalytic tetrad of residues,⁸ which reveals a chain of water molecules linked, at the interior, to Ser 214 ($-\text{Asp}-\text{His}-$

Ser 195) and at the exterior, to bulk water, ca. 15 Å removed from the catalytic Ser 195 O γ atom (Figs. 1, 2; Kinemages 2-4). The neutron diffraction study of trypsin provides a definitive description of these H-bonds (Finer-Moore et al., 1992). In all cases the Ser 214 O γ H group is in H-bonding contact to a conserved chain of water molecules (trypsin, chymotrypsin, PPE, HLE, etc.), which leads to bulk solvent at a remote site, or, due to genetic variations, is in immediate contact with bulk water (e.g., α -lytic protease). Moreover, this water channel is conserved in all structures of the trypsin-like serine proteases thus far reported in the structural literature.

With a sequence similarity of 41% to PPE, HLE is noticeably different from PPE outside the active-site region, yet HLE water molecules 564, 552, and 549 are, respectively, 1.1, 0.7, and 0.9 Å displaced from Sol 319-320-321 in PPE, when the two enzymes are superimposed (program ROTMOL; R. Radhakrishnan, unpubl.; the RMS fit of 0.36 Å for 111 common active-site backbone atoms is used for the superposition). As found in all these structural studies, the water molecules in this channel (e.g., Sol 319-320-321 in PPE) all have quite low Debye-Waller temperature factors ($B = 5.4-9.1 \text{ Å}^2$), comparable to neighboring protein (C, N, O) atoms and indicating a high degree of order and structural integrity, which suggests that they exchange less readily with bulk water (suggesting a tritiated water experiment). These internal water molecules were first identified in chymotrypsin (cf. Fig. 17, Birktoft & Blow, 1972); the authors noted the H-bonding linkage to Ser 214 and suggested that they might help "polarize or stabilize" the buried Asp 102 carboxylate group.

The charge relay mechanism was challenged by the neutron diffraction study (Kossiakoff & Spencer, 1980, 1981) of trypsin, which reported His 57 to be doubly protonated, thus obviating the necessity for further translocation of a proton to Asp 102. The small size of the monoisopropyl phosphoryl inhibitor relative to the size of the immediate active-site cavity cannot completely exclude the possibility of H-bonding linkages via water molecules, static or dynamic, from participating in the reaction. Thus, the neutron diffraction experiment, as executed in this specific case, cannot be considered to be definitive. A more recent neutron diffraction study of subtilisin BPN' to 2.0 Å resolution describes a monoprotonated imidazole group in the catalytic tetrad (Kossiakoff et al., 1991), which does not support the double protonation of His 57 imidazole in the presence of biological substrates.

Site-specific mutagenesis experiments of Craik (Craik et al., 1987; Sprang et al., 1987) showed a dramatic decrease (10^4) in enzymatic activity in rat trypsin by replacement of Asp⁻ 102 with Asn (D102N) and an equally dramatic return of activity as a function of increased pH.

⁸ The only known exceptions to this pattern are two inactive isozymes (D, E) of mouse granzymes (Jenne & Tschopp, 1988).

Fig. 1. **a:** Schematic view of the active site of PPE. **b:** Sequential arrangement of proton transfers before, during, and after formation of the acyl intermediate. Panel A represents the resting state of the enzyme, exposed to solvent at pH 7. Panel B illustrates the charge resonance-stabilized active site just before proteolytic attack (thence to panel F). Panel C represents the protonated, unproductive state below the pK_a of His 57 (ca. 3–6.5). With substrate bound and active-site residues shielded (Fig. 1a), panel D represents the rearranged order of protons leading, after flipping of the imidazole ring, to panel E, the imidazole having extracted a proton from Ser 195 O γ ; a base (B^-) can extract a proton via the chain of H-bonds. Panel F represents the resonance charge-stabilized acyl intermediate at the point of Bizzozero–Dutler inversion at the N atom of the scissile bond, permitting proton transfer from His 57 for proteolysis. Panel G represents the acyl complex and neutral His 57, as in panel A. **c:** Stereo view of the 1.65-Å resolution structure of native pancreatic elastase (Meyer et al., 1988b) (3EST), as prepared by program FRODO (Jones, 1978). Water molecules are drawn with a 3.3-Å radius to indicate channel-water contacts. The H-bonding linkage between functional groups is also indicated. The amino acids of the conserved catalytic tetrad are labeled.

Based upon limited modeling assumptions, Warshel et al. (1989) used this report to calculate the energetic necessity of Asp 102 in the catalytic machinery. Mutants of trypsin with Ser 214 replaced by Lys, Glu, and Ala have relative activities against a tripeptide substrate of 0.9%, 45%, and 184% (McGrath et al., 1992); crystallographic analysis to 2.2 Å of the inactive Lys mutant shows that the side

chain occupies the buried water channel, effectively blocking the H-bond network. Presumably, water is free to fill available volume and retain H-bonding contacts in the Ala mutant, which was not studied crystallographically. Although conclusive insights are difficult to obtain experimentally, it is clear that the mechanism of action of the serine proteases is far from simple and not nearly

Fig. 2. Schematic (a) and stereo (b) views of the 1.2-Å resolution structure of subtilisin Carlsberg (1CSE). Thr 33 has been rotated 23° from the position reported by Bode et al., (1986a); a thin line indicates the resulting H-bonding network. Two buried water molecules are drawn as spheres using the default van der Waals radius of program FRODO (Jones, 1978). The conserved residues of the catalytic tetrad are labeled, along with conserved residue Asp 60.

so clearly understood as our textbooks would have us believe. The structures of this tetrad, when superimposed by molecular modeling, present a striking example of a conserved structural homology (cf. fig. 4, Meyer & Bode, 1987) for the superposition of PPE and HLE). Moreover, enzymes of this family exhibit little structural variation upon substrate binding, e.g., the various structures of bound and free trypsin have RMS differences of ca. 0.3 Å (Marquart et al., 1983) and thus may be properly characterized as classical Fischer (1894) "lock and key" enzymes.

Let us consider the limiting case as typical for peptide-substrate interactions along the extended (ca. S4 to S3') binding site: For trypsin, PPE, HLE, etc. upon forming the Michaelis complex with a peptide-like substrate occupying both S and S' regions of the extended binding site (notation of Schechter & Berger [1967]⁹), bulk water is effectively excluded from the active site (Meyer et al., 1988a; Dauter et al., 1991), with one possible exception: χ_1 rotation of the His 57 imidazole side chain from "in" (tetrad) to "out" into solution. This alternative locus for His 57 imidazole has been experimentally observed in the pH 10 structure analysis (Sprang et al., 1987) of rat trypsin (statistical disorder between the "in" and "out" loci); however, within the pH range from 4.1 (*S. griseus*) to 8.0 (trypsin), His 57 is uniformly found to be a firmly fixed feature of the catalytic tetrad when complexed with peptide-like ligands; the "out" locus of His 57 is not observed. An adjacent water molecule in the buried active site has

not been observed crystallographically; it would create an alternative mechanism that could lead to ambiguous reaction products, and enzymes are unambiguous. Thus, water must be excluded during key microreversible steps of proteolysis.

In a number of PPE complexes (Radhakrishnan et al., 1987) and in a trypsin complex (Chow et al., 1990) with heterocyclic inhibitors, steric hindrance with the bound inhibitor forces the His 57 imidazole ring to break the tetrad linkage (after hydrolysis) and seek more favorable nonbonded interactions in the "out" location. Rapid (ca. 5 ps) ring flipping was observed in MD simulations of Asp 102 Asn PPE (Lesyng & Meyer, 1987). When His 57 flips "out," a single water molecule consistently forms an H-bonding link between Ser 195 O γ and Asp 102 O δ (Bode et al., 1989b). With this flipping, Ser 195 and His 57 become accessible to bulk water and in turn, provide access of bulk water to the otherwise "buried" catalytic machinery (Ser 195, Asp 102). This suggests the need for further physical evidence for the conformational role of His 57 in catalysis. In the "out" location, His 57 makes a snug van der Waals contact with the bound ligand (cf. fig. 4, Meyer & Bode, 1987), and, when possible, the imidazole N atoms contribute to the H-bonding pattern of the receptor + ligand complex (Bachovchin et al., 1988).

Both methylated serine proteases (Scholten et al., 1988) and neat solvent systems (Kanerva & Klibanov, 1989; Wong, 1989) depress proteolysis while leaving intact or even enhancing the esterase activity. Our crystallographic determination (Radhakrishnan & Meyer, unpubl.) of elastase (PPE) in 60% dioxane clearly indicated a displacement of His 57 imidazole into the "out" position. The disruption of the catalytic H-bonding network indicates that a different mechanism is operative in organic solvent systems than the dominant mechanism characteristic of

⁹ In the notation of Schechter and Berger (1967), receptor subsites are denoted sequentially by S and juxtaposed peptide residues are denoted by P, counting both directions from the scissile bond, *:

-----S3-S2-S1-S1'-----
*NH-P3-P2-P1*P1'-P2'-COO'.

the native enzyme. In an analogous study of methylated His 57 in PPE (Radhakrishnan & Meyer, unpubl.), the imidazole group was again found to be in the "out" orientation. These studies therefore indicate the existence of a modified structure that can be correlated with a different function (esterase vs. protease) and directly linked to the exposure of the active-site residues (Ser 195) to bulk solvent, which normally is not the case. With Ser 195 O γ exposed to bulk water, the O γ^- oxyanion could be immediately formed but would not be stabilized, readily exchanging with neighboring water molecules in an unproductive fashion. A similar phenomenon was observed for chemically modified carboxypeptidase A (Riordan, 1973).

Considering the limiting case of a totally occluded ligand-receptor complex, we have seen that while the intact catalytic machinery (tetrad) must be shielded from bulk water, a buried H-bonding network exists that links the tetrad with bulk water, buffer, etc., at a remote site (Figs. 1, 2; Kinemages 2-4). Although the spectroscopic/theoretical literature makes occasional references to the charge relay network in the serine proteases, it completely overlooks the H-bonded network immediately adjacent to it. On the basis of such a consistently conserved structural homology (these channels are spatially conserved in all serine proteases thus far studied crystallographically), it is reasonable to infer that some functional characteristic of this family of enzymes exists that thus far has been overlooked. The fundamental question in mechanistic terms is the identification of the origin or driving force acting on buried catalytic residues, whether a serine residue, Zn²⁺-water or carboxylate-water complex. The proposed mechanism for the aspartic proteases (Pearl & Blundell, 1984; Davies, 1990) employs an active-site water molecule as the specific base. Very weak acids like imidazole and alcohols are highly improbable nucleophiles without a major impetus.

Because the active enzyme + substrate complex effectively shields active-site amino acids from bulk solvent, the pH dependence of the reaction suggests that the buried active site must be linked to bulk solvent. The commonly accepted hypothesis (cf., e.g., Kossiakoff & Spencer, 1980) is that His 57 is the base, Asp⁻...His⁺...O γ^- charge separation is stabilized, and no charge relay system exists or is necessary. Above the pK_a of His 57, imidazole (ca. 6.5), is a neutral specie. However, below the pK_a, it would already be protonated and unable to extract a proton from Ser 195 O γ H, unless its proton could be disposed of. However, the arguments made for the resonance stability of the Asp⁻...His⁺...O⁻ constellation (Warshel et al., 1989) argue equally well against the role of His 57 as the proton donor to the N atom of the scissile bond. The presence of an H-bonding network thus suggests another possibility, namely that a base on the surface (Arg or ion [HCO₃⁻, H₂PO₄⁻, OH⁻]) removes a proton from within the protein and this characteristic can

be transmitted to the buried active site. The same linkage that could serve as a sink for a proton could later serve as a source in the microreversible step of H donation to the scissile N amide atom (cf. Dutler & Bizzozero, 1989). The observation that buried active-site residues are linked to the surface of the protein makes such a function topologically possible, but two questions must be addressed: (1) by what mechanism can protons be conducted within a protein? and (2) could such an operation occur within the time frame of hydrolysis (i.e., milliseconds)?

This review establishes the conserved structural presence of buried-to-surface H-bonds (frequently mediated by buried water molecules), proposes a plausible hypothesis: remote-base proton-extraction mediated by "flip-flop" H-bonds, and suggests experiments to test the hypothesis.

Flip-flop H-bonds

Betzel et al. (1984) observed that bridging H atoms were statistically disordered in careful neutron diffraction studies of cyclodextrins. They postulated concerted or consecutive H-bonding mobility to account for the experimental observations. This hypothesis was further explored by MD simulations (Koehler et al., 1987). But their (classical) model involved the rotational flipping of intact -OH groups, with no concurrent (de)protonation. As a quantum-mechanical event, the postulated "flip-flop" H-bonds offered an opportunity to consider an enhanced mechanism by which a proton extracted at one end of a chain of H-bonds can effectively be removed from the other end of the chain (Fig. 1B); this suggests, in turn, an explanation for the pH dependence of proteolysis and a suggestion as to how a bulk-solvent base is capable of affecting proton extraction from Ser 195 O γ in the step immediately prior to nucleophilic attack. This crucial intermediate step would be stabilized by its environment so that the Ser O γ^- ion would have only two alternatives (roughly equally likely), nucleophilic attack (enzymologically productive) or removal of a proton from His 57 (unproductive). This mechanism is therefore a logical extension of Blow's (1976) original description of the charge relay mechanism. Other characteristics (polarization, rigidity, etc.) may contribute to but not originate this crucial action.

Various investigators have identified "hydrogen atom transport chains" (Williams, 1961), the "proton wire" (Zundel & Metzger, 1968), "passive proton wire" (Nagle et al., 1980), the "proton channel" (Ovchinnikov, 1981), and "hydrogen bond systems" (Goodenough, 1986). These concepts are given a structural foundation by considering the flip-flop hydrogen bonds quantum mechanically (Meyer et al., 1988c). The strength and concerted properties of buried, linked H-bonds have been discussed by Scheiner (1988), Scheiner and Das (1988), and over the past 25 years by Zundel and Metzger (1968), Zundel

(1988, 1992), Zundel and Eckert (1989),¹⁰ Zundel and Brzezinski (1991), Krämer et al. (1992), and Olejnik et al. (1992). The chain of -OH groups afforded by buried water and amino acids (Ser, Thr, Tyr) would be energetically uniform (i.e., no disproportionate barriers to [de]protonation), thus facilitating the creation and activation of a "proton wire" (Zundel & Metzger, 1968). Nagle et al., (1980, cf. p. 3967) estimate that "provided n_E is less than 10 kcal, these rates exceed 10^3 /sec, which is easily fast enough for many biological functions."¹¹ Although there is sufficient information in both the spectroscopic and theoretical literature and both fields discovered a fruitful overlap in the study of bacteriorhodopsin (v.i.), they seem to have overlooked the presence of linked H-bonds in functioning enzymes, which we shall next pursue.

Subtilisin

For this hypothesis to be valid, it must apply equally well to the subtilisin family of serine proteases. Two subfamilies ("cousins") of the serine proteases may be characterized by archetypical enzymes, trypsin and subtilisin. Although their amino acid sequences show little resemblance, a similar H-bonding structure (Fig. 2a,b; Kinemage 4) is observed in the 1.2-Å resolution analysis of subtilisin Carlsberg (Bode et al., 1986a). Here, the Ser 221-His 64-Asp 32 triad is conserved and is virtually superimposable with the trypsin-like (195-57-102) members of the family. The active site is buried at the interface of two β -barrel domains. When bound to natural (i.e., peptidic) substrates, these active-site residues are also shielded from possible H-bonding with bulk water molecules, with an accessible surface area of zero for the catalytic tetrad (Dauter et al., 1991).

Wright (1972) observed the "triad" similarity in the two subfamilies of serine proteases. Although the analogous Ser 214 residue is indeed missing from the expected location, Thr 33 in subtilisin Carlsberg (Ser 33 in BPN') is a promising candidate to complete the tetrad, except that the distance of 3.5 Å from Asp O δ 1 to Thr O γ 1 is slightly too long for a proper H-bond. However, with a gentle torsional rotation of the Thr χ_1 angle of 23° (-46° to -67°), no other geometric changes are required of the time-averaged structure to permit the H-bonding pattern

to continue to Thr (Ser) 33. This H-bonding chain extends next to Asp 60 and thence to Sol 523 and on to bulk water (Fig. 2). The existence of a similar tetrad and linking H-bonding network is established. Its structural-functional role is open to experimentation: Asp 60 is a conserved residue in subtilisin and its carboxylic acid group is an integral part of the H-bonding chain; it thus is a candidate for site-specific mutagenesis to test this hypothesis (e.g., replacement with isostructural Leu). Although these relationships may be unique to the trypsin + subtilisin families (and the structures of H-bonding networks are unambiguously conserved), for the proposed mechanism to be more rigorously valid, it must be applicable to additional families of hydrolases.¹²

Other proteases (and hydrolases)

The plant thiol proteases, papain (Kamphuis et al., 1984) and actinidin (Kamphuis et al., 1985), have familial structural homologies that include structurally conserved internal water molecules. As a cysteine protease, papain has obvious differences from serine proteases at the active-site Cys 25. Instead of Asp in the catalytic tetrad, it has Asn 175; His 159 completes the triad; Ser 176 suggests a catalytic tetrad (Fig. 3a,b; Kinemage 5). A homologous "oxyanion hole" has been reported (Kamphuis et al., 1985); the inhibited papain + β -mercaptoethanol complex studied at 2.0 Å resolution has Cys 25 bound to the inhibitor and shifted away (3.8 Å) from His 159, which makes a respectable H-bond (2.7 Å) to Asn 175 O δ 1. The catalytically implicated Asn 175 is 4.2 Å removed from the nearest water molecule, H5, in the time-averaged crystallographic structure.¹³ An analogous chain exists in the active site: Asn 190 N ϵ 1-(2.6 Å)-Ser 176 O γ -(2.8 Å)-H5, which serves as the internal terminus of a channel of six buried water molecules, all with low temperature factors, and is thereby linked to a remote opening, which is flanked by two basic residues, Arg 38 and Lys 17, which could facilitate proton extraction. Thus, the leitmotif returns—a buried active site that is structurally linked to bulk water by a chain of H-bonds.

To explore the possibilities of water-mediated Cys 25 activation, several structural modifications based upon dynamic flexibility present themselves: (1) water (or OH-)

¹⁰ Homoconjugated $B^+H \cdots B \leftrightarrow B \cdots H^+B$ hydrogen bonds cause intense continua in the IR spectra . . . largely structurally symmetrical hydrogen-bonded chains or chains that are built up by the above-mentioned hydrogen bonds cause continua by collective motion . . . it is proved that a large number of hydrogen bond and hydrogen-bonded systems occurring in biochemistry show large proton polarizability (Zundel & Eckert, 1989, p. 73).

¹¹ The crystallographic method is too slow to provide structural information about individual microreversible states along the reaction coordinate (unless some tricks are used to stabilize an intermediate state). Corresponding spectroscopic measurements and MD simulations are needed to explore more fully the rate and geometry of proton exchange in catalysis. The high-resolution neutron diffraction study of natural complexes (e.g., trypsin + aprotinin) could provide definitive information (W. Bode, pers. comm.).

¹² The reader may be spared the rigor and tedium of following the individual chains enumerated here that link buried active-site residues with the enzyme surface (and strong bases). But it is exactly this tedium that is required to probe and delve into the myriad mysteries of a folded, active protein. One suspects that there are mysteries still to be explored: the structural (mechanistic) aspects of allosterism, electron transport (reaction center, etc.), protein folding in an aqueous medium, etc. Only by patient, exhaustive observation of these complex molecules in 3-D can one hope to discover new properties. It therefore is imperative that the investigator have access to facile computer graphics modeling systems with 3-D viewing capabilities.

¹³ As in the PPE structure, these water molecules have relatively low thermal motion (B factor). The authors have labeled the water molecules according to their increasing rank of thermal (B) factors: H2, H3, H5, . . . are all involved in this channel (cf. Baker & Hubbard, 1984, p. 135).

Fig. 3. **a:** Symbolic representation of the active site of papain (Kamphuis et al., 1984) (9PAP). **b:** Stereo view of a water domain (H5-H24-H2-H3-H8-H13-H15) in papain bounded by residues 13-19, 24-35, 47-50, 159-162, and 174-177. The water channel is in contact with bulk water at Arg 83 to H88.

could rearrange itself from the static positions reported in the X-ray structure to extract a proton from Asn 175 or His 159 (and thence to Cys 25); (2) His 159 could rotate about side-chain χ_1 - χ_2 angles to come within H-bonding distance of static water molecule H5, establishing an analogous chain of H-bonds; or (3) Cys 25 could itself rotate 105° about χ_1 into H-bonding position to the terminus of the water channel, H5, then back into attacking position (cf. swinging arm [piggyback] conduction; Williams, 1988, p. 94).

In each case, the buried active-site residues are linked by H-bonds to the surface of the enzyme, some 15 Å removed.

In addition to these proteases, there are other enzyme systems employing the amino acid serine as the active-site agent. They too have H-bonding patterns worth our con-

sideration. The serine hydrolase β -lactamase (Herzberg & Moulton, 1987) contains a buried active-site serine moiety with two H-bonding linkages to remote surface sites.¹⁴ Likewise, the β -lactamase structure (Oefner et al., 1990) exhibits a linked H-bonding network between buried catalytic residues and bulk water, as does penicillopepsin (James & Sielecki, 1985; Fig. 4).

Acetylcholinesterase (Sussman et al., 1991) exhibits the tetrad of catalytic residues, with the buried Ser 200 O γ

¹⁴ Ser 70 O γ (2.5) Lys 73 N ζ (2.8) Glu 166 O ϵ 2 (2.6) H62 (3.4) H15 (3.0) H20 (2.8) H14 (2.6) Asp O δ 2 ^ O δ 1 (2.8) Arg 164 N η 2 ! *Notation:* ^ is used to symbolize the branched H-bonding capabilities of His, Asp, Asn, Glu, Gln, Arg; ! is used to indicate the interface to bulk water; interatomic distances in Å units are given in parentheses. Terminal basic groups (Lys, Arg) are a frequent but not conserved feature of these H-bond chains.

Fig. 4. Schematic representation of the active site of penicillopepsin (James & Sielecki, 1985) (3APP). The surface H-bonding site is designated by the symbol !.

linked by a chain of H-bonds to the enzyme surface over a distance of 10.5 Å.¹⁵ Histidine decarboxylase (Rozwarski & Hackert, 1991) exhibits a spectacular "v-shaped pipeline" of buried water molecules (Fig. 5) linking each buried active site of the catalytic trimer with a remote surface of the enzyme.¹⁶ Two channels of water are well defined in the 1.8-Å resolution structure determination of ascorbate oxidase, linking type 2 and type 3 Cu ions (Messerschmidt et al., 1992, Figs. 16, 17): "this species must accept one or two protons to release one mol of OH⁻ or water. The protons may be provided by bulk solvent in the channels (Fig. 10) perhaps mediated by the histidine copper ligands." In yet another example, the two active sites of the triosephosphate isomerase (1TIM) dimer are

¹⁵ Ser 200 Oγ (3.09) Ne2 His 440 ^ Nδ1 (2.52) Glu 327 Oε1 ^ Oε2 327 (2.98) Ser 226 Oγ [tetrad] (2.98) HOH 643 Glu Oε1 199 ^ Oε2 (3.01) HOH 604 (2.57) Tyr 130 OH (2.66) HOH 611 (2.75) HOH 638 (2.48) HOH 647 (2.86) HOH 670 (2.96) Tyr 121 OH !

¹⁶ "The enzyme is a hexamer of identical subunits which together give an overall dumbbell shaped appearance. Each half of the dumbbell (or trimer) consists of three subunits related by a 3-fold molecular symmetry axis. Each active site is located at the interface of these 3-fold related subunits. Each of the water channels lead from the rear of the active site 15 Å out to the surface of the enzyme, giving the appearance of a cup with three leaks near the bottom. Since these channels may allow bulk solvent to access the active site after substrate has bound and that a net proton is taken-up during catalysis, we are proposing that the water channels may participate in delivering protons to the active site during catalysis. Site directed mutagenesis has been used to convert Glu 66 to a Gln. (McElroy, HE and Robertus, JD Protein Engineering (1989) 3: 43-48) At low pH, the wild type enzyme has high activity and is non-cooperative, while at high pH, has low activity and is cooperative. The E66Q mutant has low activity and is cooperative at both pHs. It appears that removing the 'connection' between subunits results in an enzyme with lowered substrate affinity since the transfer of information between subunits is now slow enough to be detected in the form of sigmoidal kinetics" (D. Rozwarski, pers. comm.).

linked by a chain of H-bonds.¹⁷ TIM has been characterized as a "perfect enzyme" (Knowles, 1987) but this H-bonding connection, which has thus far gone unreported, is consistent with the mechanisms discussed by Fisher et al. (1975). In a series of precise experiments, Rose and Kuo (1989) and Seeholzer et al. (1991) have followed the exchange of a specific H (D or T) atom between substrate, enzyme, and/or product by using deuterium or tritium labels. Their results provide a sharp experimental focus on individual proton exchanges. They concluded that a proton relay mechanism must be operative.¹⁸ Their precise experiments provide insight into the functional properties of specific protons in catalysis and thus may be applied to other enzymic systems that rely on mechanistic proton exchange.

Ribonuclease A — A neutron diffraction study

The X-ray + neutron diffraction study (2.0 Å resolution) of ribonuclease A exhibits several H-bonding chains unambiguously radiating from the active site (Fig. 6; Kinemage 6).¹⁹ Here, bridging H atoms are experimentally observed and together with the previously discussed studies of trypsin (Kossiakoff et al., 1991; Finer-Moore et al., 1992) offer the strongest proof thus far that a chain of H-bonds exists to link the buried active site of the enzyme with a remote surface and bulk water. The role of Asp 121 could be tested by a site-directed mutation to isostructural Leu. Lys 7 could be replaced by (almost) isostructural Met, but surface changes are more difficult to control than are buried groups (like Asp 121) because of the pervasive nature of water.

Metalloproteases

A thermolysin complex (Monzingo & Matthews, 1984) has two linkages radiating from the active-site Zn atom (Fig. 7).²⁰ In discussing the mechanism of action, the au-

¹⁷ ! NγA12 (4.8) OδA10 (3.4) Ne2A64 ^ Oε64 (2.6) H703 (2.4) H714 (2.9) Oδ1B78 (2.8) H675 (4.8) H704 (2.6) Ne2B64 (4.8) Oγ1A75 (4.0) NγB12 ! — either active site is open to solvent; larger distances represent cavities capable of holding more water molecules or quite mobile water molecules. The H-bonding chain is less sharply delineated in this case.

¹⁸ "Unlike hydrogen exchange with the medium, exchange between groups on a protein will not lead to a change in proton potential, and therefore, its occurrence may not be viewed as an important factor. . . . Such proton exchange networks in special cases make possible the mobility of protons for proton relays between the active site and the medium and for rapid charge rearrangement even at a distance from a signal. Given the frequency of hydrogen bonding, the presence of unshared electron pairs on amino acid residues containing N, O, and S, and the presence of bound H₂O molecules, internal proton exchange must be a constant, though silent phenomenon of proteins. "Silent" except when it provides access to a proton pool with which hydrogen from a stable position of the substrate exchanges" (Rose et al., 1990).

¹⁹ His 119N (2.9) Oδ1 Asp 121: (2.6) H239! ^ Oδ2 (3.0) H188! open channel: H160 (3.03 Å) H211! and Gln 11 Ne2 (3.0) H156 (2.51) Nγ Lys 7!

²⁰ Zn—ε2 His 146 Nδ1—Oδ1 Asn 165 Nδ2! and Zn—Oε1 Asp166 ^ Oε2 (3.1) H395 (2.8) Ser 234 Oγ (3.2) H398 (3.1) Oδ2 Asp 213 ^ Oδ1 !

Fig. 5. Stereo plot of the water channel adjacent to the active site of histidine decarboxylase, using coordinates supplied by Rozwarski and Hackert (pers. comm.). Appropriate amino acid residues are labeled; see footnote 16 for details.

thors propose that Glu 143 can activate an attacking water molecule; a "proton shuttle" (p. 5728) is postulated, but further interactions are not pursued. Carboxypeptidase A exhibits two linkages from the active-site Zn atom (Fig. 8a,b).²¹ An unidentified Arg residue was reported to be essential for proteolytic activity (Riordan, 1973; Christianson & Lipscomb, 1986, p. 7571); this hypothesis therefore suggested that Arg 127 or possibly Arg 71 is the essential residue, thus suggesting that Riordan's method could be used to disable the enzyme, which then

²¹ Zn- (2.7) Oε1 72 Asp (2.7) H587 (3.2) Nε Arg 71 !, or H587 (3.4) Nε Arg 127 (2.5) H581 and Zn- Nδ1 His 69 ^ Nε2 (2.6) Oδ1 Asp 142 ^ Oδ2 (3.0) Nη1 Arg 127 (2.5) H581 !. Although it may be presumptuous that mere speculation could contribute to the mechanistic understanding of this important enzyme, the reoccurrence of the leitmotif may stimulate further experimental investigation and insights.

Fig. 6. Schematic representation of the active site of ribonuclease A (Wlodawer et al., 1986) (5RSA). Discrete H-bonding chains are indicated by wavy lines.

could be sequenced to identify the specific Arg residue(s). Site-directed mutagenesis (Arg 127 Lys) has demonstrated that the guanidinium group plays a crucial role in stabilizing the oxyanion intermediate (Phillips et al., 1992). These two metalloenzymes were compared by Kester and Matthews (1977).

Carbonic anhydrase: The kinetic isotope effect

Proton inventory studies and the deuterium isotope effect (Schowen & Venkatasubban, 1984; Scholten et al., 1988; Schowen, 1988) provide supporting experimental evidence for two types of enzymes, the serine proteases (v.s.) and carbonic anhydrase²² (Fig. 9), which has been deter-

²² A deuterium atom (²H) exchanges more slowly than a protium atom (¹H) because of its greater atomic mass and thus helps illuminate functional aspects of catalysis. The method is inherently unable to distinguish larger (3, 4, ..., ∞) numbers of shifted protons and then only reports on the rate-limiting step in proton transfer. A chain of these should be clearly discernable experimentally, so long as no other step along the reaction pathway is rate limiting.

Fig. 7. Schematic representation of the active site of thermolysin (Monzingo & Matthews, 1984) (1TMN).

Fig. 8. **a:** Schematic representation of the active site of carboxypeptidase A (Christianson & Lipscomb, 1986) (3CPA). **b:** Stereo view of the active site of carboxypeptidase A, illustrating the coordination around the Zn atom, complexed with a dipeptidic inhibitor. The H-bonding network is shown with dashed lines leading to bulk water.

mined to 1.8 Å resolution (Eriksson et al., 1988).²³ Venkatasubban and Silverman (1980) used the kinetic solvent isotope effect to measure the change of k_{cat} and K_m and report that between 3 and 4 protons are involved, the actual number being difficult to determine, but concluding that “two or more hydrogens undergo bonding changes in the transition state of the rate-limiting step in the hydration of CO₂ catalyzed by bovine carbonic anhydrase” (Venkatasubban & Silverman, 1980, p. 4988). The active-site diagram (Fig. 9) contains two H-bonding circuits with ca. five H-bonding links. The transfer of five protons during the rate-limiting step is within the precision of the experimentally reported “3 or 4”; thus, this model offers an explanation for the experimental data. It shows the necessity of viewing the active site comprehensively and therefore the inherent limitation of computational simulations based on truncated models. This model therefore identifies the atoms involved: “the excellent fit . . . suggests that the proton transfer of eq 3 involves

water in the active site to a significant extent; a mechanism involving water bridges to relay protons is likely” (Venkatasubban & Silverman, 1980, p. 4988).

As in most other cases, mechanistic thinking has been limited to the immediate active-site region. Krauss and Garmer (1991) reviewed the treatment of carbonic anhydrase and observed that computational approaches have used an overly simplified model of the Zn coordination, without considering the possibility of a proton “shuttling between H-bonded atoms”; most models also failed to consider the H-bonding of Glu 117 to His 119. Although these authors also do not trace out to the surface the H-bonding chain presented here, they do treat the mechanism in some detail, including consideration of the fate of the proton extracted from the reactive water molecule bound to the Zn atom. They conclude, “it does show the need for a more complete treatment exists and that all the studies of CA have not properly considered the problem of the ionicity of the metal active site” nor nearby residues linked by H-bonds. Metalloproteases also must activate an inert water molecule, shuttle a proton, and stabilize intermediates. Thus, a common theme persists in a variety of enzymic systems, and the reader is encour-

²³ Several papers discuss the H-bonding linkage between the active-site Zn atom and His 64; an additional linkage is found to exist between Zn–His 119–Glu 117–His 107–Tyr 194–Ser 29–Arg 254! (Fig. 8).

Fig. 9. Schematic representation of the active site of carbonic anhydrase (Eriksson et al., 1988) (1CA2). Two H-bonding chains are indicated.

aged to inspect other enzymatic systems for the presence/absence of H-bonding linkages. Observations to date confirm that this is a general structural feature of functional enzymes.

Membrane-bound proteins

As these studies continue, attention is directed to biological systems that specifically include or exclude water and where H-bonding chains are specifically present or absent because of the spatial arrangement of amino acids²⁴ or water. An example where water is specifically excluded is the electron transport region of the photosynthetic reaction center. At 2.3 Å resolution, a number of bound water molecules have been reported in the periplasmic and cytoplasmic regions of the protein complex; water is effectively excluded from the central electron transport pathway, starting with the special pair and continuing to the final ketolonic receptor. No H-bonding chains present themselves to short-circuit the crucial electron transport process, also because this region is effectively shielded from water by the lipid bilayer and nature has made an appropriate choice of buried hydrophobic residues (Huber, 1989, Figs. 16, 20d; Deisenhofer & Michel, 1991, Fig. 1) to achieve this end.

The structure of bacteriorhodopsin has been determined by elegant electron cryomicroscopy (Henderson et al., 1990) to between 3 and 10 Å resolution,²⁵ integrating

manifold analyses ranging from spectroscopy to molecular genetics to crystallography. The retinal-binding site and adjacent residues are located in the central core of this photosynthetic protein, linked to a relatively open channel on the extracellular surface but to a tightly confined hydrophobic domain on the cytoplasmic side. A photosynthetic reaction pumps a proton from the cytoplasmic region via water-exposed Asp 96 to the (Lys 216 + retinal) Schiff base and on to the exterior of the cell via Asp 85. A crucial aspect of proton transport involves a 10–12-Å gap between the Schiff base and Asp 96.²⁶ Figure 10 depicts a model of two possible chains of bound water. Site-directed mutagenesis has been applied to specific amino acids in bacteriorhodopsin to study proton transport (Miller & Oesterhelt, 1990; Marti et al., 1991) and reactivity of photo intermediates (Subramaniam et al., 1991). An inspection of the structure reveals that although several hydrophobic residues are nearby, besides the immediately adjacent Thr 89 hydroxyl group, no H-bonding amino acid side chain intervenes in the 10–12-Å trace between the Schiff-base-linked chromophore and the crucial Asp-96 residue, which itself is located at the opening of the membrane-bound channel; thus, there is just room for water to form a network of H-bonds between Thr 89 at the the Schiff base and the Asp 96 carboxylate group. Asp 96 has an exceptionally high pK_a (10; Metz et al., 1992); water is observed in the infrared (Maeda et al., 1992) in relation to the state of protonation of the Schiff base. Free diffusion of a proton can be excluded on experimental grounds (cf. Miller & Oesterhelt, 1990). Bacteriorhodopsin studies present a more tractable experimental system, and current results support the hypothesis presented here.

In a related instance, the presence/absence of a hydroxyl group (Ser/Ala) at position 233 of the long-wave visual pigment has been shown to be the cause of our relative sensitivity to red light (color blindness). This is related to a single nucleotide difference in the X chromosome (Merbs & Nathans, 1992; Winderickx et al., 1992). The H-bonding characteristics of visual pigment thus offers another opportunity for intensive study of a physiological phenomenon. The authors report that molecular modeling methods are being explored to predict possible site-directed mutations that could specifically block this H-bonding linkage.

As explicitly enunciated by Williams (1961), a funda-

²⁴ Backbone >NH groups are less good linkages in an H-bonding chain; neighboring groups also affect the H-bonding capabilities of amides. The ionization of an amide group has been measured to be ca. 30-fold weaker than a corresponding carboxylate group (Jorgensen & Pranata, 1990; Perrin et al., 1990).

²⁵ Model based on electron diffraction amplitudes for which electron microscopy phases were measured to 3.5 Å in the x and y directions and an effective 10 Å resolution along the z direction.

²⁶ "The role of Asp96 has been established recently as that of the proton donor . . . its position 10 to 12 Å above the Schiff base . . . is entirely consistent with this function. We assume that there must be at least one or two structured water molecules in the otherwise very hydrophobic channel above the Schiff base" (Henderson et al., 1990, p. 923). "Water molecules in particular, whether in the narrow, hydrophobic channel above the Schiff base or in the wider, hydrophilic channel below the Schiff base, are certain to play a role in proton transfer. Gating . . . might be mediated by appropriate changes in hydrogen bonds without any change in the channel structure" (Henderson et al., 1990, p. 926).

Fig. 10. Stereo view of the bacteriorhodopsin structure of Henderson et al. (1990) with the z-axis vertical. The seven helices are drawn in thin lines as α -carbon traces. Functionally important amino acid are drawn and labeled. The Schiff base formed between retinal and Lys 216 is drawn in heavy lines. Atomic positions of water molecules were taken from the neutron diffraction study of ribonuclease A (Wlodawer et al., 1986) and used to create the two chains that present a plausible (within the limits of precision of the entire model) link for proton transfer. The interface surfaces between these water molecules and the hydrophobic region of the protein are drawn as a dotted surface with a double van der Waals radius. The regions above Asp 96 and below Arg 82 are reported to be open to bulk solvent.

mental aspect of Mitchell's (1961) chemiosmotic theory is that protons are selectively transported across a membrane in order to drive the reactions of oxidative phosphorylation. The specific proteins are collectively much too large to be likely to yield to high-resolution crystallography or NMR studies in the next few years, but our knowledge of H-bonding chains, mediated by "flip-flop" H-bond networks found in enzymes, suggests a mechanism by which protons could be effectively conducted across a barrier membrane. Two characteristics are essential: (1) the linkage and (2) the switch.

A linkage alone represents a short circuit, because protons would be able to flow bidirectionally. Therefore a mechanism must be postulated whereby the circuit within the membrane is first charged with a proton (analogously to bacteriorhodopsin), and then the originating circuit is broken by disengaging a switch, leaving the proton no alternative but to continue its trip along the H-bonding chain to the other side of the membrane with reverse current blocked. The switch component may be analogous to the functionality proposed in the putidaredoxin-cytochrome P-450 complex (Baldwin et al., 1991). Although the cytochrome system is employed for electron transport, it is possible to envision an analogous proton transporting/switching mechanism, employing a simple component such as a tyrosine residue, which could rotate to make/break an H-bonding chain but also could form a reversible ketonate intermediate for further covalent stabilization of the "off" state. A combination of chemical states and

geometric orientations could help stabilize the "on" and "off" levels of the switch. Although it is a leap in logic to go from the conserved H-bonding domains of enzymes to the membrane-bound machinery of oxidative phosphorylation, the analogy is inviting and therefore the hypothesis would be testable, first with some inspired molecular modeling to create plausible structures capable of forming the linkage and/or the switch and then applying site-directed mutagenesis to disrupt the linkage (no current) or remove the switch (short circuit = infinite, bidirectional flow).

This has been done: decisive experimental information (Gao et al., 1992) both of the structural presence and functional role of a proton network in the Asn 52 Ile mutant of cytochrome *c* illuminates both the conduction and switching ("gate") functions.²⁷

Dynamic structures

Whereas X-ray studies reveal a static, time-averaged picture, MD calculations offer a means of exploring a range of discrete possibilities that would be blurred in the X-ray view. MD simulations of PPE have shown a striking re-

²⁷ "The studies of cytochromes *c* help us to see how, through helix adjustments, proton H-bond networks can be made, can be energised by redox changes and can be broken (gated). A test of the opening of such a proton gate from the centre of a protein to the external water is the rate of exchange of >NH groups inside a protein. . . . This >NH lies in the proton network under discussion" (Gao et al., 1992, p. 351).

lationship: in Asp 102 Asn PPE, the His 57 side chain is quite flexible, corresponding to the disordered model of His 57 in rat trypsin (Sprang et al., 1987). A theme common to the two simulations thus far completed in this laboratory (Lesyng & Meyer, 1987; Geller et al., 1990) and studied visually by means of the facile graphics program MDKINO (Swanson et al., 1989) is that, at no time (total of ca. 275 ps simulation time) did His 57 make the expected H-bond to Ser 195 in the native enzyme (i.e., with no substrate bound). The simulation included crystallographically determined water molecules in the active site as well as ca. 600 surface water molecules. The "weakness" of this specific interaction was established in a comparison of trypsin structures (Marquart et al., 1983). When a mutant of elastase (D102N) was simulated, the rapid (<5 ps) flipping of the His 57 imidazole ring between the "in" and "out" loci was observed. Conversely, the 100-ps simulation of PPE bound to a hexapeptide substrate showed a stable tetrad H-bonding pattern with no flipping of the His 57 side chain. It also showed several very short donor-acceptor contacts (<2.6 Å) between Ser 195 O γ H and His 57, suggesting a dynamic mechanism to help lower the energy barrier for -O γ H proton extraction immediately prior to oxyanion generation and nucleophilic attack.

The mobility and permeability of ubiquitous water molecules (especially at 55.6 M) is widely recognized. Studies over the past 30 years (Hvidt & Nielsen, 1966; Englander & Kallenbach, 1984) considered the variable stability of individual protons to isotopic exchange within a protein (from seconds to days²⁸; Williams, 1988). Regarding the ability of a folding/folded protein to permit benzene to permeate into an internal cavity (Eriksson et al., 1992a,b), whereas water is excluded (presumably throughout the folding and crystallization processes), care must be taken in evaluating enzymatic studies of large vs. small substrates in the serine proteases and other comparable enzymic systems. Because these are classical "lock and key" enzymes with limited induced fit of the protein backbone (ca. 0.3 Å RMS), the size and shape of the ligand can markedly affect the access of water to the active site, not only from the "front" of the enzyme, but also from around the edges of the S1 pocket, in the case of small (e.g., Ala, Gly) side chains at P1. (In this regard, the subtilisin subfamily exhibits a more accessible active site region, e.g., via edges of the S1 primary specificity site.) The interpretation of kinetics experiments with small

model substrates likewise should consider the possibility of an altered mechanism due to invasive water that must be excluded when natural (e.g., oligopeptide) substrates are bound. Many enzymes exhibit greater structural changes upon substrate binding (e.g., the human immunodeficiency virus protease [Miller et al., 1989]); water would accompany any such structural modification and therefore must be included in the modeling process.

Current limitations

Like the weather, molecular design has become a widely discussed area of science. As pointed out by Knowles (1987), in order to design a more perfect enzyme we first must understand the intrinsic properties of the enzyme (including mechanistic roles of H-bonding). This is especially true of efforts to use immunoglobulins, macrocycles, etc. to create tailor-made molecules. It also raises the question of complexity; the folded protein not only serves as a stereospecific filter, it facilitates the creation of rare chemical/electronic states and protects them so that the reaction can proceed. The folded protein thus serves in a variety of crucial roles in living systems, and bound water molecules likewise play a variety of key roles within the folded protein. Viewed dynamically, a harmonious, functional architecture is revealed to us at the molecular, atomic, and electronic levels. Observations presented here may be taken as a paradigm for future investigations; other universal principles still await discovery before our textbooks can offer a complete picture of how an enzyme "really works."

Even as the Protein Data Bank, which received its first entry almost 25 years ago (1969), continues to grow, one may infer that additional structural concepts await discovery; a major limitation to the utilization of structural data bases is the necessity of facile²⁹ 3-D viewing. We cannot escape the fact that biological molecules are profoundly complex. Only by viewing static (i.e., crystallographic and NMR) and dynamic (MD) structures in 3-D (e.g., program MDKINO [Swanson et al., 1989]) is it possible to see beyond the most trivial atomic and molecular interactions. With the aid of 3-D graphics we can escape the confines imposed by the Flatland (Abbott, 1884) where we find ourselves (slides, blackboard, sketches, the printed page, etc.). Until real-time holographic displays (Negroponte, 1991) become available, current graphics technology is amply sufficient to reveal the architectural

²⁸ "(a) Peptide NH protons inside β -sheets or buried in β -barrel-like proteins do not exchange in 10³ sec, even in aqueous solutions of pH 7. (b) Peptide protons in α -helices show faster but still low exchange rates on the order of $\sim 10^2$ sec, even in aqueous solutions at pH 7. (c) Fast exchange i.e. <1 sec is possible from loops of peptides. (d) Rates of H-exchange from exposed amides, amines, alcohols, phenols, carboxylic acids, and indoles are fast, <10⁻² sec, when these groups are exposed to water, but the exchange rates can become slow in an aprotic solvent" (Williams, 1988, p. 92).

²⁹ Vendor-supplied "Crystal-eyes" are suitable for one or two users. A low-cost split-image stereo viewer is available (Oregon). The full-screen liquid crystal display from Tektronix (SGS625) uses low-cost (\$1.50) Polaroid spectacles, permitting groups of 20-30 individuals to see these structures in 3-D simultaneously, making this the preferable method for full-screen stereo viewing. The industry-standard open-systems PEX protocol facilitates the use of workstation software both horizontally to machines from other vendors as well as vertically to newer, faster workstations.

complexity and functional harmony of a biological macromolecule as exemplified by this survey of structures. This is because the human visual process of pattern recognition far exceeds computational approaches, but our visual skills must be trained both to recognize the known and to seek the novel.

Within the limits of the subsets of proteins reviewed, a unifying theme involving the structural presence of buried water molecules and/or H-bond chains in functional proteins has been shown to exist as a general structural feature. Compared to α -helices and β -sheets, this pattern had been more difficult to see when structures were considered individually. When similar protein structures are compared, the atomic positions may change but conserved structural patterns emerge. H-bond chains and buried domains of water molecules can now be added to this list of patterns.

A more descriptive vocabulary?

At one level, this review sought to reconcile what we know with what we do not and then sought ways to illuminate the pervasive darkness. Almost 100 years ago Emil Fischer used the vocabulary of the mechanical age to describe *how* molecules interact ("wie Schloss und Schlüssel"—like a lock and key; Fischer, 1894). What happens *when* they interact? Our prose sadly lacks descriptive terms for such functions, not to mention their 3-D juxtapositions. Whereas pictures (especially in stereo!) play a crucial role in education and communication, we continually grope for words to describe atomic and molecular interactions that form the very basis of life as we know it. Following Fischer's example, we might find useful analogies in microelectronics,³⁰ but one is quickly struck by the inadequacy of such an analogy: Not all molecules are easily adapted to Fischer's "lock and key" analogy, yet, it still has descriptive and didactic utility. Chains of H-bonds can now be recognized to be a universal feature in molecules in which protons are to be transported. Whereas buried water can act in a neutral way to conduct protons, the amino acids and cofactors can act like the components of an electronic circuit to modulate signal or energy transmissions. In most cases, only a small number of functional components is involved (but the environment plays a crucial role in fine-tuning their specific properties, dy-

namically); the remainder of the linkage being assumed by bound water molecules, where the number may be 1 (subtilisin), 2 (HLE, possibly bacteriorhodopsin), or many (His decarboxylase, papain, etc.). These similes are not meant to trivialize the complexity of the molecular structure/function but rather to seek a conceptual basis and vocabulary to link and describe various discrete steps and components, both to let us see what is there but especially to show us what is missing in our understanding.

Summary

Almost 25 years ago, the author made the first use of color raster graphics (Meyer, 1968) to create a molecular model for stereo viewing. Some 10 years passed before color raster graphics became the method of choice in virtually all fields of interactive graphics (and soon will make an even greater impact with the introduction of digital high-definition television). Two related themes span these 25 years, how better to understand the complex structures of biomolecules and how to teach this understanding to others. This review hopefully provides insight into both challenges, because they still are with us. Indeed, the dimension-gap problem raised by Abbott (1884) still pervades current biochemistry in terms of 1-D (sequences) or 2-D (diagrams) "structures." Starting with the serine proteases, this review has surveyed a number of enzymic systems and proposes two water-mediated mechanisms essential for catalysis: ligand-induced hydrodynamic displacement and H-bond-mediated proton displacement. Other types of enzymes were surveyed and uniformly found to have H-bond linkages between the buried active site and the surface of the enzyme. It then was possible to postulate a mechanistic role for chains of H-bonds in membrane transport. A mechanistic analogy to the proton gradient driving force of the chemiosmotic theory is presented. Current studies of bacteriorhodopsin and cytochrome c support this hypothesis and offer a way to test the role of bound water in proton transport.

As additional examples continue to present themselves, this review does not intend to be exhaustive and leaves to the reader the opportunity to check favorite structures and also to design experiments to test resulting hypotheses. Thus molecular design will go beyond whittling on a given molecule to the full-blown sculpting of novel structures. A key aspect of such efforts is the need for facile stereo viewing, as now best provided by workstations equipped with full-screen polarization filters. The media now used to teach structural biochemistry, being 2-D, are inherently inadequate, so more needs to be done to use available technology to bridge this gap. This review thus illustrates how one can apply pattern recognition skills to an available database to relate the obvious to the expected and thence to the unknown. Only as we obtain a deeper understanding of the structure and function of individ-

³⁰ Hydrophobic amino acids might act like resistors; His like a capacitor, Glu, Asp, Lys, Arg, and transition metal ions like transformers; Tyr and Ser (perhaps aromatic residues, purines, pyrimidines, etc.) like different types of semiconductors. Functional groups may be linked by proton "wires," which in some cases may also be switchable. The folded protein thus might be likened to an intricate electronic circuit. Another analogy might envision individual amino acids as flexible 3-D LEGO blocks, changing shape, charge, or packing property as a function of the immediate or even more distant environment. But individual LEGO blocks do not have names, so this analogy, although a curious paradigm, affords no further descriptive powers. (Perhaps writers of science fiction can come to our aid.)

ual molecular components will it be possible to understand more complex phenomena like protein folding or allostereism and use this understanding to design novel molecules for beneficial purposes.

Acknowledgments

An initial impetus was provided by Dr. Greg Cole; Miss Denise Rozwarski, Drs. Wolfram Bode, Klaus Gubernator, Osnat Herzberg, Fritz Winkler, and Profs. Marvin Hackert, Dieter Oesterheld, Joel Sussman, R.J.P. Williams, and Georg Zundel kindly provided coordinates or access to information crucial to the extension of the H-bonding patterns to other molecular systems. Critical comments were provided by Drs. Wolfram Bode, Albrecht Messerschmidt, Rosemarie and Stan Swanson, Profs. Richard Schowen and Frank Raushel, and Mr. Zhang Dachuan. Reviewers kindly pointed to research in related fields, that has been incorporated into the manuscript. This research was supported by the Robert A. Welch Foundation, the U.S. National Science Foundation, N.A.T.O., Dr. John Dinkel, Associate Provost for Computing, Texas A&M University, and also ICI Americas and the Schering-Plough Corp. The author especially thanks Prof. Sir Derek Barton for his encouragement and support.

References

- Abbott, E.A. (1884). *Flatland: A Romance in Many Dimensions*. Seley, London.
- Bachovchin, W.W., Wong, W.Y.L., Farr-Jones, S., & Kettner, C.A. (1988). NMR study of α -lytic protease. *Biochemistry* 27, 7689-7697.
- Baker, E.N. & Hubbard, R.E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44, 97-179.
- Baldwin, J.E., Morris, G.M., & Richards, W.G. (1991). Electron transport in cytochromes P-450 by covalent switching. *Proc. R. Soc. Lond. B245*, 43-51.
- Barrett, A.J. (1986). An introduction to the proteinases. In *Proteinase Inhibitors* (Barrett, A.J. & Salvensen, G., Eds.), pp. 3-18. Elsevier, Amsterdam.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The Protein Data Bank. *J. Mol. Biol.* 112, 535-542; *Eur. J. Biochem.* 80, 319-324.
- Betz, C., Pal, G.P., & Saenger, W. (1988). Synchrotron X-ray data collection and restrained least-squares refinement of the crystal structure of proteinase K at 1.5 Å resolution. *Acta Crystallogr. B44*, 163-172.
- Betz, C.H., Saenger, W., Hingerty, B.E., & Brown, G.M. (1984). Circular and flip-flop hydrogen bonding in β -cyclodextran undecahydrate: A neutron diffraction study. *J. Am. Chem. Soc.* 106, 7545-7557.
- Bieth, J.G. (1986). Elastases: Catalytic and biological properties. In *Regulation of Matrix Accumulation* (Mecham, R.P., Ed.), pp. 217-320. Academic Press, New York.
- Birktoft, J.J. & Blow, D.M. (1972). The structure of crystalline α -chymotrypsin. V. The atomic structure of tosyl- α -chymotrypsin at 2 Å resolution. *J. Mol. Biol.* 68, 187-240.
- Blow, D.M. (1976). Structure and mechanism of chymotrypsin. *Acc. Chem. Res.* 9, 145-152.
- Blow, D.M., Birktoft, J., & Hartley, B.S. (1969). Charge relay + polarization. *Nature* 221, 337-340.
- Bode, W. & Huber, R. (1986). Crystal structures of pancreatic serine proteases. In *Molecular and Cellular Basis of Digestion* (Desnuelle, P., Sjöström, H., & Noren, O., Eds.), pp. 213-234. Elsevier, Amsterdam.
- Bode, W., Mayr, I., Baumann, U., Huber, R., Stone S.R., & Hofsteenge, J. (1989a). The refined 1.9 Å crystal structure of human α -thrombin: Interaction with -Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* 8, 3467-3475.
- Bode, W., Meyer, E.F., & Powers, J.C. (1989b). Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, substrate specificity, and mechanism-based inhibitors. *Biochemistry* 28(Suppl.), 1951-1963.
- Bode, W., Papamokos, E., Musil, D., Seemueller, U., & Fritz, H. (1986a). Refined 1.2 Å crystal structure of the complex formed between subtilisin Carlsberg and the inhibitor eglin c. Molecular structure of eglin and its detailed interaction with subtilisin. *EMBO J.* 5, 813-818.
- Bode, W. & Schwager, P. (1975). The refined crystal structure of bovine β -trypsin at 1.8 Å resolution. *J. Mol. Biol.* 98, 693-717.
- Bode, W., Wei, A.-Z., Huber, R., Meyer, E., Travis, J., & Neumann, S. (1986b). X-ray crystal structure of the complex of human leukocyte elastase (PMN elastase) and the third domain of the turkey ovomucoid inhibitor. *EMBO J.* 10, 2453-2458.
- Bone, R., Frank, D., Kettner, C.A., & Agard, D.A. (1989). Structural analysis of specificity: α -Lytic protease complexes with analogues of reaction intermediates. *Biochemistry* 28, 7600-7609.
- Bott, R., Ultsch, M., Kossiakoff, A., Graycar, T., Katz, B., & Power, S. (1988). The three-dimensional structure of *Bacillus amyloliquefaciens* subtilisin at 1.8 Å and an analysis of the structural consequences of peroxide inactivation. *J. Biol. Chem.* 263, 7895-7906.
- Carter, P. & Wells, J.A. (1988). Dissecting the catalytic triad of a serine protease. *Nature* 332, 564-568.
- Chen, Z. & Bode, W. (1983). Refined 2.5 Å X-ray crystal structure of the complex formed by porcine kallikrein and the bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 164, 283-311.
- Cheng, X. & Schoenborn, B.P. (1991). Repulsive restraints for hydrogen bonding in least-squares refinement of protein crystals. A neutron diffraction study of myoglobin crystals. *Acta Crystallogr. A47*, 314-317.
- Chow, M.M., Meyer, E., Bode, W., Kam, C.-H., Radhakrishnan, R., Vijayalakshmi, J., & Powers, J.C. (1990). The 2.2 Å resolution x-ray crystal structure of the complex formed by reaction of the thrombin inhibitor 4-chloro-3-ethoxy-7-guanidino-isocoumarin with trypsin. *J. Am. Chem. Soc.* 112, 7783-7789.
- Christianson, D.W. & Lipscomb, W.N. (1986). X-ray crystallographic investigation of substrate binding to carboxypeptidase A at subzero temperature. *Proc. Natl. Acad. Sci. USA* 83, 7568-7572.
- Craik, C.S., Rozniak, S., Largman, C., & Rutter, W.J. (1987). The catalytic role of the active site aspartic acid in serine proteases. *Science* 237, 909-913.
- Dauter, Z., Betzel, C., Genov, N., Pipon, N., & Wilson, K.S. (1991). Complex between the subtilisin from a mesophilic bacterium and the leech inhibitor eglin-C. *Acta Crystallogr. B47*, 707-730.
- Davies, D.R. (1990). The structure and function of aspartic proteinases. *Annu. Rev. Biophys. Biophys. Chem.* 19, 189-215.
- Deisenhofer, J. & Michel, H. (1991). High-resolution studies of photosynthetic reaction centers. *Annu. Rev. Biophys. Biophys. Chem.* 20, 247-266.
- Drenth, J., Hol, W.G.J., Jansonius, J.N., & Koekoek, R. (1971). A comparison of the three-dimensional structures of subtilisin BPN' and subtilisin novo. Structure and function of proteins at the three-dimensional level. *Cold Spring Harbor Symp. Quant. Biol.* XXXVI, 107-123.
- Dutler, H. & Bizzozero, S.A. (1989). Mechanism of the serine protease reaction. Stereoelectronic, structural, and kinetic considerations as guidelines to deduce reaction paths. *Acc. Chem. Res.* 22, 322-327.
- Englander, S.W. & Kallenbach, N.R. (1984). Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* 16, 521-655.
- Eriksson, A.E., Baase, W.A., Wozniak J.A., & Matthews, B.W. (1992a). A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature* 355, 371-373.
- Eriksson, A.E., Baase, W.A., Zhang, X.-J., Heinz, D.W., Blaber, M., Baldwin, E.P., & Matthews, B.W. (1992b). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* 255, 178-183.
- Eriksson, A.E., Jones, T.A., & Liljas, A. (1988). Carbonic anhydrase. *Proteins* 4, 274-282.

- Finer-Moore, J.S., Kossiakoff, A.A., Hurley, J.H., Earnest, T., & Stroud, R.M. (1992). Solvent structure in crystals of trypsin determined by X-ray and neutron diffraction. *Proteins* 12, 203–222.
- Finney, J.L. (1979). In *Water, A Comprehensive Treatise* (Franks, F., Ed.), pp. 47–122. Plenum, New York.
- Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* 27, 2984–2993; n.b. p. 2992.
- Fisher, L.M., Albery, W.J., & Knowles, J.R. (1975). The nature of the proton transfer from an acid group at the active site of an enzyme, to solvent water. In *Faraday Society Symposia "Proton Transfer,"* Vol. 10, pp. 154–159. University Press, Aberdeen.
- Fujinaga, M. & James, M.N.G. (1987). Rat submaxillary gland serine protease, tonin. Structure solution and refinement at 1.8 Å resolution. *J. Mol. Biol.* 195, 373–396.
- Gao, Y., McLendon, G., Pielak, G.J., & Williams, R.J.P. (1992). Electron-proton coupling in cytochrome c studied using protein variants. *Eur. J. Biochem.* 204, 337–352.
- Geller, M., Swanson, S.M., & Meyer, E. (1990). Dynamic properties of the first steps of enzymatic reaction of porcine pancreatic elastase. Part II. Molecular dynamics simulation of a Michaelis complex: PPE and the hexapeptide, Thr-Pro-nVal-Leu-Tyr-Thr. *J. Am. Chem. Soc.* 112, 8925–8931.
- Goodenough, J.B. (1986). Proton movement in inorganic materials. *Methods Enzymol.* 127, 263–284.
- Gros, P., Fujinaga, M., Dijkstra, R.W., Kalk, K.H., & Hol, W.G.J. (1989). Thermitase + eglin c. *Acta Crystallogr.* B45, 488–499.
- Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., & Downing, K.H. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* 213, 899–929.
- Hendrickson, W.A. & Teeter, M.M. (1981). Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature* 290, 107–113.
- Herzberg, O. & Moulton, J. (1987). Bacterial resistance to β -lactam antibiotics: Crystal structure of β -lactamase from *Staphylococcus aureus* PC1 at 2.5 Å resolution. *Science* 236, 694–701.
- Huber, R. (1989). A structural basis of light energy and electron transfer in biology. *Angew. Chem. Int. Ed.* 28, 848–869.
- Huggins, M.L. (1922). Atomic structure. *Science* 55, 459–460.
- Huggins, M.L. (1971). 50 years of hydrogen bond theory. *Angew. Chem. Int. Ed.* 10, 147–152.
- Hvidt, A. & Nielsen, S.O. (1966). Hydrogen exchange in proteins. *Adv. Protein Chem.* 21, 287–386.
- James, M.N.G. & Sielecki, A.R. (1985). Stereochemical analysis of peptide bond hydrolysis catalyzed by the aspartic proteinase penicillopepsin. *Biochemistry* 24, 3701–3713.
- James, M.N.G., Sielecki, A.R., Brayer, G.D., Delbaere, L.T.J., & Bauer, C.-A. (1980). Structures of product and inhibitor complexes of *Streptomyces griseus* protease A at 1.9 Å resolution. *J. Mol. Biol.* 144, 43–88.
- Jeffrey, G.A. & Saenger, W. (1990). *Hydrogen Bonding in Biological Structures*. Springer-Verlag, Berlin.
- Jenne, D.E. & Tschopp, J. (1988). Granzymes, a family of serine proteases released from granules of cytolytic T lymphocytes upon T cell receptor stimulation. *Immunol. Rev.* 103, 53–71.
- Ji, J., Cagin, T., & Pettitt, B.M. (1992). Dynamic simulations of water at constant chemical potential. *J. Chem. Phys.* 96, 1333–1342.
- Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* 11, 268–272.
- Jorgensen, W.L. & Pranata, J. (1990). Importance of secondary interactions in triply hydrogen bonded complexes: Guanine-cytosine vs. uracil-2,6-diaminopyridine. *J. Am. Chem. Soc.* 112, 2008–2010.
- Kamphuis, I.G., Drenth, J., & Baker, E.N. (1985). Thiol proteases. Comparative studies based on the high-resolution structures of papain and actinidin, and on amino acid sequence information for cathepsins b and h, and stem bromelain. *J. Mol. Biol.* 182, 317–329.
- Kamphuis, I.G., Kalk, K.H., Swarte, M.B.A., & Drenth, J. (1984). Structure of papain refined at 1.65 Å resolution. *J. Mol. Biol.* 179, 233–256.
- Kanerva, L.T. & Klivanov, A.M. (1989). Hammett analysis of enzyme action in organic solvents. *J. Am. Chem. Soc.* 111, 6864–6865.
- Kester, W.R. & Matthews, B.W. (1977). Comparison of the structures of carboxypeptidase A and thermolysin. *J. Biol. Chem.* 252, 7704–7710.
- Knowles, J.R. (1987). Tinkering with enzymes: What are we learning? *Science* 236, 1252–1258.
- Koehler, J.E.H., Lesyng, B., & Saenger, W. (1987). Cooperative effects in extended hydrogen bonded systems involving O-H groups. Ab initio studies of the cyclic S4 water tetramer. *J. Comput. Chem.* 8, 1090–1098.
- Kossiakoff, A.A. & Spencer, S.A. (1980). Neutron diffraction identifies His 57 as the catalytic base in trypsin. *Nature* 288, 414–416.
- Kossiakoff, A.A. & Spencer, S.A. (1981). Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: Neutron structure of trypsin. *Biochemistry* 20, 6462–6474.
- Kossiakoff, A.A., Ultsch, M., White, S., & Eigenbrot, C. (1991). Neutron structure of subtilisin BPN': Effects of chemical environment on hydrogen-bonding geometries and the pattern of hydrogen-deuterium exchange in secondary structure elements. *Biochemistry* 30, 1211–1221.
- Krämer, R., Zundel, G., Brzezinski, B., & Olejnik, J. (1992). Discussion of the proton potential with proton-transfer equilibria: Thermodynamic data and infrared continua as a function of temperature. *J. Chem. Soc. Faraday Trans.* 88, 1659–1664.
- Krauss, M. & Garmer, D.R. (1991). Active site ionicity and the mechanism of carbonic anhydrase. *J. Am. Chem. Soc.* 113, 6426–6435.
- Lesyng, B. & Meyer, E.F. (1987). Energy minimization and molecular dynamics of Asn 102 elastase. *J. Comput.-Aided Mol. Des.* 1, 211–217.
- Maeda, A., Sasaki, J., Shichida, Y., & Yoshizawa, T. (1992). Water structural changes in the bacteriorhodopsin photocycle: Analysis by Fourier transform infrared spectroscopy. *Biochemistry* 31, 462–467.
- Malin, R., Zielenkiewicz, P., & Saenger, W. (1991). Structurally conserved water molecules in ribonuclease T1. *J. Biol. Chem.* 266, 4848–4852.
- Marquart, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983). The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallogr.* B39, 480–490.
- Marti, T., Otto, H., Mogi, T., Rösselet, S.J., Heyn, M.P., & Khorana, H.G. (1991). Bacteriorhodopsin mutants containing single substitutions of serine or threonine residues are all active in proton translocation. *J. Biol. Chem.* 266, 6916–6927.
- Matthews, B.W. (1968). Solvent content of protein crystals. *Biochemistry* 33, 491–497.
- McGrath, M.E., Vasquez, J.R., Craik, C.S., Yang, A.S., Honig, B., & Fletterick, R.J. (1992). Perturbing the polar environment of Asp102 in trypsin: Consequences of replacing conserved Ser214. *Biochemistry* 31, 3059–3064.
- McPhalen, C.A., Svendsen, I., Jonassen, I., & James, M.N.G. (1985). Crystal and molecular structure of chymotrypsin inhibitor 2 from barley seeds in complex with subtilisin NOVO. *Proc. Natl. Acad. Sci. USA* 82, 7242–7246.
- Meot-Ner [Mautner], M. (1988). Models for strong interactions in proteins and enzymes. 2. Interactions of ions with peptide link and with imidazole. *J. Am. Chem. Soc.* 110, 3075–3080.
- Merbs, S.L. & Nathans, J. (1992). Absorption spectra of human cone pigments. *Nature* 356, 433–435.
- Messerschmidt, A., Ladenstein, R., Huber, R., Bolognesi, M., Avigliano, L., Petruzzelli, R., Rossi, A., & Finazzi-Agro, A. (1992). Refined crystal structure of ascorbate oxidase at 1.9 Å resolution. *J. Mol. Biol.* 224, 179–205.
- Metz, G., Siebert, F., & Engelhard, M. (1992). High-resolution solid state ^{13}C NMR of bacteriorhodopsin: Characterization of $[4-^{13}\text{C}]\text{Asp}$ resonances. *Biochemistry* 31, 455–462.
- Meyer, E. (1968). *Annual Report*. Brookhaven National Laboratory, Upton, New York, p. 81.
- Meyer, E. & Bode, W. (1987). The study and design of specific inhibitors to elastase. In *QSAR in Drug Design and Toxicology* (Hadzi, D. & Jerman-Blazic, B., Eds.), pp. 247–254. Elsevier, Amsterdam.
- Meyer, E., Clore, G.M., Gronenborn, A.M., & Hansen, H.A.S. (1988a). Analysis of an enzyme-substrate complex by x-ray crystallography and transferred nuclear Overhauser enhancement measurements: Porcine pancreatic elastase and an hexapeptide. *Biochemistry* 27, 725–730.
- Meyer, E., Cole, G., Radhakrishnan, R., & Epp, O. (1988b). The crystal structure of native porcine pancreatic elastase at 1.65 Å resolution. *Acta Crystallogr.* B44, 26–38.

- Meyer, E., Takahashi, L., & Radhakrishnan, R. (1988c). A structure: function study of receptor + substrate interactions derived from high-resolution x-ray crystallography. In *Molecular Structure: Chemical Reactivity and Biological Activity* (Stezowski, J., Huang, J.-L., & Shao, M.-C., Eds.), pp. 179–188. Oxford University Press, Oxford.
- Miller, A. & Oesterhelt, D. (1990). Kinetic optimization of bacteriorhodopsin by aspartic acid 96 as an internal donor. *Biochem. Biophys. Acta* 1020, 57–64.
- Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R., Clawson, L., Selk, L., Kent, S.B.H., & Wlodawer, A. (1989). Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* 246, 1149–1152.
- Mitchell, P. (1961). Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191, 144–148.
- Monzingo, A.F. & Matthews, B.W. (1984). Binding of N-carboxymethyl dipeptide inhibitors to thermolysin determined by x-ray crystallography: A novel class of transition-state analogues for zinc peptides. *Biochemistry* 23, 5724–5729.
- Moore, T.S. & Winmill, T.F. (1912). The state of amines in aqueous solution. *J. Chem. Soc.* 101, 1635–1676.
- Nagle, J.F., Mille, M., & Morowitz, H.J. (1980). Theory of hydrogen bonded chains in bioenergetics. *J. Chem. Phys.* 72, 3959–3971.
- Negroponte, N.P. (1991). Products and services of computer networks. *Sci. Am.* 265(3), 106–113.
- Oefner, C., D'Arcy, A., Daly, J.J., Gubernator, K., Charnas, R.L., Heinze, I., Hubschwerlen, C., & Winkler, F.K. (1990). Refined crystal structure of β -lactamase from *Citrobacter freundii* indicates a mechanism for β -lactam hydrolysis. *Nature* 343, 284–288.
- Olejnik, J., Brzezinski, B., & Zundel, G. (1992). A proton pathway with large proton polarizability and the proton pumping mechanism in bacteriorhodopsin—Fourier transform difference spectra of photo-products of bacteriorhodopsin and of its pentamethyl analogue. *J. Mol. Struct.*, in press.
- Onsager, L. (1936). Electric moments of molecules in liquids. *J. Am. Chem. Soc.* 58, 1486–1493.
- Onsager, L. (1969). The motion of ions: Principles and concepts. *Science* 166, 1359–1364.
- Onsager, L. (1973). In *Physics and Chemistry of Ice* (Whalley, E., Jones, S.J., & Gold, L.W., Eds.), pp. 7–12. Royal Society of Canada, Ottawa.
- Otting, G., Liepinsh, E., & Wüthrich, K. (1991). Protein hydration in aqueous solution. *Science* 254, 974–980.
- Ovchinnikov, Y.A. (1981). Ion channels: Structure and function. *Biochem. Soc. Symp.* 46, 103–137.
- Pauling, L. (1960). *The Nature of the Chemical Bond*, 3rd Ed. Cornell University Press, Ithaca, New York.
- Pearl, L. & Blundell, T. (1984). The active site of aspartic proteinases. *FEBS Lett.* 174, 96–101.
- Perrin, C.L., Dwyer, T.J., Rebek, J., Jr., & Duff, R.J. (1990). H-bonding of primary amides. *J. Am. Chem. Soc.* 112, 3122–3123.
- Perutz, M.F. (1992). The significance of the hydrogen bond in physiology. In *The Chemical Bond Structure and Dynamics* (Zewail, A., Ed.), pp. 17–39. Academic Press, New York.
- Phillips, M.A., Kaplan, A.P., Rutter, W.J., & Bartlett, P.A. (1992). Transition-state characterization: A new approach combining inhibitor analogues and variation in enzyme structure. *Biochemistry* 31, 959–963.
- Radhakrishnan, R., Presta, L.G., Meyer, E.F., & Wildonger, R. (1987). Crystal structures of the complex of porcine pancreatic elastase with two valine-derived benzoxazinone inhibitors. *J. Mol. Biol.* 198, 417–424.
- Rashin, A.A., Iofin, M., & Honig, B. (1986). Internal cavities and buried waters in globular proteins. *Biochemistry* 25, 3619–3625.
- Reynolds, R., Remington, S.J., Weaver, L., Fischer, R., Anderson, W., Ammon, H., & Matthews, B. (1985). Structure of a serine protease from rat mast cells determined from twinned crystals by isomorphous and molecular replacement. *Acta Crystallogr. B* 41, 139–147.
- Riordan, J.F. (1973). Functional arginyl residues in carboxypeptidase a. Modification with butanedione. *Biochemistry* 20, 3915–3923.
- Rose, I.A., Fung, W.-J., & Warms, J.V.B. (1990). Proton diffusion in the active site of triosephosphate isomerase. *Biochemistry* 29, 4312–4317.
- Rose, I.A. & Kuo, D.J. (1989). The substrate proton of the pyruvate kinase reaction. *Biochemistry* 28, 9579–9585.
- Rozwarski, D. & Hackert, M. (1991). Refinement and analysis of three inhibitor complexes of His decarboxylase from *Lactobacillus* 30A. Abstract PA10. *Am. Crystallogr. Assoc.* 19, 74.
- Runnels, L.K. (1966). Ice. *Sci. Am.* 215(6), 118–126.
- Sawyer, L., Shotton, D.M., Campbell, J.W., Wendell, P.L., Muirhead, H., Watson, H.C., Diamond, R., & Ladner, R.C. (1978). The atomic structure of crystalline porcine pancreatic elastase at 2.5 Å resolution. Comparisons with the structure of α -chymotrypsin. *J. Mol. Biol.* 118, 137–208.
- Schechter, I. & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* 27, 157–162.
- Scheiner, S. (1988). Relationship between the angular characteristics of a hydrogen bond and the energetics of proton transfer. *J. Mol. Struct.* 177, 79–91.
- Scheiner, S. & Das, T. (1988). Effects of external ions upon proton-transfer reactions: Hydrogen-bonded systems containing formic acid. *Int. J. Quantum Chem. Quantum Biol. Symp.* 15, 137–147.
- Scholten, J.D., Hogg, J.L., & Raushel, F.M. (1988). Methyl chymotrypsin catalyzed hydrolysis of specific substrate esters indicate multiple proton catalysis is possible with a modified charge relay triad. *J. Am. Chem. Soc.* 110, 8246–8247.
- Schowen, R.L. (1988). Structural and energetic aspects of proteolytic catalysis in enzymes. Charge relay catalysis and function of serine hydrolysis. In *Mechanistic Principles of Enzyme Activity*, pp. 119–168. VCH Publishers, New York.
- Schowen, R.L. & Venkatasubban, K.S. (1984). Mechanistic principles of enzyme activity. *CRC Crit. Rev. Biochem.* 17, 1–44.
- Seeholzer, W. H., Jaworowski, A., & Rose, I.A. (1991). Enolpyruvate: Chemical determination as a pyruvate kinase intermediate. *Biochemistry* 30, 727–732.
- Sprang, S., Standing, T., Fletterick, R.J., Stroud, R.M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W.J., & Craik, C.S. (1987). The three-dimensional structure of Asn 102 mutant of trypsin: Role of Asp 102 in serine protease catalysis. *Science* 237, 905–909.
- Subramaniam, S., Marti, T., Rösselet, S.J., Rothchild, K.J., & Khorana, H.G. (1991). The reaction of hydroxylamine with bacteriorhodopsin studied with mutants that have altered photocycles: Selective reactivity of different photointermediates. *Proc. Natl. Acad. Sci. USA* 88, 2583–2587.
- Suguna, K., Bott, R.R., Padlan, E.A., Subramanian, E., Sheriff, S., Cohen, G.H., & Davies, D.R. (1987). Structure and refinement at 1.8 Å resolution of the aspartic proteinase from *Rhizopus chinensis*. *J. Mol. Biol.* 196, 877–900.
- Sussman, J.L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Tokar, L., & Silman, I. (1991). Atomic structure of acetylcholinesterase from *Torpedo californica*: A prototypic acetylcholine-binding protein. *Science* 253, 872–879.
- Swanson, S.M., Wesolowski, T., Geller, M., & Meyer, E.F. (1989). Animation: A useful tool for protein molecular dynamicists, applied to hydrogen bonds in the active site of elastase. *J. Mol. Graph.* 7, 223–224, 240–242.
- Takahashi, L.H., Radhakrishnan, R., Rosenfield, R.E., & Meyer, E. (1989). Crystallographic analysis of the inhibition of porcine pancreatic elastase (PPE) by a peptidyl boronic acid. *Biochemistry* 28, 7610–7617.
- Teeter, M.M. (1991). Water-protein interactions: Theory and experiment. *Annu. Rev. Biophys. Biophys. Chem.* 20, 577–600.
- Tsukada, H. & Blow, D.S. (1985). Structure of α -chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol.* 184, 703–711.
- Vedani, A. & Huhta, D.W. (1991). An algorithm for the systematic solvation of proteins based on the directionality of hydrogen bonds. *J. Am. Chem. Soc.* 113, 4860–4862.
- Venkatasubban, K.S. & Silverman, D.N. (1980). Carbon dioxide hydration activity of carbonic anhydrase in mixtures of water and deuterium oxide. *Biochemistry* 19, 4984–4989.
- Waks, M. (1986). Proteins and peptides in water-restricted environments. *Proteins Struct. Funct. Genet.* 1, 4–15.
- Warshel, A., Naray-Szabo, G., Sussman, F., & Hwang, J.-K. (1989). How do serine proteases really work? *Biochemistry* 28, 3629–3637.
- Weissbuch, I., Addadi, L., Lahav, M., & Leiserowitz, L. (1991). Molecular recognition at crystal interfaces. *Science* 253, 637–645.

- Williams, R.J.P. (1961). Possible functions of chains of catalysts. *J. Theor. Biol.* 1, 1–17.
- Williams, R.J.P. (1988). Proton circuits in biological energy interconversions. *Annu. Rev. Biophys. Biophys. Chem.* 17, 71–97.
- Winderickx, J., Lindsey, D.T., Sanocki, E., Teller, D.Y., Motulsky, A.G., & Deeb, S.S. (1992). Polymorphism in red photopigment underlies variation in colour matching. *Nature* 356, 431–433.
- Wlodawer, A., Borkakoti, D.S., Moss, D.S., & Howlin, B. (1986). Comparison of two independently refined models of ribonuclease A. *Acta Crystallogr. B* 42, 379–387.
- Wlodawer, A., Walter, J., Huber, R., & Sjölin, L. (1984). Structure of bovine pancreatic trypsin inhibitor results of joint neutron and x-ray refinement of crystal form II. *J. Mol. Biol.* 180, 301–329.
- Wong, C.-H. (1989). Enzymatic catalysis in organic synthesis. *Science* 244, 1145–1152.
- Wright, C.S. (1972). Comparison of the active site stereochemistry and substrate conformation in α -chymotrypsin and subtilisin BPN. *J. Mol. Biol.* 67, 151–163.
- Wright, C.S., Alden, R.A., & Kraut, J. (1969). Structure of subtilisin BPN' at 2.5 Å resolution. *Nature* 221, 235–242.
- Zundel, G. (1988). Hydrogen-bonded systems as proton wires formed by side chains of proteins and by wide chains and phosphates. In *Transport Through Membranes: Carriers, Channels and Pumps* (Pullman, A., et al., Eds.), pp. 409–420. Kluwer Academic Publishers, New York.
- Zundel, G. (1992). Proton polarizability and proton transfer processes in hydrogen bonds and cation polarizabilities of other cation bonds—Their importance to understand molecular processes in electrochemistry and biology. In *Trends in Physical Sciences* (Menon, J., Ed.).
- Zundel, G. & Brzezinski, B. (1991). Proton polarizability of hydrogen bonded systems due to collective proton motion—With a remark to the proton pathways in bacteriorhodopsin. In *Proton Transfer in Hydrogen-Bonded Systems* (Bonnfis, T., Ed.). Plenum, London, in press.
- Zundel, G. & Eckert, M. (1989). IR continua of hydrogen bonds and hydrogen-bonded systems, calculated proton polarizabilities and line spectra. *J. Mol. Struct. (Theochem.)* 200, 73–92.
- Zundel, G. & Metzger, H. (1968). Energiebänder der tunnelender Überschuss-Protonen in flüssigen Säuren. Eine IR-spektroskopische Untersuchung der Natur der Gruppierungen H_3O_2^+ . *Z. Phys. Chem.* 58, 225–245.