The purpose of this article is to show, using accurate illustrations, how the formation of transient linear elements of hydration in cubic conformations on the lipid surfaces of natural polypeptides drive and may assist in directing spontaneous wrapping and assembly into functional proteins.

Although great progress has been achieved in understanding how polypeptides wrap and assemble into proteins, transitions between intermediates are so rapid that it has been impossible to detect the role of water.\(^1\) Water is accepted as being involved in bridging between polar and ionic groups and in escorting them into more stable positions but not in directing them into ordered spatial positions. However, in 1944, Erin Schrödinger, in his little book *What is Life?*, proposed that water must have expressed some sort of order on natural molecules as they formed in the earliest phases of evolution to move them from randomness toward order - the opposite direction from that ordained by the Second Law of Thermodynamics.\(^2\) Since the movement of polypeptides into proteins involves same transition, the same elements of order in water must be involved.\(^3\) At the time of Schrödinger’s proposal, liquid water was considered to be a random medium with little or no structuring properties.\(^4\) However, the past ten years have brought forth dramatic changes in our understanding of its spatial-structuring properties.

As early as 1972, X-ray diffraction from the surface of liquid water revealed that most molecules are in random positions 2.9 Angstroms apart, but that two peaks, corresponding to more ridgedly-bonded trimers and tetramers are 2.76 Å apart.\(^5\) Molecular orbital calculations indicated that a trimer of that length might be the most stable ordered unit in liquid water but that the level would be so low and it would be insignificant.\(^6\) It was not until 2006 that Dr. Calder at the Stanford University Linear Accelerator Center reported that irradiation of pure liquid water at 25° C with high-energy neutrons revealed the presence of the trimer with that type of ice-like bonding. In 1999, bonding in ice had been shown to be covalent (like that of bonds between carbon atoms),\(^7\) but with the bonding electrons orbiting a central proton.\(^8\)

Thus, bonds in ice and in ordered-forms of water are fundamentally different from those between liquid molecules. Not only the bonding energy is different (2.25 kcal/mole lower than in liquid water) but the shared proton, by moving only a fraction of an Angstrom, moves the positive charge from one molecule to the next, produces a cation and an anion and permits positive charges to tunnel extremely rapidly through covalently-bonded linear elements.\(^9\)
Just as molecular orbital calculations in 1970 had forecast that the trimer would be in liquid water, calculations in 1984 indicated that linear elements of as many as six water molecules might bond together on non-hydrogen-bonding lipid surfaces.\textsuperscript{10} But, it was not until 2003 that Professor Zewail and his group at Caltech, using 4D ultra-high-speed electron crystallography, found that water, on the non-hydrogen-bonding surface of graphite and a poly-ionic surface, forms multiple layers of linear elements of hydration in cubic ice conformations.\textsuperscript{11}

As expected, the molecules are 2.76 Angstroms apart and last only about $10^{-10}$ seconds.\textsuperscript{12} The surfaces were chosen because both have surface atoms in hexagonal positions like the ends of hydrocarbon chains in lipid surfaces. Although oils and natural molecules, like fatty acids and proteins, have hexagonal patterning on their surfaces, they are too dynamic to obtain interpretable crystallographic patterns.\textsuperscript{13}

If pure water is in contact with hexagonal surfaces at 0°C, it begins crystallizing immediately. But, if it is in contact with a surface with atoms in random or regular pentagonal positions, cooling can be continued down to -30°C without freezing.\textsuperscript{14} The viscosity increases but crystallization does not occur until -40°C.\textsuperscript{15} But the ice produced is not normal - it is the cubic form, which is unstable and, on warming, rapidly isomerizes to the normal “hexagonal” form with hexagonal units above each other.\textsuperscript{16}

Thus, as water freezes, it initially produces a conformation of water molecules containing only linear elements. It is called the “kinetic” product because it is the form produced as the orbitals of adjacent water molecules overlap to form the bonds.\textsuperscript{17} Thus, whether covalent bond formation occurs at 0°C or 25°C, the same linear elements in cubic conformations will be produced initially. If formed at 0°C, ice will form - initially cubic - then hexagonal. If formed at 25°C on a lipid surface, water molecules will (within about $10^{-10}$ seconds) begin leaving the surface, absorbing quantized units of energy from molecules in the surface and moving them into lower-energy, more-ordered forms.\textsuperscript{18}

Thus, as a polypeptide is released from a ribosome, surface water must rapidly form covalent linear elements of hydration on its hydrophobic lipid surfaces. Then, as water molecules leave and move into liquid-state, the surface left behind straightens and, if even more energy is energy is withdrawn, it forms an even lower energy coil.\textsuperscript{18} Small peptides in the polypeptide chain, like glycine and serine, which have no lipid side-chains to shield them, continually form dynamic point-charge bonds with surface water; exchanging energy to permit mobility in the chain. Although cubic structuring of water on ordering surfaces lasts only an instant, it is long enough to define preferred angles for bending and for lengths of chains and coils.\textsuperscript{3}

As polypeptides formed at random in the earliest phases of natural molecule formation, it appears to have been linear segments and coils of preferred lengths which folded together to produce the anhydrous cores of the proteins. Polypeptides which could not spontaneously assemble to produce stable functional forms, were resorbed by lytic enzymes.
**Insulin Assembly**

One of the first proteins to have its spatial structure revealed by X-ray crystallography was the hormone insulin. Prior to Dorothy Hodgkin’s report of the structure in 1971, it was known that insulin is produced as a single linear segment in beta cells of the pancreas and that several regions spontaneously wrap into coils, as A and B, and then assemble into the molecule as shown below. Segment C is extremely mobile and serves as a tether to guide A and B together. D contains a number of hydrocarbon peptides and serves to transport the insulin molecule through lipid membranes. C and D are cleaved enzymatically before the insulin molecule and the C protein are released into the blood-stream.

The A and B regions fold rapidly into coils and assemble so tightly together that the central core region of the molecule is completely anhydrous. However, as the molecules combine to form crystals, they carry water with them into channels which have been studied extensively because they permit the water to adopt cubic symmetry. In fact, the molecule is an ideal candidate to study based on the transient linear hydration hypothesis. As its polypeptide is released from a ribosome, the short linear covalent linear elements of hydration which form in cubic conformations on hydrocarbon surfaces, must immediately, as they leave, begin withdrawing energy from those surfaces to form linear segments and coils. Then, by losing even more covalently-bonded surface water, fit A behind B to form the anhydrous core of the molecule. Most of the surface is left covered with point-charge hydrogen-bonding polar and ionic groups to provide for dynamic hydrogen-bonding and water solubility.

Since outer surfaces of most water-soluble proteins are covered with small hydrogen-binding and ionic peptides which bind surface water in multiple orientations, linear-order in surface water is disrupted and they do not display exterior cubic patterning. However, insulin has a sufficient number of order-disrupting peptides to be soluble in water, but is a hormone which binds into receptor sites in membraneal proteins to regulate the uptake of glucose into cells. Thus, it must have an exterior structure which permits it to displace transiently-ordered water in binding sites. As illustrated above, it displays cubic patterning in multiple orientations. In fact, as will be pointed out on page 6, recent studies indicate that it is the flat planar lower right-hand face of the molecule, as displayed in the middle Front View above, which binds most tightly to the ordering surface of the receptor proteins.
**Insulin B-Chain**

It is not surprising that unstable covalent linear elements of hydration, which form rapidly on both sides of the hydrocarbon surfaces between glycines 8 and 20, absorb energy as individual water molecules leave and move peptides in that segment into a lower-energy coil. Remember, by moving from covalent to point-charge hydrogen bonding, even a trimer can absorb about 5 kcal/mole of quantized energy from a chain.

Although coil-formation reduces covalent bonding, hydration order continues to be induced on three sides of coil. With surface water bonding on almost the entire right-hand side of the coil and B chain, the upper section is folded over to the left into cubic hydration stabilized region to release additional ordered water and bring the aromatic ring of phenyl alanine 1 next to valine 18. Although ordered water is released on the left side of the coil, it still covers the hydrophobic back side.

With the unstable left side of the coil covered, the lower segment of the chain bends at glycines 20 and 23 and folds into position on the back side of the coil as shown in the Front and Top Views. However, the hydrophobic region shown by the shaded area and linear element of water in the Top View is so large that the terminal linear segment ending in alanine 30 may not fold into place behind the coil until the A-unit polypeptide is in place.
**Insulin A-Chain**

Instead of forming a single coil, the A chain forms two coils, each initiated by serines at positions 9 and 12. Another difference is that the chain contains relatively few peptides with hydrocarbon side chains - instead, peptides with alpha and beta methylenes shield the central chain from binding with surface water and force water above and below the chain to form unstable covalent hydrogen-bonding.

Just as the B chain had hydration-ordering side-chains on the left and disordering on the right, distribution is the same on the A coils. Thus, they would be expected to fold together rapidly to release all water between them to form the tight assembly shown on the upper right. Notice in the Front View that a disulfide forms between the two coils to tie them tightly together.

**Insulin C-Chain**

As you can see, the C chain is quite different from A and B. By including series of hydration order-disrupting glycines, asparagines and prolines, the chain is so dynamic that it can exist in numerous conformations and associations.
It is not difficult to believe that early polypeptide synthesis on crude ribosomes was by trial and error, but it is difficult to understand how those molecules could have assembled spontaneously to produce the living cell if they formed independently. Although some believe that assembly involved a “Plan,” it is possible that, once crude ribosomes were produced, that polypeptides, nucleic acids and regulator molecules may have been produced in such a simultaneous manner, using the limiting order of cubic hydration patterning for stability, that synthesis, assembly and selection occurred simultaneously and cooperatively.

For example, the insulin molecule binds to multiple receptor proteins within body to regulate a multitude of functions. As the A, B and C units were forming to satisfy the cubic patterning of external water, receptor proteins must have been forming with cavities filled by covalent linear elements of hydration in the same cubic configurations as those simulated by the insulin molecule. Of course, most proteins with such cavities did not satisfy requirements of charge distribution, stability and function - but some did.

Although the C chain is extremely flexible, five glycine peptides in the center of the C chain are in precisely the proper position to form a $180^\circ$ loop, assemble six hydrocarbon peptides together to form a linear complex with polar peptides on one side and two leucine peptides on the other side - the same length of surface covalent linear elements of hydration as on unit A. By flipping over, unit A can form a complex with C, as shown in the center above, and be carried in an integrated manner into B. The terminal segment of B then folds into the back of A, leaving C trailing off up above. The two circled prolines in C play a critical role by being able to rotate around the bond between them to permit multiple steps in assembly while pairs of cationic peptides on the ends of the C peptide provide sites for enzymatic cleavage and release of insulin and C.
Hormonal Activity

If we look closely at both Top and Front Views of the insulin molecule relative to the cubic lattice, you can see that its surfaces are covered with ionic and polar groups which hydrogen bond directly with surface water to disrupt hydration order and increase both stability and solubility. Notice that the layers in the cubic lattice have been numbered in the Front View so they can be identified in the hexagonal overlay in the Top View.

Although, the receptor-binding region of the insulin protein is not visible in the above orientation, if the molecule is rotated 30 degrees to the right, as shown below, the flat planar surface on the lower right of the Side View formed by the rings of phenylalanines 24 and 25 and the methyls of valine 12 has been identified as the surface which binds to glucose-uptake receptors. As you can see, it is a face which, in binding, would release linearly-ordered water from both surfaces to drive the binding forward.

Hopefully, receptor binding studies will continue to experimentally determine whether or not the role proposed above for surface water is valid or not. Also, it is hoped that this presentation of the possible role of surface water in the folding and assembly of insulin and carboxypeptidase A will stimulate those with sufficient computer capability to incorporate transient covalent linear elements of hydration into their mathematical simulations of folding, assembly and function to provide more accurate representations of the role of surface water.
References


The Author

Dr. Collins received his degrees in Chemistry from Wayne University and the University of Wisconsin. After employment at General Motors Research, Central Research at E. I. Dupont in Wilmington, Delaware and Sterling Winthrop Research Institute in Rensselaer, NY, he accepted a position as Associate Professor and Chairman of the Chemistry Department at Illinois Wesleyan University. In 1967, he returned to Sterling Winthrop to direct Medicinal Chemistry Research and then handle Technical Improvements for the Corporation until 1987 when he retired to devote full time to his study of the role of water in the living cell.

He has a number of publications and patents to his credit and, while at Illinois Wesleyan, developed a technique for selectively oxidizing primary alcohols to aldehydes which was subsequently named The Collins Reagent. However, it was during his first employment at Sterling Winthrop that he began constructing permanent models of hormone and neurotransmitter molecules and found that distances between polar atoms on the ends corresponded those of water molecules hydrogen-bonded together as they are in ice. Could it be that water, in orderly low-energy forms, occupies binding sites in receptor proteins when regulator molecules are not there? Since bonding between water molecules would be unstable in binding sites above 0°C, it would explain why water has never been detected there. However, when models of a number of water-soluble enzymatic proteins were examined, it was found that the polar and ionic atoms of peptide side-chains leading into catalytic binding sites were in positions to support transient linear elements of hydration.

Furthermore, when these enzymatic proteins were examined more closely, it was found that the geometries of their anhydrous cores corresponded to units of water molecules in cubic ice. Since cubic ice is composed entirely of linear elements of water molecules and is the initial form produced on lipid surfaces as water freezes, it might well be that they form as unstable linear elements in cubic forms on the lipid surfaces of polypeptides as they are released from ribosomes. On the other hand, small peptides which hydrogen-bond directly with surface water are in precise positions in polypeptides to provide for bends and turns. While regions of polypeptides which form adjacent linear elements of hydration lose water and produce anhydrous cores, regions which hydrogen-bond directly with surface water end up on the surface, increasing the solubility and stability of finished proteins. Only small surface regions continue to induce hydration order to provide sites for binding other proteins, regulator molecules and substrates for reactions. These were the concepts presented in 1991 in my first book, The Matrix of Life,

Although two more books and a number of web sites were published after that, only two comments were received regarding the concepts. One was from Linus Pauling who received a copy of the Matrix book. “You are on the right track, but I think your concepts are too simple.” The other was from Dr. Michael New, a lead investigator at NASA who was asked to review a preprint of “Biomolecular Evolution from Water to the Molecules of Life.” His comment was: “Your concepts of Transient Linear Hydration and Cubic Hydration Patterning are valuable contributions to our understanding the unique role of water in origin of life research.” What do you think? Send your opinion to jcbjh2o@aol.com.