



Futures in Biotech, 41: Modeling Life With The World's Fastest Computer

Leo Laporte

Bandwidth for Futures in Biotech is provided by Cachefly at cachefly.com.

Marc Pelletier

This is Futures in Biotech, episode 41: Modeling Life With The World's Fastest Computer. Futures in Biotech is brought to you by audible.com – the Internet's leading provider of spoken word entertainment. Get a free audiobook download of your choice when you sign up today. Log on to audible.com/biotech today for details; and by squarespace.com the fast and easy way to publish a high quality website or blog. For a free trial and ten percent off a new account go to squarespace.com/biotech.

[Music]

Welcome back to Futures in Biotech. I am Marc Pelletier. Today's guest is Dr. Vijay Pande, Stanford professor and creator of folding@home. We had him on eighteen months ago in December 2007 to talk about his work which involves modeling the processes of life down to the atomic level. To do so he requires an enormous amount of computational power and at that time his system was running at 1.3 petaFLOPS. That means 1.3 quadrillion floating point operations per second. Well, as I record this introduction and post here, his system is running at 8.5 petaFLOPS. During the interview, it was at 8.2.

He is really beating Moore's Law and it's based on the graphics processor primarily of the PlayStation 3 and in terms of competition, there is not a whole lot out there that can match this kind of computational power. IBM has a machine called the Sequoia that is dated to be released in 2012 and they claim that it will be able to run at 20 petaFLOPS, but I suspect that Dr. Pande will have them beat by then, by 2012 probably be at 28 or 30, I am not sure.

I also invited Randal Schwartz who's the host of FLOSS Weekly on this network so that we could get a number of perspectives on Dr. Pande's work. So without further ado, on to the interview. I think Randal is a little... not distracting, he is raising the level of fun in this too.

Randal Schwartz

Yes, it's my job. It's really my job, come on.

Marc Pelletier

So thank you, Vijay, for coming on, I really, really appreciate it.

Dr. Vijay S. Pande

Sure, it's my pleasure.

Marc Pelletier

And one of the reasons why I really want to bring Randal on, we had Steve Gibson last time, is because well, well I sort of get an understanding of the protein folding aspects but the protein folding is really just the surface of what you are doing because from my understanding on your website and the stats, you are running the world's fastest computer or the world's fastest computer system. Is this...?

Dr. Vijay S. Pande

Yeah. I think that's based on, in terms of apps that we would want to run I think that's definitely the case. People use different metrics for things and even what is one FLOP really varies from person to person. But by anything that we would want to run I think it's definitely the fastest, there are certain apps that would not run well on our system, apps that need a really lot of tight communication between processors. But there are many apps that aren't in that category and certainly, having some creative algorithmic

development means you can often get around that as it is, so... But within those caveats, I think yeah, we're on five petaFLOPS, the next biggest machine is around one petaFLOP, so it's quite a big gap so far.

Marc Pelletier

Well I see your stats, right now you are at 8.263 petaFLOPS and you are running, just in PlayStations alone you are running 780,000 CPUs through PlayStation 3.

Randal Schwartz

Wow...

Dr. Vijay S. Pande

Yes.

Marc Pelletier

So you are being very kind by saying to the other folk, well for some apps, it's not necessarily appropriate, as you said the ones where the CPUs need to talk with each other but you are flying. Yeah, there is no question. This is really amazing stuff. How does it feel to be at the helm of this and how did, maybe you can give us a little background. How about that? Again because we did cover this in a previous episode but I think it's really important to see the timeline here. If you can give us the timeline, when you started folding@home and how you have seen this logarithmic growth and why did you see logarithmic growth?

Dr. Vijay S. Pande

[5:28] Sure, sure, yeah. So I started at Stanford almost exactly 10 years ago. So it's coming on, I started in like July of 1999. So it's getting pretty close to the 10-year anniversary of that. And when I started here I wanted to do something that wasn't just incremental. I wanted to do something really which could sort of make a major advance over what was going on. And I think one of the things that was clear to everybody, or to many people, was that the computational power involved in things like drug design or protein folding simulations was really orders of magnitude below what it really needed to be in order to make these [ph] technical advances. And people had different ways to work around this with simpler kinds of applications and things like that, simple models. But I didn't find those models actually all that appealing, especially if you really want to make these direct connections experiments really address something like Alzheimer's disease.

And so I started thinking of what we could do to really make an impact here and, you know, at this time things like distributed.net and GIMPS and SETI@home were already out. But none of those apps actually applied this type of approach to any sort of what I call scientific calculations, not a mathematical calculation or a computer science calculation. And part of the reason why is that I think most of the feeling at the time was that you wouldn't be able to run anything useful. To make a rough analogy, you might think that a research lab would need maybe 10 or 20 really great scientists, Ph. D level scientists but if you gave them 10,000 3-year-olds, that wouldn't necessarily make progress any easier or any faster. It probably would slow things down. And so the challenge is that, you know, if you're thinking about a traditional supercomputer as being a powerful machine and you are substituting that with lots of PCs I think the kneejerk reaction was that there is no way that I would be able to do anything useful, far less for something that I wouldn't be able to leapfrog over what people were doing.

And so the first thing we had to do was to figure out how we could make something useful out of that. And before we did any coding of folding@home or anything like that, we came up with a series of ideas that we could run that we felt pretty confident would be able to make an impact. And then roughly, maybe in the summer of 2000, maybe even spring or summer of 2000, we had to sort of make a decision, do we want to actually just sort of – we had the papers ready to say that this could do something neat but did we want to actually do it.

Marc Pelletier

Right.

Dr. Vijay S. Pande

And doing it was diving in, in a sense because since then a lot of it, even just my personal life has been tied up with keeping folding@home running whether that means coding, restarting machines, getting corporate partners on board, pushing the signs, pushing the experiments, all of it. And so it's something that, especially at the time there wasn't a lot of infrastructure to help us out for these types of things. And so, but we decided to do it and I really feel strongly that I like to not just talk about something but I actually hopefully come up with some solution.

And so in 2000 we started out and we had, I think relatively low expectations just because I wanted to not get overexcited too early. But even early on we got to 5,000-10,000 CPUs pretty quickly, and if you think about what even just a 10,000 CPU cluster costs. That's a huge cost in people; in academia, research groups don't even have anything like that, typically for their own use. And we have been gradually growing as the years went on and the key thing is that I have really have tried to push us at the envelope of what we could do. I've read books like the Innovator's Dilemma and things like that and I could see very much that distributed computing could be a really disruptive technology. But I've always been thinking about what would disrupt distributed computing and you can imagine things like GPUs, Intel processors and PlayStations, those types of things would be very disruptive and so I have tried to incorporate that along the way, just trying to really push us at the leading edge of this.

Marc Pelletier

What do you mean by disruptive?

Dr. Vijay S. Pande

Oh yes, so I think, it's a term that is not my own but it's commonly used. It's basically an idea that you might have some technology A that looks pretty successful like a mainframe, and basically, it seems great but then there is a technology B that looks at first to be not really competition for it, like a minicomputer. But as time goes on the competition B starts looking better and better and eventually its capabilities exceed what the mainframe could do and as many people pointed out, went from mainframes to minis, minis to workstations, workstations to PCs and in terms of just regular computing laptops and netbooks obviously now are playing a big role. And what I see as an interesting platform for us...

Marc Pelletier

Are you saying – netbooks are playing a big role, is this in your distributed computing or...?

Dr. Vijay S. Pande

Not necessarily in distributed computing, I'm think just in like, what is being sold right now.

Marc Pelletier

Okay.

Dr. Vijay S. Pande

But for us in distributed computing, I think the jump will be from desktops to like gamer desktops, which you can't get like a multi GPU setup in a netbook or other – or a laptop or something like that very easily. But also I think the move to consoles like PS3s and things like that is also potentially the next part of this evolution.

Randal Schwartz

Was, as you were thinking about distributed computing obviously the size of a bite-offable chunk or work unit, did that influence the kinds of tasks you are looking at?

Dr. Vijay S. Pande

[10:50] Yeah, definitely. No, that's a great question. I think we tried to sort of – part of the problem here is all the algorithms for using multiple processors basically assume that you have really fast communication between them and that's something that we just didn't have. And so we had to really rethink how we just design these algorithms. And, but this was something that was sort of familiar to us and to me because I

actually was involved relatively early on when massively parallel supercomputers came out, and things like the Connection Machine, and things like that.

And that – what I really wanted was one computer that was, let's say a hundred or a thousand times faster and instead I got a hundred or thousand computers. And they are really relatively tightly connected and very homogeneous and all and stuff but still we had to really just completely rethink how we are going to do our algorithms for those machines. I think that sort of rethinking spirit was on my mind when distributed computing was sort of on the horizon, because I had no issues trying to rethink how we would do things.

And part of it, in terms of your question, is can we think of algorithms in which we can fit within these constraints of what could be done on a couple of days on a PC.

Randal Schwartz

Right.

Dr. Vijay S. Pande

And so far that hasn't been a huge problem. I think the biggest problem is not so much the CPU time as much as the bandwidth to, like, if we want to run a million-atom simulation, it is just sloshing back and forth let's say 100 megabytes a day and that can sort of cause problems for people's bandwidth caps and things like that. So I think that sort of the – it's not even the latency as much as just sending the data back and forth after a while.

Randal Schwartz

So there's useful science to be computed in something that's going to run on my PlayStation for 12 hours?

Dr. Vijay S. Pande

Yeah, absolutely. I mean especially the key thing here is that it's, what happens after those 12 hours, it's not that that calculation ends but rather it comes back to us and then we either send it back to someone directly or maybe analyze it then send it back to someone directly such that what really happens is that people are running calculations for three months or six months or twelve months. It is just being passed around in terms who is doing the next step.

Randal Schwartz

Oh, I had no idea that folding@home, I wasn't actually completing something at the end of 12 hours, so I am just, I'm moving some calculation forward each time.

Dr. Vijay S. Pande

Yeah, exactly. It's, occasionally it's something where you would, where one work unit would be everything we would need to do but that's actually pretty rare. And it's more likely that you are part of a long chain of people helping us out.

Marc Pelletier

Let me just interject right here to say that I don't want to discourage people from folding@home feeling that they are not accomplishing anything here by doing one piece of a large puzzle. To me, I see this as the – the molecular dynamics experiments alone, as a way to really accelerate drug development. And I can't wait to, I am not at the point where I can really grab this and collaborate yet, certainly. But in about a year from now in our drug development program, I'd love to run some simulations with you. And so for people out there don't get discouraged that these are small packets. These are a very, very important piece of a puzzle where, within the next 15, 20 years, we are going to see some major changes in how people develop drugs that save millions of lives, let me just say millions. It is not an understatement.

Dr. Vijay S. Pande

I would...

Marc Pelletier

Go ahead.

Dr. Vijay S. Pande

I would emphasize one thing one thing I've done which is even if we had a big massive supercomputer like a trillion dollar machine that could replace folding@home we would still run the calculations the same way because you never run something for a month without saving data along the way because things can crash and so on. So this idea of checkpointing after 12 hours and then restarting is just this way things are done. And so the key thing is that, and also I think I would stress that the fact that calculation takes a year sounds like long time, but what folding@home can do is that, it's only a year instead of ten thousand years, or something like that.

Marc Pelletier

Right.

Dr. Vijay S. Pande

So...

Randal Schwartz

Before the heat-death of the universe.

Dr. Vijay S. Pande

Yeah, so and also, finally in terms of drug design, we have a lot of I think results that we are sort of still testing and going to lab but there's been direct small molecules that have been predicted from our simulations, that then we test experimentally either in my lab or a collaborator's lab and that's moving along. As you guys know really well there is a long route between this looks cool in the simulation to this is working in the test tube, to this is working in animals to something that someone can take as a drug. But I think the whole idea about this is to really accelerate this process and to come up with leads that you wouldn't be able to guess by other means. And then, in the end toxicity and things like that, those are difficult problems but hopefully we can greatly accelerate getting to the point where this gets into animals, gets into the lab.

Randal Schwartz

What kind of security do you have in terms of me not just faking results and sending it back?

Dr. Vijay S. Pande

[15:52] Yeah there is a variety of levels of things there. I think some of them are computer science level security like digital signatures and things like that. And these signatures are big enough that you would need a folding@home like machine to crack it if you were to try to do that.

Randal Schwartz

Yeah.

Dr. Vijay S. Pande

I mean, or even bigger than that. And, but then on top of that, we also have a sort of chemistry kind of things where you would have to fake it, you would have to beat the computer science first and you would have to sort of somehow come up with something that is chemically plausible and then on top of that if something was, if you, if something looked odd, we would rerun it and see and so then also that final check we test it. So....

Randal Schwartz

So you don't automatically send everything out to two people then. We chatted with the BOINC people and that was one of their solution – was to send everything out to two people and thereby cutting the power in half, you don't automatically do that then?

Dr. Vijay S. Pande

No, not at all. And I think what we, there are things that we might do adaptively where there is, if there is something suspect, we would retest it. But it is not like, we don't retest everything.

Randal Schwartz

Okay. So what's the total – actually I am curious about how you got this on my PlayStation. Is, was that a difficult process? Was that easy? Was it sort of in the plans from the beginning knowing that we had these great Vector processors in there?

Dr. Vijay S. Pande

Yeah, when the PlayStation 3 specs came out, which was some time ago, I actually really hoped that that would be a neat machine and we'd been doing a lot with GPUs but GPUs at that time were really, really early stage. You know, this is before something like CUDA existed where if you wanted to get a GPU calculation you would have to express your calculation in DirectX or OpenGL, essentially as a graphics operation. And the Cell processor compared to that at that stage seemed really appealing because while it was complicated and so on it felt like a real computer. And so I have been wanting to put it on there and I have been trying to sort of find the right person to talk to at Sony and as you can imagine, Sony is a big company and they don't have like a website where it has the people's names and phone numbers for where you want to go to find the right person. And so – but I was trying to work my connections for people who knew people within Sony and there is several people who gave me a lot of good advice. But in the end actually after all that work they contacted me. So...

Randal Schwartz

Really?

Dr. Vijay S. Pande

Yeah, so that made life really easy and so it was one of these things where they contacted me more or less exactly the same time when I was trying to sort of find the right person and so Noam Rimon, one of the senior developers there, he runs a development team, contacted me and the first thing that we talked about would be what is the best part of the folding@home calculation to put on. Because for Cell processors and GPUs, there is a certain sweet spot in terms of what we do that I think we have been really able to exploit and on cell processors we can get like a 40X speed up over a fast CPU and on GPUs we can get somewhere in the 100s of X speed up, over a fast GPU.

Marc Pelletier

How do you do that? How do you do that?

Dr. Vijay S. Pande

So, Yes I think the whole thing about both processors is that they have just lots of floating point units especially GPUs. You look at like a GPU die, just the transistor layout compared to like a CPU transistor layout. On the CPU transistor layout, you see fields and fields of cache and on a GPU you see lots of FPUs and lots of processing cores. And so you just have a lot of raw processing power. The challenge is can you take advantage of it? The CPUs have all that cache for a reason, which is that it's hard to sort of keep these things fed considering the fact that memory is usually much, much slower than CPUs external memory. And so the trick is can you construct a calculation where now almost like the FLOPS are free, the math is free, what is expensive is moving data around. And can you, sort of again, sort of we were forced to rethink the algorithm and think about what we need to do, in a way that within now a given box could we distribute the calculation on multiple GPU cores? And this was a very different style of distributing it and it was more akin to sort of the more traditional parallelism but, and it took us many years. We have been working on GPU stuff easily for four maybe five years. And we are on our third generation code and I think the second generation code was a thing that got really exciting in terms of showing that this was break even in all the stuff. The third generation code is, already had two scientific papers that had been produced and I think it's chugging along.

Randal Schwartz

It is really ironic I think. I just got to point this out that the hardware that was built so that people can do shoot 'em up games better and faster and cheaper is being to use to save lives now.

Dr. Vijay S. Pande

Yeah, I mean, what is neat about is it that yeah I totally agree with you. Especially if you think about even a PlayStation 3, no one expected a PlayStation 3 to run drug design calculations or anything like that. The thing that is neat about it though and I think what is going to be our, will help us is for consumer electronics to really be pushed. Because if someone designed a GPU to do drug design calculations they would sell it to maybe – they would sell ten thousand units or a hundred thousand units if not tens of millions of units and the cost per unit would be maybe ten times or a hundred times more. And it would suddenly become not worth doing. And so we really have this strong symbiosis with consumer electronics because one of my colleagues, Michael Levitt used to joke that he wished Microsoft Word required a 16-core machine to get any sort of performance.

Marc Pelletier

It kinda does.

Dr. Vijay S. Pande

Yeah and it does, perhaps, yeah but the joke being that as consumer software requires all this more performance, it brings a need for that type of hardware along that we can take advantage of for these other applications. And the, what you see right now on CPUs and multicore is something that concerns me in terms of it's, right now, that I don't know if there is the software that would justify someone needing a 32-core machine. And I think it will be neat to see in that side of things where people come up with new apps and it might be a new version of Quake or it might be something else. Yeah what, I mean, so I think what's nice is that the gamers want fast GPUs and that's going to continue for quite a long time and so that's part of the reason to – that's part of the reason why it shifted to GPUs.

Marc Pelletier

We've got about a half-second lag by the way. I can let the audience know I am not going to truly edit this out but there – we're on Skype and we've got about a half-second lag right now, so it is really tough – go ahead.

Randal Schwartz

I'm just going to make a stupid joke, if only YouTube required 16 cores!

Dr. Vijay S. Pande

Yeah, absolutely. Yeah or Facebook or you know. I am sure those guys will come up with some good use of it, you know maybe 3D YouTube or something like that. But yeah, and I think it will come. I mean it just, there needs to be a really compelling reason and for now GPUs I think are one of the most compelling reasons and so it is a great time for us to be on that platform because I think those processors will, there will be a continued need for them and they will get cheaper and cheaper.

Randal Schwartz

Do you have much overlap with BOINC at all? Is there any sharing of technology, I mean you are approaching sort of the same idea distributed data and they are applying it to a number of projects. You've got your own independent code-base – I recognize that but is there any overlap or correlation?

Dr. Vijay S. Pande

[23:20] Yeah, we chat a lot. I've chatted a lot with David Anderson, and I really respect all the things he's done over the last 10, 15 years. And every once in a while we look and see if there is something that could be appropriate for BOINC and I think in many ways it looks like it could map reasonably well but I think our style of doing things on the backend has been different enough that my team I think is really happy with our backend software. And so there hasn't been a compelling reason to switch over. And also part of it is that we have been trying to push the limits. So when we went on to the PS3, we didn't need BOINC to be ported to the PS3 first. And David I think has been wonderful though about getting all that type of stuff done. So he would be and I think has been a great type of collaborator for that type of thing but I think being a little smaller, without requiring all that infrastructure, has made it a little easier for us to

go into new platforms. We could do it ourselves without sort of having to convince other people that these are interesting things to do.

Marc Pelletier

Cool. I'd like to thank audible.com for sponsoring Futures in Biotech. I think they have over 51,000 audio titles right now. For the book of the week, I apologize the clip is a little bit long but it is extremely relevant to the discussion that we are having today. Because the book is entitled, *Always Looking Up* by Michael J. Fox and narrated by Michael J. Fox and he is living with Parkinson's, a protein folding disease, so here is the clip with Michael J. Fox.

Michael J. Fox

That was the original title for the memoir written eight years ago. On the second and third page of the earliest draft, I made reference to myself as being a lucky man. After a few edits, I kept going back to those two words and eventually they found a way onto the cover of the book. They fit then and they still do now. As a title for this new book, *Always Looking Up* works on a couple of levels.

First off, let's just get this one out of the way, it's a short joke, at a fraction of an inch under 5 foot 5, much of my interaction with the world and the people in it, has required that I tilt my head backward and direct my gaze upward. However this isn't the manifesto about the hardships of the vertically challenged. Frankly, my height or lack thereof never bothered me much. Although there is no doubt that it's contributed to a certain mental toughness. I made the most of the headstart one gains from being underestimated. And that's more to the point of it.

Always Looking Up alludes to an emotional, psychological, intellectual and spiritual outlook that has served me throughout my life and perhaps saved me throughout my life with Parkinson's. It is not that I don't feel the aching pain of loss, physical strength, spontaneity, physical balance, manual dexterity, the freedom to do the work I want to do when I want to do it, the confidence I can always be there for my family when they need me, all of these have been if not completely lost to Parkinson's, at least drastically compromised.

The last 10 years of my life, which is really the stuff of this book began with such a loss. My retirement from *Spin City*. I found myself struggling with the strange new dynamic, the shifting of public and private personas. I had been Mike the actor, then Mike the actor with PD. Now was I just Mike with PD? Parkinson's had consumed my career and in a sense had become my career. But where did all of this lead me? I had to build a new life when I was already pretty happy with the old one. I had been blessed with a 25-year career and a job that I loved, I had a brilliant, beautiful, funny, supportive wife and an expanding brood of irrepressible kids. If I had to give up any part of this, how could I possibly protect myself from losing all of it? The answer had very little to do with protection and everything to do with perspective.

The only unavailable choice was whether or not to have Parkinson's. Everything else was up to me. I could concentrate on the loss, rush in with whatever stopgap measures my ego could manufacture. I could rely on my old friend from the '90s – denial, or I could just get on with my life and see if maybe those holes started filling in themselves. Over the last ten years they have, and in the most amazing ways.

What follows is a memoir of this last decade but unlike *Lucky Man* it is thematic rather than chronological. Work, politics, faith and family – these are the struts of my existence. These are the critical supports of my life. While not a strict narrative, *Always Looking Up* describes a journey of self-discovery, and reinvention. The story is a testament to the consolations that get me through and give meaning to every area of my life. For everything this disease has taken, something with greater value has been given. Sometimes just a marker that points me in a new direction that I might not otherwise have traveled. So sure, it may be one step forward and two steps back but after a time with Parkinson's, I have learned that what's important is making that one step count, *Always Looking Up*.

Marc Pelletier

So if you'd like to download *Always Looking Up* by Michel J Fox for free, all you need to do is sign up for a 14-day free trial. And you will get a credit towards a free audio book. Now if you decide to cancel, you still get to keep the audio book. So it's a win-win situation. So head over to audible.com/biotech.

[28:20] So we recorded the last episode in December 28, it was released on December 28, 2007. So we recorded it in November of 2007. So we're talking 18 months, and at that point we were at 1.3 petaFLOPS. And now you are quickly coming up on 10 petaFLOPS. And, while the rest of the world, I mean, it's still the same PlayStation3s and consumer electronics is really tapering out is – it's as if now computers are taking the same old business model as cars, just different models and with the same power. But you are really pushing us forward and is this mainly because you guys went to the next level of software. And you mentioned this. But I'm sort of, as I said, the computer dumb-dumb here on the show.

So is it really the GPU that's – and is that – so you're – that's how you see this moving forward?

Dr. Vijay S. Pande

You know, the big difference between now and when we last chat really was the GPU, the PS3s got us a petaFLOP, an additional petaFLOP. So we were like under a petaFLOP, like a third or a half, and PS3 has added one whole petaFLOP pretty fast. And then the GPUs added multiple petaFLOPS. And so, and if you think about it, just one GPU, the peak speed of one GPU, which no one can really realize, the peak speed, but the peak speed is a teraFLOP per GPU. And so even if you can just realize let's say a third of that, that's 300 gigaFLOPs per GPU or 0.3 teraFLOPs. You just have 3,000 of those and that's a petaFLOP you know. So it's something where it's – I think the math works.

Marc Pelletier

I'm counting here, you've got 46, 60, 66,000 teraFLOPs running just on GPUs, ATi and video on PlayStation3.

Dr. Vijay S. Pande

Yeah, yeah. I mean the one thing is that the numbers I just gave are for like the highest end GPUs and so the GPUs that are running folding@home, not all of them are the highest end and not all of them are running 24/7 and all that stuff. But just a few thousand GPU cluster would get you to a petaFLOP pretty fast. And so I think that's where as GPUs get faster and faster and as people upgrade from these older machines to the newer generation, I think we are going to continue to see something there. And so I think consumer, in terms of consumer electronics, people are not going to be buying 32 processor machines I think – 32-core machines super soon. I mean there is only going to be some people who want that. But I think people with multiple GPUs, gamers with multiple GPUs, they'll be around for a while. So I think even in the economic downturn I think there could – this could still work reasonably well. But I think we talked a lot about FLOPS and things like that. I think, really the part that will be exciting in the end for let's say my colleagues that do drug design, will not really be the FLOPS but be the molecules. What can you really do...

Marc Pelletier

Oh, for sure.

Dr. Vijay S. Pande

And that's where I am, we have all these results in the lab that are unpublished that I am itching to get out. And I think it may be another 18 months we'll be past 10 petaFLOPs but I'll be talking about our recent papers that describe these interesting molecules and the experiments that were done on them and hopefully even in animals and so on. So that'd be really exciting.

Marc Pelletier

Before we get to the biology and I definitely want to end on the biology. I know that Randal has to head out in about 15 minutes.

Randal Schwartz

Yes.

Marc Pelletier

And he is got a hard leave and I don't want to keep you.

Randal Schwartz

Okay.

Marc Pelletier

But so...

Randal Schwartz

I just had...

Marc Pelletier

Go ahead. And then we'll get to the biology...

Randal Schwartz

I have one pedantic point I want to make. There is no such thing is a – no such a thing as a FLOP, there's no such thing as a FLOP.

Dr. Vijay S. Pande

Absolutely.

Randal Schwartz

Stop saying FLOP, it's FLOPS. There's one FLOPS. It's FLOating Operations Per Second.

Marc Pelletier

And that's why Randy is here. Randal.

Randal Schwartz

Yeah, that's why I'm a writer, yes.

Dr. Vijay S. Pande

Fair enough. I mean but even though, even with that grammatical distinction I think just even how you define a FLOP is really tricky. How you define one FLOPS?

Randal Schwartz

It's such a horrible acronym.

Dr. Vijay S. Pande

It's weird to say. Yeah, yeah. It is tough to use the singular articles with a plural sounding word. But yeah how do you define one FLOPS?

Marc Pelletier

Is it one FLOPS, two FLOPS, or one FLOPS, two FLOPSes?

Randal Schwartz

No.

Marc Pelletier

How do you define it? How do you define it?

Dr. Vijay S. Pande

Yeah FLOPS are like rpm, right? I mean you have 1 rpm or 2 rpm or something like that. It's certain amount of things per certain amount of time.

Randal Schwartz

Yeah.

Dr. Vijay S. Pande

So it's a certain amount of calculations per second.

Randal Schwartz

Right.

Marc Pelletier

It's floating, I'm trying to get the acronym here, Floating-point Operations Per Second but F-L-O-P-S. So...

Randal Schwartz

Right.

Marc Pelletier

So Floating, oh FL from floating...

Dr. Vijay S. Pande

Yes.

Marc Pelletier

And PS is per second.

Randal Schwartz

Yes. Right.

Dr. Vijay S. Pande

Yeah.

Randal Schwartz

Because MIPS was the other one that's like that. I remember people were always talking about a one MIP machine and I went, what? This is 1 MIPS. Millions of instructions per second, right? You've got to have the second in there again.

Marc Pelletier

Well, you know, this isn't slowing down, right? And the world out there is sort of migrating towards netbooks. But as you said gamers are not migrating towards netbooks. They are really pushing it, and with the popularity of shows like PC Perspective, with Ryan Shroud on the TWiT network, he is constantly reviewing the best hardware, and there is definitely a population out there. But I want to know what's happening over the next 50 years with distributed computing. And if you think about today and folding@home as being the Apple1 that Steve Wozniak brought into the home brewing club, where do you see, and this is kind of how I see it because not everybody has a setup running at 8.2 petaFLOPSes.

Randal Schwartz

Yeah.

Marc Pelletier

So where do you see this going in the next like even 20 years? Will everybody have access to this? Are you going to become the next Google with it to allow people to do everything from banking, are you licensing this out, what, how do you see this growing in the next 20 years, let's say?

Dr. Vijay S. Pande

[34:40] That's a great question and it's always hard to tell but I can give you my guess. At least since we've been doing this for 10 I feel like we have a license to talk at least about the next 10. But we'll see, I think I – you know one thing that seems clear is that computers are sort of blending into things, they are sort of appearing everywhere. I was really impressed with even my 5-year-old daughter. Her Elmo talks and something like that. I asked her, how does Elmo talk? And she said there is a computer inside. And so she knows that and obviously it's not a very fast computer so we're not going to be running on 10,000 Elmos or anything like that. But these machines are basically going to be everywhere. You can imagine that people talked about putting Cell processors in TVs and I think as these – as computers move from being in the early days, the thing on the accountant's desk to something which is just sort of disappearing into the background, such that they are everywhere, including in your children's toys and you are not thinking about it. And everything is networked. You can imagine that you could choose to give some fraction of your toaster or TV or whatever it is, computer time. The toaster joke often comes up because these things will be hot enough that probably you could build a toaster off of one of these GPUs relatively easily.

Randal Schwartz

Right.

Marc Pelletier

Compute and cook your toast at same time.

Dr. Vijay S. Pande

But – yeah absolutely why not take advantage, if you –

Marc Pelletier

And the fact that you look towards your daughter's Elmo, you looked at her – your daughter's Elmo and said there is a computer – she said that there is a computer inside and you thought, well, I'm not going to use...

The fact that you were thinking about using Elmo. So there is no question you're thinking around the right lines here. But how about iPhones?

Dr. Vijay S. Pande

Yes, so I – the big problem with all these mobile devices is – and I have an iPhone, I love it, is that my battery life is bad enough as it is. And so even if, let's say, the next generation iPhone had floating point capabilities, which the current does not, if the next one did, and let's say it had like a pretty good like GPU even, let's say. You would burn up that so fast that you'd never be able to get a useful calculation out of it.

So I think that's where – it has to be something high end and I can imagine that graphics is such an important part of – humans are so visual, that GPUs will become parts of just everyday things, in the background. And so I can imagine that those are devices that we could run on. And I think people – if you look at let's say GPU technology, which you could do five years ago, it doesn't look particularly visually appealing, it looks very clunky and not even classic, like Tron might look classic, or something like that, it just looks clunky. And so I can imagine people will demand a more and more sophisticated look on things and GPUs could be really everywhere. So that's part of my other motivation for things all running on GPUs, is that I could imagine that it has a long term need in consumer technology that's not going to go anywhere. Whereas I, at the moment I don't see us having 128 processor machines serving a use.

You want obviously avoid the Bill Gates scenario. Nobody needs 640 processors to do anything. But I think – at least right now the app doesn't exist. And maybe there will be some social network thing, that it has agents all over the place looking for things and you need to have a 100-processor machine to get any sort of speed on that. But until that app exists, I think that GPU stuff seems like a good bet.

And so 20 years in the future I can imagine computers are just everywhere and we would have to careful what we run on and so on and people will have to understand that they would have to let us run on – like

on the PS3, it comes installed on the PS3 but you have to turn it on so people would have to agree to turn it on. But the hope is that it could be some – there could be a lot of machines just in places you wouldn't normally think of as being computers.

Randal Schwartz

You know your mention of the toaster getting warm or actually doing this because the toaster will warm up, it sort of has me curious actually and I sort of was curious about the BOINC project as well. But you are burning a lot of electricity to make these calculations happen. Is there any thought that maybe you ought to do some carbon offset for all this computation?

Dr. Vijay S. Pande

[38:50] Yeah, no, that's a really good question and good point and something that we thought a lot about. It is something my team actually thinks a lot about. There is a couple of ways to think about this, one of which is that we are running on a lot of PS3s and you could worry about that much electricity. But there is a lot more PS3s out there running, doing things anyway, there's a lot more computers. So in terms of all the computers in the world, it's – or let's see all the computer time in the world, this is a relatively small fraction.

Another way to think about it is that if we were to run this in a supercomputer center, all these machines would be sitting in one room and so you'd have to have this massive air conditioning and you'd spend almost – you'd double or triple your electricity bill because of just having to deal with all the intense heat that comes out of that. And so it's something where, I think, this – it could be argued that if you were to do this calculation, this is probably amongst the cheapest ways to do it, because you are distributing everything all over the place.

But whether humanity should do this calculation or not is something that I guess is left to the free market kind of thing. There is no counsel saying we should do this or not do this. But if we were to do this calculation I think it's a pretty, an efficient, energy efficient, ecologically reasonable way to do it.

Randal Schwartz

Yeah. Because it may be raising, my room – my house may be raising my temperature here maybe like a half a degree that I have to offset for because my PS3 doesn't generate that much heat. But I was just thinking, collectively this must be, we have to trade off burning this and burning the carbon that it takes to make this calculation happen versus say some other calculation. But I like your conclusion there, which is if this is all happening in a supercomputer room, it would be even probably more offsetting because you'd have to thrust a whole lot of heat out of the space.

Dr. Vijay S. Pande

Yeah, exactly. And the hope is too that – I mean the one thing about a supercomputer you imagine building it somewhere where electricity is cheap and cooling is cheap, like Alaska or something like that. Or something like that –

Marc Pelletier

Or heating is important.

Dr. Vijay S. Pande

Yeah.

Marc Pelletier

It's absolutely important. I heard on one of the shows, I think it was Maxwell House, they were saying that electric heating is 100% efficient. So whether it comes out of your PS3 or out of the electric heating in the Canadian home, in the middle – dead of winter, it doesn't matter. Right?

Dr. Vijay S. Pande

Right

Marc Pelletier

You're going to want northern hemisphere to really crank out for you.

Dr. Vijay S. Pande

We actually do see this, that I have seen anecdotally people on the folding@home forums talks about how when it's summer for them depending on which hemisphere they are in, they will tend to run a little less than in winter where they don't mind cranking it up more, because it is just like their heater plus the calculations. So – but yeah these things are on my mind I do worry about it. And I think that's the best I've come up with so far that I think that if we were to do this calculation it's a pretty efficient way to do it. So.

Marc Pelletier

So Randal has to head out, do, would you have one last question, Randal about..?

Randal Schwartz

I just actually, just a thank you actually, to really – to take the vision that you have to see that there is these idle CPUs sitting around and saying what can we do with that that's not just letting it display a nice screensaver all day. That to me, I appreciate people like you that have taken action to further humanity with the technology that we've created. It's very cool. Thank you for doing the work that you're doing.

Dr. Vijay S. Pande

Yeah, thanks. Hopefully, I mean in the end I think we are going to be judged on the results and one of the things on my mind is the ten year anniversary of folding@home coming up October 2010. And we are pushing hard to have actually some real concrete stuff that, not just sort of what we've done so far, which actually is a lot of things. If you go to our website it takes you to a lot of results. But I'm hoping to have drugs to show off. And that I think will really I think crystallize the vision and hopefully get people excited for the next ten years.

Randal Schwartz

Especially if you have drugs to pass out. No, no, I'm just kidding. It they're really good designer drugs, they will be really well appreciated.

Dr. Vijay S. Pande

Yeah, that's a whole another distributed computing project, Randal.

Marc Pelletier

Well if you think about it right the top human killers, killing prematurely 100 million a year around the world, cancer and heart disease. Heart disease they are some drugs out there. But they are certainly not the best. Some, many of them have side effects. There is still a drive for pharmacology to save hundreds of millions of human lives. So this is a – the platform that will enable the improvement certainly in the next while, especially in the post genomic era, right? We're seeing all new kinds of drug targets available and we need to crank out really specific stuff. But before we go, one last computer question. Do you see the platform, PlayStation, do you see Sony's success or failure as detrimental to you? Are you – what happens if PlayStation 3 is the last of the great PlayStations? Are you, do you have a plan B to switch Xbox or Nintendo if they become the next superpower machine?

Dr. Vijay S. Pande

In the current generation of consoles the PS3 is appealing to us because of the Cell processor. What's on our mind is actually we could even run on GPUs, because PlayStation has an NVIDIA GPU. It's just not a CUDA capable NVIDIA GPU, but it could have been. I mean, essentially it's very similar in many ways. And what I can imagine is that if the next Xbox, or the next PS3 is like the current Xbox, where you have a pretty powerful processor, but not like a hyper powerful one like a Cell, like a four-core processor, but you have to have some kick ass GPU on there. Like something equivalent to GTX280 or something like that. We could just run on the GPU. So I think, if – so, that's, another thing is that I think people will demand graphics more than they'll demand faster Microsoft Word. And so as long as we're running interesting stuff on the GPU, then we just keep on going there.

Randal Schwartz

You just need Wii Folding, that would be kind of cool

Dr. Vijay S. Pande

Yeah. I mean, I think, I think it will be interesting to see – the Wii is a great box and it is unfortunate that, and well it makes sense that their design specs are exactly opposite of what we need, graphics, but –

Marc Pelletier

My six year old will argue with you on that one.

Randal Schwarz

You can have like the controllers being able to shape the molecule while you are playing around and then have it actually be the... that would be cool.

Dr. Vijay S. Pande

Yes, absolutely.

Marc Pelletier

That would be cool. We got to get one in here. Okay, we'll definitely let Randal go, because he said – it's 10 o'clock out there right? I don't want you to miss your appointment

Randal Schwartz

Yep. That is right.

Marc Pelletier

But thank you very much, Randal, for coming on, I appreciate it.

Randal Schwartz

Thanks for inviting me, it was a lot of fun and thanks, Vijay, for the work you're doing, really, definitely. Thank you.

Dr. Vijay S. Pande

Thank you, it was great chatting with you.

Marc Pelletier

Thanks a lot. And any – those people that want they can tune in to Randal's show, he does FLOSS Weekly on the this WEEK in TECH network. I would also like to take a minute to thank squarespace.com for sponsoring Futures in Biotech. Squarespace.com is a fast and easy way to publish a high quality website or blog.

Now, Futures in Biotech.com, the blog associated with this podcast was over on Blogger for years, but I was never happy with the design. There was always elements in the template that you just could not get rid of and it was the same thing for WordPress. So after hearing about squarespace.com on MacBreak Weekly, I decided to give it to try.

And it is really easy to use, it has a drag and drop Ajax interface, which means you can login to your website and while looking at you website, just move things around. It's as easy as that. It's absolutely the coolest experience with respect to designing a website or blog that I've ever had. So if you are running website for your lab or your company, I urge you to give it a try. Go to squarespace.com/biotech for a free trial.

Coming back to the biology now, I suppose, so that people know where the CPU power is going, could you describe I suppose what are the kind of projects folding@home are folding?

Dr. Vijay S. Pande

[47:11] Yeah, sure. So the, our primary interest is in sort of, as you could guess from the website and so on, is protein folding, especially protein misfolding. And so, just for people who aren't familiar with those terms; protein folding is the act in which proteins, these fundamental molecules in biology, they play as the enzymes or antibodies, they are the molecules that get things done in biology. Now, proteins before they can act as enzymes, which are essentially like, enzymes are like little nano machines that do things, before this machine can do anything, it's not surprising that the machine has to be assembled. And so protein folding is the act in which these proteins actually assemble themselves. Which in a sense is kind of neat too, because there is no little construction workers there that can sort of put these nano machines together, they actually have to assemble themselves. And it's amazing that whenever you look at biology, it's always breathtaking that it actually works. And so therefore it's just really neat to just understand protein folding and there is lots of interesting implications in nano technology, implications in fundamental biology and protein folding is at the key part of the central dogma.

So there is all this interesting stuff there. But there is all this relevant – biomedically relevant aspects too in terms of disease. That – diseases like Alzheimer's, Huntington's, Parkinson's, ALS, CJD, Mad Cow Disease in cows, all these diseases are believed to be related to the case where proteins don't assemble themselves correctly. Instead of folding; they misfold. And there it's interesting to understand what's going on and if we could understand the process, could we suggest small molecule drugs that could stop the process? And such that would hopefully be eventually therapeutics for things like Alzheimer's.

Due to sort of my expertise and also our collaborator's expertise, Alzheimer's has been sort of the key singular symptom we've been going after, although we have seed projects in Huntington's disease and in the design of novel antibiotics. Both of those are interesting areas, and the biophysics and biology of protein folding has implications in many other diseases as well. But in terms of going after small molecules, there's only so many researchers in my lab, it's like 20 people. It makes sense to focus on one thing and really try to make an impact on one, that one thing. And so far that primary focus in terms of a small molecule drug design has been Alzheimer's.

Marc Pelletier

Are you tackling this in like a two pronged approach, trying to see the – do shotgun experiments on the major biological problems from like a comprehensive look versus what is water doing?

Dr. Vijay S. Pande

Yeah I mean I think we are interested sort of at both scales. What often happens is that like when you have any sort of tool, you want to understand how well the tool works and what it can be used for, but you don't want to spend too much time studying that and never get anything done. So like if you are a painter, you want to get a good technique for brushes, but that doesn't mean you spend 20 years learning about brushes and then you never do anything else. And so kind of in parallel, we're understanding sort of what the strengths, limitations of various simulation methodologies would be.

And even just questions of like what is water doing in these biological contexts, this is really relevant since water is often handled really approximately by almost all other sort of drug design methodologies. And I think there are some really interesting surprises about how water is really different in a cell than it is in a test tube. That in a cell there is really no, not a lot of room for bulk water. Water is always up next to something and when it is next to something it behaves differently than when it does when just surrounded by just by other water molecules. And I think there are some interesting lessons from where just looking at water could teach us a lot about what's important about how these things work.

But on the same token we, I think it is important not all only do the basic biology as much as that, I think is sort of the seed corn for the ideas for the next ten years, but try to make an impact in specific areas as well.

Marc Pelletier

So how many atoms like, in a classic experiment that you are doing, how many atoms can you model at one time? So this is, from what I understand, you put in a whole like a three-dimensional framework of a

model of what these atoms are placed amongst each other and then you launch them into a pseudo environment? Is that how it works or...?

Dr. Vijay S. Pande

[51:43] Yeah that is how it works and the simulations really vary. Sometimes we look at proteins that are relatively small. Like a typical small size for us is like only about 40 amino acids, which is the size of a single A beta chain, that's involved in Alzheimer's. So it's a very small size, because many proteins are hundreds of amino acids, but it's a very relevant size because of the relevant proteins of Alzheimer's and Huntington's are – often the relevant parts of them are on those small length scales.

On the other hand if we are interested in antibiotics, sometimes we're, we do simulations with a ribosome, which is a common target for antibiotics. Those simulations might involve thousands of amino acids and RNA bases, but lots of water and co-solvents. So all in all those simulations might be millions or multi millions of atoms.

So there is a big range and depending on what we need to learn, we don't want simulate in any more detail than we have to. That you want to have model, which has everything you need but not more than that. And so sometimes for some questions we can get by with simpler models and sometimes we need these really detailed ones. And obviously, since we're generally impatient about this, we are doing folding@home because we don't want to wait a long time to get these results. That's another reason why we try to sort of build models which are very streamlined ideally. But then you may have to know what has to go into it, which is the real question.

Marc Pelletier

So are you screening like potential compounds, potential structures that are out there. You've got these small molecules with maybe about 40, 50 atoms and you are looking for small molecule to go bind to the ribosome, which is a bacterial ribosome, unique to bacteria. So to create this antibiotic, are you going with known structures, chemical structures? Or are you throwing in some randomness to say, here is the building blocks of organic molecules, throw them in there into that ribosome, see what sticks? What's the way to go on that?

Dr. Vijay S. Pande

So we rely heavily on all the great work that crystallography and crystallographers have done. So when we do the ribosome we start from one of the key ribosome structures that has the work of, it's been the work of other people that have done amazing things.

Marc Pelletier

Ironically, we met at Yale, when you came and gave a talk, right?

Dr. Vijay S. Pande

Yeah. Absolutely.

Marc Pelletier

And Tom Steitz was in the audience and he was responsible for one of the structure, I think the first structure of the ribosome and...

Dr. Vijay S. Pande

Yeah. Yeah. Steitz and Moore especially. And there are a few others that have interesting structures as well but yeah, those really pioneering groups in those areas. We kind of in a sense, I don't know how crystallographers would feel about what I'm about to say, but I feel like what they do is critical for us to have a starting point. But I kind of view crystal structures as a starting point rather than as the end. Because I think you could learn a lot from a crystal structure, but a crystal structure is inherently static. And a lot of biology involves motion, it involves change and ribosomes are great examples, there's a well known ratcheting motion that goes on and people like Jamie Cate at Berkeley and others are trying to capture that ratcheting motion I think through crystal structures. The crystal structure that Jamie has

already published has two ribosomes in it that differ and that's kind of, that's really neat. But it's just, crystallography is not a method designed for dynamics. And you could – we could back up a layer.

Marc Pelletier

Yeah, maybe we could break that down, because, so what a crystal structure is and I guess I'll explain, is where you get a protein or a biological molecule that's attached together rank and file like being in a military march, where everything is perfectly lined up. And by having that perfect lineup you can then throw in X-rays and by the diffraction of the X-rays, get an image of how these things are, because there is so much symmetry and repetition that you can amplify the image and then model, right? But those things can't move, they can't wiggle in the crystal. They are stuck like carbon atoms in a diamond.

Dr. Vijay S. Pande

Yes. Especially since also most of these crystals are cryocrystals, they are at liquid nitrogen temperatures or something like that. So you know even if they could –

Marc Pelletier

So that can't include biology, right? You can't, that's not biological?

Dr. Vijay S. Pande

I mean, I think there is, sort of just to play devil's advocate there, even though I kind of see what you are getting at and I agree with it in many ways, but just to play my own devil's advocate I think some crystals actually, the proteins do function in the crystal. So I think there are some examples where I think those structures are probably biologically relevant, but the question is whether that – it's sort of like saying you can understand how a tiger works by looking at a single picture of a tiger. And you could learn a lot about tigers if you never knew anything about a tiger, you could look at a tiger and you could say oh, it's like a cat, other cats we've seen. It has four legs, it has big teeth, it is probably a predator, you could look at the eyes and say well the eyes make it look like a predator. So looking at just the structural biology of tigers from that point of view, you could probably guess a lot.

But there is a lot you couldn't guess, unless – especially like how do tigers hunt, what do they hunt, what do they do? All of the things that are dramatic, that you see on TV about tigers. It'd be difficult to guess that from a static picture. And so I think that's why I think crystallography is something which is critical to us and is critical to biology. Structural biology I think really is founded on things like crystallography. But I view it very much as a beginning for what hopefully could be done either by other experimental methods, like single molecule, spectroscopy or things, I think hope to be a middle ground to bring dynamics and things. But computational methods, if they are sufficiently accurate, I think it is the ideal situation there, because we have the best temporal resolution and best spatial resolution you could hope for, within the caveat of is this – how good is this simulation, how predictive it is?

Marc Pelletier

In your – the projects that are going on in your lab and you are looking at preventing mis-folding of proteins, you're looking at potential small molecules that go stabilize those, I suppose that's the best way to do it, right? When you get an amyloid, which is a fibril that gets stuck in your brain, it's too late, right? You have to find the molecule that stabilizes it and find people that have the likelihood of developing, because of family history or because of the genes that they carry, right?

Dr. Vijay S. Pande

[58:10] Yeah, the Alzheimer's story is really kind of an interesting complicated mess, because there is even I think still disagreement in terms of what aspect of the story is the toxic aspect. So as you were referring to, the final endpoint is the formation of these fibrils which instead of having one A beta peptide, which is – each peptide is 42 amino acids. But instead of one chain, there might be thousands or hundreds of thousands of chains or more that form this really large structure. And for the longest time people have assumed that these fibrils are the toxic element. And there are some data that is basically agreeing with it, but I think it was early on and I wouldn't say there was overwhelming data to suggest it.

About 10 years ago, several researchers suggested an alternative hypothesis, that it's not the fibrils, these really big things that you can almost see with your eye in a sense, or at least with a microscope, but instead it's really small things. It's the things that might lead up to fibrils eventually. These are oligomers, these really small aggregates, instead of thousands or millions of chains maybe just like 4 or 24 chains.

Marc Pelletier

Intermediate stages.

Dr Vijay S. Pande

Yeah and so – and I think now 10 years after those original hypotheses were made, I think the pendulum has shifted to suggest that it's more of these intermediate stages that are toxic and there is examples of even drugs that came on the market to break up fibrils.

Marc Pelletier

Wow.

Dr Vijay S. Pande

And if you think the fibril is the toxic element, you definitely want to break it up. And if you think the oligomers are the toxic agent, you don't want to break it up. It's – imagine like you had something that in one form it's, so I mean, by now you get the idea is that – and so in these drug trials, I think it was somewhat inconclusive, because the trial was stopped prematurely. But basically the fibril blockers had disastrous results in clinical trials. And I think that really soured a lot of people in Alzheimer's therapeutic sector for a while from the biz dev point of view.

But I think that seems to be very consistent with the idea that the oligomers are the toxic agent and so under that hypothesis the fibrils actually might be even fairly inert. And if you break up something that is inert, and makes something that's toxic in one huge like splash, you can imagine that that could be really nasty. And so since then I think other people have really been pursuing this oligomer hypothesis and that's I think kind of consistent with what we have seen as well. Then there's a debate of how big are the oligomers that are relevant? Is it 4 chains, 24 chains, 100s of chains?

But what it does put on the forefront is that these intermediate steps is the thing that matters. It's not the final place that these chains go, but the journey.

Marc Pelletier

So you can have plaques in your brain and it's not the plaques that are killing you, it's the protein that's in transition to those fibres that form plaques?

Dr. Vijay S. Pande

Yeah, yeah, and that's a current hypothesis that is still under a lot of debate. But I think that is probably what is going to prevail and I would say roughly by now maybe the majority of people agree with that, although it is not a huge majority. And so what that means though from the point of view of experiments is that the fibrils are really the easy thing to look at in a relative sense. I mean, I should say "easy", because the experiments done there, beautiful solid state NMR experiments by people like Rob Tycko and others that these experiments are very difficult and very exciting from a structural point of view. But if this oligomer hypothesis is true then those are not the toxic elements. It's something along the way. And these intermediate states are really, really difficult to capture experimentally, because...

Marc Pelletier

How do you do it? It's impossible, right?

Dr. Vijay S. Pande

Yeah, yeah, they are fleeting. And so it's something where it's easy to capture the beginning and the end but it's really hard to capture the middle, especially since as you talked about for crystallography, you couldn't crystallize it, because these things would probably want to crystallize into fibrils, if you could even crystallize them. So that becomes a real challenge and that's where simulation I think has, that's a sweet

spot for simulation. Because it's something where if the experiments were trivial then there is no reason for us to do a big simulation on it, right. I mean, we want to go after something that is an important problem and Alzheimer's is clearly an important problem and I think in the next 10 or 20 years, it will become even more of an important problem.

So it's an important problem and it's a place where experiments, really, there's not much you can really do to these oligomers in terms of structural experiments. And so we had a recent paper come out which I think, I hope is a tip of the iceberg for Alzheimer's work, where we predicted the nature of the structure of these oligomers. We've then been doing subsequent experimental testing to look into it and then that's the stuff that is not published yet, that I am excited to talk about maybe in the next 18 months or so.

Marc Pelletier

Sure. Just to go back one quick second though, when we refer to oligomer, this is – you got your single protein, it forms small little aggregates or small little clumps of maybe did you say 24, up to 24?

Dr. Vijay S. Pande

Yeah, up to 24. I mean the magic number is also unknown. You might guess that the small oligomers might not be toxic and that the big ones, if any of them are, the bigger ones might be toxic. And so I think other people have extracted the oligomers and would suggest that the 24 ones are the key toxic element.

But this – Alzheimer's is such a difficult field and I feel the pain even of the people in my lab who do the experiments on it. Because normally when you do protein experiments, you can do lots of different protocols and your protein, if it's a well behaving, well folding protein, it will fold. You can change the salt concentrations, you can change the temperature and no matter – if you are not doing anything crazy it will just behave for the most part. And usually if you have a protein that is misbehaving, that aggregates all the time, that's usually a time to throw it away and start on a new project, work on a new protein.

And so the Alzheimer's experiments can be very frustrating and I think people in my lab have had, even just one interesting story is that they were doing construction in the building next to the lab that they are in and the construction probably put in a lot of dust and that dust really alone can really, even just some really small, almost humanly imperceptible I think amount of dust, it was sort of noticeable but wasn't huge. That alone just having in the ambient air can really do strange things and our controls were all failing and things like that and so we had the natural sirens that went off to say that there was a problem.

But it's a very frustrating system to work with and I can imagine, that, that challenge also leads to the fact that people have different protocols, different protocols lead to different results and then people get different results and academically it's hard to know which protocol's the right protocol to be predictive of things physiologically and so on. So it's a really challenging area and that's why I think simulation, it's just yet another reason where sample preparation and all that stuff is very easy to do reliably in simulation. We can sort of avoid a lot of those pitfalls, but in the end what it comes to is for the results that we have predicted in that first paper that came out, it all comes out to how good is those predictions and can it yield something interesting.

Marc Pelletier

You are trying to understand the – how a disease, the molecular pathology is working down to the atomic level, right?

Dr. Vijay S. Pande

Exactly

Marc Pelletier

And this is definitely a goal and I certainly think it's achievable, but where are the most difficult challenges have been, have been protein folding. And if you think about cystic fibrosis, right, CFTR gene, it was cloned in 1989. It was a Cell paper and yet 20 years later, we understood that it's – exactly what's happening to that protein, but yet there is still no molecule out there that can restore the defect found in that protein.

Dr. Vijay S. Pande

Yeah, absolutely. And cystic fibrosis is a great example, because it's another example where it really is – in that case the protein doesn't even misfold. It just folds a little too slowly.

Marc Pelletier

And it's like having a pair of pants but the button is missing on the back of the pants, right. And it is preferable to have pants on than not have any pants at all.

Dr. Vijay S. Pande

Yeah, exactly. And I think there is going to be other cases like that. I think there is cases where there is mutations in P53, one of the key proteins for cancer and this is something that we looked at in folding@home some time ago, where there is interesting hypotheses where P53 has a mutant, that doesn't affect its activity but it just slows its co-translational folding a little bit.

That P53 is believed to fold while – so this is a little more biology than some people might want to hear but that those ribosomes make proteins and you can have two ribosomes or multiple ribosomes on a single piece of RNA. And it is kind of cool, the RNA is sliding along through multiple ribosomes, through multiple machines that are turning that single piece of RNA into multiple copies of a protein. And on the multiple ribosome or polysome you could have P53, this key protein for fighting cancer. 50% of all cancers I think involve a mutation in P53. P53 is folding probably somewhat co-translationally and it needs to form a ligamer, like with four copies of P53 and you can imagine that if you fold just a little too slowly, you are not going to fold on the polysome and so these things will fall off the polysome a little too early and then they get lost and it's hard for them to find their copies again with any sort of speed.

And then you can imagine P53 production might just go down a bit and these things are so sensitive that it's sort of another story that people hypothesize where just even subtle things unfolding could have big impacts, potentially even in cancers.

Marc Pelletier

[67:46] So you know if people were wondering before what folding@home is doing, this is exactly what it's doing. It's trying to solve these problems. And these diseases are so difficult. As a molecular biologist I see them as things that are – the things I'd put on the backburner until we get the technology to be able to tackle them. But that backburner is getting closer and closer and closer when you think about your system which is just the most powerful system out there. It's doing biology, right. So that's kind of neat. Maybe the military, do you know if the military has a system that is running at your speeds?

Dr. Vijay S. Pande

I have no idea.

Marc Pelletier

Do you really? Or are you advising them...

Dr. Vijay S. Pande

It wouldn't surprise me at all. I mean because one way to put it is that if you could buy, especially at military-like budgets, you could buy something, you could build something this scale with military-like budgets...

Marc Pelletier

Okay, so that would mean buying 800,000 PS3s. Or just the components, just the GPUs and then line them up?

Dr. Vijay S. Pande

Yeah, the key thing right now is that is to look at the active CPU. So they could build something if they had interesting GPU code, they could buy higher GPUs and just buy 10s of 1000s of them. So it's not even that much money. And build a big rack machine full of it. I can imagine actually there's interesting

NSA applications that run on GPUs. I'm sure they're already on top of that. So I could imagine that they have something very powerful. There's always been the suspicion that NSA has probably the most powerful supercomputers in the world. So, I can imagine that they have something. I wouldn't be surprised if they had something ten times more powerful. And what would be interesting is to see how similar the architectures are and I can imagine that they could be very similar that they would have, in order to achieve that type of stuff, as long as they can port their stuff to GPUs, they probably would be running on something like a GPU.

Marc Pelletier

Why would they need such CPU power or GPU power? I can understand trying to predict the position, basically the quantum positions of a million atoms in a simulation, is somewhat more demanding than necessarily doing so, something like image analysis on satellite images.

Dr. Vijay S. Pande

Yeah I could imagine – again I should stress, I know nothing about computers. I'm sorry. This is all guessing from watching 24 and things like that.

Marc Pelletier

Fringe, yeah.

Dr. Vijay S. Pande

Fringe, yeah, so, but my understanding is that one of the things that they would love to do is to do machine learning on all these conversations and whatever. So – perhaps including this one. I mean that type of thing probably could be fairly CPU intensive. They probably have huge datasets.

Marc Pelletier

All right. Well, I am certainly glad that you are really cranking out – as we will start calling the petaFLOPS and using it for biology, because it's truly inspirational. And for me it's, when I see some of the stuff I could be doing and right now, we are in the lab, we are doing the wet bench stuff and the wet bench stuff as you said is – can be easier sometimes and we're at that stage. But when we are trying to understand how the drugs that we are developing here are fitting and binding to those proteins so that we can make subtle minor changes to the electrostatic interactions between those drugs and the proteins, I would certainly love to be able to get in touch with you and put...

Dr. Vijay S. Pande

It would be fun to chat. We're always looking for neat experimental collaborators. I mean one of the big challenges is that in the '80s people pitched computational drug design as the future and I think a lot of people were really burned by that. A lot of drug companies got into it and even – there is a great book called the billion dollar molecule, talking about the rise of Vertex and other companies that really were designed around this idea. And you can imagine the computer power available to a pharma in the '80s compared to folding@home now. We are talking easily maybe a factor of a million or something like that. Easily maybe more.

Marc Pelletier

More, more, more.

Dr. Vijay S. Pande

More. Maybe tens of millions or billions. And so, it's not a surprise to me that the '80s calculations really failed and – but because of that there is a lot of skepticism in the pharma community about computation and I think rightfully so. They've got – it's a hard job to do things and I think what is my hope is, is that in addition to actually coming up with interesting small molecules is that we can really demonstrate that this technique is a really viable alternative or viable additional tool to have in your arsenal.

Because pharma right now is in a tough spot. It's in a tough spot with pipelines getting kind of short. You know are not very full and I think – when I talk to, many people in pharma I think they are willing to take any tool that works. They just want to see that it works and I hope we could – but there, and there's this

chicken and egg, because they are not going to buy a billion dollar computer until they know it works and until you have the big machine, you don't know if it works.

And so we could perhaps be a solution to that chicken and egg and the key thing there is that the solution comes from the really generous donations of essentially hundreds of thousands and millions of people throughout the world, that because of their donations, hopefully we can solve that chicken and egg. Along the way make an impact on some of these diseases. But my broader goal is to really change the perception for how drug design could be done. And that could be a probably some of the biggest implications.

Marc Pelletier

It will eventually become easier. It will be just anybody should be able to run a quick experiment and make a prediction and save the two years I've been at the bench, right and slaving. I worked 270 hours in the last three weeks, right, so. At the bench.

Dr. Vijay S. Pande

You can also imagine that there is all this slaving you do and you get to a point where you have a molecule that looks great and it works well either in tissue or in animals and then finally it fails in human tox.

Marc Pelletier

Right.

Dr. Vijay S. Pande

And another area of our – interest of ours, that we're not running at folding@home but could eventually get there, is to do much more with tox. Because I think if we could head off tox earlier to tell you that that compound, even if it is going to work in terms of efficacy in vitro, it's not going to be a good compound to go into clinical trials. That alone will also save peoples.

Marc Pelletier

It could – it's not just time. But if you think about it, getting to the point of tox, right, in a drug development program would cost several million dollars, right. And that money doesn't come back. Because once dead end, it's finished. What you can learn from a drug that fails is not a whole lot other than it fails. It means go to the next one, and you hope you have a few in the pipeline.

Dr. Vijay S. Pande

Yes

Marc Pelletier

Fortunately the structure to the drug target that I am working on came up yesterday down to 1.8 angstrom. So I think once I finish the stuff that can be done on the bench and we narrow it down making those final adjustments it will be really, really cool. So I will definitely knock on your door.

Dr. Vijay S. Pande

Yeah it'll be fun to chat offline, I am curious what the system is.

Marc Pelletier

All right. Okay I'll let everybody go now I've got one of the world leading experts on molecular dynamics here. I'm going to let you guys go and I will stay online with Dr. Vijay Pande from Stanford University and Director of folding@home. Thank you very much for coming on.

Dr. Vijay S. Pande

Thank you. It was a pleasure as always.

Marc Pelletier

I would really like to thank Dr. Vijay Pande, Director of folding@home and Associate Professor of chemistry and of structural biology at Stanford University. I would also like to thank Randal Schwartz, host of FLOSS Weekly on this WEEK in TECH network. So you can find his podcast at twit.tv/floss.

I would also like to thank the folks at Pods in Print for the transcripts to the show, which you can get at futuresinbiotech.com. Pods in Print are fantastic. If you have some transcripts that you need doing, even if they are technical, they'll find the experts that can transcribe them without mistakes, they're really great. If they can handle biotech, they can handle anything.

I would also like to thank Phil Pelletier and Will Hall for the opening and closing themes and also Dane Golden and Leo Laporte at twit.tv for coproducing this podcast.

For comments and suggestions, send me an email at marc@twit.tv. Many thanks for those that have. I really appreciate them and they have been extremely valuable in helping me guide where this podcast is going. There is – a little backlogged, but I'm in grant season and I will be done on May 1. So I look forward to answering those emails soon after that. For Futures in Biotech, I am Mark Pelletier. Thanks for listening.

Randal Schwartz

I am going through a tunnel, I am sorry.

Marc Pelletier

That was a great April Fools on Leo by the way.

Randal Schwartz

Yeah the little sliding video that was so fun yeah. Leo it must be the cable, then I didn't realize he actually had analog cables in the loop, so it actually could have been the cables. That was good.