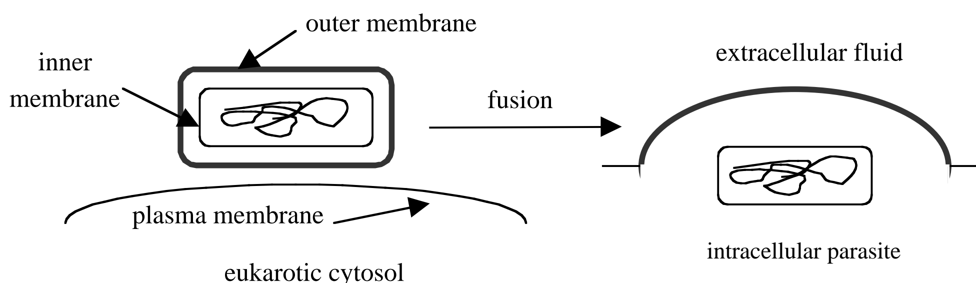
May 20, 2002

ID number _________

Biochemistry Cumulative Exam

This is a thought problem. You will be given imaginary data on a nonexistent organism and fashion a reasonable model to explain the data. Have fun with it!

Thyrobacillus rochanhors is pathogenic bacterium that undergoes an unusual life cycle. It can be found often in a non-virulent form on the skin, but it undergoes an amazing transformation if it enters the body via a cut or other injury. Soon after entering the bloodstream, it sheds its cell wall and remodels its outer membrane. (Like most eubacteria, *T. rochanhors* has an outer membrane and an inner membrane.) Somehow the remodeled outer membrane has the ability to fuse with the plasma membrane of an endothelial cell lining the capillaries. The remainder of the bacterium (inner membrane with intact cytosol within) enters the cytoplasm of the eukaryotic cell, while the bacterial outer membrane remains incorporated with eukaryotic plasma membrane.



In this form, it is an intracellular parasite, although a relatively mild one. It does not kill the cell in which it hides; it replicates relatively slowly. It has been hypothesized that this behavior is part of a repertoire of responses that keeps the immune system from recognizing the parasite. If the infected cell is recognized by the immune system and forced to undergo apoptosis, the intracellular bacteria are released upon cell lysis.

It should be obvious that this transition from a free-living organism to an intracellular parasite is going to involve large changes in gene expression. For one thing, there would disastrous consequences if the cell were to jettison its cell wall prematurely. In this problem, you will be given (imaginary) experimental data, from which you will fashion a model of how the transition to a pathogenic form is regulated.

1) Researchers reasoned that the *T. rochanhors* cells had to know when they were in the bloodstream. If they are added to blood serum (*i.e.* no blood cells), the bacteria jettison their cell walls and start to remodel their outer membranes. Searching for the cue read by the bacteria, they radiolabeled blood proteins by iodination with ^{125}I and looked to see if any were bound to the bacteria. One major protein of 62 kDa

was bound by the *T. rochanhors* cells. The position of the 62-kDa protein on a 2-D gel identified it tentatively as serum albumin. To test this, they added serum albumin to *T. rochanhors* cells growing in artificial medium. They found that a significant number of the cells would lose their cells walls. They could not trigger cell wall ejection with other proteins that they tested. Perhaps the most important experiment was to use serum from which albumin had been depleted (using an antibody raised against the albumin) -- this was ineffective.

Someone wonders if the albumin detection signal involves an intracellular second messenger. Using similar techniques to those used to find "magic spot" (ppGpp or guanosine tetraphosphate, the signal used in the stringent response), the researchers allow the bacteria to take up ^{32}P-phosphate for several minutes, and then wash it out and replace it with "cold" medium. They break open the cells, extract the small molecules and separate them by HPLC to see the distribution of nucleotides and other small molecules. If, after labeling, they treat the cells with serum albumin for several minutes and break them open, they see a prominent new molecule that they call "TRA-spot" (for "Thyrobacillus response to albumin"). They do some simple analyses of the molecule after purification by HPLC. If they treat with base, they can convert the molecule to cytidine and phosphate. If they treat it lightly with base, they see that there are two possible intermediates on the way to cytidine: cytidine 5'-monophosphate and cytidine 3'-monophosphate.

What is TRA-spot (draw a structure, if you can)? How is its synthesis triggered? Does this make sense?
Do you think this observation might explain why people on very low fat diets appear to be more resistant to infection by *T. rochanhors*?

2) Meanwhile, the geneticists have found mutants that are not capable of becoming pathogenic; they call them *pat* mutants. They fall into different classes based upon which stage in the life cycle they can proceed to before failing to go further. There are a few mutants that fail very early on, and appear not to respond to serum. They test these and find that 3 of them do not respond to serum albumin. Because of the medical important of *Thyrobacillus*, its genome has been sequenced by now (these days, a new bacterial genome is sequenced every week), which facilitates the identification of the genes affected by the mutations. Antibodies are raised against all of the cloned *pat* gene products.

The *patD*$^{-}$ and *patF*$^{-}$ mutants cannot make TRA-spot after challenged with albumin. The *patD* gene is similar to *E. coli* adenyl cyclase gene, while the *patF* gene is similar to the genes encoding the *tar/tsr* class of bacterial receptor proteins. In eubacteria, these are involved in chemotaxis and other signal transduction process; an extracellular domain binds the ligand, which activates the kinase activity of the intracellular domain. After exposure of cells to albumin, the PatD protein becomes phosphorylated; this does not happen in a *patF*$^{-}$ mutant. If exposed to labeled serum, *patF*$^{-}$ cells do not bind albumin; *patD*$^{-}$ and *patC*$^{-}$ mutants can bind albumin.

The *patC*$^{-}$ mutant can make TRA-spot in response to albumin, but still does not eject its cell wall. The predicted polypeptide displays some similarity to the CRP/CAP protein of *E. coli*. The anti-PatC antibody recognizes a 33-kDa protein. It does not seem to be phosphorylated in response to albumin treatment: its position on a 2-D gel does not change when albumin is added.

The *patC* gene is cloned into a plasmid expressed at high level in *E. coli* and the PatC protein is purified. It behaves as a DNA-binding protein when exposed to *Thyrobacillus* genomic DNA, but only in the presence of TRA-spot. Without TRA-spot, it binds with low affinity to many sites on the DNA; with TRA-spot, it binds with high affinity to a relatively small number of sites (<1000 sites/genome). They cut the genomic DNA with the EcoRI restriction endonuclease to completion, and expose the fragments to pure PatC protein in the presence of TRA-spot. They purify the DNA bound to PatC using the anti-PatC antibody. Running the purified fragments on an agarose gel, they can estimate that there are 50-100 unique DNA fragments bound by PatC. (They also clone the EcoRI fragments for later studies.)

Explain briefly the function of PatD, PatF, and PatC.

PatF:

PatD:

PatC:

3) A new variant of *T. rochanhors* called R-1 is on the rise. Interestingly, it will eject its cell wall when exposed to whole serum but not to serum albumin alone. However, it does make TRA-spot when exposed to albumin. 2-D gel analysis indicates that PatC from *T. rochanhors* R-1 is more negatively charged; it has the same apparent molecular weight, but its pI is lower. Addition of albumin by itself does not change this, but addition of whole serum results in a shift of the PatC spot to the same position as observed in the original *T. rochanhors* strain (called N-1 now to distinguish it from R-1).

Comparison of the N-1 and R-1 strains by genetic analysis reveals that the difference is due to one gene, called *rsgA* ("R-1-specific gene"), present in R-1 but not N-1. Transformation of N-1 cells with the *rsgA* gene will make them behave like R-1 cells. Mutation of the *rsgA* gene in R-1 cells makes them behave like N-1 cells. The predicted RsgA polypeptide is similar to PatF; it is also a membrane protein with the

same predicted topology. Hypothesizing that RsgA must bind something as well, they do the same experiment as before: expose R-1 cells to iodinated serum. They find a new protein that binds to R-1 cells but not N-1 cells. It is 32 kDa and appears to be carbonic anhydrase, another major serum protein. If R-1 cells are exposed to both albumin and carbonic anhydrase, they will eject their cell walls.

How do you think RsgA is exerting its effect?
What is its function? One of the researchers hypothesizes that RsgA represents a "double check" on its status -- because inappropriate ejection of the cell wall while out of the body would be tantamount to suicide, it pays to make sure that one is really in the bloodstream. What do you think of this idea? Is it plausible? How would it work?

4) Meanwhile, the researchers who cloned the EcoRI fragments that bound to PatC (remember them?) have been hard at work. They have been using them to identify genes that might be activated by the albumin response. There is a site that is present in all of the fragments at least one time. The consensus sequence is ATGGANNNNTCCAT, where "N" represents any base. It turns out that a few of the mutants affected in later stages of infection have mutations in genes that are downstream of these so-called "PatC sites".

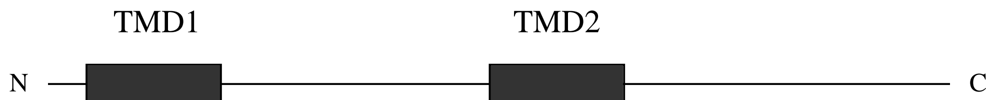
Quick question: do you think PatC binds DNA as a monomer or a dimer?

The genes downstream of PatC sites fall into several classes:

a) Enzymes to remove the cell wall and remodel the outer membrane: hydrolases, glycosyltransferases, enzymes involved in phospholipid synthesis, *etc*.
b) Permeases: probably to take advantage of the rich source of nutrient in the blood
c) Three outer membrane proteins, called PomA, PomB, and PomC (for "pathogenicity-related outer membrane" protein).
d) Two inner membrane protein, both of which belong to the same class of proteins as PatF and RsgA; they are called PimA and PimB (for "pathogenicity-related inner membrane").
e) A protein with similarity to sigma factors, called PsgA.

These genes are known as "stage 1" genes -- they come on when exposed to serum, but are not expressed once the cell becomes an intracellular parasite. At that point, almost all are turned off (with the exception of the permeases, which stay on), and the "stage 2" genes come on. The stage 2 gene products appear to be more involved in co-habiting with the eukaryotic cell, and we will not consider them in more detail, except to mention one detail: their promoters are different — unlike the stage 1 genes, they do not have the consensus site of a σ^{70}-type promoter. (The permeases seem to have an upstream sequence that looks like this alternate promoter.)

PimA resembles PatF and RsgA — they all have a similar primary structure:



In the tar/tsr chemoreceptors, TMD1 (transmembrane domain 1) is also a non-cleaved signal sequence. The sequence after TMD2 in PimA is almost identical to the same region in RsgA, but the rest of the protein is somewhat more divergent. Based on what you know about how signal sequences and translocation work, what is the predicted topology of PimA? Is the region similar to RsgA inside or outside the cell?

5) By deleting the genes for these proteins, they get some clues to their functions:

pomA$^-$ & *pomC*$^-$ mutants: cannot fuse with an endothelial cell

pomB$^-$ mutants: have a weak response to albumin; all of the stage 1 genes are expressed at low level in the presence of albumin. They infect very poorly.

pimA$^-$ mutants: can fuse with the endothelial cell, but do not subsequently shut down the stage 1 genes. As a consequence, they are very poor intracellular pathogens.

pimB$^-$ mutants: inappropriately express stage 2 genes before fusion. They infect very poorly.

psgA$^-$ mutants: can fuse with the endothelial cell, and shut down the stage 1 genes, but they cannot turn on the stage 2 genes.

They can distinguish between the *pomA*$^{-}$ and *pomC*$^{-}$ mutants by a simple test: the *pomA*$^{-}$ mutants can attach to the endothelial cells but cannot subsequently fuse. The *pomC*$^{-}$ mutants cannot attach to the endothelial cells, but if they are forced into proximity by a high cell density, some limited fusion with the endothelial cells can be seen. The extracellular domain of the PomA polypeptide is predicted to have several amphipathic α-helices, while the extracellular domain of PomC is quite large. PomB has a large domain as well, but it is predicted to be localized to the inter-membrane space (*a.k.a.* the periplasm); the portion on the other side of the outer membrane is very short.

If they express PimA under the control of an r-protein promoter (*i.e.* a strong basal promoter) without the PatC site, they can cause it to be expressed without adding albumin. In this case, they find that the cells will still make TRA-spot when exposed to albumin, but the stage 1 genes are not expressed. When they examine the PatC protein, they find that it migrates in a 2-D gel to exactly the same position as it does in an R-1 strain, whether or not albumin is present. This is despite the fact that this is an N-1 strain. (They add carbonic anhydrase as well, because they have nothing better to do, but it has no effect.) If they co-express PimA and PomB in the same strain, then PatC migrates normally, and the cells respond to albumin normally by expressing the stage 1 genes. One more thing: although PimA is similar to RsgA, cells expressing PimA do not bind carbonic anhydrase.

If cells exposed to albumin are loaded with ^{32}P-phosphate, the cells are lysed with detergent, and the PsgA protein is immuno-precipitated with an anti-PsgA antibody, then PsgA is found to be labeled (*i.e.* it is phosphorylated). The following experiment is performed: ^{32}P-labeled cells are allowed to fuse with endothelial cells for an hour, all the unfused cells are washed away, and the endothelial cells (along with their intracellular guests) are lysed with detergent. Now they find that PsgA is still present, but it is no longer phosphorylated. Phosphorylation of PsgA does not occur in *pimB*$^{-}$ mutants. PimB is not very similar at all to RsgA, although it has the same general topology. Subsequent work reveals that the extracellular domain of PimB binds tightly to the Hsp70 chaperone protein.

What do you think are the functions of these proteins?

PomA:

PomB:

PomC:

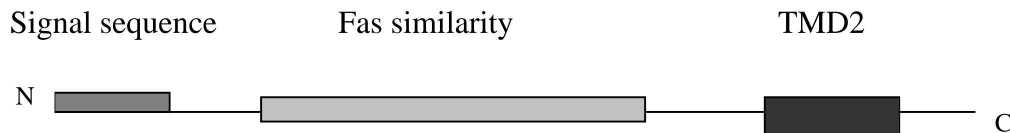
PimA:

PimB:

PsgA:

OK, put it all together. Write out a concise model explaining how *T. rochanhors* can induce the stage 1 genes after getting into the bloodstream, and then after fusion subsequently repress them and induce the stage 2 genes. Include any details you may not have included in your answers above. Speculate on how the how the *rsgA* gene arose.

6) Another new strain is on the rise, called S-1 or sometimes the "sneaky" strain, so called because it is able to elude the immune system better than the normal strain. As it turns out, its superior ability to hide from the immune system is only active when it is an intracellular parasite. While in the bloodstream, it is not any more able to resist the immune system than the N-1 strain. Researchers perform a similar genetic analysis as they did with the R-1 strain and they find that S-1's phenotype is due to a single gene, which they call *ssgA* (yep, you guessed it – for "S-1-specific gene"). The predicted SsgA polypeptide is also a membrane protein, but its primary structure looks like this:



There is an N-terminal signal sequence (with a consensus site for signal peptidase) and a single predicted transmembrane domain near the C-terminus. Between them is a region with strong similarity to the extracellular domain of the Fas receptor. This protein binds FasL ("Fas ligand"), a protein on the plasma membrane of immune cells that can trigger apoptosis of infected cells. Like many so-called "death receptors", Fas has a "death domain" (FADD) in the cytoplasmic side, which can recruit and activate caspase precursors. However, unlike Fas, the SsgA polypeptde lacks the FADD domain. In this regard it resembles the so-called "decoy" proteins thought to reduce the apoptosis response. These decoy proteins are often over-expressed in certain aggressive tumor cells. Analysis of the *ssgA* mRNA demonstrates that it is a stage I gene – it is heavily expressed in the blood, but turned off after fusing with the eukaryotic cell.

Make some educated guesses about SsgA:

- In what membrane would you expect to find it after exposure to serum?
- In what membrane would you expect to find it after cellular fusion?
- What is its predicted topology in those membranes above? Draw a picture.
- How do these decoy proteins work?
- Where did the *ssgA* gene come from?