MOLECULAR PRESENTATIONS

CUBIC HYDRATION OF INSULIN AND ITS POLYPEPTIDE

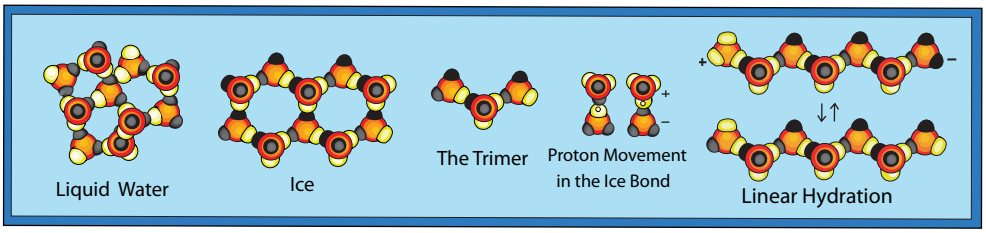
J. C. Collins, PhD

Dedicated to the late Professors Carl Djerassi and William S. Johnson of Stanford University.

The purpose of this article is to show, using accurate illustrations, how the formation of transient linear elements of hydration in cubic conformations on lipid surfaces of the insulin polypeptide may assist in directing and driving the assembly and functions of proteins.

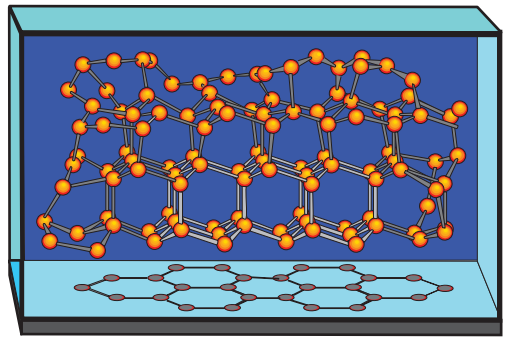
Although great progress has been achieved in understanding how polypeptides wrap and assemble into proteins, transitions between intermediates are so rapid that it has not been possible to detect the role of water.[1] Although water is accepted as being involved in bridging between polar and ionic groups and in escorting them into more stable positions, it is not believed to be involved in directing them into ordered spatial positions. In 1944, Erin Schrödinger, in his little book *What is Life?,* proposed that water must have expressed some sort of spatial-order on natural molecules as they formed in the earliest phases of evolution, to have 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 today involves the same transitions, 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 structure.[4] However, the past twenty years have brought forth significant changes in our understanding of the structuring properties of liquid.

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 minor peaks corresponding to more ridged ice-like trimers and tetramers are also present with oxygens 2.76 Å apart,.[5] In 1969, molecular orbital calculations suggested that a trimer, with that bond-length, might be the most stable ordered unit in liquid water,[6] but it was not until 2006 that Dr. Calder, at the Stanford University Linear Accelerator Center, reported that the trimer had been detected in pure water at 25°C.[7]



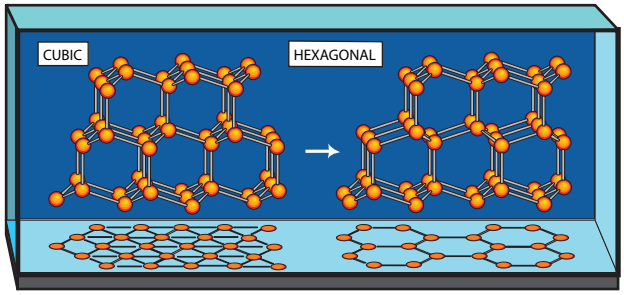
As illustrated above, liquid bonding is extremely dynamic, while X-ray analysis indicates that bonding in ice is covalent, with electron orbitals of adjacent water molecules overlapping in in the center and the proton close to the overlap. Not only is the bonding energy 1.5kcal lower in ice, but the proton, by moving only a fraction of an Angstrom, can charge adjacent water molecules.[8] If the molecules separate, ions are produced, but, if positive charges are produced in linear elements, as those shown above, like those that form on inner surfaces of nerve fibers, it can cascade (tunnel) through end-to-end at extremely high speeds.[9]

Just as molecular orbital calculations had forecast that the trimer would be in liquid water, in 1984 calculations suggested that linear elements of as many as six water molecules might bond together on non-hydrogen-bonding lipid surfaces.[10] But, it was not until 2009 that Professor Zewail and his group at Caltech, using 4D ultra-high-speed electron crystallography, found that water, at subzero temperatures on the non-hydrogen-bonding lipid surface of graphite and a poly-ionic surface, forms multiple layers of linear elements of hydration in cubic ice conformations.[11]



As expected, the molecules are 2.76 Angstroms apart and last about 10^{-11} seconds. The surfaces were chosen because both have surface atoms in hexagonal positons to seed ice-formation and are solid. Surfaces of proteins and nucleic acids also seed ice-formation but their surfaces are too dynamic to permit visuallization.[13]

If pure water is in contact with hexagonal surfaces at 0°C, crystallization seeding is immediate, but, if it is in contact with a surface were atoms are in random or regular pentagonal positions, cooling can be continued down to -30°C without freezing.[14] Viscosity increases,[15] but crystallization does not occur until -40°C.[14] But the ice produced is cubic: a form which also occurs first at 0 C, but reverts to the normal hexagonal form so rapidly that it cannot be detected.[16]

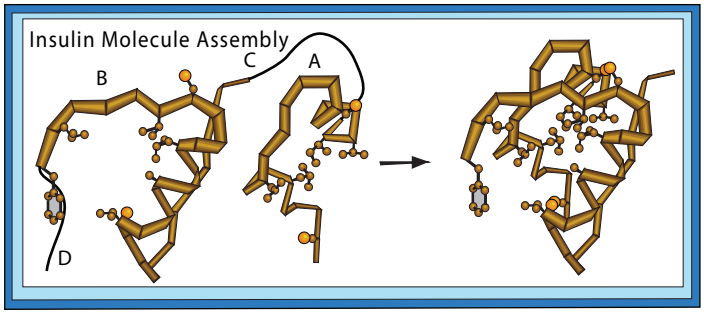


Thus, as water freezes, it initially produces a cubic conformation of water molecules containing only linear elements. It is called the "kinetic product" because it is the form produced as the orbitals of adjacent water molecules overlap to form the bond.[17] Whether the ice-bond forms at 0°C or 25°C, the same linear elements in cubic conformations are produced initially. However, if formed at 25° C on a lipid surface or between ions, water molecules (within 10^{-10} seconds) begin leaving the surface, absorbing quantized units of energy from adjacent surfaces and moving molecules within them into lower-energy, more-ordered forms.[18]

Thus, as a polypeptide is released from a ribosome, surface water must rapidly form low-energy linear elements of hydration on its hydrophobic lipid surfaces, and then, as the water leaves and moves into liquid-state, surfaces left behind straighten and, if enough energy is withdrawn, form even lower-energy coils![19] Small peptides in the polypeptide chain, like glycine and serine, which have no side-chains, continually form dynamic hydrogen bonds and absorb energy from surface water to permit mobility in the chain. Although linear and cubic structuring of water on lipid surfaces lasts only an instant, it is long enough to define preferred angles of bending and lengths for combining with other chains and coils.[3]

As polypeptides formed at random in the earliest phases of natural molecule formation, those with lipid regions of preferred hydration lengths that could folded together, release ordered water and produce anhydrous cores, survived. Those which could not spontaneously assemble to produce stable functional forms, were resorbed by lytic enzymes.

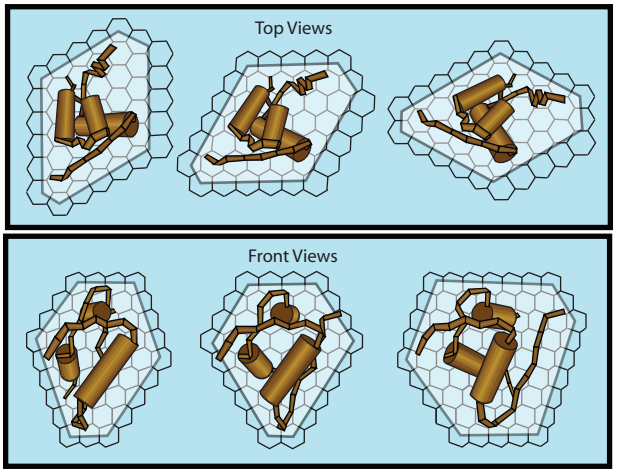
One of the first proteins to have its spatial structure revealed by X-ray crystallography was the hormone, insulin.[20] Prior to Dorothy Hodgkin's report of the structure in 1971, it was known that insulin is produced as a single linear segment in beta cells of the pancrease and that several regions spontaneosly 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 hydro-carbon peptides that serve 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.



As its polypeptide is released from a ribosome, short linear ice-like linear elements of hydration which form in cubic conformations on lipid surfaces,[18] must immediately begin leaving those surfaces and, as they leave, remove energy from units A and B and move them into lower energy linear and coil forms.[3] As ice-like elements form on those new surfaces, they assemble and then fit together to produce the anhydrous core of the molecule with the hydrocarbon side-chains packed tightly together. Most of the surface is covered by polar and ionic groups which dynamically hydrogen bond with surface water to provide solubility and stability. However, as the molecules combine into crystals, they carry water with them in linear layers in cubic forms, just as they were when they began to assemble.[3]

Since outer surfaces of most water-soluble proteins are covered with small hydrogen-bonding 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.

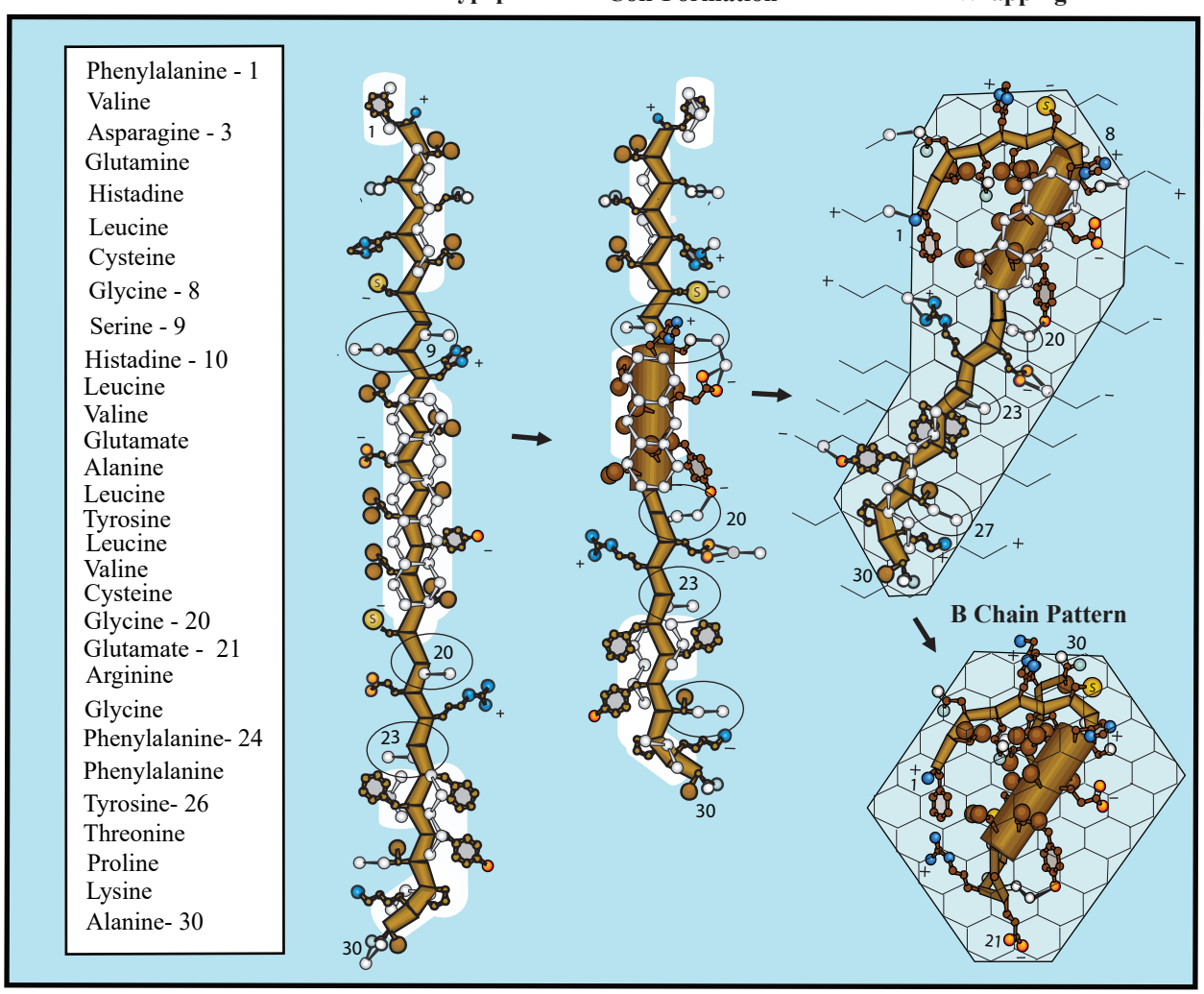
However, insulin has a sufficient number of order-disrupting peptides to be soluble in water, but, as a hormone which binds into receptor sites in membranal proteins to regulate the up-take of glucose into cells, it has a spatial struc ture which permits it to displace transiently-ordered cubic water from binding sites. As illustrated on the right, it displays cubic patterning in multiple orientations. In fact, recent studies indicate that it is the linear lower right-hand face of the molecule as displayed in the middle Front View which binds most tightly to the receptor site.[21]



Although it is difficult to portray three-dimensional structures in two dimensions, an attempt will made to present the assembly of insulin proyein in a manner that will provide a more detail view of the role of water.

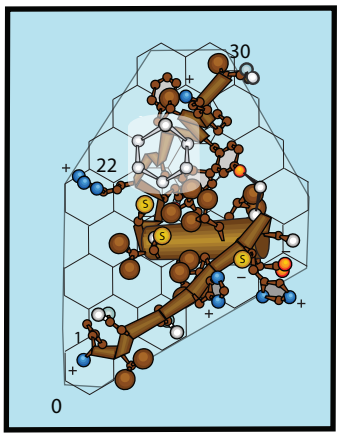
Insulin B-Chain Folding

In the schematic figure below, it should not be surprising that the unstable covalent linear elements of hydration which form rapidly on both lipid sides of the polypeptide chain between glycines 8 and 20, absorb enough energy as they leave, to produce the lower-energy coil.[19] By moving into dynamic hydrogen bonding, about 2.5 kcal/mole of quantized energy is removed from a chain as each covalent bond in linear hyration is lost.[3]



The third figure illustrates how energy flows into the small circled glycine and serine at 8 and 9 to turn the chain and follow hexagonal patterning over to position the aromatic ring of phenyl alanine next to the methyl groups of valine 18 on the chain. Extensions of linear elements are shown extending out from the molecule to provide ties to adjacent molecules. Although the front side of the upper section is covered with order-disrupting polar groups, the back side of the coil and the back of the upper section are lipophilic. Glycines at 20 and 23 then permit turns so that the lipid side of the lower chain folds into the lipid back of the coil.

Top View



On the other hand, the shaded area in the top view on the right is so large that the segment ending in alanine 30 may not fold up behind the coil until the A-unit is in place.

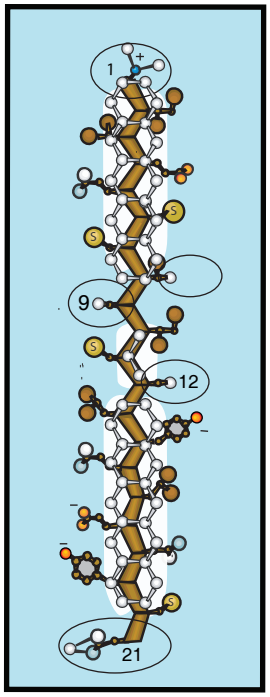
Insulin A-Chain

Instead of forming a single coil, the A chain forms two coils, each initiated by serines at positions 9 and 12. Another difference in A is that the chain contains relatively few peptides with hydrocarbon side chains - instead, peptides with alpha and beta CH_2 methylenes shield the central chain from binding with surface water and force water above and below the chain to form unstable low-energy covalent-bonding.

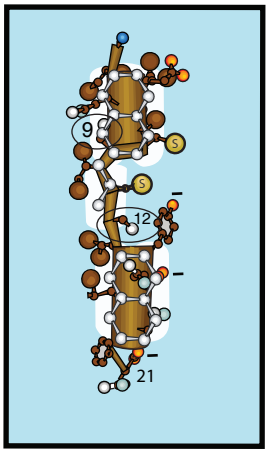
A Chain

Glycine -1
Isoleucine
Valine
Glutamic Acid
Glutamine
Cysteine
Cysteine
Threonine
Serine- 9
Isoleucine
Cysteine
Serine - 12
Leucine
Tyrosine
Glutamine
Leucine
Glutamic Acid
Asparagine
Tyrosine
Cysteine
Asparagine - 21

Initial A Polypeptide

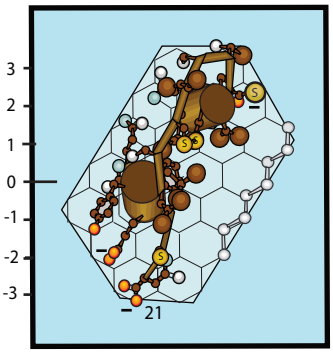


Coil-Formation

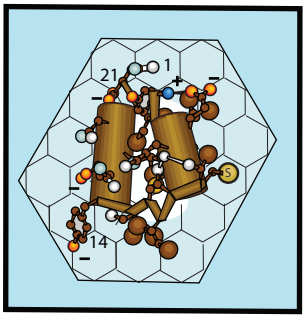


By bending at serine 12 and cysteine 11, the lower coil can fold up next to the upper coil, release unstable water and form the unit on the right with more unstable water on the right

Front View



Assembly - Top View



With unstable ordering water on the left on the B chain and unstable water on the right on the A coil unit (as shown on rhe right), the two units are prepared to fold together, release the water between them and form the tight anhydrous core-unit assembly. Notice in the Front View that a disulfide forms between the two coils to tie them together.

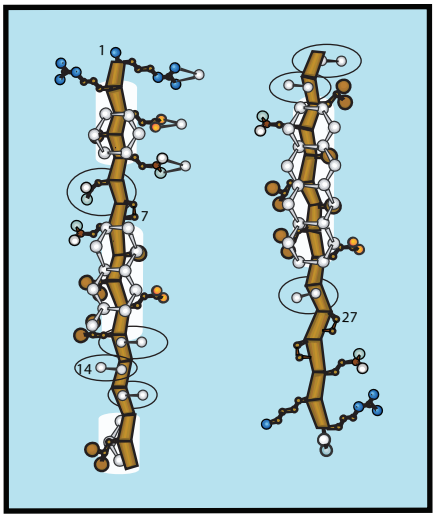
Insulin C-Chain

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 assemblies and, as shown below, bring A into B.

C-Chain

Arginine- 1
Arginine
Glutamic Acid
Alanine
Glutamine
Asparagine
Proline - 7
Glutamine
Alanine
Valine
Glutamic Acid
Leucine
Glycine
Glycine -14
Glycine
Leucine
Glycine
Glycine
Leucine
Glutamine
Alanine
Leucine
Alanine
Leucine
Glutamic Acid
Glycine
Proline - 27
Proline
Glutamine
Lysine
Arginine

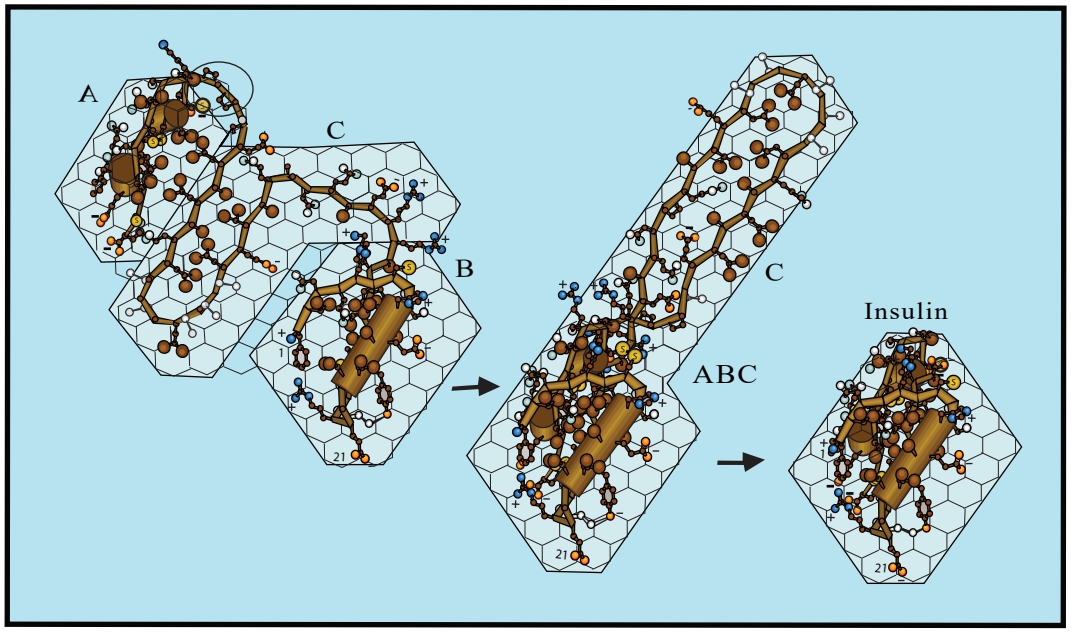
C-Polypeptide



Chain C and Assembly

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, 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 the body to regulate a multitude of functions. Thus, as the A, B and C units were forming to satisfy cubic patterning of external water, receptor proteins must have been forming with cavities filled by linear elements of hydration in the same cubic configurations as those simulated by the insulin molecule. Although most small molecules did not satisfy the spatial and charge-distributions of ordered water in open spaces in proteins, some did. In the figures shown below, the A and B units are shown closely-tied together, with C guiding A into the back of B. Initially, the units may have been significantly far apart connected by hydrated linear segments of C.

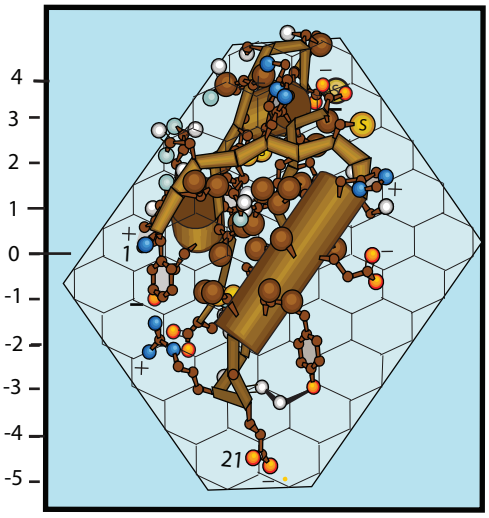


Although the C chain is extremely flexible, five glycine peptides in the center of the chain are in precisely the proper position to form a 180° loop assemble of six hydrocarbon peptides together to form a linear complex with polar peptides on one side and two leucine peptides on the other side. Since the lipid side is same length as the lipid side of the A unit, by flipping over, the complex shown in the center can be formed to carried the A unit in an integrated manner to B to slide into place. As the terminal segment of B folds into the back of A, a hydrated loop of C is left trailing 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 from C.

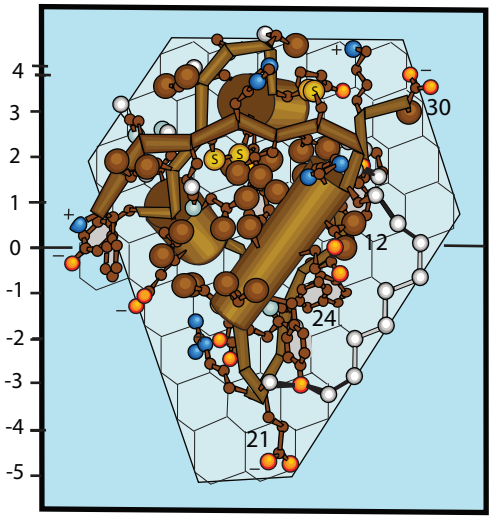
Hormonal Activity

If we look closely at the Side View of the insulin molecule, relative to the cubic lattice, we can see that its surfaces are covered with ionic and polar groups which hydrogen-bond directly with surface water disrupting hydration order and increasing both stability and solubility.

Side View



Front View



Although, the receptor-binding region of the insulin protein is not visible in the left illustration, by rotating the molecule 30 degrees to the right, as shown on the right, the flat planar surface on the lower right, with the rings of phenylalanines 24 and 25 forming a nonpolar center, provide a linear surface for binding to glucose-uptake receptor protein sites. The are sites which are occupied by the linear elements of hydration and release as binding occurs.[3]

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 four other proteins 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 way molecules and ions interact in the living cell.

References

1. K. A. Dill and J. L. MacCallum, *Science* **338**: 1042-1046 (2012). The Protein-Folding Problem, 50 Years On. See also, D. Chandler, *Nature* **437**: 640 (2005). Interfaces and the driving force of hydrophobic assembly. Also, B.J. Berne, J. D. Weeks and R. Zhou, *Ann. Rev. Phy. Chem.* **60**: 85-103 (2009). Dewetting and Hydrophobic Interaction in Physical and Biological Systems. Also, R. L. Baldwin and Rose, *Proc. Natl. Acad. Sci. USA* **113** (44): 12462-1466 (2016). How the hydrophobic factor drives protein folding.

2. E. Schrödinger, *What is Life?*, Cambridge University Press (1944). *What is Life? with Mind and Matter*. (Cambridge University Press 1967).

See also, R. Lumry and S. Rajender, Biopolymers **9**: 1125-1227(1970). Entropy-Enthalpy Compesation Phenomena in Water Solutions of proteins and Small Molecules: A Ubiquitous Property of Water.

3. J. C. Collins, *The Matrix of Life, Second Edition* (Molecular Presentations Inc., 2023). www.cubichydration.com.

4. F. Franks (ed.) *Water- A Comprehensive Treatise.* (Plenum Press, 1972). See also D. Eisenberg and W. Kauzmann, *Structure and Properties of Water* (Oxford University Press, 1969).

References

5. A. H. Narten and H. A. Levy, *Water - A Comprehensive Treatise* pp. 311-332 (Plenum, 1972). Surface of Liquid water: Scattering of X-rays.

6. J. R. Hoyland and L. B. Kier, *Theor. Chim. Acta.* **15**: 1-11 (1969). Molecular orbital calculations for hydrogen-bonded forms of water. See also, J. Del Bene and J. A. Pople, *J. Chem. Phys.* **52:** 48-61 (1970). Theory of Molecular Interactions. Molecular Orbital Studies of Water.

7. Y. Zubavicus and M. Grunze, *Science* **304**: 974-976 (2004). New Insights into the Structure of Water with Ultrafast Probes. See also, N. Calder, www.slac.stanford.edu/structureofwater.html.

8. E. D. Isaacs, et al., *Physical Rev. Letters* **82**(3): 600 (1999). Covalency of the Hydrogen Bond in ice: A Direct X-Ray Measurement. See also, M. W. Fryereisen, D. Feller and D. A. Dixon, *J. Phys. Chem.* **100**(8): 2993-2997 (1996). Hyrogen Hydrogen Bond Energy of the Water Dimer.

9. A. J. Horsewell, *Science* **291**: 100 (2001). Evidence for coherent proton tunnelling in a hydrogen bond network. Also, J. Lin, H. A. Balabin and D. H. Beratan, *Science* **310**: 1311 (2005). The Nature of Aqueous Tunnelling Pathways. Also, U. S. Raghavender, et al., *J. Phys. Chem. B.* **115**(29): 9236-9243 (2011). Entrapment of a Water Wire in a Hydrophobic Peptide Channel.

10. C. Y. Lee, J. A. McCammon and P. J. Rossky, *J. Chem. Phys.* **80**(9): 4448 (1984). The structure of liquid water at extended hydrophobic surfaces.

11. D-S. Yang and A. H. Zewail, *Proc. Natl. Acad. Sci. USA* **106**(11): 4122-4126 (2009). Ordered water structure at hydrophobic graphite interfaces observed by 4D ultrafast electron crystallography. See also, C-Y. Ruan, V. A. Lobastov, F. Vigliotti, S. Chen and A. H. Zewail. *Science* **304**: 80-84 (2004). Ultrafast Electron Crystallography of Interfacial Water

12. A. Tokmakoff, *Science* **317**: 54-55 (2007). Shining light on the rapidly-evolving structure of water.

13. J. L. Ranck, L. Mateu, D. M. Sadler, A. Tardieu, T. Gulik-Krzywicki, and V. Luzzati, *J. Mol. Biol.* **85**: 249 (1974). Order-disorder conformational transitions of hydrocarbon chains of lipids.

14. A. K. Soper, *Science* **297**: 1288 (2002). Water and Ice. See also, T. Sun, et al., *Science* **343**: 795-798 (2014). An Antifreeze Protein Folds with an Interior (Pentagonal) Net-work of More Than 400 Semi-Clatherate Waters.

15. H. E. Stanley et al., *J. Phys. Condens. Matter* **21** (50): 504105-504118 (2009). Heterogenities in Confined Water and Protein Hydration Water. Also, P. Gallo et al., *Chemical Reviews* **116**: 7463-7500 (2016) Water: A Tale of Two Liquids.

16. E. Mayer and A. Hallbrucker, *Nature* **325**: 601 (1987) Cubic ice from liquid water. See also, B. Kamb, *Structural Chemistry and Molecular Biology* pp. 507-542. (Freeman,1968). Ice polymorphism and the structure of water.

17. P. Sykes, *A Guide to Mechanisms in Organic Chemistry* (Pearson Prentice Hall, 1986). See also, M. A. Fox and J. K. Whitsell, *Organic Chemistry*, 3rd. Ed. (Jones and Bartlett, 2004).

18. F. Mallamace, C. Cosaro, D. Mallamace, P. Baglion, H.E. Stanley and Sow-Hsin Chen, J. Phys. Chem. B 115(48): 14280-14294(2011). A Possible Role of Water in the Protein Folding Process. See also, J. D. Weeks et al., *Ann. Rev. Phys. Chem.* **60**: 85-103 (2009). Dewetting and Hydrophobic Interaction in Physical and Biological Systems.

19. K. Lindorff-Larsen, P. Stefano, R. O. Dror and D. E. Shaw, *Science* **334:** 517-520 (2011). How Fast-Folding Proteins Fold.

20. T. L. Blundell, J. F. Cutfield, S. M. Cutfield, E. K. Dodson, G. G. Dodson, G. G. Hodgkin, D. A. Mercola and M. Vijayan, *Nature* **231**: 506 (1971). Atomic positions in rhombahedral 2-zinc insulin crystals. See also, F. Sanger and E. O. Thompson, *The Biochemical Journal* **53**(3): 366-367 (1953). The amino-acid sequence in the glycyl chain of insulin.

21. J. Badger and D. L. D. Casper, *Proc. Natl. Acad. Sci. USA* **88**: 622-626 (1991) Biophysics. Water structure in cubic insulin crystals.

22. C. W. Ward and M. C. Lawrance, *BioEssays* **31**(4): 422 (2009). Ligand-Induced activation of the insulin receptor.

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 Departmnt 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^{o}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. Micharel 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 Hyhdration and Cubic Hydration Patterning are valuable contributions to our understanding the unique role of water in origin of life research." Check out molecularcreation.com and cubichydration,com for more information.