Research Games in Structural Biology

Nic Vigilante

Oberlin College

Follow this and additional works at: https://digitalcommons.oberlin.edu/fol_research_awards

Part of the Chemistry Commons

Recommended Citation

https://digitalcommons.oberlin.edu/fol_research_awards/7

This Paper is brought to you for free and open access by the Oberlin College Libraries at Digital Commons at Oberlin. It has been accepted for inclusion in Friends of the Libraries Excellence in Research Awards by an authorized administrator of Digital Commons at Oberlin. For more information, please contact megan.mitchell@oberlin.edu.
Research Games in Structural Biology

Nic Vigilante
Oberlin College, Department of Chemistry and Biochemistry, Oberlin, OH 44074

Crowdsourcing, citizen science, research games, bioinformatics, structural biology, protein folding, RNA structure prediction, multiple sequence alignment, public resource computing, Foldit, Phylo, EteRNA

ABSTRACT: As problems in structural biology and bioinformatics have become increasingly complex, innovative new computational methods have been proposed and implemented. Of these methods, research games are unique in how they privilege human intuition over algorithmic verification in certain steps of the research process; researchers in structural biology have developed a number of research games since 2008 to utilize human pattern recognition and spatial manipulation skills. Players are not expected to have any scientific knowledge, as the game interfaces abstract the relevant problems from their biological contexts, and they are not reimbursed or incentivized in any way apart from methods such as in-game leaderboards and earning in-game points. Games such as Foldit (protein structure prediction), EteRNA (RNA structure prediction), and Phylo (DNA sequence alignment) have proven incredibly successful, solving some longstanding problems which had puzzled the scientific community for years.

INTRODUCTION

Structural biology is the area of scientific research concerned with understanding, predicting, and designing the structures of biological macromolecules such as enzymes, RNA, and DNA. A multidisciplinary field, structural biology draws from biochemistry, molecular biology, biophysics, bioinformatics, computer science, and applied mathematics.

Beginning in the late 1960s, techniques from computational chemistry began to be applied to problems in structural biology. In the intervening years, as computational resources have become cheaper and more advanced, researchers have been able to investigate complex systems such as protein folding and macromolecule solvation. Bioinformatics as a field addresses these problems, attempting to apply statistics and mathematics to biological systems using computational methods.

Computational modeling in structural biology is crucial to areas including drug design, disease detection, and vaccine development. Yet even with modern computational resources, some avenues of research remain fundamentally challenging. Three problems which are of central importance to biology and pose unique computational difficulties are protein structure prediction, RNA structure prediction, and multiple sequence alignment.

PROTEIN STRUCTURE PREDICTION. The large size of proteins makes atomic-level modeling near impossible, but coarser approaches often disregard important electronic and thermodynamic interactions. Protein folding modeling is often based upon Monte Carlo algorithms or similar search algorithms; it is difficult for these algorithms to make large structural adjustments, however, and as such it is easy for them to get stuck in local energy minima. The success of these algorithms is also highly dependent on what force field (i.e., method of evaluating various interparticle forces) is being used, and trade-offs must be made between computational expense and accuracy. In order to provide benchmarks for the field, a semi-regular international competition called the Critical Assessment of Techniques for Protein Structure Prediction (CASP) allows labs to test their best methods against one another.

RNA STRUCTURE PREDICTION. RNA structure prediction methods have improved significantly in recent years, with secondary structure prediction becoming more accurate and efficient. Tertiary structure prediction, however, remains computationally challenging. Many of the challenges inherent in protein structure prediction problems are also true of RNA structure prediction problems, including the need to consider trade-offs between efficiency and accuracy.

MULTIPLE SEQUENCE ALIGNMENT. Multiple sequence alignment (MSA) refers to both the act of aligning sequences in DNA and RNA and the end-product of that alignment. MSA is a mathematically complex problem, as both the generation of a MSA and the scoring of that MSA are NP-hard. NP-hard is a designation in computational theory wherein a decision problem (essentially a yes/no question) is at least as difficult as every problem which can be solved in polynomial time. NP-hard problems are very complex, and the only certain method of solving them is through an exhaustive search of every possible solution, which of course is near impossible when examining large structures such as biological macromolecules. While some algorithms have been developed to approximately solve MSA problems, these remain some of the most difficult challenges in computational biology.
HISTORY

In order to understand the role that research games play in structural biology today, it is helpful to trace the history of their emergence in scientific research. The sections below outline the evolution of crowdsourcing and citizen science into public resource computing and eventually into research games.

CROWDSOURCING AND CITIZEN SCIENCE. Although “crowdsourcing” as a term has existed only since 2006, it has been practiced for at least a century, and possibly significantly longer. Perhaps the best-known and earliest example of crowdsourcing is the National Audubon Society’s Christmas Bird Count, which has run since 1900. “Citizen scientists” – that is, members of the general public who may or may not have any specific training or disciplinary expertise – track the type and number of birds that they observe during this event. The National Audubon Society then collects this data as a “census” to be used for conservation efforts.

This practice of engaging the lay public in scientific research has increased dramatically since the advent of the internet. Interested individuals with a computer, an internet connection, and some free time can classify galaxies, assess infrastructural earthquake damage, assist in urban planning, create crisis maps to aid in disaster relief, and more. As social media have become more commonplace, crowdsourcing has continued to gain traction as a viable and sometimes necessary method for gathering publicly distributed information or analyzing large amounts of unprocessed data.

Crowdsourcing traditionally relies upon an individual’s goodwill, curiosity, and/or excitement to be involved in research, with no reward or payment for work performed. The emergence of online marketplaces such as Amazon’s Mechanical Turk has challenged this approach, instead linking companies in need of cheap online labor with individuals looking to make money (usually very little) by performing simple tasks such as image categorization. Using a similar model, new programs such as MobileWorks have instead foregrounded a social mission to provide “a fair baseline wage” to “marginalized populations in the developing world.”

PUBLIC RESOURCE COMPUTING. Meanwhile, during the growth of crowdsourcing practices in other disciplines, fields such as structural biology and computational chemistry were beginning to require more than the computing power which was available to them at the time. Experimental techniques were being developed which allowed for the characterization and quantification of analytes in very small quantities, in highly complex matrices, and with a high degree of structural specificity. There was a demand for similar advancement in computational techniques, but the computing hardware available to many researchers did not allow for such complex modeling without significant simplifications.

In the early 1980s, distributed computing was proposed as an alternative to supercomputing: instead of having one extraordinarily powerful supercomputer running a complex program, distributed computing breaks this program down into many small pieces that can be run on a group of standard computers before having the results “reassembled.” This was first implemented using computers housed within individual universities, but soon this process was outsourced to public computers as well. This led to the growth of public resource computing, also known as peer-to-peer computing or global computing.

Public resource computing was first implemented in 1996 with the Great Internet Mersenne Prime Search (GIMPS). This program used personal computers to search for a specific class of prime number, sometimes with humorous and illegal results. This software was designed to be downloaded by any interested member of the general public with a personal or work computer. The program runs as a background process, using excess computing time that is not claimed by other programs running on the machine. In this way, small chunks of processed data can be collected from a distributed host of public computers without negatively impacting the performance of those computers for user tasks.

This model was soon taken up by researchers in the natural sciences. In 1999, the University of California, Berkeley, released SETI@home, an extension of the Search for Extraterrestrial Intelligence (SETI). In 2004, SETI@home provided more than double the computing power of the most powerful supercomputer in the world at the time. The Berkeley Open Infrastructure for Network Computing (BOINC) was launched just a few years later, introducing a central hub which allows users to browse and engage with a range of different public resource computing projects, from asteroid dynamics to human cognition modeling to unsolved World War II Enigma Codes.

RESEARCH GAMES. One project that can be run through BOINC is Rosetta@home, which performs protein folding calculations using the Rosetta force field. Like many of

Figure 1. Rosetta@home screensaver (from https://boinc.bakerlab.org/rosetta/rah/rah_graphics.php, accessed April 4, 2018).
the public resource computing projects available through BOINC, Rosetta@home runs in the background while the personal computer is in use. When the computer is not in use, a screensaver appears showing visualizations of the energy minimizations being performed (Figure 1).

The inclusion of this visualization method led to a complete rethinking of the role of the “citizen scientist”. The makers of Rosetta@home, the Baker lab at the University of Washington, began to receive letters from the program’s users. While users seemed to enjoy watching the calculations which their computers were performing, many seemed upset at the program’s slow progress. After some time spent watching the computer “try out” different protein folding patterns, many users became frustrated that certain conformations which to them seemed intuitive were not being found by the program.

To test whether human users without any formal training in microbiology would in fact be more efficient at solving protein folding problems than a standard modeling algorithm, the Baker lab developed Foldit in 2008. Similar to BOINC, Foldit is a program which can be downloaded by any user interested in contributing to scientific research. Unlike BOINC, however, Foldit does not passively run in the background; instead, it invites users to actively participate in manipulating a protein in order to search for a low-energy conformation. Designed with a user-friendly interface, Foldit replaces a search algorithm with innate human spatial reasoning.

After some startling and significant successes by the Foldit player community, other labs in structural biology began to take note and develop their own research games. These include Phylo (DNA sequence alignment) and EteRNA (RNA structure prediction).

Foldit, Phylo, and EteRNA are each discussed in turn below, with special attention to their gameplay, methods, and results. Finally, there is a brief discussion on the drawbacks and ethical concerns of research games.

FOLDIT

Foldit was released in 2008 and can be accessed at http://fold.it.

GAMEPLAY. Foldit is playable as an executable download which does not require any manual compilation, available for Windows, Mac, and Linux. As of 2016, Foldit had over 500,000 registered players. It runs as a windowed program, aesthetically falling between a browser-based Flash game and a full-screen high-budget video game.

Gameplay in Foldit consists of physical manipulation of protein backbones and sidechains, actions which players perform directly with their cursor and indirectly using a range of built-in tools. Players can zoom in and out and rotate the full protein in any direction. Using the cursor, sidechains can be manually rotated and the position of backbone segments can be shifted. Further manipulations of the protein structure are possible using the tools discussed below. The visualization methods used in Foldit are clearly reminiscent of and influenced by molecular dynamics interfaces, grounding it in the traditions and aesthetics of computational chemistry and biology (compare to, for example, Visual Molecular Dynamics and Molden).

Various pop-up visualizations are used to give players immediate feedback on the positive or negative consequences of their actions; these include indications of strong hydrogen bonds, exposed hydrophobics, and steric clashes. Examples of these can be seen in Figure 2. Players are also given a lengthy tutorial which teaches them how to navigate the UI and how the game is scored. This tutorial serves both to teach players about the game interface and to provide instruction on some basic chemical concepts (such as hydrogen bonding and hydrophilicity).

TOOLS. Perhaps the most useful tools available to players are “shake” and “wiggle.” Each of these is an automated coarse energy minimization, with “shake” affecting sidechains and “wiggle” affecting the protein backbone. The player observes this process and can choose to stop the minimization or take manual control of it at any point. Players can also “mutate” the protein, changing the identities of specific amino acids. This can be automated through a built-in Monte Carlo-based algorithm or controlled manually by players.

Many other tools are also available to players, from large-scale alterations, such as changing secondary structure characteristics, to small-scale alterations, such as freezing specific sidechains before initiating automated energy minimizations. The Baker lab has sought to give players the ability to accomplish anything that they could possibly consider doing in the puzzles. Since the explicit goal of this program is to harness the spatial cognitive abilities of humans over the algorithmic approaches of computers, giving players as much freedom as possible is crucial.

RETENTION. Once players download and begin playing Foldit, several methods are used to engage them and keep them interested in continuing to play.

The first, already discussed above, is the inclusion of immediate visual feedback whenever an alteration is made to the protein folding. By allowing players to immediately see the consequences of their actions, Foldit creates a sense of investment in personal learning and growth. This is further emphasized through the inclusion of Achievements, which can be earned by reaching specific gameplay goals (e.g., solving a certain number of puzzles).

Secondly, Foldit includes an extensive set of mechanisms for inter-player competition. Leaderboards are included within each individual puzzle, showing one’s own score and rank as well as the score and rank of all other players who have completed the puzzle. The UI draws further attention to this feature by flashing a large “Rank Up 1” whenever a player increases their score enough to move up on the leaderboard. In addition to these leaderboards on individual levels, the Foldit website shows full leaderboards combining scores from all levels played, creating a further achievement level of “global leader.” As a kind of Holy Grail of player achievement, players could be given a
coauthorship on a peer-reviewed paper (though most papers which involve Foldit simply list “Foldit Players” as an additional author).

Foldit also includes mechanisms for team play, in which multiple players work together to solve difficult puzzles. This type of communal space both increases retention by connecting players with each other and allows more complex problems to be solved. It is also possible to use the in-game chat client to send messages to people in your team or anyone else online concurrently.

**RECIPIES.** One of the most unique aspects of Foldit as a research game is its “recipes” mechanic. Players can create recipes which string together manipulations that they find to be often helpful in tandem. Thus, for example, if a player begins with the same initial 15 moves for every protein, they could create a recipe which would instead automate all those steps.

Players are also encouraged to work in teams to write recipes, thus creating a set of “best practices” from the most successful players in the community. These serve both as helpful teaching aids for newer players and useful algorithms for research. In fact, it was demonstrated that a popular communally-constructed recipe called Blue Fuse “showed superior performance to the algorithm developed by professional structural biologists.”

---

**Figure 2.** Foldit screenshots (taken April 4, 2018). **a)** Title screen; **b)** main menu; **c)** first tutorial level; **d)** example of visual feedback text showing results of player’s move; **e)** example of visual feedback symbols (blue/white striped rods indicate well-aligned hydrogen bonds; red spheres indicate internal vacancy in protein; red spheres with spikes indicate steric clashes between sidechains; yellow spheres with small knobs indicate exposed hydrophobics); **f)** example of a full Foldit level (top-center ~ personal score, score rank relative to other players who have completed this level, and level details; mid-left ~ available player generated recipes; bottom-left ~ protein manipulation tools available to players; bottom-right ~ inter-player chat box; top-right ~ leaderboard highlighting highest player and team scores on this level).
METHODS. Under the hood of Foldit, so to speak, the Rosetta force field is running constantly. Player scores are directly linked to energies calculated using the Rosetta force field; every time a player manipulates the protein, the total intramolecular energy of the protein is calculated using Rosetta and a score is output to the player’s UI. Visual aids such as hydrogen bonding indicators and exposed hydrophobics indicators are simply extensions of normal Rosetta calculations.

A force field is a method of describing the inter- and intramolecular potential energy levels of a molecule or set of molecules. A force field is not an algorithm; that is, it does not perform any calculations itself. Instead, it provides the mathematical scaffolding on which the relevant algorithms are run. Essentially, a force field enables a researcher to communicate with a computer and tell it the identity of the particles being examined (depending on the size of the model, this could be anything from molecules to atoms to, hypothetically, sub-atomic particles) and how those particles interact with each other.

There are many different force fields in use by different corners of the scientific computation community. They differ widely in their efficiency (how long it takes to run an algorithm using that force field) and their accuracy (how well experimental results or full ab initio calculations can be matched by an algorithm running the force field). Unfortunately, efficiency and accuracy in force fields are often inversely related to one another: the more exact a calculation is, the longer it takes for the computer to perform it. Thus, tradeoffs are made based on the needs within a specific lab. If a lab is investigating large systems involving macromolecules such as proteins on long time scales, they might use a coarse-grained force field such as AMBER; alternatively, if a lab is performing small molecule molecular dynamics on short time scales, they might use a fine-grained force field such as OPLS.

Rosetta resembles AMBER more than it does OPLS in many ways; that is, it is a coarse-grained force field specifically designed for complex biological macromolecules. Important potentials such as van der Waals interactions and the free energy of solvation are calculated, but significant approximations are made. For example, bond lengths and angles are kept fixed, conformational entropy is not calculated, and torsional potentials are adapted from experimental data rather than being explicitly calculated. All of this makes Rosetta a useful energy function for proteins: it incorporates enough atomic-level specificity to account for important intramolecular interactions such as hydrogen bonding, but is approximate enough to use up relatively little computational time.

In this way, Rosetta force field calculations are constantly running in the background of Foldit. Most players will likely not be aware of these calculations, as there is essentially no lag between manipulating the protein and seeing the resulting score. This is an important characteristic of Foldit, as it allows for continuity between actions taken by the player without having to wait for the energy calculations to catch up.

RESULTS. Foldit has been very successful since its release in 2008. Player generated “recipes” have been shown to outperform professionally developed algorithms in certain circumstances, and results from the game held their own when used as part of submissions to CASP10 (see Introduction) in a consortium called WeFold. In addition, in 2012 Foldit players were able to develop a “helix-turn-helix” motif that enabled researchers to successfully remodel a computationally-designed enzyme.

Foldit’s most important success was determining the crystal structure of a protein crucial to understanding and treating HIV/AIDS, the Mason-Pfizer monkey virus retroviral protease. This problem had remained unsolved despite fifteen years of active research in the scientific community; Foldit players determined the structure within ten days.

PHYLO

Phylo was released in 2010 and can be accessed at http://phylo.cs.mcgill.ca/.

GAMEPLAY. Phylo is a browser-based, Unity-powered game which runs in a windowed screen but can be set to a cinematic full-screen mode. No download is required. As of 2011, seven months after its release, Phylo had over 12,000 registered players; unfortunately, no more recent usage statistics are available. The active scaling of the game to fit your screen, the pop-jazz background music (which Dr. Jérôme Waldispühl of the developing lab at McGill himself helped compose), bright and clear sound effects, and bright, modular colors gives Phylo the feel of a smartphone game- and sure enough, it is also available for download on Apple and Android devices.

Puzzles in Phylo are broken up into several different stages, where each stage involves manipulating different rows of colored blocks. Once a row is introduced into the puzzle, it stays there until the end of the last stage; thus, by the end of the puzzle, players must work with many rows at once, often revising their solutions to prior stages of the puzzle. On the left side of the screen is shown an abstracted phylogenetic tree, and every time a new row is introduced so is the clip art of an animal accompanying it. Players are responsible first for matching closely related nodes on the phylogenetic tree, and later for relating more disparate parts of the tree to one another.

RETENTION. Like Foldit, Phylo’s strategies for retaining players include a significant emphasis on inter-player competition. Within each puzzle, players are shown three large numbers: their score, the score goal, and the top player score. While players are given the option to move on to the next stage of the puzzle once they have reached the score goal, they are encouraged to keep playing in an attempt to reach the top player score (which is often much higher than the score goal). At the end of an entire puzzle, players are shown not only their score but also their percentile ranking compared to other users and between one and
three stars based on that ranking. A global leadership board based on total score can also be accessed from the Phylo’s homepage.

Another similarity between Foldit and Phylo is the inclusion of achievements. Like Foldit, achievements in Phylo are unlocked not only by reaching personal milestones in gameplay but also by competing with other players.

One important difference from Foldit is how Phylo seeks to retain players through increased portability. As a browser-based game, players can log into their account from any computer (unlike Foldit, where the game must be downloaded and installed on each computer on which it is run). In addition, the availability of a smartphone version gives players an additional method by which to play. All player data is cloud-based, and thus games played on a laptop, a desktop, and a smartphone all contribute to your total score and player ranking.

Finally, Phylo emphasizes repeatedly throughout the game that players are contributing to disease research. Beneath the “Phylo” logo on the home screen is a subtitle: “Solve a puzzle and help genetic research.” When selecting a puzzle, players choose one of eight categories which highlight the puzzles’ biological underpinnings (such as “Cancers” and “Metabolic Diseases”). While there are no differences at all in gameplay or aesthetic design between these categories, delineating puzzles in this manner forces players to more directly acknowledge and engage with Phylo’s research function. Every “Level Completed” screen...
reiterates what disease category the puzzle was classified within, and there is an achievement for completing a puzzle in every disease category.

**METHODS.** As was discussed above, Phylo heavily abstracts MSA problems in order to make them accessible and fun to the general public. Each colored block within the game refers to a specific nucleotide, though players are never explicitly told this. The puzzles included in Phylo incorporate abstractions of small portions of whole genomes; where mammalian genomes might be billions of nucleotides long, the puzzles solved by players are only about 24 nucleotides long. This makes the puzzles solvable in a fairly short period of time.

Phylo is based upon the results generated by Multiz, a powerful program designed to tackle large 44-species MSA problems. The Phylo team extracts portions of Multiz’s results which are likely to be misaligned and opens these problems up to the player community. Once sufficient data are gathered, that portion of the MSA is improved based on player results and integrated back into the full MSA (see Figure 4). The score goal value that is shown to players is in fact the result of the Multiz program, and any score higher than that is an improvement upon the results of the program.

Scoring is based upon a simplified version of the BLASTZ algorithm, a tool which assigns scores for pairwise comparisons between nucleotides. Phylo simplifies these scores by making them round integers, as these are quicker to calculate in response to the player’s actions and are also easier for the player to understand. The developers claim that the “slight accuracy loss” that results from this approximation is compensated for by the increased speed of calculation.

**RESULTS.** One year after Phylo’s release, it was found that player-generated alignments outperformed Multiz results 70% of the time, though it was unclear how they compared to standard realignment algorithms. Although there was improvement over Multiz a significant majority of the time, the magnitude of these improvements was often relatively small. It was also seen that when these extracted sequences were reintroduced into their genomic context score improvements dwindled and in some cases disappeared. Overall, Phylo appears to be a promising supplement to traditional computational MSA methods, though it has not yet achieved any of the monumental breakthroughs that Foldit has.

**EteRNA**

EteRNA was released in 2011 and can be accessed at www.eternagame.org.

**GAMEPLAY.** EteRNA is a browser-based Flash game which does not require a download and has over 250,000 registered players. The aesthetic design of the game falls in-between Foldit and Phylo; it is not designed to replicate molecular dynamics visualizations as Foldit is, but neither does it attempt to completely abstract the scientific problems under consideration as Phylo does. Its UI resembles the “computer game” atmosphere of Foldit more than the “app game” atmosphere of Phylo, though its design is sleeker and less cluttered than Foldit’s. This design includes small bubbles constantly appearing and floating upwards in the background, creating a calm, soothing feel.

In EteRNA, players switch the identities of amino acids within RNA molecules, symbolized by different colored circles. By controlling the identities of these amino acids, players seek to create a secondary structure which minimizes energy. Players are given continuous feedback in the form of a total energy display and a map of the molecule indicating where unfavorable interactions are occurring.

Players must progress through a lengthy tutorial before contributing to real research problems. This tutorial focuses less on the (few) in-game tools available to the players and more on strategies for creating amino acids sequences to achieve a desired structural or energetic outcome.

In addition to crediting “EteRNA players” as a coauthor on scientific papers, EteRNA developers are now seeking to involve players in the paper review process as well. Players are invited to view and comment on manuscripts before they are submitted for publication, in what the EteRNA team is calling “community review” and “democratized writing.”

**RETENTION.** Like Foldit and Phylo, EteRNA uses inter-player competition to boost player investment in the game. A global leaderboard ranks players based on their “wealth,” measured by in-game currency earned through completing puzzles. Global rank is also prominently shown on each player’s profile, and players can earn badges by completing specific tasks.
The truly unique aspect of EteRNA lies outside of the game itself. Each week, the eight top-scoring RNA structures are synthesized and tested for stability in an actual laboratory setting. This both leads to a heightened sense of competition within the player community and allows for a more direct connection to the laboratory science which this game is supporting. Each player’s “synthesizes” score (the number of times one of their structures has been chosen for synthesis in the lab) is also shown on the leaderboard.

METHODS. EteRNA uses structure mapping scores to give feedback to players and evaluate submitted structures. These scores are based on single-nucleotide pairing and reactivity, where each individual nucleotide contributes points based on its immediate surroundings.

RESULTS. EteRNA players have consistently and significantly outperformed many standard computational models. Recently, EteRNA has taken steps towards developing a fast and sensitive RNA-signature-based detective method for tuberculosis.

It was observed that many players were successfully designing unusual RNA features. To evaluate these, experimental tests were paired with a Monte Carlo algorithm called EteRNABot which includes some of these features. While EteRNABot was not quite able to match player scores, it was able to outperform both NUPACK and RNAinverse, two standard computational RNA folding models.

DRAWBACKS OF RESEARCH GAMES

While research games have a wide range of potential benefits, such as low operating cost and extensive “computational power” based on human spatial reasoning abilities, there are drawbacks associated with their use. The data generated by a research game are dependent upon the actions of that game’s players, and the actions of a game’s players are controlled and limited by the architecture of the game. In other words, a game’s players are only as good as the design of the game itself. This can play out in two ways: a player being unable to complete a given action and a player being unable to consider a given action. If a player wanted to accomplish a certain action that was outside the scope of what was pre-programmed into the game, that player’s human cognition is being limited by the architecture of the game. The second potential situation is best described with an example.

In all three of the research games described above, players are instructed to maximize their score (or minimize their energy). The players are given differing amounts of information on how that score is calculated. Phylo outlines its scoring system very clearly for players, while Foldit never describes how score is calculated. Regardless of how cognizant players are of the scoring system, however, they are affected by the limitations of that scoring system. If the Rosetta force field is unable to accurately calculate a specific kind of sidechain interaction, for example, players will receive positive reinforcement for incorrect actions and inaccurate data will be generated. Game designers must balance speed of calculation, which is neces-
sary for providing immediate feedback to players, and accuracy of calculation, which provides players with feedback that is useful.

Additionally, some ethicists have argued that there are serious ethical issues associated with crowdsourcing and research games, and they have proposed that all future work of this kind be reviewed by Institutional Review Boards (IRBs)\(^{46}\). They argue that potential negative outcomes of research games include contributing to internet addiction and preventing players from spending more time with their families. They also argue that since playing games often results in dopamine release in players, research games should be regulated in the same way as studies wherein human subjects would be injected with dopamine.

**CONCLUSION**

Research games such as Foldit, Phylo, and EteRNA provide researchers with a novel analytical tool: human spatial reasoning abilities. By abstracting problems in biology into puzzles accessible to non-experts, researchers can gather large amounts of crowdsourced data, and these data have in many cases been shown to be more accurate than those generated by traditional computational algorithms.

Many other research games exist beyond the three profiled here. In biology alone, examples include Ribo\(^{47}\), Dizeez\(^{48}\), The Cure\(^{49}\), Fraxinus\(^{50}\), and Nanocrafter\(^{51}\). Researchers are also beginning to create more collaborative methods for crafting research games, such as Open-Phylo, a system allowing any research team to submit sequences into the Phylo game\(^{52}\).

Games as a research tool are still in their infancy; while there have been some notable and well-publicized successes, much remains to be done to better integrate these games into existing laboratory procedures. Player communities have continued to grow since the introduction of the first research games a decade ago, and an increase in federal funding (National Science Foundation and National Institutes of Health in particular) and non-profit funding (from organizations such as the Bill and Melinda Gates Foundation) indicate that research games continue to be an exciting frontier in the sciences.

All research games and distributed computing projects mentioned in this paper are available online free of charge.

**AUTHOR INFORMATION**

**Corresponding Author**

*Nic Vigilante, Oberlin College, Department of Chemistry and Biochemistry, Oberlin, OH 44074 nvigilante@oberlin.edu*

**ABBREVIATIONS**

BOINC, Berkeley Open Infrastructure for Network Computing; CASP, Critical Assessment of Techniques for Protein Structure Prediction; GIMPS, Great Internet Mersenne Prime Search; IRB, Institutional Review Board; MSA, multiple sequence alignment; SETI, Search for Extraterrestrial Intelligence; UI, user interface.

**REFERENCES**
