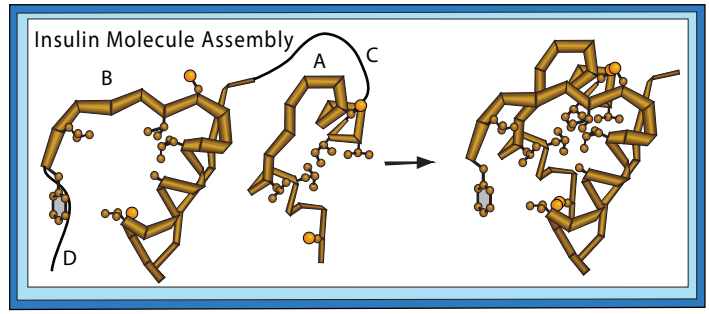
MOLECULAR PRESENTATIONS

CUBIC HYDRATION OF INSULIN AND ITS POLYPEPTIDE

J. C. Collins, PhD

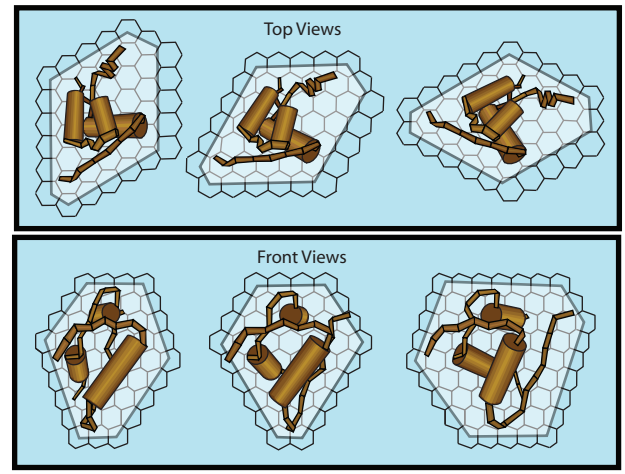
Dedicated to the late Professors Carl Djerassi and William S. Johnson of Stanford University.

One of the first proteins to have its spatial structure revealed by X-ray crystallography was the hormone, insulin.[1] Prior to Dorothy Hodgkin's report of the structure in 1971, it was known that insulin is produced as a single linear polypeptide in beta cells of the pancrease and that several regions spontaneosly wrap into coils (as shown in 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, unstable linear and cubic units of hydration must begin forming rapidly on lipid surfaces,[2] and then on leaving, withdrawing quantized units of energy from those regions. By forming on both sides of the central section of B and two short regions in A, enough energy is withdrawn to convert those regions rapidly into coils.[3] Since the back side of B and the front side of A are lipid and covered with water in unstable linear elements, as the water leaves, segment C guides the two exposed surfaces together to form the insulin structure with no water trapped inside, but with enough bound by dynamic liquid-bonding to oxygen and nitrogen atoms on the surface to permit it to be soluble in water. Enzymatic cleavage of C and D release the molecule into the blood stream. Only the linear surface on the lower right-hand side of the molecule is sufficiently lipophilic to continue to form adjacent transient linear elements of hydration.

Although most water-soluble proteins have so many hydration-order-disrupting peptides on their surfaces, that it is impossible to see the cubic patterning which directed the formation of the molecule, insulin is a hormone which binds into receptor sites in membranal proteins to regulate the up-take of glucose into cells. It has a spatial structure 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 lower linear right-hand face of the molecule, as displayed in the middle Front View on the right, which binds most tightly to the receptor site.[4]

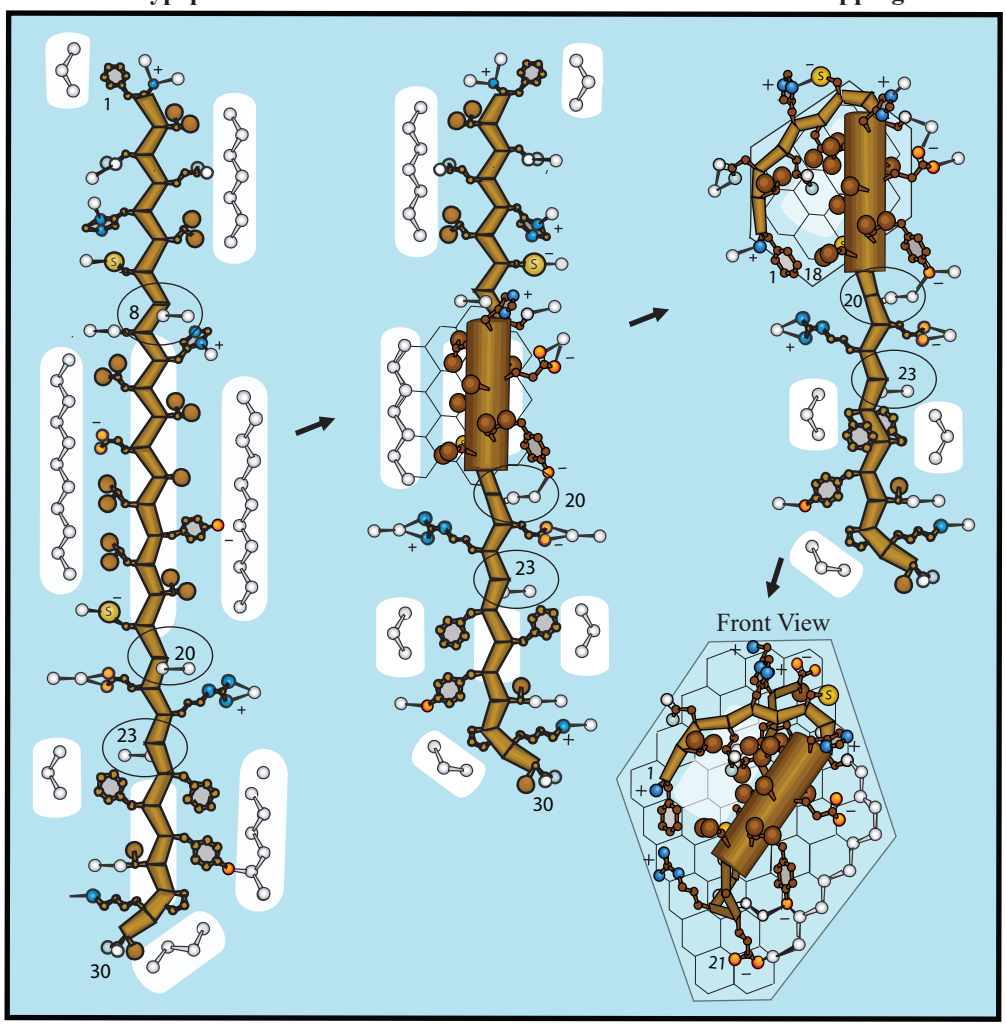


Insulin B-Chain Folding

It should not be surprising that 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 the water leaves, to produce a lower-energy coil.[2] By moving from ice-like to point-charge hydrogen bonding, even a trimer can absorb about 3 kcal/mole of quantized energy from a chain.[3]

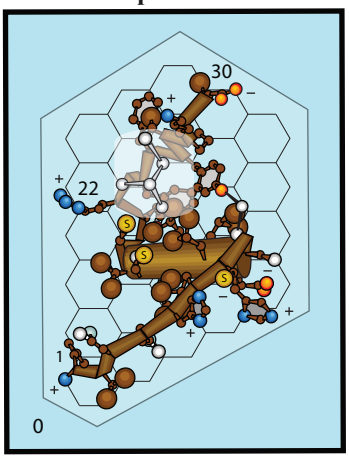
B Chain

Phenylalanine - 1
Valine
Asparagine - 3
Glutamine
Histadine
Leucine
Cysteine
Glycine - 8
Serine - 9
Histadine - 10
Leucine
Valine
Glutamate
Alanine
Leucine
Tyrosine
Leucine
Valine
Cysteine
Glycine - 20
Glutamate - 21
Arginine
Glycine
Phenylalanine- 24
Phenylalanine
Tyrosine- 26
Threonine
Proline
Lysine
Alanine- 30



Although coil-formation reduces ordered surface hydration, order continues to be induced on three sides of coil. Transient linear order adjacent to methylenes on the left-hand side of the coil and the upper segment of B, by leaving, draws those surfaces together to bring the aromatic ring of the terminal phenylalanine next to the lipid methyl groups of valine 18. Although unstable ordered water is released on the left side of the coil, it still covers its hydrophobic back side and, as it leaves, glycines at 20 and 23 permit the lower segment to fold into the back side of the coil as shown in the Front and Top Views.

Top View



However, the hydration-ordering region shown by the shaded area on the right 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.

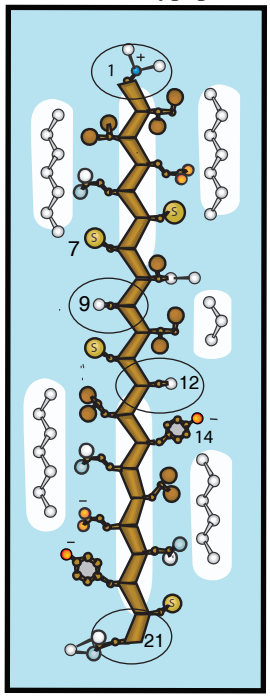
Insulin A-Chain

Instead of forming a single coil, the A chain forms two coils, each initiated by serines at positions 9 and 12. Another difference 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.

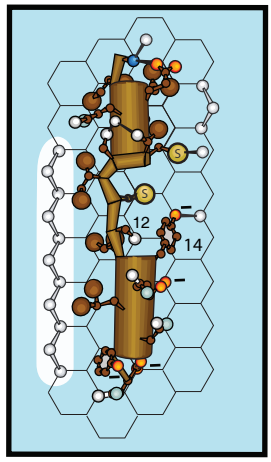
A Chain

Glycine -1
Isoleucine
Valine
Glutamic Acid
Glutamine
Cysteine
Cysteine - 7
Threonine
Serine- 9
Isoleucine
Cysteine
Serine - 12
Leucine
Tyrosine
Glutamine
Leucine
Glutamic Acid
Asparagine
Tyrosine
Cysteine
Asparagine - 21

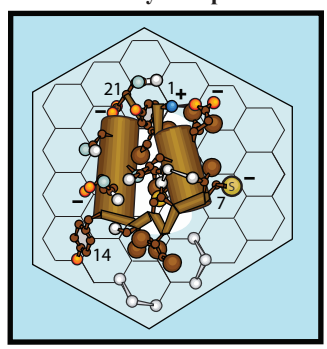
Initial A Polypeptide



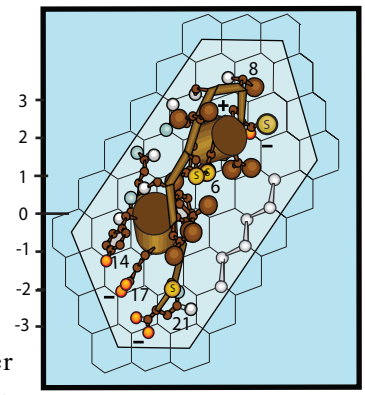
Coil-Formation



Assembly - T op View



Front View



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

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 of insulin. Notice in the Front View that a disulfide forms between the two coils to tie them together.

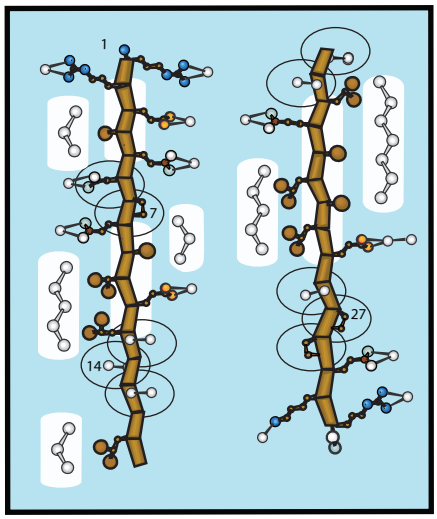
Insulin C-Chain

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 bring A into B.

C-Chain

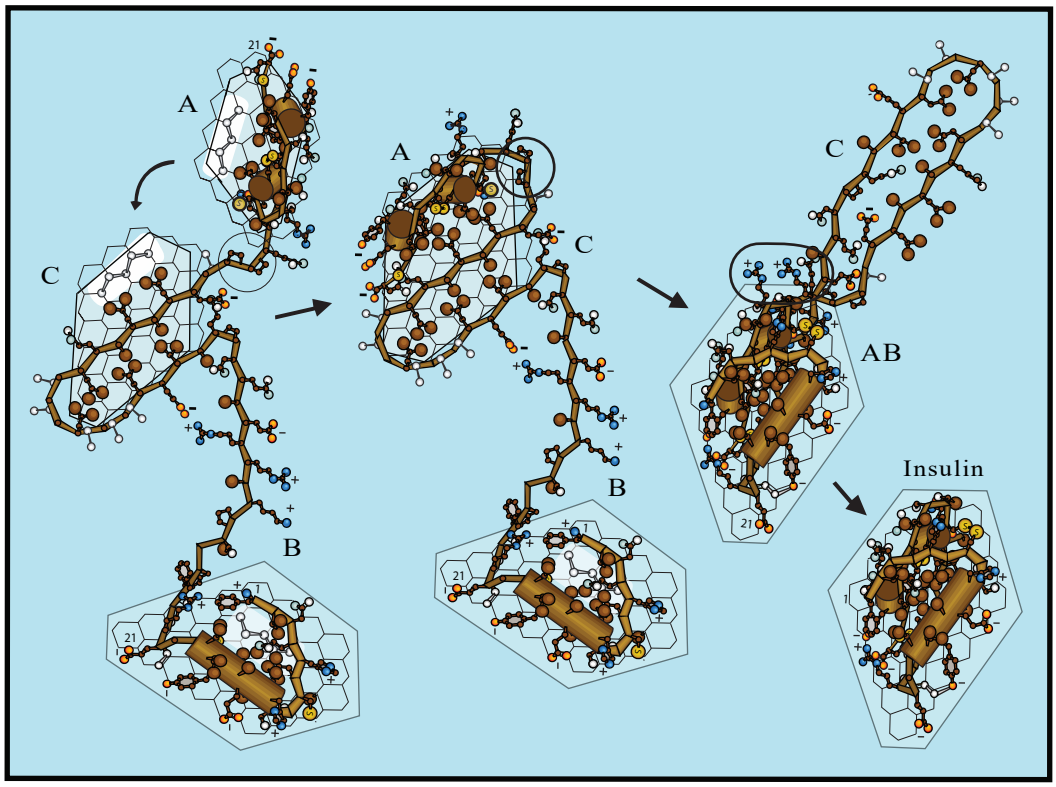
Arginine- 1
Arginine
Glutamic Acid
Alanine
Glutamine
Asparagine
Proline - 7
Glutamine
Alanine
Valine
Glutamic Acid
Leucine
Glycine
Glycine -14
Glycine
Leucine
Glycine
Glycine
Leucine
Glutamine
Alanine
Leucine
Alanine
Leucine
Glutamic Acid
Glycine
Proline - 27
Proline
Glutamine
Lysine
Arginine

C-Polypeptide



Chain C and Assembly

Although the process of folding A into B is complex, the C unit has hydration-ordering regions in its amino-acid sequence that make it possible to imagine how it may occur. based on the same order/disorder principles tha assembled A and B.



As shown on the upper left, the C segment, separated by two prolines from A, has two segments of lipid peptides, separated by 5 glycines, which can fold to form a unit with five linearly-ordered water molecules, the same as the back side of A. By flipping over and releasing unstable water from both surfaces, a central three-layered unit can be formed. Then, by sliding the front side of A to the back side of B and releasing the lower side of the central unit to form a lipid layer on the end of the tether, the insulin molecule unit can be formed.

Enzymatic cleavage at the two cationic arginines, releases the insulin molecule. In order simplify the assembly, the D segment, which extents from the other end of B and transports the insulin molecule through the outer membrane of the pancreatic cell, is not shown. It is cleaved as the insulin molecule is released into the blood stream.

When polypeptide synthesis first began on crude ribosomes, most-likely they produced random sequences of peptides that immediately became coated with water molecules which, either were attached to the small peptides or formed linear elements. Lipid and ionically-charged segments which could combine and release water to form firm anhydrous unions, formed new spatial structures. Those which could not form new spatial structures, accummulated at the exit site of the ribosome and by a process called “feedback inhibition,”slowed the production of those polypeptides and eventually stopped production. Those that could assemble spontaneously continued to produce new forms yielded the complex functional units in cells today. Amazing how nature works.

References

1. 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.

2. 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.

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

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