

(With embedded Q&A)

Last updated: Tuesday, September 21, 2010 01:53 AM

© Copyright 2010 Lawrence Chasin and Deborah Mowshowitz Department of Biological Sciences Columbia University New York, NY:

Bio C2005/F2401x 2010 Lec.5 L. Chasin September 221, 2010

Handout: Protein structure

Props: 3 ropes etc ...

Paper electrophoresis

Paper chromatography

Fingerprinting: 2-dimensional electrophoresis + chromatography of small peptides

Sickle cell disease: an altered sub-peptide of the globin polypeptide

3D structure is reproducibly formed.

Denaturation, renaturation, dialysis

Primary (1^o) structure = AA seq

Constrained rotation of the backbone

Secondary Structure

Alpha helix

Beta pleated sheet

Tertiary structure (3^o) = overall 3-D conformation of a single polypeptide

Protein domains

Chaperonins

Quaternary (4^o) = association of multiple polypeptides

Protein domains

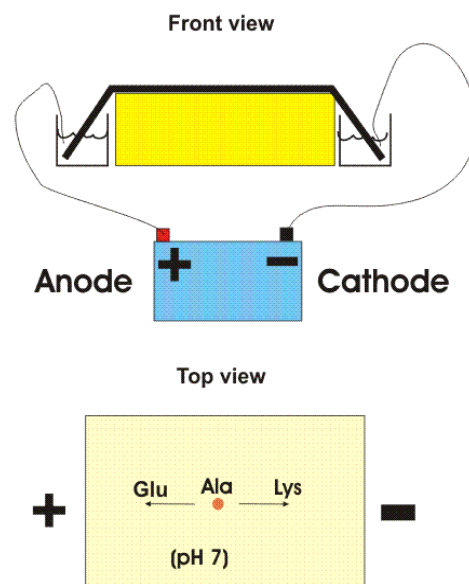
Prosthetics groups

Membrane proteins

One way is based on the migration of amino acids in an electric field. In **PAPER ELECTROPHORESIS**, an amino acid mixture is spotted onto a sheet of filter paper, the paper is wet with a buffered salt solution and placed between two electrodes and high voltage (e.g., 2000 volts) applied. At neutral pH, the acidic amino acids (asp and glu) will have a net negative charge and will migrate toward the ANODE (+ pole) while the basic amino acids (arg and lys) will migrate toward the CATHODE (- pole). **{Q&A}** Electrically neutral amino acids will not migrate much, unless the pH is made acidic or basic (as it is in some problems in the problem book) .

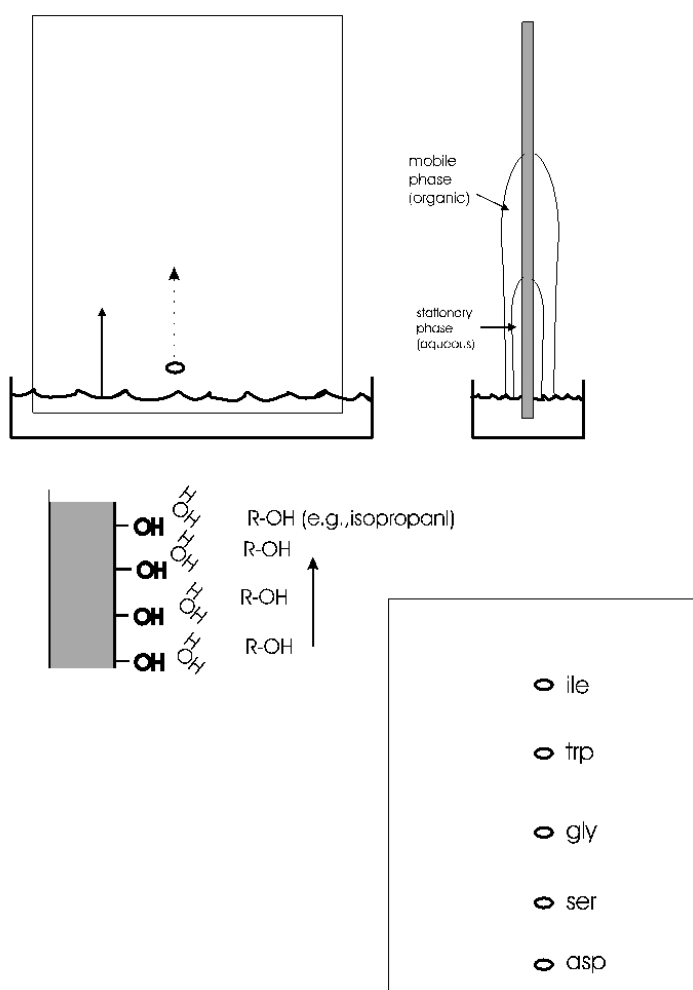
Viewed after application in the center followed by electrophoresis.

Paper electrophoresis of amino acids



Try problem 2-1.

A more versatile separation method is **PAPER CHROMATOGRAPHY**. This method is based on the differential solubility of the different amino acids



in organic (non-polar) solvents, which in turn is determined by the nature of the side group. The amino acid mixture is spotted onto a filter paper; one edge of the paper is immersed in a mixture of aqueous and non-aqueous solvents. (See [handout](#).) The liquid will be drawn up the paper by capillary action. As it rises, the water in the liquid mixture is bound by the paper (cellulose, with its many OH groups), forming a stationary water layer, or stationary phase. The organic solvent (e.g., propanol) moves up without as much interaction with the solid cellulose; it is considered the mobile phase. The amino acids will be constantly equilibrating between being in the mobile organic phase or the stationary water phase. The more polar the side chain, the more time the amino acid will spend in the stationary phase. The more

hydrophobic the side chain, the more time it spend in the mobile organic phase. By using a series of different solvents, all 20 amino acids can be separated in this way. It works for many other organic molecules as well. The distance that an organic molecule moves in a particular chromatographic system is called the R_f , which stands for mobility Relative to the Front, that is, the distance the organic molecule in question migrated divided by the distance that the front of liquid has risen on the paper at that time. R_f 's in a particular solvent and at a given temperature are reproducible and are published for many organic compounds, including all the amino acids. [{Q&A}](#)

Small PEPTIDES [I emphasize peptides here, that is, oligopeptides, not *polypeptides*], like the sub-peptides produced by trypsin digestion of the

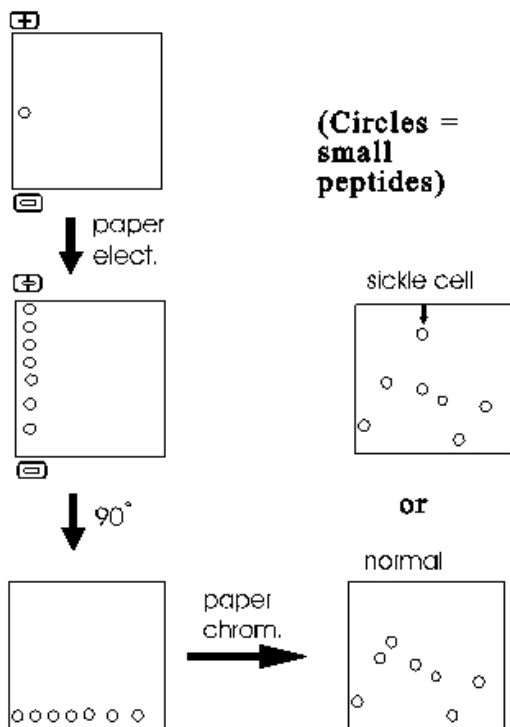
polypeptide pictured above, can also be separated by both of these techniques; the properties of the peptides will be a **COMPOSITE** of the properties of the constituent amino acids. That is, each of the amino acid side groups in a particular peptide will contribute some hydrophobicity or charge or polar quality; the resulting peptide will reflect the combination of all these effects. For example, a peptide with 2 arginines and one glutamic acid as the only charged residues will have a net charge of +1 at pH7 and so migrate toward the cathode in paper electrophoresis. [{Q&A}](#)

To review chromatography, try problems 2-9 A & B & 2-10.

One of the most famous examples of the use of these methods to analyze peptides rather than single amino acids was in the study of sickle cell disease, a genetic disease. Sickle cell disease is caused by an abnormal hemoglobin protein and results in abnormally shaped red blood cells that can clog capillaries and so prevent blood flow. Hemoglobin is made up of several components, one of which is a polypeptide called beta-globin. The sequence of amino acids in beta-globin from sickle cell hemoglobin was found to differ from that of normal beta-globin. In fact, the difference between sickle cell beta-globin and normal beta-globin was the first case in which the molecular basis of a genetic disease was defined. This was first done by Vernon Ingram in the 1960's. Ingram chopped up the sickle cell beta-globin into small peptides. This digestion can be done using enzymes (endo-proteases, that cleave polypeptides in a specific way, that is, at specific amino acid residues. For example, as mentioned above, the protease trypsin hydrolyzes polypeptides after (i.e., on the carboxyl side of) lysine and arginine. So first treat the beta globin protein with trypsin to break it into pieces (sub-peptides) by hydrolysis after lys and arg residues.



The resulting mixture of sub-peptides (outside of this lecture they are just called peptides, the "sub" being understood from the context) are then first separated along one edge of a filter paper sheet by paper electrophoresis. Note that these are peptides that are migrating, not free amino acids. The sheet is then turned 90 degrees and subjected to paper chromatography.



The result is a series of spots (after staining to visualize their positions) representing all the sub-peptides.

One peptide migrates differently in sickle cell globin compared to normal globin. This peptide can then be eluted from the paper and sequenced. Comparison with the normal counterpart peptide shows that the sickle cell globin carries a single amino acid substitution. In place of glutamic acid, it has a valine at one position in the peptide. How could such a single change have such a large effect? The answer lies in the 3-dimensional shape of proteins, to which we will soon discuss.

Most proteins can be separated into characteristic patterns of spots this way. No two proteins have the same primary sequence, and so each protein will yield a

different set of sub-peptides after trypsin digestion. Most of the sub-peptides from any two polypeptides will migrate differently, so the total pattern of spots will be different for each protein. The procedure is called **FINGERPRINTING a protein**, since the migration patterns are so characteristic. I use the term fingerprinting to refer to the entire process of digesting the polypeptide into sub-peptides, separating it in 2 dimensions and then visualizing and comparing the spots. [{Q&A}](#)

To review fingerprinting, try problem 2-11.

Protein 3-dimensional structure

Now let us return to polypeptide structure.

Each polypeptide has a particular sequence of amino acids. Thus if we could

examine several molecules of the protein albumin we might find:

Molecule #1: N-met-leu-ala-asp-val-val-lys-....

Molecule #2: N-met-leu-ala-asp-val-val-lys-...

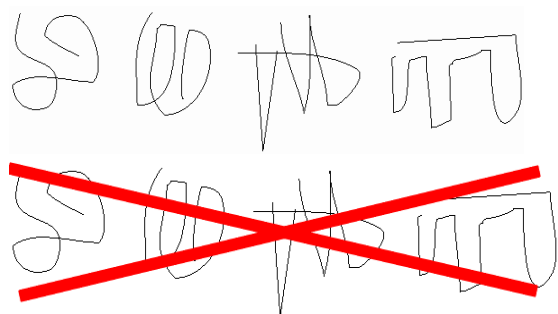
Molecule #3: N-met-leu-ala-asp-val-val-lys-... etc.

So they have the same primary structure. But as always, we must consider structure in 3-dimensional space for a real picture of the molecule.

While the linear structure is the same, the 3-D structure for each molecule must surely be different in solution, no? After all, thermal motion will be buffeting this rope of strung-together amino acids all about, so that each molecule will be expected to take on a random configuration, no? Look at this scale model of a POLYPEPTIDE OF 500 amino acids, a CLOTHES LINE. The dimensions are about right, but the side chains have been left out. I have put colored parts of the rope **red** to indicate *polar* side chains, the *white* parts being apolar or hydrophobic [board]. At 37 degrees, you might imagine this clothesline in a Jacuzzi, constantly taking on new shapes, with its hydrophilic side chains constantly forming new hydrogen bonds to water.

This is the wrong picture. A more appropriate picture is a bundled up rope, folded into a compact structure that withstands this thermal motion at body temperatures [bundled rope].. red on outside ...white hydrophobic on inside (which makes sense based on the weak bond behaviors we discussed).

OK, maybe this molecule could collapse on itself .. after all the hydrophobic side chains will tend to aggregate. But if we took another molecule, another linear chain, it would probably fold a different way, after all, 500 amino acids, there must be many many ways to get the hydrophobics inside. I could stuff the white parts of the rope together and put them on the inside in many different ways: leucine-2 pushed up against valine 25 in one molecule but associated with leucine-346 in the next molecule, so we might expect each of these protein molecules to have a unique structure in 3-D space:



But in fact if we look for a second folded up example of this molecule, it looks like this [second rope bundle], **exactly the same** as the first (note loop count, etc.). Protein molecules exist as precisely defined 3-dimensional structures in solutions, each molecule like the next, super imposable.

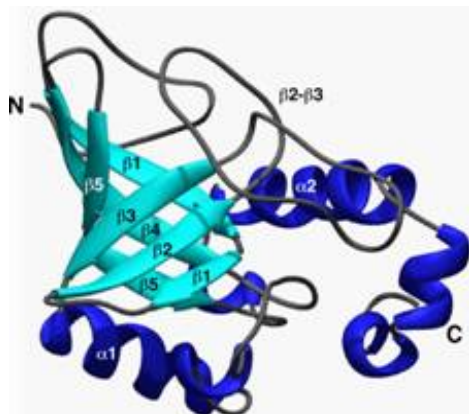


That is, a typical polypeptide chain, having some 10,000 atoms linked together, is folded up so that these 10,000 atoms all have the same position relative to each other in each and every molecule you examine. This still amazes me. How could this be? How could each molecule find just the right structure as it folds up, each and every time? Let's see.

Well, what is holding the molecule in this shape? The four weak bond types we discussed earlier, plus one new bond to be described in a few minutes.

So now we can see that one polypeptide molecule can be folded into a compact structure and we can understand what holds it together, but why is it that there is only one structure formed and not many? Is there only one solution to the folding problem for a particular polypeptide chain?

Perhaps all possible conformations are tried in the course of folding, and only the most stable one accumulates. Can we predict the conformation from first principles? If we plug in the properties of all the amino acid side chains, how hydrophobic they are, what is the strength of an ionic bond, etc. we can ask a computer to try and try many combinations, many interactions. This is a very difficult computer problem, even for today's supercomputers, because the number of possibilities for a good-size polypeptide of say 500 amino acids is enormous (20^{500}). But it has been tried, and so far usually the wrong structure comes out if one relies solely on first principles. (more accurate predictions can be made by comparing an unknown protein to known examples). The right structure is determined by examining crystals of the proteins, beaming X-rays through the protein crystals and calculating how they are refracted by the atoms in the crystal. Perhaps we really don't know the right properties of the side chains. Or perhaps there is some guide to folding that being imposed on the polypeptide as it is being polymerized in the cell, some outside influence, even a template of sorts, one can imagine a plaster mold analogy, for example. That is not the case, as we shall see next.



(Denaturation, renaturation)

Well, if it is true that the folded structure of a particular protein is unique simply because it is the most stable, then if we UN-fold the polypeptide, it should be able to RE-fold into a its unique structure. How could we unfold a protein, let's say one with only weak bonds holding the protein i its 3-D shape. We could consider egg albumin, for an everyday case of polypeptide denaturation. Raw egg white is a concentrated solution of this single ~500 amino acid



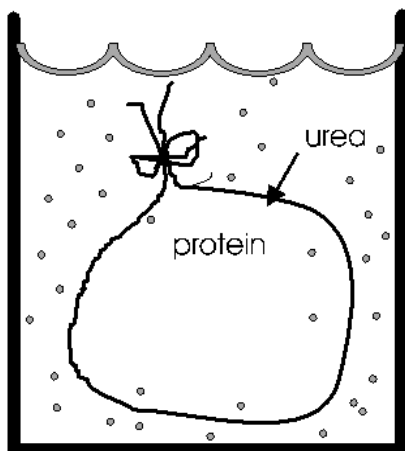
polypeptide, that exists folded into a roughly spherical shape.

How can we denature it?...The 3-D structure is being held together in most proteins lust by the weak bonds we have discussed (H-bonds, ionic bonds, van der Waals bonds and hydrophobic forces). Heating proteins typically unfolds them: thermal motion becomes too great for the weak bonds.

Let's heat it to denature it (boil it). The sphere is now subject to faster and faster thermal motions, until finally it starts to unravel. What has happened to the egg white (the albumin polypeptide)? It has become **denatured**. [Purves6ed 3.9; 7th 3.11](#). No longer **native**, which is the structure in the cell. No covalent bonds have been broken by this 100°C temperature. The bundled up rope became the open randomly coiled rope in the Jacuzzi, and this allowed many wrong bonds to form, it exposed the hydrophobic groups normally hidden in the interior of the protein. In this concentrated solution a tangled mass of interacting polypeptide chains was produced, which resulted in a gel, a hot hard-boiled egg. So while folded up polypeptides are stable enough in their native environment inside the cell, the 3-dimensional structure is typically rather fragile: most proteins are easily denatured by heat and other treatments that can affect these weak bonds. This bundled rope in the Jacuzzi exists on the verge of becoming unraveled.

So now let's reverse the denaturation, let's cool down this hard-boiled egg and return it to normal temperatures. The gel seems to stay. We do not get back our runny egg white. A case of irreversible denaturation. But not a very fair experiment, letting all those molecules, present at a very high concentration, get SO tangled with each other. Let's try a denaturation - renaturation in a more gentle, gradual way.

A fellow named Christian Anfinsen did this experiment in the 1950's. He took a protein called ribonuclease, a protein that is a digestive enzyme, a protein that helps break down the macromolecule RNA. It must be in its native folded structure to do this job.



DIALYSIS

(Dialysis)

Anfinsen placed the polypeptide in a plastic sack, and added urea,



to the solution outside the sack. At high concentrations (e. g., 8M), urea will break hydrogen bonds [{Q&A}](#) [{Q&A}](#). The sack is made of a semi-permeable plastic material with pores big enough to allow small molecules like urea and water to pass through but not macromolecules like albumin or ribonuclease. This process of allowing the concentrations of small molecules to change while holding

the concentrations of large molecules constant is called **dialysis**. After allowing time for diffusion, the concentration of urea inside the sack should be the same as the concentration outside. [{Q&A}](#) He then checked that the protein had become denatured (e.g., by

ultracentrifugation, see below). And it no longer do its job catalyzing the hydrolysis of RNA.

Now he gradually dialyzed *out* the urea (by changing the solution outside the sack to stepwise lower and lower concentrations of urea). A dilute solution of the protein was used, and the gradual removal of the urea gave time for the polypeptide to re-fold.

He got back native ribonuclease. It checked out physically, and also functionally, by the fact that it regained its ability to digest RNA.

This type of experiment has been now been repeated many times for many different proteins. It works for many, fails for some. But the positive results are very important, for they prove that for many or even most proteins, *all the information that is necessary for the complex and unique 3-dimensional structure is present in the primary sequence of the polypeptide chain.*

The discovery of this fundamental tenet of biochemistry earned Anfinsen the Nobel Prize. It is difficult to overestimate the significance of this finding, as it implies the information we need for life is carried in the sequence of amino acids in our proteins.

Chaperonins)

That said, it must be added that in the past 20 years, it has become apparent that some special proteins, called **chaperonins**, can help certain other proteins to fold with in the cell. It seems that these chaperonins may be needed for initial folding, perhaps to keep the as yet unfolded molecules from interacting with each other in non-productive ways, such as aggregating via exposed hydrophobic groups, as I suggested as occurring during the solidification of a hard boiled egg. See Purves 7th ed. Sadava 8th Fig. 3.12 for one complex chaperonin. The chaperonins are also used when proteins denature inside the cell: for instance, after they have traversed a membrane, with its hydrophobic environment. Or after cells have been exposed briefly to slightly elevated temperature (called heat shock), when a few of the least stable proteins may start to denature. The role, the generality, and the mechanism by which these proteins aid other proteins in folding correctly is not yet well understood. However, these cases do not really detract from the general principle that primary sequence CAN determine all higher order structures.

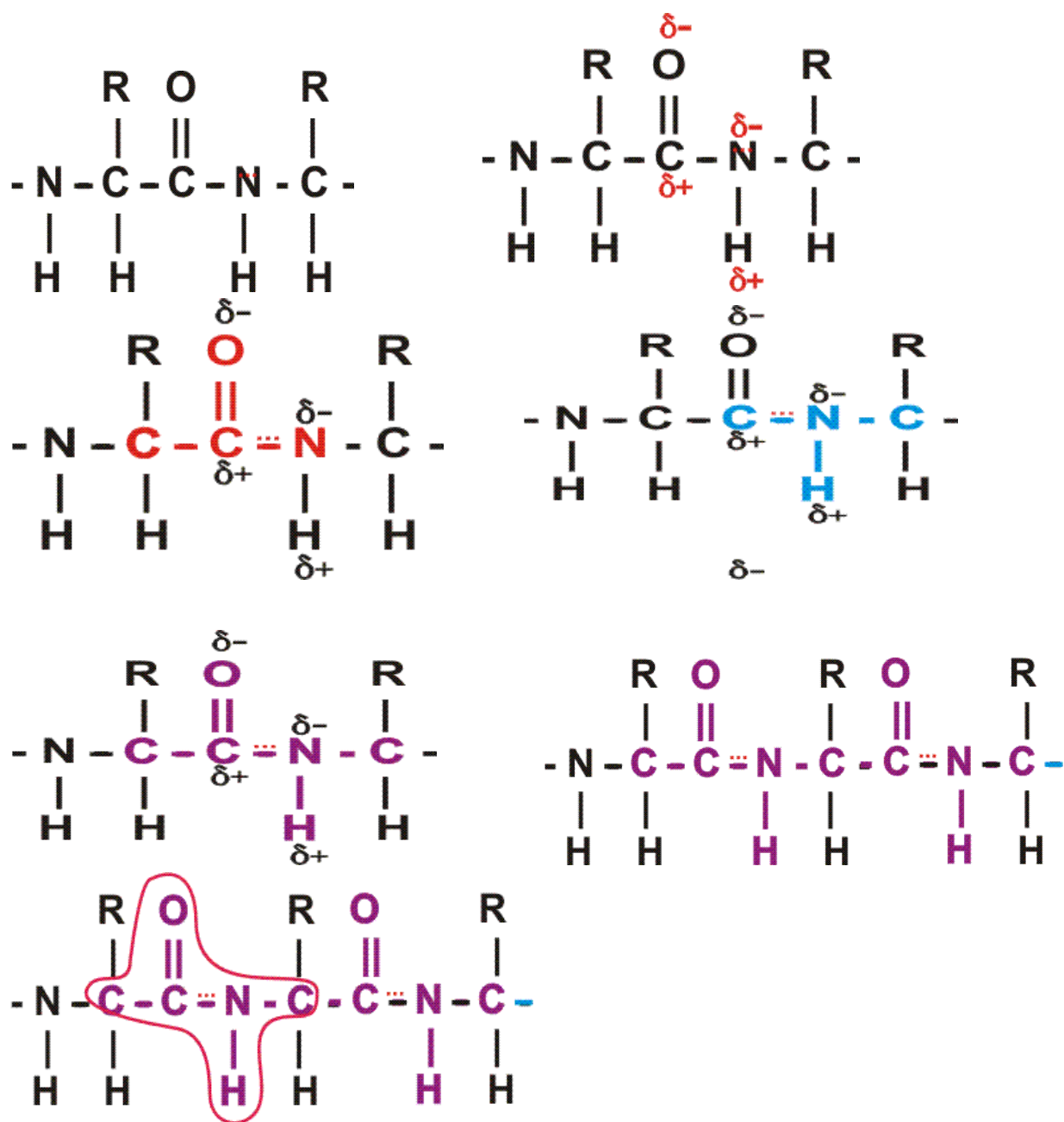
Well, what is holding the molecule in this shape? The four weak bond types we discussed earlier, plus one new bond to be described in a few minutes.

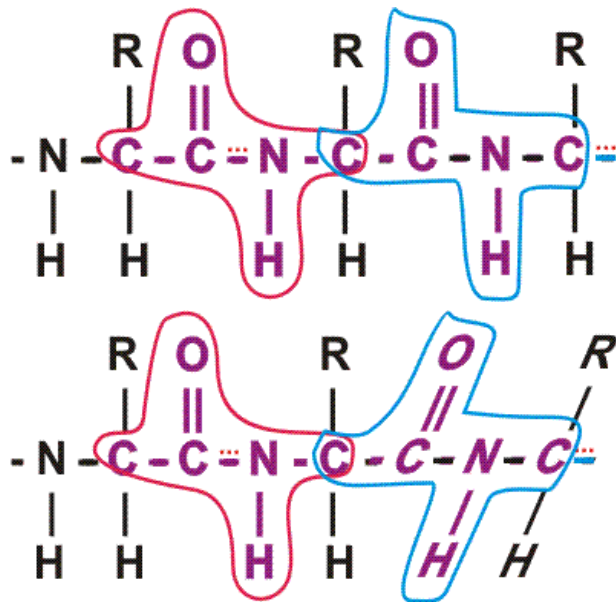
Let's consider how this folding looks in more detail:

First, a flexible rope is not a good representation of even the backbone, because the peptide bond itself imposes some constraint on structure. The peptide bond itself has a property that influences all polypeptides regardless of the side chains. Because of the electronegativity difference between C and O or N, there is a partial separation of charge, one you could have predicted. What you may not have realized is that the partial + charge on the C and the partial - charge on the adjacent N, imparts a partial extra bond between those 2 atoms, and thus a

partial double bond character to the C-N bond. This partial double bond is sufficient to stop free rotation about the C-N bond. Thus the backbone is not free to rotate around all connections, but rather *each repeat contains 6 atoms confined to one plane*:

[The four red atoms (third figure) lie in one plane. The four blue atoms (fourth figure) lie in one plane. The 4 red atoms and the 4 blue atom have 2 atoms in common (the C-N), both of which lie in the same plane. So all six atoms must lie in the same plane]

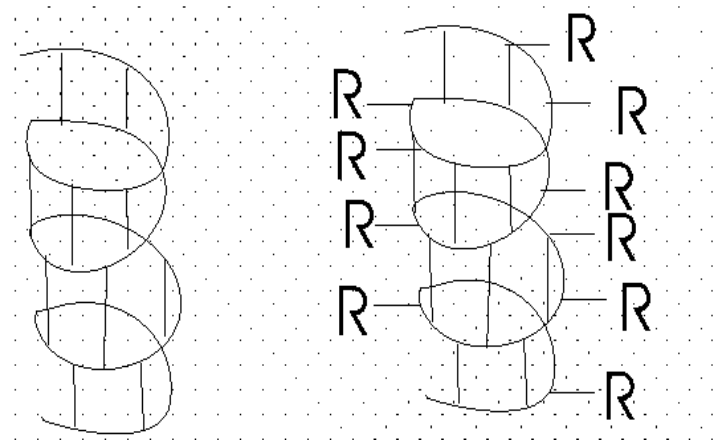
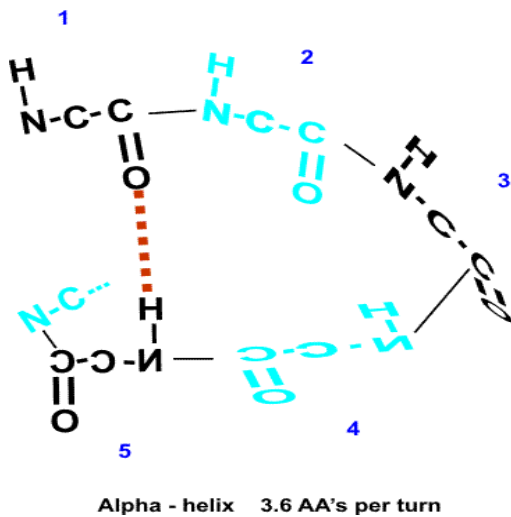




The polypeptide can be visualized as having a series of planes, each able to rotate about one another. So a chain would be a better representation than a rope. See [handout](#).

(Secondary (2°) = alpha-helix, beta sheet)

This partial separation of charge also means that the O and the NH of the peptide bond can hydrogen bond... to water for example. Since the NH is a hydrogen donor and the O is a hydrogen acceptor for a hydrogen bond, we should consider the possibility that these groups can H-bond to each other. But H-bonds require a linear orientation of the 3 atoms involved, so certainly the NH of the very next residue cannot H-bond to a C=O preceding it. But what about the next residue? No, still can't make it. But by the fifth residue down you are able to line up an NH to the O: $-C=O..H-N-$. i.e., there are 3 complete residues 3 in between these two interacting amino acid residues. [{Q&A}](#)

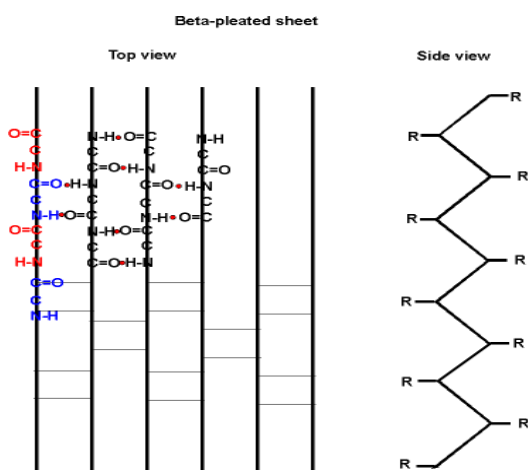


So the C=O of #1 can H-bond to the HN of #5. But then also the C=O of #2 should be able to H-bond to the HN of #6, and so on. This twisting and H-bonding can hold the backbone in a HELIX, the so-called [alpha-helix](#). {Q&A}

The alpha-helix is an example of **secondary structure**, which is (my definition): **structure produced by regular repeated interactions between atoms of the backbone**.

We might expect **all** the amino acid backbone atoms to be in an alpha-helical conformation, but we have left out consideration of the side chains, which can greatly influence the folding, as we will see in a minute.

The alpha-helix is not the only form of secondary structure, there is another, the beta-pleated sheet. In this case we once again have the C=O and the NH of the backbone forming H-bonds to each other, but in this case two sections of the polypeptide are aligned side by side (each vertical line represent a different region of the same polypeptide):



Several sections of polypeptide can line up like this, to produce a sheet of strands. The chains are usually anti-parallel, but parallel alignments are also possible.

See your texts for better pictures of these structures. B: 49 and [Purves6ed 3.5a](#).

Once again, side chain interactions play a major role in allowing or disallowing such secondary structures to form. But in fact, most proteins do have extensive regions folded into alpha-helices and beta-pleated sheets.

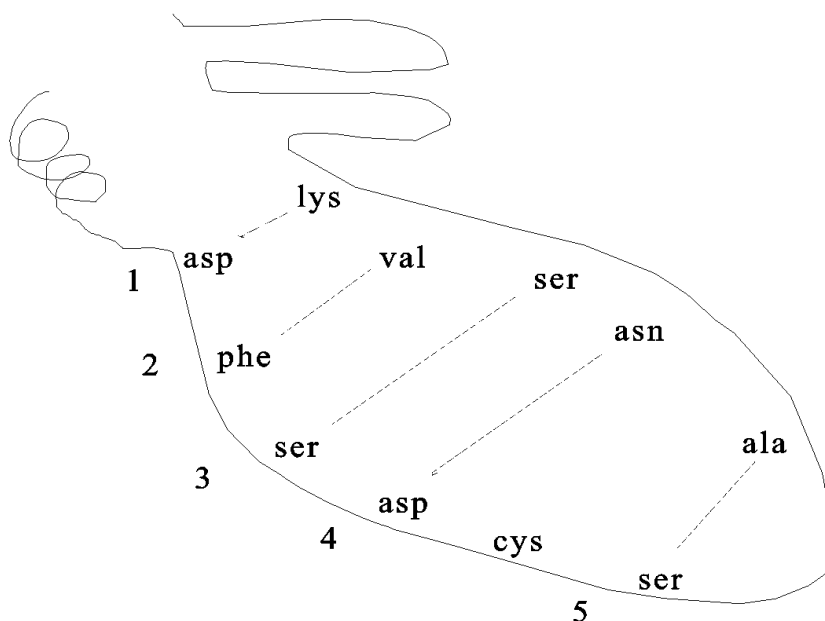
Secondary structure consists mostly of these 2 structures. {Q&A}

To review secondary structure, try problem 2-3 part C.

(Tertiary (3^o) = overall 3-D of a polypeptide)

Tertiary structure means the overall 3-dimensional folding of a single polypeptide chain. For this overall shape, interactions between side chains are very important, as are interactions between side chain and water. A **generality** is that the hydrophilic groups are folded to be on the outside where they can interact with water via hydrogen bonds, while the hydrophobic side chains are collected in the inside of the structure, pushed together by hydrophobic forces. This rule is not at all 100% true, and most proteins have side chains that deviate from this generality. That is, there are hydrophobic side chains on the surface,

but they are intermingled among the hydrophilic groups. And there are hydrophilic groups on the inside, where they are usually interacting with other hydrophilic groups. In fact, it is this interaction of side chains with each other that confers most of the overall 3-dimensional shape on a given polypeptide.



Pictured here are the weak bonds that were introduced earlier. The side chain interaction indicated in the diagram illustrate examples of these various interactions. Consult your text for the exact nature of the side chains:

1. ionic (lys - asp)

2. hydrophobic and VDW (phe - val)

3. H-bond (ser - ser)

4. ion dipole interaction (asp - asn)

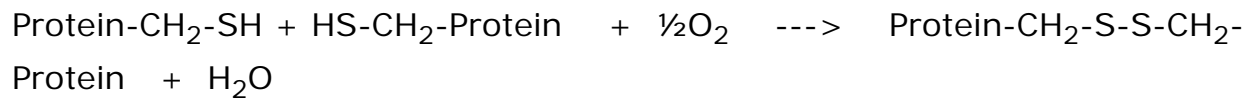
5. van der Waals (ser - ala)

Most proteins fold into a roughly globular shape (enzymes, Hb, antibodies--see [pictures](#) of proteins in your texts: a space-filling model, or showing just the backbone connections, or a ribbon model), but many take on an elongated or even fibrous shape (collagen, myosin [in muscle], fibroin [in silk]).

These are weak bonds, but **in the aggregate**, they are strong enough to hold the polypeptide together at least under the thermal motion conditions of physiological temperature (37 deg C) ([Purves6ed 3.8](#)). [{Q&A}](#)

(Sulfhydryls, disulfides)

But there is one strong bond that contributes to the folding of some proteins. This is the DISULFIDE BOND, and it differs from all these other bonds in being a covalent bond. It can only be formed between the side chains of two CYSTEINE residues. The side chain $-CH_2-SH$ contains a SULFHYDRYL group: $-SH$. Two sulfhydryls can react with oxygen to lose their 2 hydrogen ATOMS (H with its electron; not H^+ ions, not protons) and become bound to each other in the process:



So now the two sulfur atoms are sharing electrons in a strong covalent bond. This bond cannot be broken by mere thermal energy, and so disulfide bonds hold the two parts of the polypeptide chains that had contained the two cysteines firmly together. Not all proteins have disulfide bonds, but many do.

This reaction is an example of an oxidation-reduction reaction: the sulfhydryls are getting oxidized (here losing hydrogen atoms is oxidation), while the oxygen is getting reduced, (or gaining hydrogen atoms) and ending up as water. This reaction can take place rapidly with no further help from catalysts.

(Note that it is not a hydrogen ION (proton, H⁺) that is being moved about here, but the hydrogen ATOM with its electron. Actually, it is the electrons that accompany the hydrogen atoms that are fundamental to the definition of oxidation/reduction, rather than the hydrogen atoms as a whole, as we shall see later. That is, oxidation is the loss of electrons, and reduction is the gain of electrons, with or without an accompanying hydrogen ion.)

[Add another cys and -S-S- bond in folding diagram above]

The net result is tertiary structure, or the overall 3-dimensional shape of a folded-up single polypeptide. Note that there will be many regions of secondary structure within this overall tertiary structure. It is the interactions of the side chains that are to a large extent responsible for preventing the whole polypeptide from simply becoming 100% alpha-helix or 100% beta-sheet. (see [picture](#) of the tertiary structure of proteins represented by a folded ribbon, a space-filling model, a charged surface model, and by simply tracing the backbone).

(Show tubular model ball with alpha-helices and beta-sheets within it, simulate a Jacuzzi environment)

To review tertiary structure, try problem 2-3 part D.

((Quaternary (4^o)= association of multiple polypeptides)

Tertiary structure describes folding of a single polypeptide, and while many proteins do consist of a single chain, most are composed of **several** distinct polypeptide chains. The association of these separate chains is known as **QUATERNARY STRUCTURE**.

The number of polypeptides in a protein can be 2, 4, 8 or higher. Or 3 (rarer).

These chains are folded up in 3-dimensions, assuming a tertiary structure, and then are stuck to each other. What keeps them stuck together? The same answer as usual: those weak bonds we keep discussing, and more rarely, the covalent disulfides.

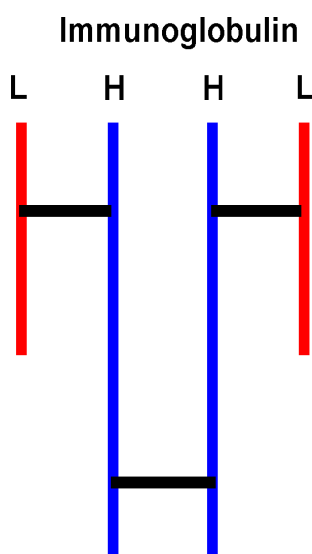
Proteins with quaternary structure are called **MULTIMERIC** proteins. Individual polypeptides are called **SUB-UNITS** (of the protein). [{Q&A}](#)

One polypeptide chain can be considered a monomer, relatively speaking. A protein with 4 chains a tetramer. Etc. The subunits can be identical (called **HOMOPOLYMERIC**) or they can be different polypeptides (or **HETEROPOLYMERIC**).

Now we can distinguish a "**protein**" from a "**polypeptide**". In its **native** form, the macromolecule is called a protein, and may consist of one or more polypeptides, depending on the protein.

E.g., Hemoglobin, Hb, has the structure $\alpha_2\beta_2$, consisting of 4 polypeptides, 2 alpha chains and 2 beta chains, of MW 16,000 each. So the MW of the Hb protein (a tetramer) is 64,000. [\[Purves6ed 3.7\]](#)

If you denature a multimeric protein, the MW will change (unless the subunits are held together with disulfide bonds and you don't disrupt them), e.g., the MW changes from 64000 to 16000. [{Q&A}](#)



The subunits of some multimeric proteins are held together by disulfide bonds (in addition to the usual weak bonds). For example, the antibody molecule, immunoglobulin, is a tetramer of two identical "heavy" chains (H) and two identical "light" chains (L), or H_2L_2 and it includes S-S bonds between the H and L chains. You must denature AND reduce the disulfides to get the individual subunit polypeptides dissociated from each other.

So the surfaces of polypeptides have also evolved to allow interaction with other particular subunits but not with other proteins in general.

Consider now Sickle Cell Disease again. Hemoglobin is a tetramer of 2 pairs of identical sub-units: $\alpha_2\beta_2$. Glu --> val was the a.a.

change comparing normal Hb to sickle Hb (HbS). The result is that the tetramers inappropriately interact, presumably via hydrophobic interactions that in normal Hb is precluded by the charged glu. In HbS this position is valine and now has a more hydrophobic patch of surface. The result is that these patches can now get stuck together by hydrophobic forces, aided by the fact that each HbS molecule has two such patches (one for each beta chain) and that the concentration of Hb molecules inside a red blood cell is very high (they can almost be viewed as bags as Hb). You get long chains of tetramers, and these long arrays can distort the shape of the RBC (red blood cell) into a sickle shape. This shape is not as hydrodynamic as the original shape, and the RBCs can now get clogged in small capillaries, the manifestation of the disease. One a.a. out of 250 was responsible. Once again we see that proteins are fragile, are often only on the brink of stability.

(Domains)

DOMAINS: One additional aspect of protein structure is protein **domains**. The overall shape of most proteins is roughly globular, but if one looks more closely, one can see that most proteins can be divided into sub-regions that are folded more or less independently of the rest. These folded up globules are called domains. An interesting feature of many domains is that homologous domains can often be found in many different proteins. Many of the individual amino acids in the primary structure are different, but many others are the same, and the overall shape of the domains in different proteins can be very similar. A recognizable domain in a protein can often be associated with a particular function, often the ability to bind a particular small molecule. In the hemoglobin example, each subunit can bind a molecule of oxygen and an oxygen-binding domain of similar structure might be found in several different proteins, each of which needs to bind oxygen to carry out its function. Thus we might also have a glucose binding domain, a phosphate-binding domain or an RNA-binding domain in several proteins whose function requires them to bind these molecules.

PROSTHETIC GROUPS: There are some NON-amino acid components of proteins that are so tightly bound they are considered part of the protein. These small molecules are called prosthetic groups, and in general they are not covalently attached to the protein but rather are bound tightly by a series of weak bonds. These small molecules are usually essential for the function of the protein. For example, in hemoglobin, the "heme" groups are actually organic ring compounds with an iron atom at their center, and it is this iron atom that actually binds the oxygen that is carried by the hemoglobin protein. Some of the vitamins become prosthetic groups (e.g. riboflavin). Metal ions, especially divalent metal ions such as magnesium or zinc can also be considered prosthetic groups if they remain bound during the purification of a protein. See Becker for heme structure (also in the protein structure handout).

(Membrane proteins)

As an introduction into an example of the function of proteins, let's consider first a special class of proteins that do not follow some of the rules we have seen that govern protein structure in general. These are the **MEMBRANE PROTEINS**, the proteins that reside, not in the aqueous environment of the cytoplasm or the nucleus, but rather inside the hydrophobic environment of the membranes of the cell, including the cell membrane.

In these proteins, the hydrophobic side chains are on the outside, as there are no hydrophobic forces present to force them to coalesce. These proteins can usually diffuse laterally in the lipid bilayer of the membrane they can aggregate with each other with specificity, and they can become anchored via attachment to structural cytoplasmic fibers. They can be nearly completely enveloped by the lipid bilayer, or they could be partially immersed, with a more conventional half (i.e., hydrophobics on the inside, hydrophilics on the outside) sticking out into the cytoplasm or on the outside of the cell. See [[Purves6ed 5.1](#); 7th 5.1; Sadava 8th, 5.1], [[Purves6ed 5.4](#); 7th 5.2; Sadava 8th 5.2].

One class of membrane proteins act as channels through the membrane. The channel proteins are formed into cylinders, with a hydrophobic exterior, but with hydrophilic groups lining the hole through the interior of the cylinder. Small molecules can pass through the cylinder, or channel, but large molecules cannot fit through (note: macromolecules generally CANNOT diffuse onto cells). See [Purves6ed 5.9](#); 7th 5.9; Sadava 8th 5.10 and 5.11.

To consider how peptides fit into a membrane, try problem 2-13, part D.

At this point you should be able to do all the problems in problem set #2.

So we have a cellular function for a protein, and a function that shows some selectivity on the basis of size there. How about other criteria of selectivity, like charge?

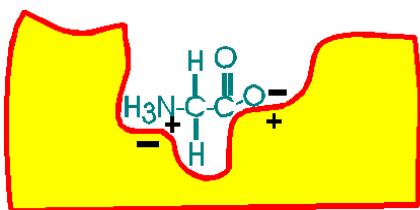
Some channels can distinguish charge: + repels +, and attracts -, so a channel that is lined with positive charge could bind a negatively charged ion and, if there's a higher concentration outside than inside, eventually pass these ions along from the outside of the cell to the inside. A positively charged ion on the other hand would be repelled by the channel and so would not get into the cell by this route.

You can imagine the same sort of selectivity based on hydrogen bonding, for example.

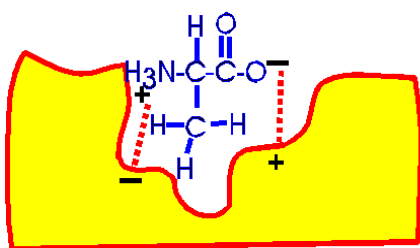
So a protein can detect charge and surface electrical properties, but how about shape?

(Proteins bind molecules with great specificity)

Consider a pocket on the surface of a folded protein. As I've drawn this surface pocket, the free amino acid gly can fit in and bind using the electrical attraction of ionic bonds. However, the closely related amino acid alanine, with a methyl group as a side chain, cannot fit into this hypothetical protein A drawn here (yellow). Ionic bonds and van der Waals bonds (VDW) are responsible for the binding in this example so far (using different colors for protein and aa)



Protein A + glycine:
tight binding

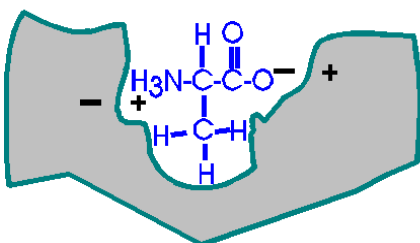


Protein A + alanine:
poor binding

But a similar pocket on the surface of another protein (gray protein B) could be built to accommodate the methyl (the -CH₃ group) of alanine and supply some more V.D.W. bonds there in the process.

So Protein A binds glycine, but not alanine.

And Protein B binds alanine (but gly not so tightly).



Protein B + alanine:
tight binding

So you can get specific binding ... and this binding is critically dependent upon the structure of the protein, the shape of the binding

site. This binding is also a critical part of the *function* of the protein, and we will soon see.

Specific binding at a protein surface is not restricted to interactions between a macromolecule and a small molecule. There is also specificity in the interactions between two macromolecules, as exemplified many times by quaternary structure: the complementary surfaces of the two correct subunits fit together with great specificity; just the right subunits polypeptide specifically associate to form a multimeric protein. For example, the subunits of immunoglobulin (Ig) never associate with the subunits of hemoglobin (Hb).

© Copyright 2010 Lawrence Chasin and Deborah Mowshowitz Department of Biological Sciences Columbia University New York, NY: