

Increasing Public Involvement in Structural Biology

Seth Cooper,^{1,*} Firas Khatib,² and David Baker²

¹Department of Computer Science

²Department of Biochemistry

University of Washington, Seattle, WA 98195, USA

*Correspondence: scooper@cs.washington.edu

<http://dx.doi.org/10.1016/j.str.2013.08.009>

Public participation in scientific research can be a powerful supplement to more-traditional approaches. We discuss aspects of the public participation project Foldit that may help others interested in starting their own projects.

It is now easier than ever for the public to get involved in science. The Internet has made it feasible for research groups to easily connect with people all over the world. Personal computers have also become powerful enough to run computationally intensive programs, giving the public the opportunity to contribute to scientific research. Volunteer computing allows the public to share their spare CPU cycles with structural biology researchers in projects like Folding@home (<http://folding.stanford.edu>) and Rosetta@home (<http://boinc.bakerlab.org>).

Actively involving the public in scientific research—often referred to as citizen science—provides a means of engaging people's skills, rather than just their computational resources. This approach has been successfully used in astronomy, for example, to locate celestial objects with the Galaxy Zoo project (Lintott et al., 2009). But how can we increase the ability for the public to get actively involved in structural biology research? Citizen science projects can require a lot of work to set up and maintain, but there is the potential to yield great results if more labs start them.

For the past several years, we have been running the Foldit project (<http://fold.it>), which allows participants to directly manipulate proteins in an online video game (Figure 1). In that time, participants of Foldit have made contributions to a number of scientific publications. They were instrumental in the solution of the crystal structure of the Mason-Pfizer Monkey Virus Retroviral Protease (Khatib et al., 2011) and the design of a novel synthetic enzyme for the Diels-Alder reaction (Eiben et al., 2012).

We're very excited about the possibility for games and other forms of public involvement in science to help advance the field. To our knowledge, there have been a few other projects actively involving the public in structural biology, and we look forward to many more in the future. Structural biology problems involving the analysis of existing molecules and the design of new ones are promising areas for citizen science as structural problems both are amenable to human spatial-reasoning skills and can make enjoyable puzzles. In the spirit of helping others who might be interested in starting their own such projects, here are some elements that were important in the development of Foldit.

Single Quantitative Metric of Success

Foldit is based on the principle that proteins fold to their lowest free-energy states. In structure prediction problems, Foldit participants seek to find the structure with the lowest Rosetta energy (Leaver-Fay et al., 2011) for the sequence they are given. Participants collaborate and compete to find the lowest free energy (in the case of Foldit, the highest scoring) structure and, in design problems, the lowest energy sequence. Biology problems that can be posed as global optimization problems, with a single metric such as energy, are more amenable to gamification than those that are not.

Having software like Rosetta already in place for computing energies, and generating and sampling alternative structures and sequences, gave us a starting point for building on a gaming interface.

Openness to Collaboration in a Variety of Forms

The core of the project has been a very fruitful collaboration between the Computer Science and Engineering Department and the Biochemistry Department at the University of Washington. Both departments were able to bring their knowledge and skills together to make a successful team. We have also been lucky to work with scientists who are interested in collaborating with the public and allowing them to share credit for the discoveries they have made, not fearing that their thunder will be stolen by Foldit participants.

We have also found that participants may prefer to be credited differently than academics—citizen scientists are a lot less interested in coauthorship. The three Foldit participants whose model led to the solution of the MPMV-PR all declined coauthorship, choosing instead to publish under their Foldit team's name. The solution of the previously unsolved crystal structure, arguably our most prominent result, also came out of a collaboration with another biochemistry lab at the Adam Mickiewicz University in Poland. Using the Foldit-generated model, the lab was able to solve the crystal structure via molecular replacement.

Community Support and Fostering a Connection to Science and the Project

None of the results produced by Foldit would have been possible without the hard work of its participants. Because of this, a large effort is put into supporting and fostering the development of the community. Participants can

communicate with each other and the team through a variety of channels, including an in-game chat and forums.

In addition to the usual support, we strive to give participants a connection to the science behind the game. Blog posts, videos, and podcasts from scientists help us to communicate what the participants are working on and why it's important. We also have regular scientist chats and developer chats where participants can communicate directly with the people working on the project to get a view of what is going on behind the scenes and have any burning questions answered. This effort has grown over time with the project, and we have added more channels of information flow between the participants and project team. We've even recently hired a community liaison to help manage this and improve the flow of information.

When the project started, we were fortunate to be able to draw on an existing community to bootstrap ours rather than starting entirely from scratch. Some of our earliest participants came from the Rosetta@home community, a group already interested in protein folding. New projects might look to existing communities of related interest when starting out.

Support for Social Play and Sharing of Expertise

We wanted to apply the power of multiple people's minds working together to solve problems. Foldit participants are not just working on structures in isolation; they have the option to form groups, in which they can share structures to get ideas and build upon each others' work. Our best results have come out of multiple participants working together. It would be interesting to explore other formats for information sharing, and we hope to add functionality for multiple participants to work together simultaneously on the same structure.

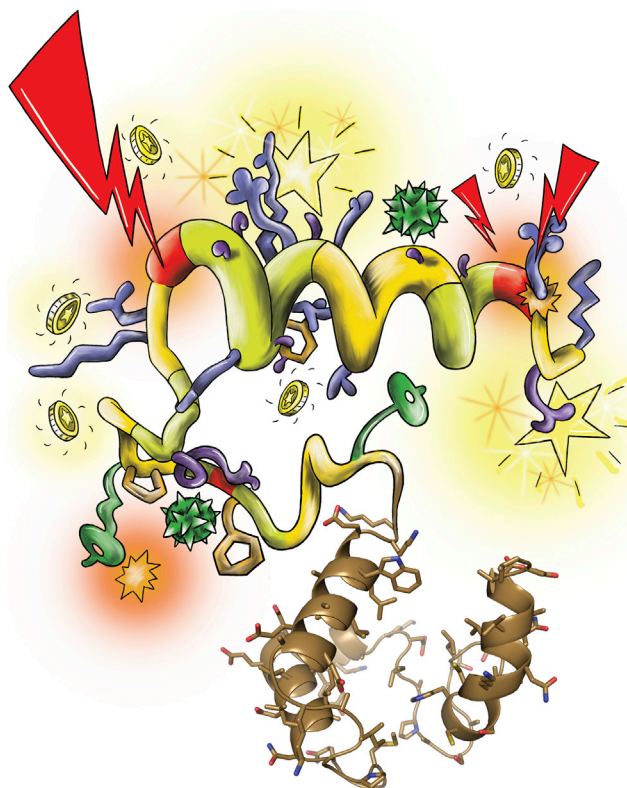


Figure 1. Involving the Public in Science Can Be Both Rewarding for Researchers and Fun for Participants

Giving more experienced participants the opportunity to share what they've learned and help beginners has also been useful. The participants have come up with many interesting ways to play the game, so we also support the sharing of their expertise. Initially, the participants themselves created a wiki (<http://foldit.wikia.com>) to share their folding strategies. We've also added the ability for participants to encode their strategies and share them in the game, and we've seen that they use a variety of techniques.

A Cycle of Continuous Refinement Based on Data

From the beginning, we didn't plan on getting everything right the first time. So we set the project up with the infrastructure and mindset that it would evolve.

Every week, the project team meets to look at structural data gathered in the preceding week, evaluate the data, and feed that back into the plans for what to do next. This iterative process has

allowed us to continually refine the game's utility as a problem-solving tool even as the participants themselves become better problem solvers.

This cycle also means that there is always something new going on with the project. Every week there are new puzzles to solve, and we frequently post updates to the game itself with new features, fixes, and responses to participant feedback. We try to keep the project refreshing and alive and so that when participants come back there is something new and interesting.

Development of an Accessible Interface to Complex Structures and Problems

Protein structures can be complex to look at. Rather than looking at them in a more traditional way, we wanted to take a new look at proteins to cast problems in a manner that can involve

the public. We sought to abstract away many details and highlight areas of the structure where action could be taken.

Having people try out this new interface as early as possible was of the utmost importance. Early versions of the interface went through many rounds of playtesting. Even having just a few novice users try out the interface can illuminate how people want to interact with it and highlight the most critical areas for improvement: where participants would get stuck, confused, or frustrated.

We have found that the game's interface can be useful for protein structure modeling in general. Therefore, we have released a version that is free for noncommercial use, called Foldit Standalone. It can be downloaded from the University of Washington Center for Commercialization website (<http://tinyurl.com/academic-foldit>).

In the last few years, other successful games in structural biology have emerged. Eterna (<http://eterna.cmu.edu>) is allowing participants to help design a library of new RNA shapes.

Participants in EyeWire (<http://eyewire.org>) are helping to map the neurons of the retina. Phylo (<http://phylo.cs.mcgill.ca>) participants are working to solve the multiple sequence alignment (MSA) problem and were recently shown to be able to improve on automated algorithms for MSA (Kawrykow et al., 2012).

We encourage anyone working on difficult structural problems to reach out to the public for help. Certainly, involving the public is not a silver bullet to solve every problem, but public involvement can be a powerful means to complement existing methods. We have consistently

been impressed by Foldit participants' creativity and ingenuity. For the right types of problems, cast in the right way, there is the potential to involve a new group of creative problem solvers. Conversely, games such as Foldit can be a powerful means for popularizing and disseminating structural biology research to the general public.

REFERENCES

- Eiben, C.B., Siegel, J.B., Bale, J.B., Cooper, S., Khatib, F., Shen, B.W., Players, F., Stoddard, B.L., Popović, Z., and Baker, D. (2012). *Nat. Biotechnol.* **30**, 190–192.
- Kawrykow, A., Roumanis, G., Kam, A., Kwak, D., Leung, C., Wu, C., Zarour, E., Sarmenta, L., Blanchette, M., and Waldispühl, J.; Phylo players. (2012). *PLoS ONE* **7**, e31362.
- Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., Popović, Z., et al.; Foldit Contenders Group; Foldit Void Crushers Group. (2011). *Nat. Struct. Mol. Biol.* **18**, 1175–1177.
- Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P.D., Smith, C.A., Sheffler, W., et al. (2011). *Methods Enzymol.* **487**, 545–574.
- Lintott, C., Schawinski, K., Keel, W., van Arkel, H., Bennert, N., Edmondson, E., Thomas, D., Smith, D., Herbert, P., Jarvis, M., et al. (2009). *Monthly Notices of the Royal Astronomical Society* **399**, 129–140.