Important question: In the protein engineering paper, do they use a different potential?

Naveed, Jackups and Liang's predictions of protein-protein interfaces in outer membrane proteins.

They did it by looking for high energy! Easy.

High energy... of pairwise contacts, and solvent interaction!

If it's high energy of solvent interaction, then *any* surface could bind to it, right?  
Are their results because of high energy of solvent interaction, or high energy of pairwise contacts? Which si it?  
My *own* results tell me... I don't know. Moments aren't the same as what he's doing, and are usually right anyway.

And are there heterooligomerization sites that they've failed to predict? Why can they predict the OMPLA interface and not the FepA interface?

THEY DO! Not FepA, but FhuA... what're the data here again?

Boulanger et. al 1996, in Biochemistry: FhuA is a monomer.  
Liu et. al 1993, in PNAS: FepA has both monomeric and trimeric forms.  
1FEP structure determination paper: the molecules do not associate to form trimers in crystal (how did they crystalize two different forms of OMPLA? Did they need to include a particular lipid or was the calcium enough?)  
Locher et. al 1997 in European Journal of Biochemistry: FhuA forms dimers and trimers *in vivo.*

I really want to know what kind of prediction they gave for 1FEP and 1KMO. Later, though, if it becomes important.

Oh, and here's a partial answer... their method *didn't* predict 1QD6! They're not trying to predict oligomers at all... they seem to be trying to predict proteins that are *unstable as monomers!*

But what's the evidence that these porins are unstable as monomers? How can they be validating their method, when they don't say what their method *detects*, and say which proteins in the dataset *have that property*?

Alright. So the trimeric porins have high scores on this interface metric. What is it, exactly? It's related to energy...

NaJaLi include in the supplementary information histograms that count strands. They are binned by difference between strand energy and protein-specific average strand energy.  
The monomer histogram has a peak and then a decay, but the oligomer histogram has *two* peaks, and a higher average.  
Average deviation from protein-specific-average strand energy, then is the metric? Where deviation is always positive - absolute distance.  
That's different than variance - variance is average *squared* deviation from the average. So what's their metric?

But this has to be a typo, because they wouldn't write that it's the square root of a square. They must have meant

Which is the .... um. Hmm.

Without the square root sign, it would be the variance - the sum of the squared distances from the mean, divided by the number of squared distances from the mean. The average squared distance from the mean. Maybe it's supposed to be

Then it would be the standard deviation.

But, I'm not sure what the square root covers. And it's not *n*, the number of strands: it's *nu*, the number of *unstable* strands. Though "unstable" is not defined.  
It says to look in *Methods* for more details, but there *are* no more details, anywhere. That makes me really mad but I'm not going to let them waste any more of my time trying to figure out their stupid index.

Vik says that every computational paper is about an automation of some aspect of chemical intuition.  
I guess the intuition here is that, in their native, folded forms, proteins are as stable as you can get without aggregation. So, if you make an arbitrary disturbance, separate two parts, taht can only *destabilize* the protein.  
But also, interactions in proteins, physical interactions, are local. So before you give the protein time to move - before you let it find a new, mroe stable conformation - it's going to be unstable *right in the part where you made the change*.  
So if you picture a permanent oligomer - you figure it's in its most stable conformation, there as an oligomer. Then you just *remove* most of the protein. If that was a *more* stable conformation, you would think it would be the native conformation. If it was equally stable, you would expect to see both (and you do... so it's not obvious that this would work... what's up with those PorA monomers?)

But if there's no monomers in vivo, then if you correctly simulated the in ivo environemnt, then separating the protein would have to be, would have to be destabilizing. And before you let the proteins adapt and find a new conformation, the destabilized part would be the part *atg the interface*. If it was more stabl;e, it'd be the native state.

Which brings us to the question - well - what about oligomers that are monomers in crystal? They'd ahve a a weaker signal that oligomers that are oligomers in crystal, because... because because. They'd have basically had energy minimization run on them, thjey'd have adapted. THey'd be more stable then if you just rigidly separated them from the oligomer, and the *in*stability wouldn't *necessarily* be local. The extreme case is that of a tube-shaped protein that polymerizxes art two ends, and if you separate it rigidly, it'd be unstable at the ends... but ifyou ran moleculear dynamics on the individual protein, its ends would reach around and touch each other, becoming perfectly stable, with all the instability at the twisted middle. The monomers in crystal have had molecular dynamics run on them by the greatest molecular dynamics sim ever, the universe, the original that we mimic with computers.

In fact... does anything come out of this method that wasn't put in? Some of the members of the dataset are oligomers in crystal (like 1QD6, which was prediced monomer I think)... is the method just picking up on hte fact that a protein is taken from a crystal structure where strong contacts have been removed? Would slight shifts eliminate the signal?

The answer is "probably not" since their energy function is based upon taking note of neighbors and rough position (only 3 bits of position info!), not exact geometry.

And actually the answer is "definitely not" because the method works from sequence too... though I don't entirely understand how.. or even sort of understand how, really.

But they do, and they can identify *oligomers* pretty well. This chapter isn't really *about* the prediction of which are oligomers, though, is it? And then it comes down to their prediction of interfacial *strands* from sequence, which they *don't actually give in their paper*.  
Although... the prediction of the protein being an oligomer is very important. *Very* important. Their method of discerning an interface assumes the existence of an interface; it will never give the result "no interface", not ever.

*All* the trimeric porins have inplugs away from their interface. So the energy at the interfaace is even higher than the energy for missing in-plugs. *That's* interesting.

And it's interesting that their strand variance metric doesn't come up with the proteins with in-plugs. Why wouldn't it? The chart *I* want to see is: a histogram of the strand deviation metric, for proteins with in-plugs. (also: doe s it depend on size? How do youmake decisions just based upon these figures in papers?)

The question is... WHAT DOES THIS PAPER'S FINDINGS REFLECT, THAT WE KNOW TO BE TRUE ABOUT THE PROTEINS, THROUGH EXPERIMENTATION?

Alright. The thing is... the other proteins, that there's some evidence of them having oligomeric forms... those studies report *mixed* oligomer and monomer. Are there studies that indicate *solely* oligomeric forms of the ones labeled oligomer in Naveed's dataset? There are proteins that are stable as oligomers and stable as monomers, and proteins that are stable as oligomers and *unstable* as monomers, and *that's* what Naveed is looking for.

And... as for the idea of being unstable as a monomer... why would we expect that? 1. Finding no monomer when we look. 2. Look up a review on instrinsically disordered proteins... this is something we see *elsewhere.*