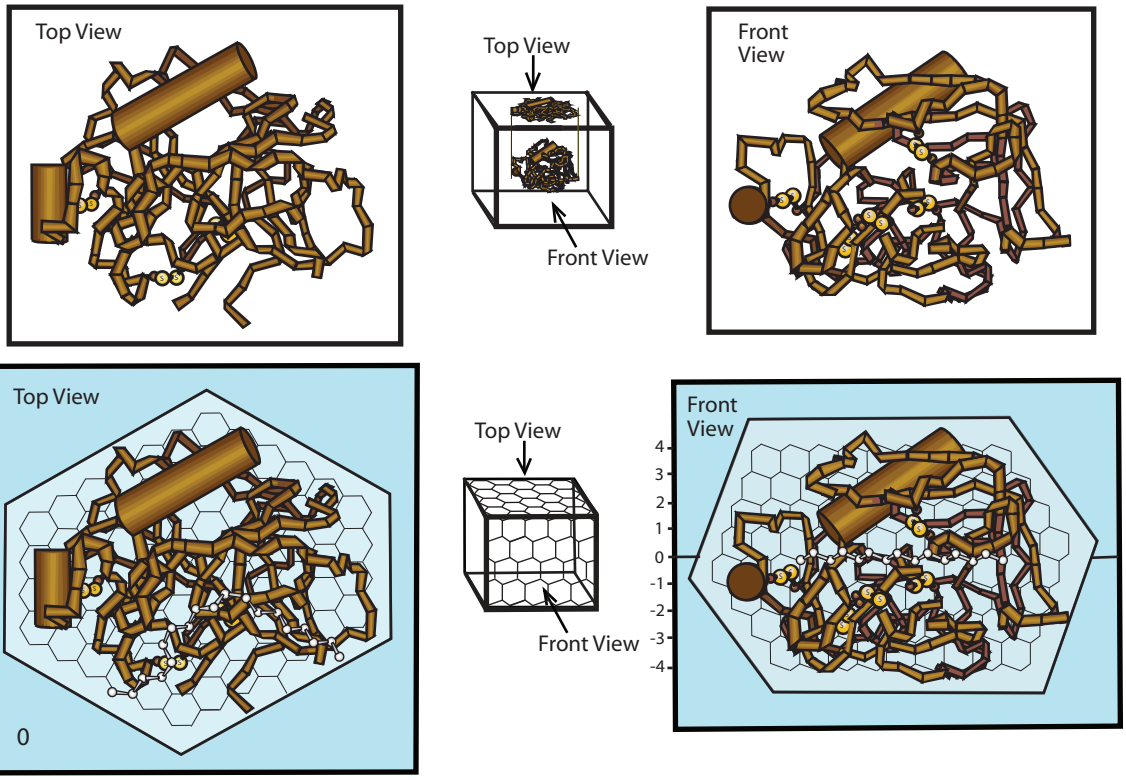
MOLECULAR PRESENTATIONS

Alpha-Chymotrypsin - Hydration-Directed Assembly and Catalytic Function [1]

J. C. Collins, PhD

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

Chymotripsin is synthesized on a ribosome as a single polypeptide strand of 245 amino acid residues which fold spontaneously into a finished protein called chymotripsinogen.[2] That form of the protein is inactive as an enzyme but enzymatic cleavage of peptides 15 and 16, followed by removal of several other peptides, produces the active stable enzyme, alpha-chymotrypsin.[3] Since the body of the enzyme is not affected by proteolytic activation and our analysis of folding begins with peptide 245, at the acid end, the presentation of wrapping and assembly is valid for both forms of the protein. An interpretation of the role of surface water in the catalytic mechanism of action of the enzyme is also included.

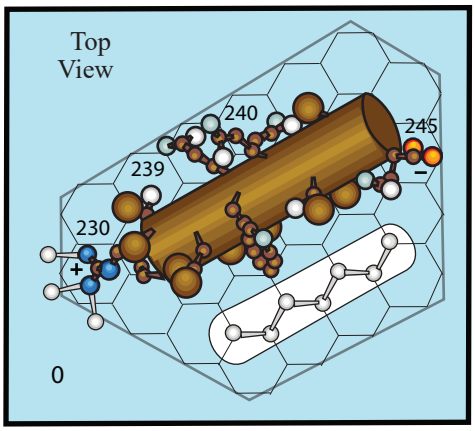


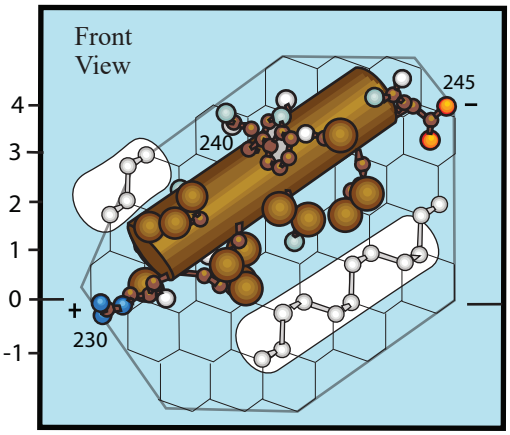
In order to gain a perspective of the spatial structure of the enzyme, relative to the 2.76A cubic lattice of hydration, both Top and Front Views are presented above with hexagonal layer 0 illustrated behind the Top View.[4] As usual, the analysis of this protein will begin with the longest coil with all but its upper surface hydrophobic. The coil is not parallel with any of the major axes of its X-ray crystallographic coordinates but is parallel with transient linear elements in the cubic lattice. Thus, coordinates reported for the crystalline form of the enzyme were used for the comparative analysis without changing the location of any atoms to satisfy hypothetical cubic hydration positions. Although analyses of most proteins required rotation of X-ray coordinates to provide for best fit of coils and sheets within the cubic lattice, this enzyme was analyzed directly as reported by Professor Blow.[3]

Substrate binding in this enzyme is complex because cleavage is in the center of polypeptide chains at aromatic peptides like phenylalanine and tyrosine. The linear elements of hydration shown in both Top and Front Views above are where covalently-bonded water molecules extend out from the enzyme to attract substrate polypeptides into the reaction site for cleavage. As you will see, the reaction site is surrounded by aromatic rings and sulfur atoms to provide for the formation of charge-transfer complexes with the aromatic rings of substrate polypeptides as they approach and bind in the site.

Alpha-Chymotrypsin (245-215)

Sometimes, as polypeptides emerge from ribosomes, they remain as linear units but the 15 peptides on the end of chymotrypsin are so hydration-ordering that they wrap rapidly into a coil.[5] However, the cation in arginine 230 produces so much hydration disorder that it produces an abrupt change in chain direction. With a positive charge at one end and an anionic charge at the asparagine end, a dipole is produced which assists in subsequent assembly.

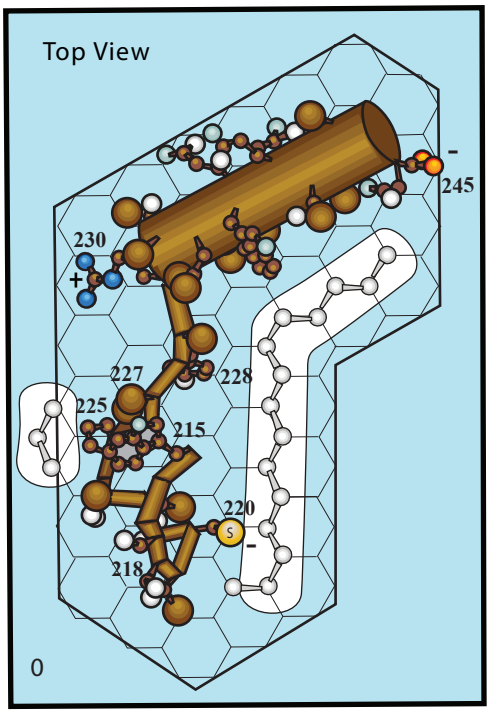




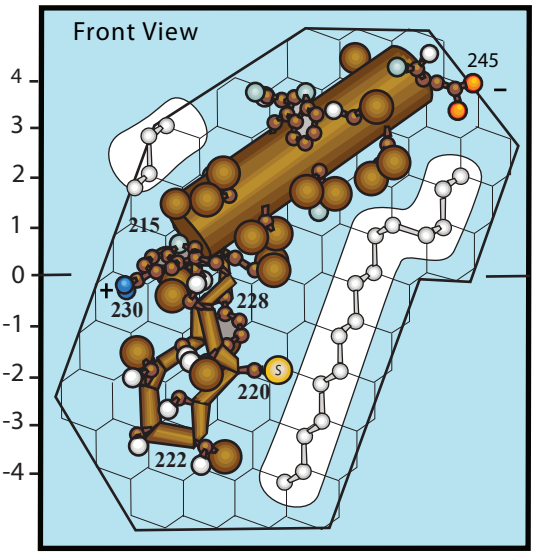
Asparagine 245
Alanine
Alanine
Leucine
Threonine
Glutamine 240
Glutamine 239
Valine
Tryptophane 237
Asparagine
Valine
Leucine
Alanine
Threonine
Valine
Arginine 230

Glutamines 239 and 240, with their amide oxygen and nitrogen atoms in spatial positions that do not support cubic patterning, hydration order is disrupted above and behind the coil. Conversely, hydrophobic methyls of valines and leucines on the lower and front sides continually induce the formation of covalently-bonded "ice-like" linear elements of hydration which will be displaced by complimentary lipid surfaces.

Since,the continuing polypeptide chain, as shown below, does not contain a linear segment of ordering peptides of proper length to displace the water next to the coil, it follows transient linear elements down from the coil with proline 225 initiating a dynamic loop of hydration order-disrupting serines, threonines and glycines, with their polar oxygens pointing outward and their lipids and the sulfur atom pointing inward. The loop is terminated by tryptophane 215 with its aromatic ring next to proline 225.

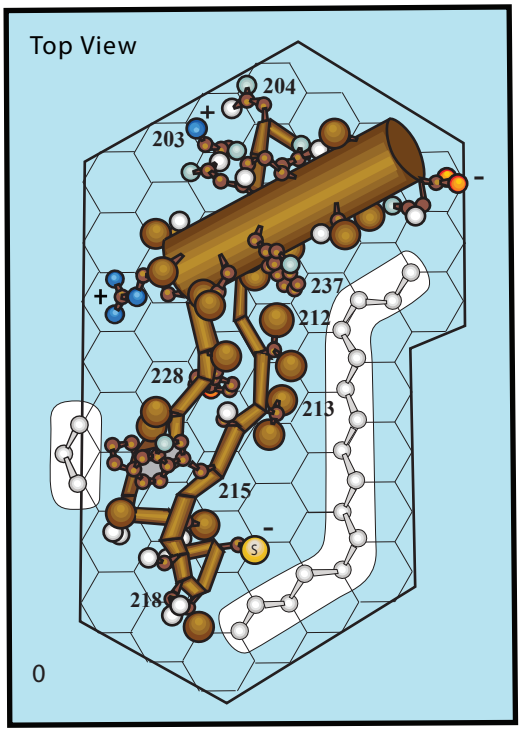


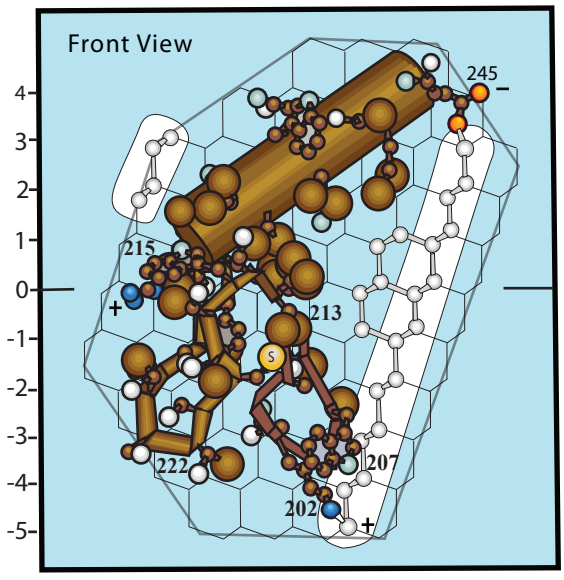
Alanine 229
Tyrosine
Valine
Glycine
Proline 225
Threonine
Serine
Threonine 222
Serine
Cysteine 220
Threonine
Serine - 218
Serine
Glycine
Tryptophane 215



Alpha-Chymotrypsin (245-202)

With the methyls of valine 213 positioned nxt to tyrosine 228 and those of isoleucine 212 below tryto-phane 237, the chain drops down at glycine 211, goes under the coil and forms a loop below it.





Tryptophane 215
Serine
Valine 213
Isoleucine 212
Glycine 211
Valine
Leucine
Threonine
Tryptophane 207
Alanine
Glycine
Asparagine
Lysine 203
Lysine 202

In the Front View, the lower loop can be seen following cubic patterning with the highly hydrated cation of lysine 202 turning the chain around at plane 5.

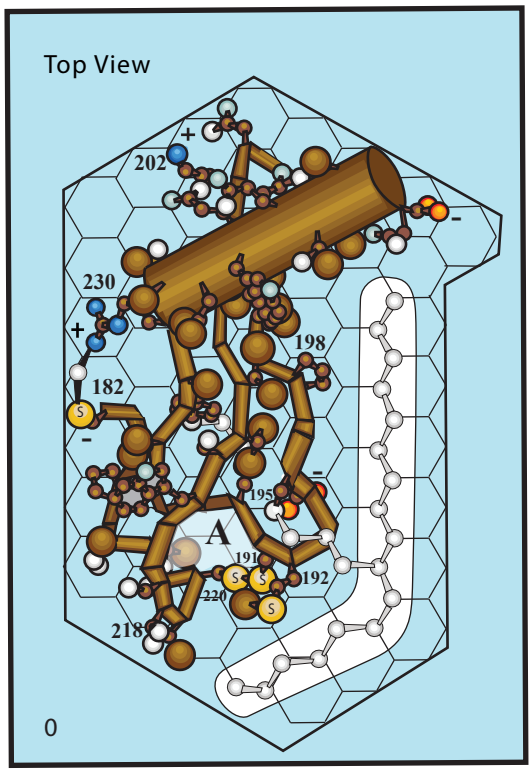
Again, we must be reminded that, even though surface water is displayed as extended linear elements, at an instant, they are present only as short linear segments like the trimer. Only when viewed over time, would water molecules be in the probability positions which generate cubic patterning.[4]

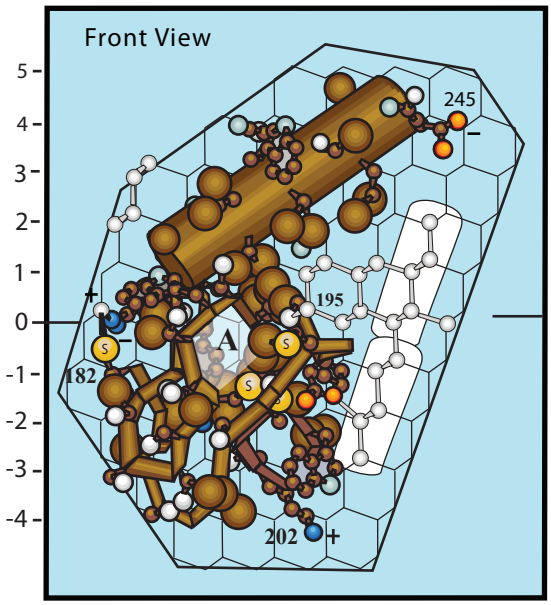
With anionic asparagine 245 at the top and lysine 202 at the bottom, a dipole of transient linear hydration is generated permitting proton tunnelling to stabilize the charges.[6] However, the hydrophobic cavity produced on the right side is so large that the lower loop might initially fold behind the first loop and move into the position shown as more space is filled by the continuing chain. Water trapped as droplets in crude oil, even above 0°C as it is pumped from deep sea-wells, must be warmed to prevent crystallization and blockage of oil flow.[7]

Since studies have never been reported on the folding of this short terminal segment polypeptide shown above, we can only speculate that it will be the same as in the finished protein. However, polypeptides pack together extremely tightly in cores, with little or no water to produce changes in location, so there is a high probability that they form in a stepwise fashion and that they are in the same same or similar positions during assembly as they are finished proteins.

Alpha-Chymotrypsin (245-182)

Once the chain has turned at 202 and passed back under the coil, it travels upward to proline 198 then through two glycines to serine 195, which is held tightly in plane 0 by a disulfide bond between cysteines 191 and 220.





Lysine 202
Cysteine
Valine
Leucine
Proline 198
Glycine
Glycine
Serine 195
Aspartate
Glycine
Methionine
Cysteine 191
Serine
Serine
Valine
Glycine
Serine
Alanine
Glycine
Alanine
Cysteine 182

To complete this section, the chain turns 120 degrees, passes below the first two linear segments to form a water coupling between anionic cysteine 182 and cationic arginine 230.

In Top and Front views, it can be seen that the polypeptide chain, by wrapping back and forth, has filled a rectangular space, displacing all of the water. It is as if serine 195, which positions its hydroxyl precisely in a cubic water position at the corner of a sold block of atoms, might serve some important function.

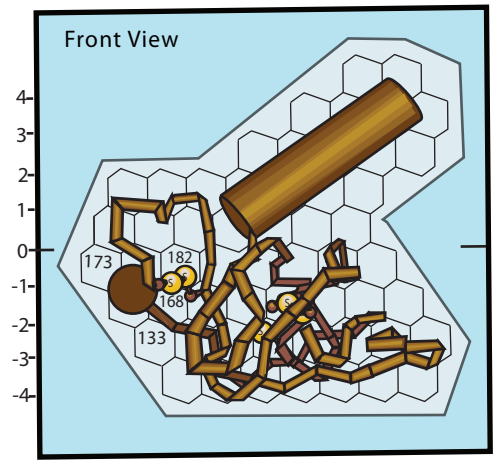
And, indeed it does - as protein assembly continues on the right side, a linear segment of substrate polypeptide will replace water shown leading into the hydroxyl group of serine 195 for binding and cleavage. Serine 195 must be held by the block of atoms in precisely the proper position to serve a primary role in the catalytic hydrolysis of the amide portion of aromatic peptides in the center of polypeptide chains. The open space, designated as A, is the cavity into which the aromatic rings of peptides are inserted and tightly bound.

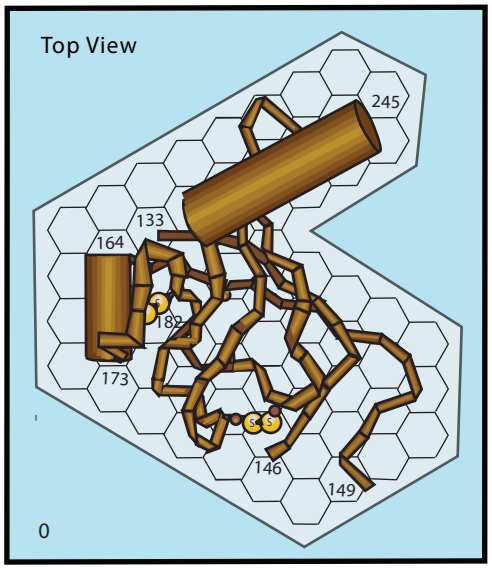
As you know, enzymes increase reaction rates thousands of times faster than normal. To do that, binding sites must hold reacting regions in precisely the proper positions and at precisely the proper angles for ordital overlap and the formation of new bonds between atoms with minimal activation energy.

The left side of this enzyme, as shown above, holds reacting aryl peptides in precisely the proper position to react with the serine 195 hydroxyl. The right-hand side, which will be shown on page 6, is involved in directing polypeptides with aromatic rings into the reaction site. Based on analyses of a number enzymes, it appears that it is covalently-bonded linear elements of hydration which are induced to form adjacent to specific sets of ordering peptides which drive and direct substrates, in multiple steps, into reaction binding sites. Selection of substrates for reactions often occurs at entrances to those sites as well as within those sites.

Alpha-Chymotrypsin (245-133)

Since some proteins do not contain coils, alpha-carbon plots may be used to select the best orientation to view transient covalent linear elements and cubic patterning.[4] Viewing the plot of alpha-chymotrypsin from asparagine 245 to glycine 133, it can be seen that the chains, in wrapping back and forth, appear to follow linear elements in cubic hydration patterning.

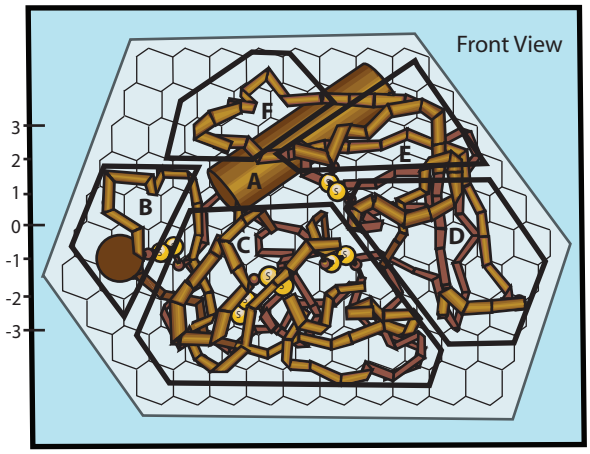




Isoleucine 181
Methionine
Alanine
Aspartate
Lysine
Isoleucine
Lysine
Threonine
Glycine 173
Tryptophane
Tyrosine
Lysine
Lysine
Cysteine 168
Asparagine
Threonine
Asparagine
Serine 164

Following the placement of cysteine 182, the chain moves up vertically to plane 1 and then down at glycine 173 to form a coil, including cysteine 168, which bonds with cysteine 182 to tie the coil horizontally in Plane -1. In the Top View, it can be seen that peptides 147 and 148 have been removed catalytically to activate the enzyme.

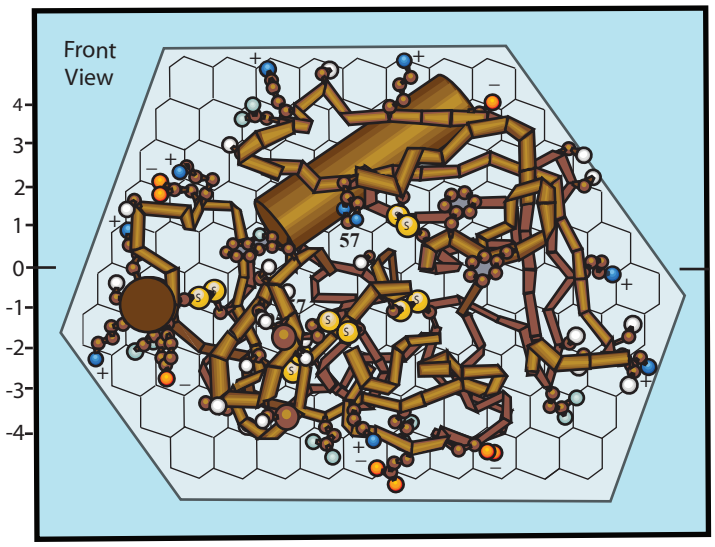
In the Front View of the total protein shown below, the molecule is divided into six spatial units which correspond to units of cubic patterning. It would be interesting to see if natural proteins are composed of a finite number of cubic space-units.

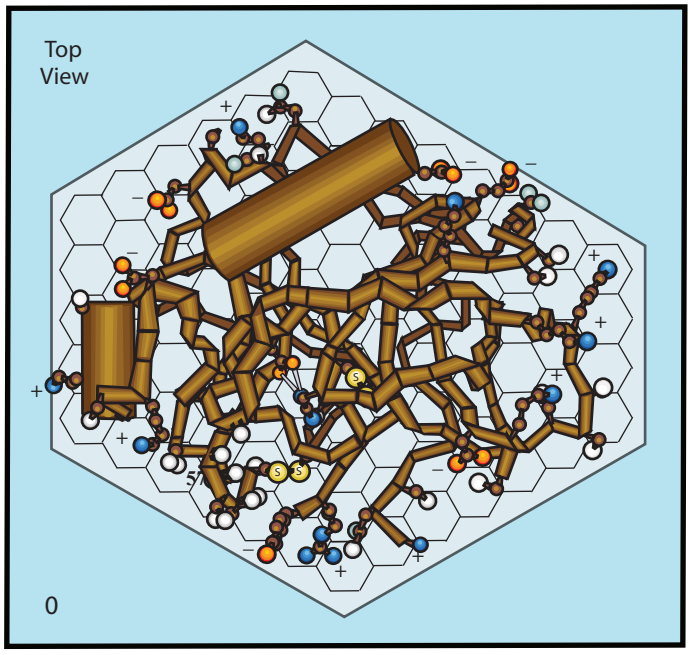


As illustrated in the Front View above, individual chains do not necessarily follow the orientations of ordering surface water but the assemblies which develop in the internal regions of proteins do reflect cubic patterning. Whenever linear chains change direction, there is usually a glycine, serine or proline at the position of change to hydrogen bond into surface water, disrupt order, increase hydration entropy and drive the chain in an alternate direction. Although an understanding of the role of transient linear hydration might assist in providing more viable interpretations of folding and assembly, the question of whether or not it will assist in providing more accurate predictions of final protein structures from polypeptide sequence information must wait until appropriate simulation models are available. In 2020, researchers at DeepMind, a U.K.- based-company, were been able to predict the finished structure of a protein using an AlphaFold program and, in 2021, Baker and Baek were able to predict conformations of proteins based on amino-acid sequences using a program they developed.[8] However, both programs used structural units in known proteins for the derivations - neither program included the role of surface water.

Alpha-Chymotrypsin - TLH Model

In these views of the total protein, surface groups are included to illustrate that most of them are not in positions to support cubic hydration patterning but, as mentioned before, in positions to disrupt the formation of transient linear hydration and coordinated cubic hydration in order to increase both stability and solubility.





Although surface groups in the crystalline state, as illustrated here, are held in relatively ridged positions, in queous solution, they have a good deal of freedom. For example, in the Front View, the loop above the coil has a degree of freedom to move upward but the one bearing histadine 57 in front of the coil is tied by a disulfide link to the lower unit and is relatively ridged. The histadine ring can rotate as it participates in the hydrolysis reaction, but the peptide itself is in a relatively ridged position to provide conformational control.

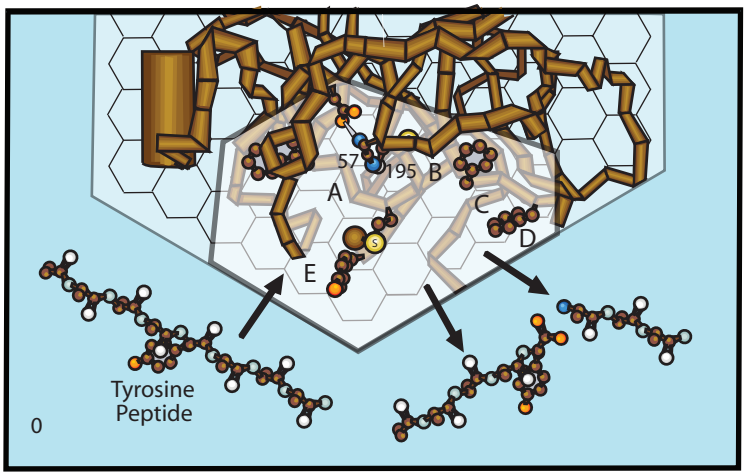
Since this molecule is composed of only two coils, it is the disulfides, formed by the oxidation of cysteine sulfides during assembly which provide for much of the stability. As will be illustrated on subsequent pages, it is the probability positioning of transient linear water in and around the binding site, which directs substrates into it.

Remember, the cubic lattice is displayed behind the protein only to illustrate the orientation of linear elements of surface water involved in directing folding and assembly. Several proposals have been presented that water forms clatherate-type cages around proteins but they don't seem viable because of large losses in entropy that would occur in surface water.[7] Instead, as mentioned before, it is most likely the formation of short covalently-bonded linear elements of hydration, like the trimer, which form kinetically, in an integrated fashion, on hydrophobic and poly-ionic surfaces. They last only about 10^{-10} seconds, by forming and degrading, generate a three-dimensional matrix of cubic hydration around the molecules.[3] Some oxygen atoms in serines and threonines are in proper positions and at proper angles to bond into transient linear elements and reinforce cubic patterning while others usually disrupt local hydration order.

Alpha-Chymotrypsin - The Catalytic Reaction Site

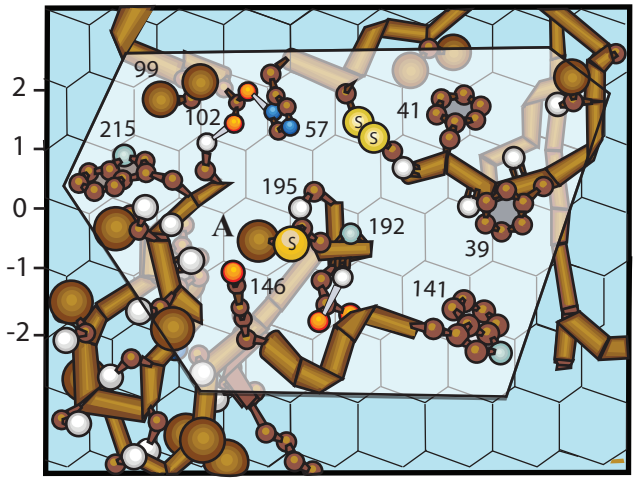
As illustrated below, the primary role of alpha-chymotrypsin in digestion is to hydrolyze polypeptides at aromatic peptides like tyrosine and phenylalanine. The cavity designated as "A," which was left open and hydrated during the folding process, is precisely the proper size to fit the aromatic rings of bound peptides as they are held in position for hydrolytic cleavage.[3]

Top View of the Catalytic Binding Site.

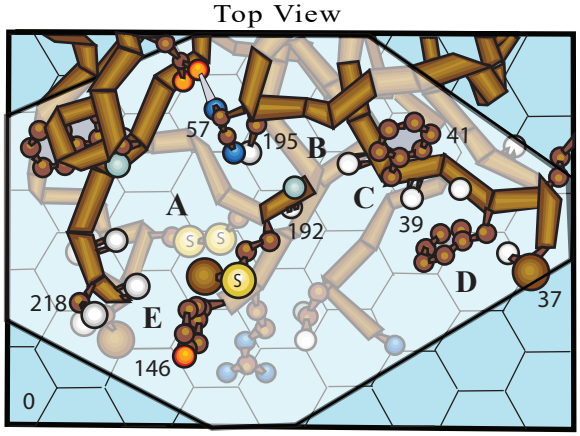


In the Front View on the lower left below, it can be seen that the reaction site is surrounded by peptides with aromatic rings at 215, 41, 39, 141 and 146. This increases hydration order around the binding site and permits electron-coupling with the aromatic rings of substrate polypeptides. Methyl groups on two corners of the site provide even greater order in surface water. As you will see, histamine 57 and serine 195 are intimately involved in the catalytic cleavage reaction.

Front View



Top View

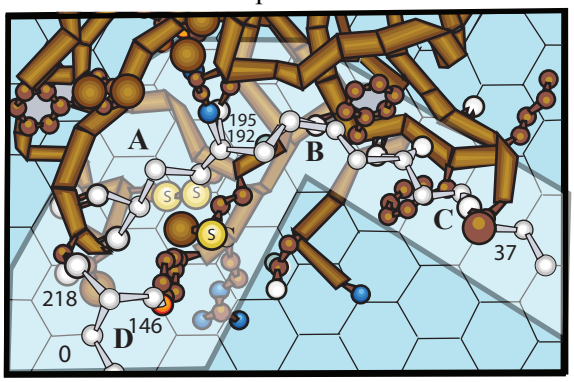


In the Top View on the right, a number of peptide carbonyl oxygens, as well as the amide nitrogen of peptide 192 are identified because they provide specific sites for binding substrate polypeptides. Again, Region A will bind aromatic rings, Region B will bind the peptide to be cleaved; C, D and E will bind adjacent peptides in the chain. The phenolic ring on tyrosine 146, which can rotate toward and away from A is involve in binding to the polypeptide chain and bring the substrate in and out of the binding site.

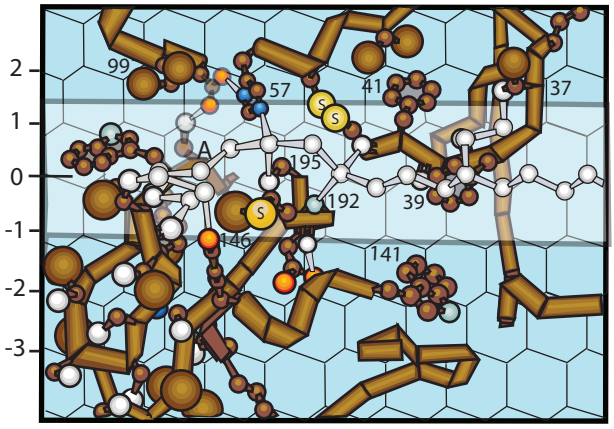
Alpha-Chymotrypsin - Reaction Site Hydration

The X-ray diffraction pattern of the crystalline enzyme reveals a number of water molecules in binding cavity A, but their high energy makes it difficult to define precise locations.[2] In spite of that, the water molecules most likely have preferred probability locations of lower energy than liquid water.[4] Since hydration locations are important in determining binding sites for substrate molecules as they move into the site, physical and computerized models were used to determine probability positions of water moleecules at acceptable angles and distances apart. The distribution shown is one of several alternative arrangements.

Top View



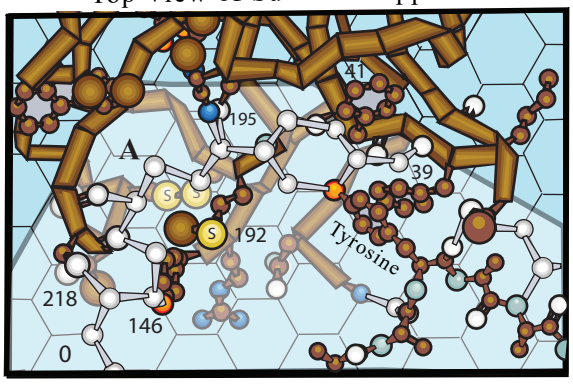
Front View



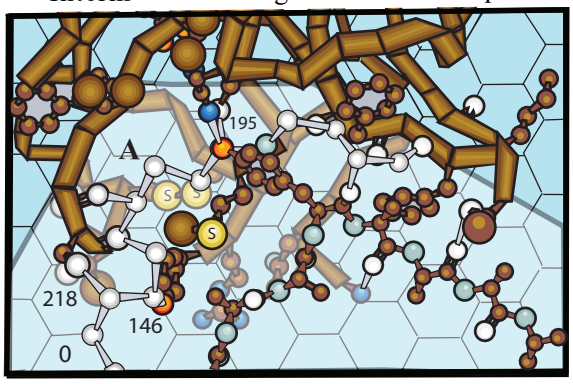
Many of the bridging water molecules are in cubic hydration positions but some are not. Most bridging hydration is in levels 0 and 1 with threonine 37 at level 2, extended forward above hydration in level 1. Also, note the positions of the aromatic rings of phenylalanines 39 and 41; they are in positions to couple with the aromatic rings of substrate polypeptides as they enter the site for cleavage.

What we often forget is that, as substrates approach an enzyme, each step must provide, not only acceptable binding, but a degree of selection for the peptide that will be cleaved. Thus, as illustrated on the left below, the aromatic ring of tyrosine in the polypeptide chain, most likely couples with the aromatic ring of phenylalanine 39 as the polypeptide chain begins to move into the binding site. Since the polypeptide chain may enter the site in many different orientations, aromatic rings on the polypeptide chain may form charge complexes with the rings at 41 and 39, as well as with the sulfur atom on methionine 192.

Top View of Substrate Approach



Intermediate Charge-Transfer Complex

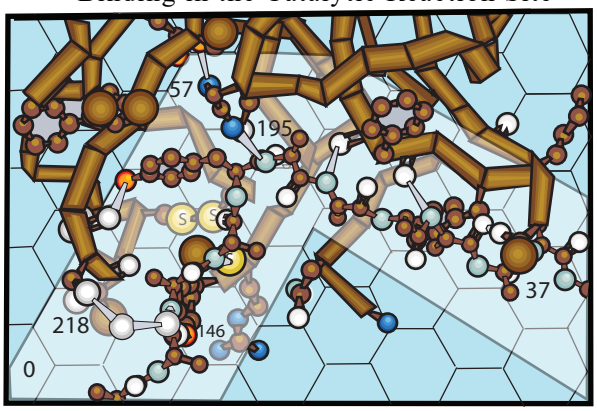


On the right above, the phenolic oxygen of tyrosine is shown bound to serine 195 with additional bound water displaced to increase hydration entropy and spontaneously move the substrate into the reaction site. The aromatic rings of phenylalanines and tryptophanes, could bind as shown on the left, but would bind over the methionine sulfur atom instead of serine 195 as shown above.

Alpha-Chymotrypsin - Binding and Reaction

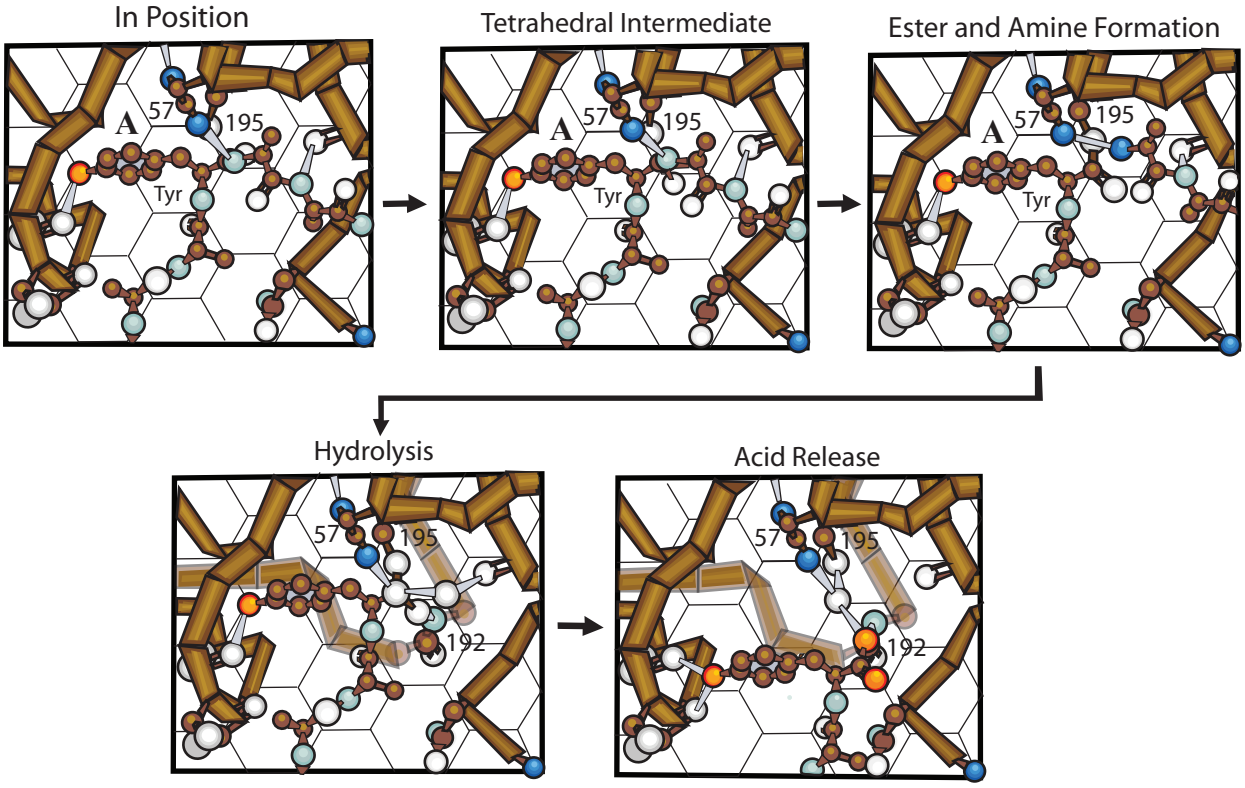
Once again, tyrosine must bind in many ways as it approaches the reaction site, but, once in the site, it must bind tightly in one position and displace most, if not all, of the water. As illustrated on the right, the aromatic ring fills cavity at A, with the chain on the right hydrogen-bonded to the 39-41 biaryl site while, on the left, the chain is hydrogen-bonded to the phenolic oxygen of tyrosine 146.

Binding in the Catalytic Reaction Site



The carbonyl of the peptide to be cleaved is held next to serine 195 at a 90-degree bend in the chain by a hydrogen-bond between the histadine 57 ring and the nitrogen of the tyrosine carbonyl.

The tyrosine peptide is bound in such a perfect position in the site that the electron orbitals of the serine oxygen and carbonyl spontaneously form a tetrahedral intermeditate, break the C-N bond, form the ester and release the right-hand portion of the polypeptide chain as a free amine.

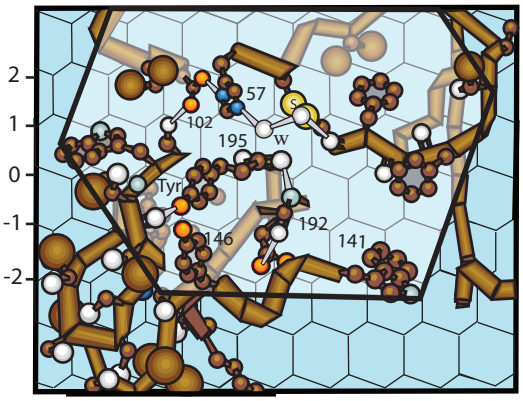


Peptide 192, which is now shown below the ester in the Hydolysis Figure above, holds the carbonyl oxygen of the ester in precisely the proper position that the water molecule, w, above it forms a second tetrahedral intermediate, breaks the ester bond and releases the tyrosine end of the polypeptide chain as the acid. By holding the molecule in precisely the proper orien- both reactions occur rapidly with minimal loss in energy: that is effective catalysis.

Alpha-Chymotrypsin - Summary

In this Front View of the ester hydrolysis, it is clear that water molecule w, which is hydrogen bonded to the nitrogen of histadine 57, is close enough to the carbonyl carbon of the ester to bond with it, form the acid and release the left side of the chain.

Front View



What is not clear in the illustration is that, in forming the initial reaction complex, stress is imposed on the bonds which is relaxed in forming the ester and conversion to the acid.[2] However, it must be remembered that a driving force in the formation of the reaction complex is the release of ordered water and the increase in hydration entropy.

Based on the current concept that the distribution of water around the molecules in living cells is totally random, it is impossible to understand how they can function with such extreme levels of order and efficiency. Furthermore, it is impossible to understand how they could have evolved spontaneously from random distributions of small molecules which were present on the early earth to yield the orderly self-replicating phenomenon we know as life without some sort of guiding order. However, natural molecular evolution did not occur in the randomness of air, it occurred in an environment which possesses the properties of both randomness and order - randomness to provide for motion and change, order to provide for efficiency and reproducability.

References

1. For background information on the hydration concepts used in this article, check out www.linearhydration.com.

2. W. Appel, *Clinical Biology*. **19**(6):317-322 (1986). Chymotrypsin: molecular and catalytic properties.

3. D. M. Blow, *Acc. Chem. Res.* **9**: 145-152 (1976). Structure and Mechanism of Chymotrypsin. See also, D. M. Blow and T. A. Steitz, *Ann Rev. Biochem.* **39**: 86 (1970). Mechanism of Chymotrypsin,

4. J. C Collins, *The Matrix of Life*, (Molecular Presentations, 1991). www.cubichydration.com

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

6. A. J. Horsewell, *Science* **291**: 100 (2001). Evidence for coherent proton tunnelling in a hydrogen bond network. 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.

7. D. Greaves et al., *Chemical Engineering Science* **63**(18): 4570-4579 (2008). Hydrate formation from high water content crude oil emulsions.

8. R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **111**(36): 13052-13056 (2014). Dynamic hydration shell restores Kausman's 1959 explanation of how hydrophobic factor drives protein folding. See also, S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, V. Heugen, M. Garebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. USA* **104**(53): 20749-20752 (2007). An extended dynamic hydration shell around proteins.

9. D. Baker and M. Baek, *Science* **373** (6552): 262-263 (2021). Protein structure prediction now easier, faster.