

Signaling¹

Brief Summary:

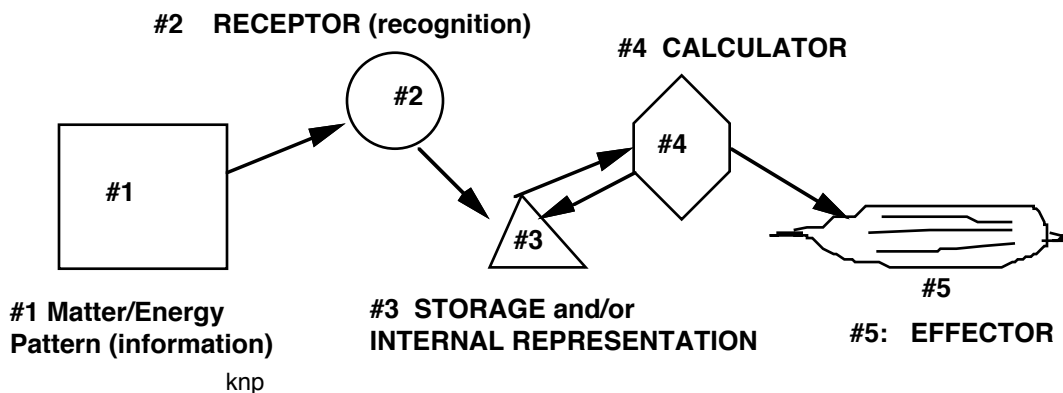
I. An Introduction to Information: You are familiar with matter and energy as the basis of the physical world. However, by themselves they are incomplete. Without sounding metaphysical, another phenomenon, perhaps even more difficult to define than matter and energy but which is just as present and just as fundamental to the physical world is **information**. In this section we will see that understanding physiological systems is in large part a question of understanding how information is handled. At a fundamental level, physiological and behavioral systems are shaped more by the rules of information handling than by the types of information they must handle and the actions they must initiate.

Take a few minutes and read the discussion about information and its many definitions at the following URL: <http://en.wikipedia.org/wiki/Information>. You will see that I present a slightly different but heavily overlapping take on the subject in these notes. You need only read down through the section on information in physics. There is a lot to think about there.

II. INFORMATION THEORY

A. What is information? There are a number of precise definitions involving entropy. However, for our purposes, we can use a more qualitative, biologically useful definition. Information is a particular state, pattern or distribution of matter and/or energy.

1. At no level is information independent² of matter and energy.
2. A particular pattern (information) can be represented by another pattern as long as some other pattern/process is capable of making that recognition. Likewise, the "meaning" of that particular pattern depends on what the recognition system (itself an information collection) does with the information:



B. "Meaning" and Information. Meaning is determined by some sort of mechanism that initiates action based on certain information state(s). Thus, the same pattern of matter or energy (information, we'll call it pattern "x") may have very different meanings and produce very different actions depending on the information content of the system that operates on the

¹ Copyright ©2015 by K.N. Prestwich. These are not for commercial use but their free use and modification is encouraged. Please contact me if you make use of them kprestwi@holycross.edu.

² At least as far as science is concerned -- remember that we are interested in materialistic explanations.

information state "x" in the cartoon on the last page, meaning is determined largely by systems #2 and #4.

? Interesting philosophical questions -- is there information without "meaning systems"? For instance, one might guess that the simplest system would be one that determined that some information state exists.

C. Representations of Information:

1. The simplest such system would be a condition where there are:
 - **two possible states or patterns of matter or energy**
 - these states are **not random**

! Notice that the objects or the energy are not themselves (for the purposes of this example) the information. The pattern they are in is the information; they merely carry or represent it.

This simple pattern, which can be fairly represented as **yes vs. no** (or 0 vs. 1, or + vs. - , hi vs. low -- you get the idea) -- is referred to as a **binary** system. (a) notice that without further modification, this system can only contain two pieces of information.

(b) however, if we add the additional concept of **combinational rules** where the method by which bits are combined has a certain meaning, then, in theory it is possible to represent any information with just two alternative entities (e.g., 0 and 1) so long as they repeated according to some sort of rule. The most **common combinational rules** are simply ones of:

- (i) **relative spatial position** or
- (ii) **temporal order**

(c) Let's assume that position in left to right sequence has meaning. Lets assume that we need to represent 4 pieces of information using a binary system. Note that as before, information is made up of unique states or patterns. Therefore, to represent 4 different things, we need four different patterns. Using just 1 and 0 as symbols, we can generate 4 different patterns:

00, 01, 10, 11.

Each of these is an information state itself and therefore each can be used to represent information.

The case above is an example of what is called a **two-bit**³ system. A **bit** is simply a pair of alternative states. Thus a one-position system with two alternative states contains one bit of information⁴; if two positions are involved, it is a two-bit system and it has the potential of 4 informational states. More generally, the number of informational states in a binary system is given as:

eq. 1. # of states = 2^n

³ Two-bit does not mean "cheap" in this case!

⁴ The term "bit" often is also applied to the two alternative states.

where **n** refers to the number of positions or places. Thus, a 16-bit system can potentially represent 2^{16} or 65,536 states.

! Notice that the exact meaning of these 16 place representations is up to whatever decodes them. For instance 0000000000000001 could just as well represent 1, 61113, or strawberry shortcake. Don't get hung up on the meaning of these symbols according to our usual rules of numbers, counting and order. All that matters at the moment is that there are some rules -- not which ones!

2. "**Information density**" -- the amount of space or energy needed to write a piece of information. So far we have worked on the assumption that unique pieces of information are only built up with various combinations of one symbol (or two but where one of the symbols may simply be the lack of a symbol in a certain place). However, patterns can be based on any number of unique informational representations that are organized spatially and/or temporally. You are reading this page by using one such system that consists of well over 50 unique states (letters, punctuation marks and spaces, and numbers). Where more states occupying a certain extent of space or time can be produced, then potentially more informational states can be generated in a given amount of space. The representation of the information is denser.

a. note however, that in all cases each piece of information still has **one unique representation**, e.g., eight, 1111, and 8 (we have decided that these stand for the same things; in each system the concept eight has one unique physical state).

b. note also that a system with any number of different possible states at one position, whether that is 2 or a very large number, can in theory be used to represent any possible information state.

A more generalized form of the equation that predicts the number of possible states (see eq. #1 above) is:

eq. 2. # of unique states = S^n

where S is the number of states at any time or location and n is the positions or iterations included in the description. Thus, using 2 positions with 10 possible shapes at each gives 100 possible unique states

? How many unique shapes are possible in a position in DNA? What is the role of the sugar phosphate backbone in the storage of information in nucleic acids. Does the backbone contribute directly (in terms of its shape) to the information content? How about the nitrogenous bases?
Assume that amino acids and nucleic acids are about the same size (in fact n.a. are a bit bigger). On a positional basis only, which chemical could represent more information in a given amount of space?

D. Computation:

1. When we deal with information systems, it is customary to refer to any sort of transformation or alteration as a **computation**.

2. When used this way, however, a computation does not necessarily need to be some sort of numerical operation. **Any change of state due to some sort of interaction is referred to as a computation.** The interaction can be viewed as a mathematical function of some sort and can generally be approximated by a function.

3. Thus, any chemical reaction involving the change of one or more substances into a new structure or new energy state is analogous to such a function.

We can write the **reaction of A to give B (or vice versa)** as:

eq. 3. $A \leftrightarrow B$

we can show the energy transformation equivalent as:

eq. 4. $PE \text{ state of } A \leftrightarrow PE \text{ of state } B + \Delta \text{entropy}$

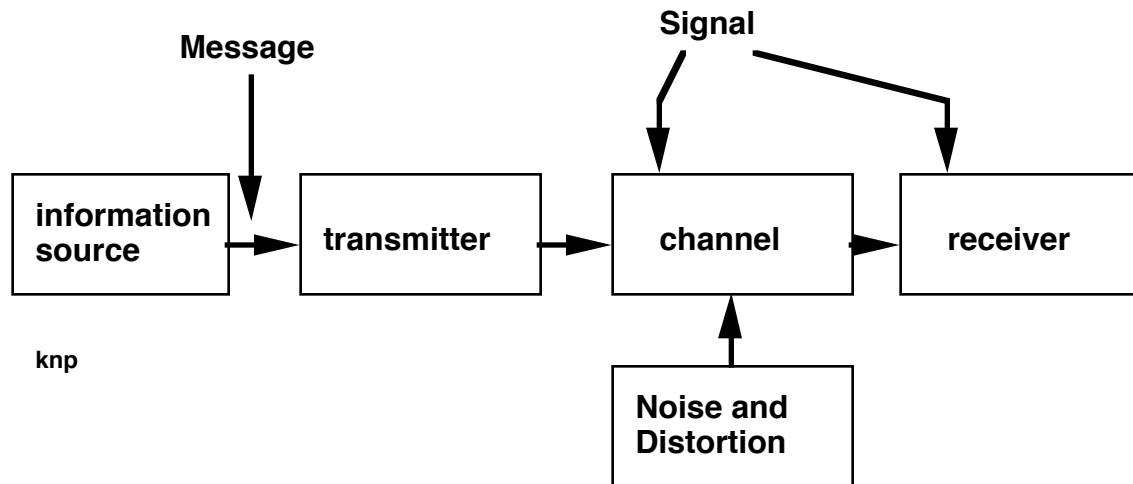
The important thing to realize is that **any process involving matter also involves energy and involves a change in information (pattern).**

? Very Important: Does B contain more information than A? Explain your answer.

Communication: A concept that is closely related to information and that shares many central concepts is **communication**. Since information gathering, storage, computational and effector elements of a physiological or behavioral system are always (at least some extent) physically separate from each other (even when they all occur in a single cell), there must be some way for these systems to communicate with each other.

! Communication is a very important aspect of biology, philosophy, and psychology. Entire lives are devoted to its study. Language studies, rhetoric, etc. are all vitally important endeavors. **What follows** is in no way meant to simplify the richness of communication anymore than the previous section was meant to do so with information. But be aware that it is in some ways a very **simple treatment**. For instance, **it supposes that all communications are honest** -- they are meant to correctly convey information. **This ignores the fact that some communications are meant (at some level) to be deceitful.** The reason that we will ignore the possibility of dishonesty (which, let's face it is one of the most important parts of communication) is that we are only concerned in this course with communication of one part of an organism with another. No advantage is to be gained by dishonesty -- if one part gains some other part loses or worse yet, the entire individual loses. It is an entirely different story when whole organisms are considered.

A. Information theory is about the process of communication. A simplified generalized diagram of such a process is given below:



Note: the elements in this system do not necessarily correspond to distinct cells, proteins or pathways. We will see a number of examples where the same physical entity, at least some level, appears to take on two or more functions.

B. Now let's look at the elements of the system:

1. In some ways the most important element of the system is the **RECEIVER** since its capabilities are what allow information to be identified as such. Receivers filter through the vast amount of potential information and focus on relatively specific manifestations.

2. The **INFORMATION SOURCE** and **TRANSMITTER (XMITTER)** are often the same or are certainly closely connected entities. If they are the same, they are simply something that produces at least two discernibly different (to a receiver) states. More formally,:

a. the **SOURCE** creates the information as something we refer to as a **MESSAGE** (information that can potentially be received and acted upon) and passes it on (internally) to the xmitter.

b. The **TRANSMITTER** takes the message and **TRANSDUCES** (changes) it to a different form and then "broadcasts" the information into some portion of the surrounding environment.

3. At this point, the message is now referred to as a **SIGNAL**, with the difference being that the message can be thought of as information located in some place while the signal is a representation of that information that moves.

4. The particular physical or chemical means by which the message propagates is referred to as its **CHANNEL**. For instance, messages that are converted into changes in concentration of some chemical move through the **CHEMICAL CHANNEL** (hormones, neurotransmitters, etc.) or those coded into different patterns of membrane voltage changes as in excitable cell communication, move through the **ELECTRICAL CHANNEL**.

5. Notice that the signal may be **DEGRADED** -- that is, the information contained in it can be lessened or even changed. There are two general ways this can happen:

a. **NOISE** -- random, unpredictable signals that mask the signal of interest to sender or receiver.

b. **DISTORTION**: specific operations that in a potentially knowable manner alter the signal.

c. the process of removing or minimizing noise and distortion by the receiver is termed **FILTERING** This is a very important process and one that generally needs to

be minimized within an organism if possible -- it is far better (i.e., more **reliable**) to generate signals for internal use that are reliably passed than to try to reconstruct them at the receiver.

Example of Noise and Distortion: Using the chemical channel, a noise would be other chemicals that would potentially interfere with the signals detection by a receiver; an example of distortion would be a situation where a signal that was supposed to persist a certain period of time (example -- a neurotransmitter) persists for a longer or shorter period.

C. Given these elements, let us now reconsider and refine what we know about information:

1. The presence of a suitable receiver is required for information to have any real meaning -- without it, we can talk about potential information at best.
2. In order for information to be transmitted:
 - a. There must be at least two different states produced in the channel.
 - b. **Difference is, in part, a property of the receiver.**

? Can you think of any less complex system capable of storing and transmitting information?
? Does the transmitter need to be "purposeful" to send information -- in other words, does the signal (and message) need to be designed by the transmitter in some sense (even if the sense is simply that of a system designed by natural selection).
? Why does the receiver seem to have such a central role?
? Can receivers create information or merely exploit what is already present?
? What is the difference between random and regular? Irregular?

D. Discrete vs. continuous information revisited: Simple information systems that are built entirely around **DISCRETE STATES** (nothing in-between has any meaning) are referred to as a **DISCRETE INFORMATION SYSTEMS**.

1. The simplest discrete systems are built around **two states**, generally one is random background (the absence of a signal) and the other is some well-defined perturbation. Such a system is referred to as being **BINARY** (see above) and can in theory be used to convey any type of information.

2. **MULTIPLE STATE SYSTEMS** can also be devised.

Nevertheless, the key to any discrete communication system is that some set of rules exist that make certain that either discrete signals are produced and/or continuous signals are converted by some rule into discrete signals by the receiver.

3. By contrast, **CONTINUOUS INFORMATION** implies more of a shade of meaning or change and any assignment to some category of "meaning" (for example, one of two binary states) by a receiver is somewhat artificial and results in some loss of information in the signal.

? Which is more susceptible to noise and distortion -- a discrete or continuous system?
? What are the principle advantages and disadvantages of multiple state vs. binary systems? Can either one represent information that the other cannot?? Give examples of things that you believe are best represented as continuous entities. What are things that we experience that are certainly discrete?

? Can continuous information types be represented by discrete information systems? If that is possible, what would determine how good the representation is?
? You often hear the terms digital vs. analog used as synonyms for the two terms we are discussing. Which goes with which? Are they appropriate synonyms?

III. Chemical Systems as Giant Computational Systems:

A. Introduction. We will start with chemical communication systems since they are central to understanding the operation of nervous systems. They are probably also the **oldest form of communication and computational system used in living systems.** **Ultimately, all biological communication is chemical.** Furthermore, as we will soon see, it is the proteins that are doing the signaling and computing!

? Why do you suppose that chemical communication be considered the most primitive form of communication and computation?

B. Overview of chemical computers:

1. Any chemical system above absolute zero is constantly in flux (in fact, such systems are even in flux at absolute zero due to quantum effects, but the flux is minimal).
2. Each second, myriad interactions (operations in our informational terminology) occur.
3. Many of these interactions are totally random and tend to reverse immediately. They represent the noise of such a system and are not our focus. These are the **entropy component** of the system.
4. However, many of the interactions are in some sense more directed and their results may feed into other operations some time later.
 - a. All of these interactions (including the entropic ones) produce local results (informational changes) as the interaction occurs. And there are potentially many such changes in each molecule. On the other hand, random changes produce no macroscopic change in the overall system and so are not associated with information at that level.
 - b. Thus, we can speak of **any chemical system as being a potentially very complex, real time, massively parallel (and serial) computer.** It is our challenge to see how given all of the potential information that could be handled by such a system, certain factors are ignored and others are amplified and operated on.
5. **Consider (again) the information process/chemical reaction.**

A <---> B

This process above can be viewed as a process of accumulation (B accumulates and A decreases) and it is possible to attach meanings to accumulations of amounts of B. Note that "B" could be anything -- a simple molecule, a protein in a particular conformational state or whatever. But regardless the chemical nature of B, we will always see that proteins are crucial to this process. Hopefully, this will become clearer below.

Please realize that in information computations/transmission, it is not necessarily the amount of "B" (or whatever) that is of primary interest. Rather it is the information that the amount of B represents.

B. Proteins and Chemical Computation -- A Review of the Relevant Features of Proteins

1. Recall from other courses that **protein structure determines function** and it is a result of the interplay between the interactions of amino acids making up the protein and the environment.

a. The structural gene that codes for the linear sequence of amino acids making up a particular protein blindly "expects" a certain environment in order to produce a protein with a certain three dimensional shape.

b. The function of the protein is determined by the exact shape of the protein -- produced by this interaction of the primary amino acid sequence and the environment.

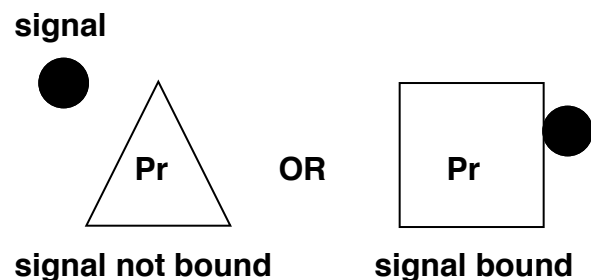
Note: for the discussion below, imagine that we are considering a type of protein whose role it is to respond to a signal when it reaches a certain level.

2. Functions of individual members of a protein species vs. function of the population of that protein species.

a. A **protein species** can be defined as all molecules with the **same primary protein structure** (same amino acid sequence). Put another way, they are all derived from the same allele (or multiple copies of that allele).

b. We need to be familiar both with how individual molecules work and how a protein species functions as a whole or population in processing information.

(i) **Individual function:** each molecule of a particular type of protein can exist in several conformational states. To make this simple, we will assume that the protein we are interested in has two possible shapes (functional levels) that correspond to whether or not the signal molecule is bound:

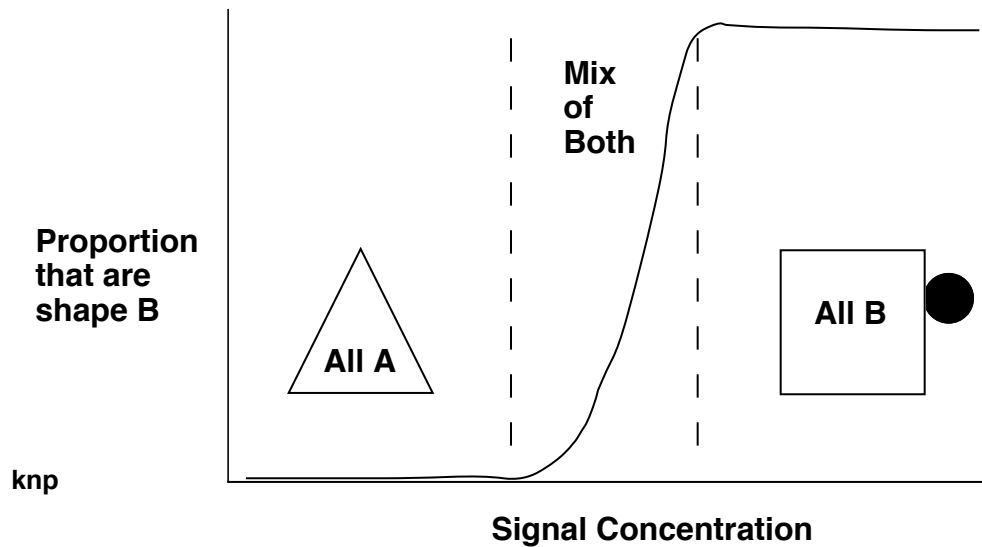


Thus, the protein exists in **one of two shapes** and these shapes register the presence of a signal. You can think of these states as No and Yes in terms of information.

Note that in the example above, the protein Pr is essentially monitoring the concentration of signal. In other systems these same types of differences could be achieved by something that acted directly and specifically on protein Pr and caused an allosteric change -- for example and

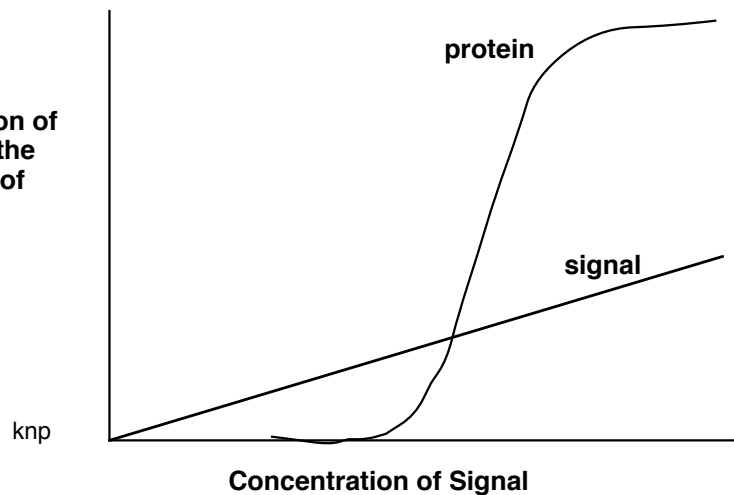
enzyme that cleaved off part of Pr or one that attached a small molecule such as ATP or phosphate.

(ii) **Population function:** Normally there are multiple copies of the same type of protein present in a cell. Thus, the protein responds as a population to the signal and since not every protein will bind with the signal and change shape at exactly the same time (local concentrations differ as do local conditions that affect the ability of the proteins to bind), the behavior of the population is a curve:



! In the example above, the protein would need to bind several signal molecules to be fully converted from one shape to the other (the S-shaped curve is associated with multiple, cooperative bindings).

Notice that the smooth continuous (analog-like) increase of the signal (x-axis) has been converted by the protein population into a much more discrete signal:



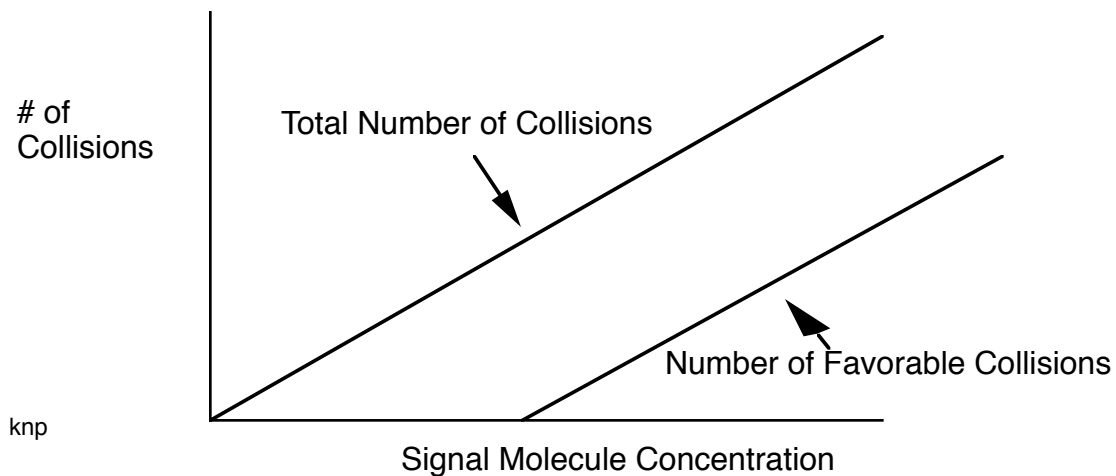
The shape of the curve is related to two important characteristics of proteins and protein populations: affinity and number.

3. Proteins and affinity – let's reiterate what we just saw

a. In order to have an interaction between a protein and some other substance, they must come into contact. Furthermore, the contact must last for some period of time.

i. As concentration increases, favorable collisions always become more common. This would be true in any system.

ii. If this was all that mattered, then a simplistic (but illustrative) way to view this would be to envision a linear relationship between concentration and the probability of binding signal molecule to protein where the number of favorable collisions is some subset of total collisions:



b. However, it is not this simple because when a favorable collision occurs there is a certain probability that the protein will hold on to the signal long enough for a shape change to occur.

! Keep in mind that the binding between ligand (signal) and protein is usually (but not always) weak (H and vander Waals) -- it is generally not covalent. As a result, ligands are relatively easily displaced from the binding site.

i. This probability function is synonymous with a property we call **affinity**.

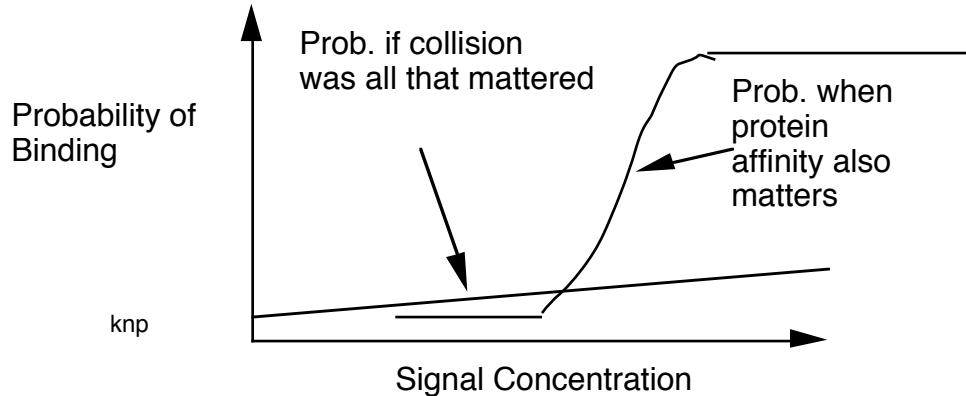
ii. Affinity itself is simply a measure of the strength of binding between the protein and the ligand (signal).

(a) thus, for a given favorable collision rate, a high affinity protein is **more likely** to possess a bound ligand at any moment in time than is a low affinity protein.

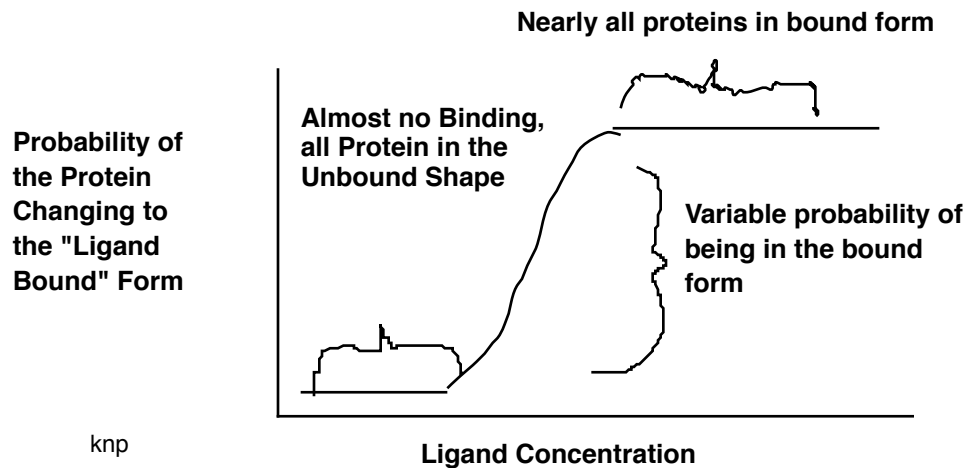
(b) Viewed as a population, in a high affinity population a greater proportion of the protein molecules will possess bound ligand for a given concentration (favorable collision rate) than will a low affinity protein population.

! Note that the preceding discussion only applies to concentrations that are low enough that not all of the binding sites would be occupied (saturation, see below).

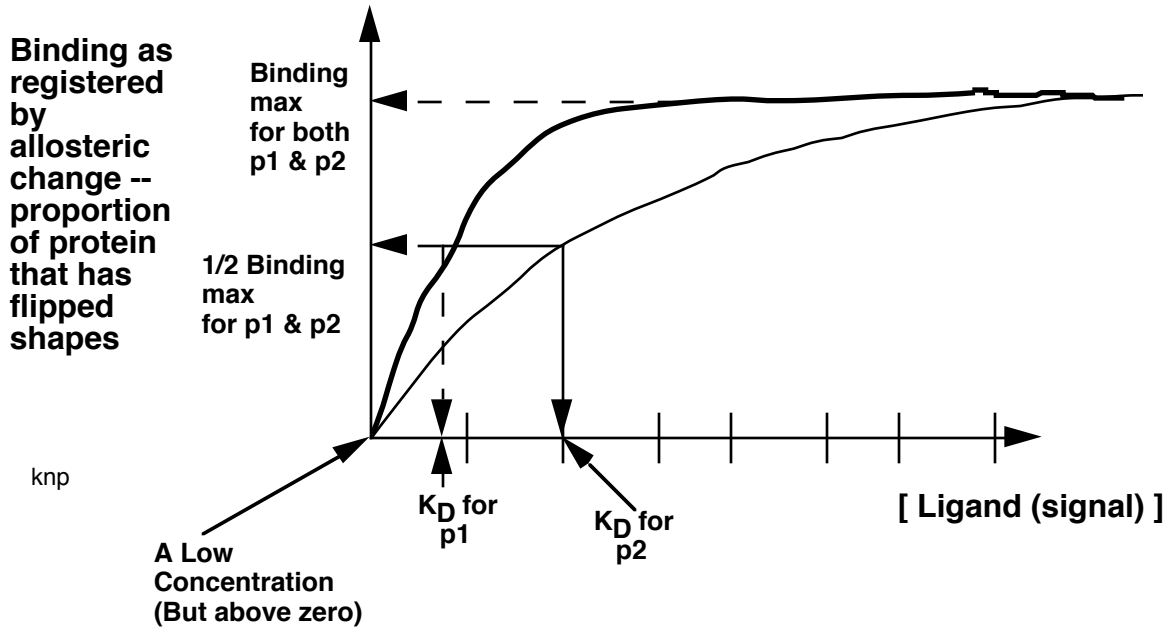
iii. Thus, at relatively low concentrations, not only is there a low chance that the signal will attach to the protein correctly, but also, since the signal does not bind covalently to the protein there is a chance that it will be broken loose before any shape change occurs. The more strongly the protein binds the signal, the greater the chance that it will remain long enough for a shape change to occur. The strength of this binding therefore determines the affinity and the affinity therefore helps to determine the likelihood that a protein will be in one shape or another at a given concentration of substrate:



Taking these factors into account we see that a population of proteins will behave such that at low concentrations very little ligand will bind long enough to induce and continue a shape change; at high concentrations, ligand is nearly always bound and therefore nearly all of the protein is in the shape associated with binding and at intermediate concentrations there is a rapid transformation of the protein population from one shape to the other:



Saturation is the point where, even though more collisions are still occurring, we see no further increase because none of them involve the binding site. The graph below shows the binding curves for two different proteins that both bind the same signal. Notice that at concentrations below saturation that the higher affinity protein always has a greater proportion of its population flipped into the shape consistent with binding:



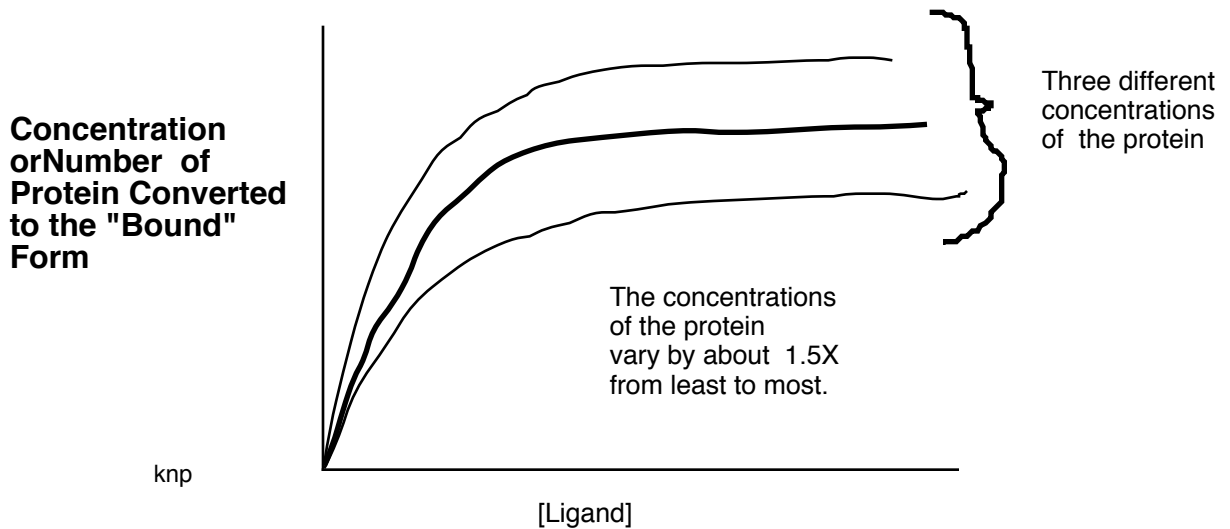
Note the parallels between this curve and the typical curve for enzyme kinetics. Notice that the dissociation constant, K_D , is an analog of the K_M and numbers of proteins that have changed shape is the analog of V_{max} . In "chemical" biological computation both enzymes and binding proteins will be important.

? What is the role of affinity in biological information processing and communication?
 From what you know about proteins, how can the response properties (at the protein level of organization) of a biological computer be modified on both the short and long terms?
 In what sense does a protein act to "digitize" a continuous signal into a binary?

4. Signal Size and Proteins:

a. We have just seen that if a protein is being used to register the existence of some state in a cell, the decision on when to register an event is largely tied to the affinity of the protein for the signal.

b. Given that a protein has registered a change, an important factor is the **magnitude of the registration signal**. This is primarily determined by the concentration of the protein since the greater the concentration, the greater the potential number of signal molecules registering a change:



c. This is another **population aspect**. The maximum potential size of the signal is controlled:

- i. at the **level of protein synthesis** where the control of transcription and translation results in a certain number of proteins being produced
- ii. at **regulatory levels associated with the protein molecules themselves whereby certain cellular conditions, essentially signals, change the functionality of the proteins.**

C. MOLECULES, INFORMATION AND COMMUNICATION: A BINARY MODEL.

1. For the simplest model we will assume that **each element** in a computer or communication system can exist in **one of two STRUCTURALLY DISCRETE STATES**

- a. One state corresponds to the **ABSENCE** of the signal.
- b. The other corresponds to when a **SIGNAL EXISTS AT SOME CRITICAL**

LEVEL.

NOTE: These states are different shapes (proteins) or numbers of certain shapes (any elements). In the case of proteins the shapes can be associated three different important actions:

- (i) Changed catalytic properties
- (ii) Changed abilities to bind to other molecules such as proteins, in a non-catalytic way.
- (iii) Changes in shape that alter the ability of other substances, such as ions, to move from one place to another via the protein. In this case the protein is a special type of ion channel called a gate. (this feature will be important when we consider what we will arbitrarily designate "bioelectrical systems")

2. The elements **TOGGLE** between these states as a result of **recognizing** something as a signal.

a. Thus, with chemical signals, a protein must **BIND** or be **INFLUENCED**⁵ by some **signal**.

b. **Recognition is registered by a shape change. Shape changes are synonymous with function changes.**

! Note the first feature here -- **selectivity**. Of all possible information, given elements of an informational system only pay attention to certain environmental events. As a result they ignore non-essential information and further, they can be optimized to deal with certain types of information. Think about how many things you ignore -- most of which you do without ever even knowing that they exist! Related terms are **threshold** and **stimulus filtering**. This is a fundamental feature of all living systems -- it is attributable to protein structure.

c. Clearly not only selectivity but also affinity and signal size are important at this step.

3. The result of the toggling of structure can be to:

a. produce a signal that symbolizes that a certain state has been reached (see earlier discussion of proteins responding to ligand signals); this signal could be used in further computations or stored (next item).

b. store information (arbitrary signals) -- e.g., the presence of a particular structure might be used to mark that something has happened.

c. directly trigger a process that responds to the signal, for instance by activating some biochemical pathway(s).

D. A bit about the evolution of signaling and computational systems

1. In the simplest situation, for a signal to evolve all that is needed is for some protein to be able to bind it. If that happens, then, in one sense there has been a form of chemical communication since the protein recognized the signal.

2. However, it is hard to see how selection would increase the frequency of the allele for a protein that recognized some chemical unless it somehow altered the organism's function in a way that increased its fitness. Thus, the affected protein, the receiver, will need to fit into or affect some larger physiological system. If and when this happens selection can evaluate and perhaps improve the receiver.

3. In general, if a communication system based on some protein receiver becomes established we will expect the following to happen:

a. There may well be an increase in receptor **AFFINITY** AND **SELECTIVITY**. Both of these would be achieved by a modification of the amino acid sequence -- therefore they both depend on mutations to structural genes.

? How long would affinity continue to increase? What would determine whether or not selectivity increased?
Try to frame your answer in terms of optimality -- are greater and greater affinities and selectivities always better? Come up with a graphical way to illustrate your answer as far as affinity is concerned -- use "improvement in affinity" vs. whatever variables you thought were important in establishing an optimality (one or two examples).

⁵ An example of influenced would be something like pH or temperature that that could affect protein function without actual binding.

b. Improvement in the degree of response by the individual receivers -
- how unambiguous is its response? (conformational change eventually leading to other effects). If the original conformational change is small and results in a relatively small affect that does not necessarily always trigger other changes with which it is paired, it is reasonable to assume that there will be selection:

- i. on the receiver to improve the "distinctiveness" of its response to the signal
- ii. on the component(s) cell system(s) directly affected by the receiver to make them better able to recognize the change. (Notice that this is exactly analogous to the process occurring between the receiver and the environmental signal. Only here, since it is internal communication we are talking about, it is not only the receiver (the affected process) that is acted on by selection, but also the signal (the receiver aimed at the environment).
- iii. both of these would be achieved by modification of the structure of the two types of receivers. this could be achieved either by different amino acid sequences or different internal environments.

c. Improvement in the overall magnitude of the response to a signal. If the receiver protein was originally rare, there may well be selection to increase its numbers and thereby generate a greater receiver response for a given size of signal. This is largely a question of gene expression -- rate of transcription and translation vs. rate of removal.

? How does this relate to signal to noise ratio?
Can you think of times when it would not be important to increase the total number of receivers even if they were relatively rare?

We will apply these principles to the analysis of real systems over the rest of the semester!

Note that the description of the evolution of signaling within an organism has strong parallels to the ways that communication evolves between organisms. The selection pressures with respect to information transfer are the same at either scale.

D. Rapid Intracellular Communication -- Electrical Fields: You are certainly aware of nerve impulses. We are now about to learn a lot about their nature. First, let's look at what they really are in terms of the communication model just considered. The signal in this case will be a significant change in electrical potential (voltage across the membrane). So, just like with the chemical example above, the change must be of at least of a certain size in order to matter. This change in itself means nothing except that something (who knows what!) has happened. Other mechanisms will be needed to attach meaning. The receiver in this case will be a population of proteins that can themselves help to generate electrical potentials. Or, we will see that in some cases, the receivers are proteins that generate chemical signals. But in any case, the receiver proteins will once again act in a binary fashion --either they will respond to the signal and do what they are supposed to do, or they will ignore the signal and do nothing. Once again, they will therefore exist in one of two states.

To understand this type of signaling, we are going to need to understand the ionic system that is at the heart of forming the signals. To do that, we are going to go back to our knowledge of the Donnan equilibrium. Here we go, along with some history. As you read this,

try to think about it not only as a series of events you need to understand and learn, but also as a signaling system.