hhmi Howard Hughes Medical Institute

Published as: Nature. 2010 August 5; 466(7307): 756-760.

Predicting protein structures with a multiplayer online game

Seth Cooper¹, Firas Khatib², Adrien Treuille^{1,3}, Janos Barbero¹, Jeehyung Lee³, Michael Beenen¹, Andrew Leaver-Fay^{2,4}, David Baker², Zoran Popovi¹, and >57,000 Foldit players⁵ ¹Department of Computer Science & Engineering, University of Washington

²Department of Biochemistry, University of Washington

³Department of Computer Science, Carnegie Mellon University

Abstract

People exert significant amounts of problem solving effort playing computer games. Simple image- and text-recognition tasks have been successfully crowd-sourced through gamesi, ii, iii, but it is not clear if more complex scientific problems can be similarly solved with human-directed computing. Protein structure prediction is one such problem: locating the biologically relevant native conformation of a protein is a formidable computational challenge given the very large size of the search space. Here we describe Foldit, a multiplayer online game that engages nonscientists in solving hard prediction problems. Foldit players interact with protein structures using direct manipulation tools and user-friendly versions of algorithms from the Rosetta structure prediction methodologyiv, while they compete and collaborate to optimize the computed energy. We show that top Foldit players excel at solving challenging structure refinement problems in which substantial backbone rearrangements are necessary to achieve burial of hydrophobic residues. Players working collaboratively develop a rich assortment of new strategies and algorithms; unlike computational approaches, they explore not only conformational space but also the space of possible search strategies. The integration of human visual problem-solving and strategy development capabilities with traditional computational algorithms through interactive multiplayer games is a powerful new approach to solving computationally-limited scientific problems.

While it has been known for over 40 years that the three dimensional structures of proteins are determined by their amino acid sequencesv, protein structure prediction remains a largely unsolved problem for all but the smallest protein domains. The state-of-the-art Rosetta structure prediction methodology, for example, is limited primarily by conformational sampling; the native structure almost always has lower energy than any non-native conformation, but the free energy landscape that must be searched is extremely large

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Correspondence to: Zoran Popovi , David Baker, and Adrien Treuille. Requests for materials should be addressed to S.C. or F.K. (scooper@cs.washington.edu or firas@u.washington.edu).

⁴present address Department of Biochemistry, University of North Carolina

^{5&}lt;sup>r</sup>Worldwide

All authors contributed extensively to the work presented in this paper.

Supplementary Information accompanies the paper on www.nature.com/nature.

—even small proteins have on the order of 1000 degrees of freedom—and rugged due to unfavorable atom-atom repulsion which can dominate the energy even quite close to the native state. To search this landscape, Rosetta uses a combination of stochastic and deterministic algorithms: rebuilding all or a portion of the chain from fragments, random perturbation to a subset of the backbone torsion angles, combinatorial optimization of protein sidechain conformations, gradient based energy minimization, and energy-dependent acceptance or rejection of structure changesvi, vii, viii.

We hypothesized that human spatial reasoning could improve both the sampling of conformational space and the determination of when to pursue suboptimal conformations if the stochastic elements of the search were replaced with human decision making while retaining the deterministic Rosetta algorithms as user tools. We developed a multiplayer online game, Foldit, with the goal of producing accurate protein structure models through gameplay (Fig. 1). Improperly folded protein conformations are posted online as puzzles for a fixed amount of time, during which players interactively reshape them in the direction they believe will lead to the highest score (the negative of the Rosetta energy). The player's current status is shown, along with a leaderboard of other players, and groups of players working together, competing in the same puzzle (Fig. 1, arrows 8-9). To make the game approachable by players with no scientific training, many technical terms are replaced by terms in more common usage. We remove protein elements that hinder structural problem solving, and highlight energetically frustrated areas of the protein where the player can likely improve the structure (Fig. 1, arrows 1-5). Sidechains are colored by hydrophobicity and the backbone is colored by energy. There are specific visual cues depicting hydrophobicity ("exposed hydrophobics"), interatomic repulsion ("clashes"), and cavities ("voids"). The players are given intuitive direct manipulation tools. The most immediate method of interaction is directly pulling on the protein. It is also possible to rotate helices and rewire beta sheet connectivity ("tweak"). Players are able to guide moves by introducing soft constraints ("rubber bands") and fixing degrees of freedom ("freezing") (Fig. 1, arrows 6-7). They are also able to change the strength of the repulsion term to allow more freedom of movement. Available automatic moves-combinatorial sidechain rotamer packing ("shake"), gradient-based minimization ("wiggle"), fragment insertion ("rebuild")—are Rosetta optimizations modified to suit direct protein interaction and simplified to run at interactive speeds.

To engage players with no previous exposure to molecular biology, it was essential to introduce these concepts through a series of introductory levels (Fig. S1 and Table S1): puzzles that are always available, and can be completed by reaching a goal score. These levels teach the game's tools and visualizations, and certain strategies. We have found the game to be approachable by a wide variety of people, not only those with a scientific background (Fig. S2); in fact, few top players are professionally involved in biochemistry (Fig. S3).

To evaluate players' abilities to solve structure prediction problems, we posted a series of prediction puzzles. Puzzles in this series were blind, in the sense that neither the target protein nor homologous proteins had structures contained within publicly available databases for the duration of the puzzles. Detailed information for these 10 blind structures,

including comparisons between the best scoring Foldit predictions and the best scoring Rosetta predictions using the rebuild and refine protocol⁷, is given in Table 1. We found that Foldit players were particularly adept at solving puzzles requiring substantial backbone remodeling to bury exposed hydrophobic residues into the protein core (Fig. 2). When a hydrophobic residue points outward into solvent, and no corresponding hole within the core is evident, stochastic Monte Carlo trajectories are unlikely to sample the coordinated backbone and sidechain shifts needed to properly bury the residue in the core. By adjusting the backbone to allow the exposed hydrophobic residue to pack properly in the core, players were able to solve these problems in a variety of blind scenarios including a register shift and a remodeled loop (Fig. 2a-b), a rotated helix (Fig. 2c), two remodeled loops (Fig. 2d), and a helix rotation and remodeled loop (Fig. 2e).

Players were also able to restructure beta sheets in order to improve hydrophobic burial and hydrogen bond quality. Automated methods have difficulty performing major protein restructuring operations to change beta sheet hydrogen-bond patterns, especially once the solution has settled in a local low-energy basin. Players were able to carry out these restructuring operations in such scenarios as strand swapping (Fig. 3) and register shifting (Fig. 2a). In one strand swap puzzle, Foldit players were able to get within 1.06 Å of the native, with the top scoring Foldit prediction being 1.36 Å away. A superposition between the starting Foldit puzzle, the top scoring Foldit solution, and model 1 of the native NMR structure 2kpo are shown in Fig. 3b. Rosetta's rebuild and refine protocol, however, was unable to get within 2 Å of the native structure (Fig. 3a, yellow points). This example highlights a key difference between humans and computers. As shown in Fig. 3c, solving the strand swap problem required substantially unraveling the structure (Fig. 3c, bottom), with a corresponding unfavorable increase in energy (Fig. 3c, top). Players persisted with this reconfiguration despite the energy increase because they correctly recognized the swap could ultimately lead to lower energies. In contrast, while the Rosetta rebuild and refine protocol did sample some partially swapped conformations (Fig. 3a, leftmost yellow point), these were not retained in subsequent generations due to their relatively high energies, resulting in the top Rosetta prediction being further from the native than the starting structure (Fig. S5).

Human players are also able to distinguish which starting point will be most useful to them. Fig. 3d-e shows a case where players were given ten different Rosetta predictions to choose from. Players were able to identify the model closest to the native structure, and to improve it further. Given the same 10 starting models, the Rosetta rebuild and refine protocol was unable to get as close to the native as the top scoring Foldit predictions.

Foldit players performed similarly to the Rosetta rebuild and refine protocol for three of the 10 blind puzzles (Fig. S6). They outperformed Rosetta on five of the puzzles (Figs. 3, S5, and S7), including the two above cases where players performed significantly better. A larger set of successful solutions for similar, though non-blind, puzzles are described in Figs. S8, S9, and S10. For two of the 10 blind puzzles, the top Rosetta rebuild and refine prediction was numerically better than the Foldit solution (Table 1) but still basically incorrect (RMSD to native structure > 5.7 Å) (Fig. S11).

Despite the promising results described above, there still exists room for improvement. For one particularly difficult class of problems, players are only given an extended protein chain to start from. Although the Foldit tools are sufficient to reach the native conformation from this unfolded start (Fig. S12), players can have trouble reaching it from so far away (Fig. S11a). This indicates the need to find the right balance between humans and computational methods; players guided by visual cues perform better in resolving incorrect features in partially correct models than "blank slate" de novo folding of an extended featureless protein chain.

As interesting as the Foldit predictions themselves is the complexity, variation and creativity of the human search process. Foldit gameplay supports both competition and collaboration between players. For collaboration, players can share structures with their group members, and help each other out with strategies and tips through the game's chat function, or across the wiki. The competition and collaboration create a large social aspect to the game, which alters the aggregate search progress of Foldit and heightens player motivation. As groups compete for higher rankings and discover new structures, other groups appear to be motivated to play more (Fig. S14a), and within groups the exchange of solutions can help other members catch up to the leaders (Fig. S14b).

Humans use a much more varied range of exploration methods than computers. Different players use different move sequences, both according to the puzzle type and throughout the duration of a puzzle (Fig. 4a). For example, some players prefer to manually adjust sidechains; some will forego large amounts of continuous minimization at the beginning of a puzzle, but increase it as the puzzle progresses; and some prefer a more direct approach and use more rubber bands when the puzzle begins from an extended chain. Within teams, there is often a division of labor; some players specialize in early stage openings, others in middle and end game polishing. Our informal investigation revealed a fascinating array of thought processes, insights and previously unexplored methodologies developed solely through Foldit gameplay (see Supplemental Text, Player Testimonials section and Table S3 for more information).

In designing Foldit we sought to maximize both engagement by a wide range of players (a requirement common to all games), and the scientific relevance of the game outcomes (unique to Foldit). We fine-tuned the game through continuous iterative refinement based on observations of player activity and feedback, taking approaches from players who did well and making them accessible to all players. Most of the tools available to players today are a product of this refinement. They either did not initially exist or have undergone major revision. The introductory levels were also iteratively tuned to reduce player attrition due to difficulty or lack of engagement. Just as Foldit players gained expertise by playing Foldit, both individually and collectively, the game itself adapted to players' best practices and skill sets. We suspect that this process of co-adaptation of game and players should be applicable to similar scientific discovery games.

To attract the widest possible audience for the game and encourage prolonged engagement, we designed the game so that the supported motivations and the reward structure are diverse, including short-term rewards (game score), long-term rewards (player status and rank),

social praise (chats and forums), the ability to work individually or in a team, and the connection between the game and scientific outcomes. A survey of Foldit players (Fig. S4) revealed that while the purpose of contributing to science is a motivating factor for many players, Foldit also attracts players interested in *achievement* through competition and point accumulation, *social interaction* through chat and web-based communication, and *immersion* through engaging gameplay and exploration of protein shapesix. We expect generally future scientific discovery games will also benefit from varied motivation sets.

There is still much to be learned about the basis for human achievement with Foldit, which will require more specific analysis of how players acquire domain expertise through gameplay, and can discover promising solutions. Such insights could also lead to improved automated algorithms for protein structure prediction.

The solution of challenging structure prediction problems by Foldit players demonstrates the considerable potential of a hybrid human-computer optimization framework in the form of a massively multiplayer game. The approach should be readily extendable to related problems, such as protein design and other scientific domains where human 3D structural problem solving can be leveraged. Our results suggest that scientific advancement is possible if even a small fraction of the energy that goes into playing computer games can be channeled into scientific discovery.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank David Salesin, Kathleen Tuite, Joshua Snyder, Daniel Suskin, Philipp Krahenbuhl, Alex Cho Snyder, Hao Lu, Lu Sien Tan, Alex Chia, Ming Yao, Eric Butler, Chris Carrico, Philip Bradley, Ian Davis, David Kim, Rhiju Das, Will Sheffler, James Thompson, Oliver Lange, Robert Vernon, Bruno Correia, David Anderson, Yutong Zhao, Sayer Herin, and Benjamin Bethurum for their help. We would like to thank Nobuyashu Koga, Rie Koga, and Ashley Deacon and the JCSG for providing us with protein structures before their public release. We would also like to acknowledge all the Foldit players who have made this work possible. Players whose solutions were used in figures can be found in Table S4. This work was supported by NSF grants IIS0811902 and 0906026, DARPA grant N00173-08-1-G025, the DARPA PDP program, the Howard Hughes Medical Institute (D.B.), Microsoft, and an NVIDIA Fellowship. This material is based upon work supported by the National Science Foundation under a grant awarded in 2009.

References

- i. von Ahn, L.; Dabbish, L. CHI '04: Proceedings of the 2004 Conference on Human Factors in Computing Systems. ACM Press; 2004. Labeling images with a computer game; p. 319-326.URL http://dx.doi.org/10.1145/985692.985733
- ii. von Ahn, L.; Liu, R.; Blum, M. CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM; New York, NY, USA: 2006. Peekaboom: a game for locating objects in images; p. 55-64.URL http://dx.doi.org/10.1145/1124772.1124782
- iii. Westphal AJ, et al. Non-destructive search for interstellar dust using synchrotron microprobes. Proceeding of ICXOM20, 20th International Congress on X-ray Optics Microanalysis.
- iv. Rohl, C.; Strauss, C.; Misura, K.; Baker, D. Methods in Enzymology. Vol. 383. Department of Biochemistry and Howard Hughes Medical Institute, University of Washington, Seattle; Washington 98195, USA: 2004. Protein structure prediction using Rosetta; p. 66-93.URL http:// dx.doi.org/10.1016/S0076-6879(04)83004-0

- v. Anfinsen CB. Principles that govern the folding of protein chains. Science. 1973; 181:223–230. URL http://view.ncbi.nlm.nih.gov/pubmed/73220860. [PubMed: 4124164]
- vi. Das R, Baker D. Macromolecular modeling with Rosetta. Annual Review of Biochemistry. 2008; 77:363–382. URL http://dx.doi.org/10.1146/annurev.biochem.77.062906.171838.
- vii. Qian B, et al. High-resolution structure prediction and the crystallographic phase problem. Nature. 2007; 450:259–264. URL http://dx.doi.org/10.1038/nature06249. [PubMed: 17934447]
- viii. Bradley P, Misura KMS, Baker D. Toward high-resolution de novo structure prediction for small proteins. Science. 2005; 309:1868–1871. URL http://dx.doi.org/10.1126/science.1113801.
 [PubMed: 16166519]
- ix. Yee N. Motivations of Play in Online Games. Journal of CyberPsychology and Behavior. 2007; 9:772–775. [PubMed: 17201605]
- x. DeLano, WL. DeLano Scientific. Palo Alto, CA, USA: 2002. The PyMOL Molecular Graphics System. URL http://www.pymol.org
- xi. Rosato A, et al. CASD-NMR: critical assessment of automated structure determination by NMR. Nature Methods. 2009; 6:625–626. URL. 10.1038/nmeth0909-625 [PubMed: 19718014]

Cooper et al.



Figure 1. Foldit screenshot illustrating tools and visualizations

The visualizations include a clash representing atoms that are too close (arrow 1); a hydrogen bond (arrow 2); a hydrophobic sidechain with a yellow blob because it is exposed (arrow 3); a hydrophilic sidechain (arrow 4); and a segment of the backbone that is red due to high residue energy (arrow 5). The players can make modifications including bands (arrow 6), which add constraints to guide automated tools and freezing (arrow 7), which prevents degrees of freedom from changing. The GUI includes information about the player's current status, including score (arrow 8); a leaderboard (arrow 9), which shows the scores of other players and groups; toolbars for accessing tools and options (arrow 10); chat for interacting with other players (arrow 11); and a cookbook for making new automated tools or "recipes" (arrow 12).



Figure 2. Structure prediction problems solved by Foldit players

Examples of blind structure prediction problems in which players were successfully able to improve structures. Native structures are shown in blue, starting puzzles in red, and top scoring Foldit predictions in green.

(a) The red starting puzzle had a register shift and the top scoring green Foldit prediction correctly flips and slides the beta strand.

(**b**) On the same structure as above, Foldit players correctly buried an exposed Isoleucine in the loop on the bottom right by remodeling the loop backbone.

(c) The top scoring Foldit prediction correctly rotated an entire helix that was misplaced in the starting puzzle.

(d) The starting puzzle had an exposed Isoleucine and Phenylalanine on the top, as well as an exposed Valine on the bottom left. The top scoring Foldit prediction was able to correctly bury these exposed hydrophobic residues.

(e) Another successful Foldit helix rotation that correctly buries an exposed Phenylalanine. Images were produced using PyMOL softwarex.



Figure 3. Puzzles in which human predictors outperform the Rosetta rebuild and refine protocol Panels a, b, and c show puzzle 986875. Panels d and e show puzzle 986698.
(a) Comparison of Foldit player solutions (green) to the low energy structures sampled in Rosetta rebuild and refine trajectories (yellow) for blind Foldit puzzle 986875 based on the recently determined structure and sequence of 2kpo. The x-axis is the all-atom RMSD to 2kpo, and the y-axis is the Rosetta energy. The starting Foldit puzzle was 4.28 Å away from the native structure (shown by the black dot on the plot); Foldit players sampled many different conformations, with the top scoring submission (the lowest scoring Rosetta energy) 1.4 Å away from the native, while the automated Rosetta protocol did not sample below 2Å. The blue dots and lines correspond to the trajectory of a single Foldit player in c.
(b) Superposition of the top-scoring Foldit prediction in green with the experimentally determined NMR model 1 in blue. The starting puzzle is in red, where the terminal strand is incorrectly swapped with its neighbor, 8% of all Foldit players were able to correctly swap these strands (Table S2).

(c) A score trajectory with selected structures for the top scoring player in puzzle 986875 over a two hour window, showing how the player explores through high energy conformations to reach the native state. The y-axis is the Rosetta energy and the x-axis is the elapsed time in hours. The starting structure had a Rosetta energy of -243. Each point in the plot represents a solution produced by this player. The first structure (c1) is near the starting puzzle structure, shown as the black dot in **a**. The following structures (c2-6) are shown as blue dots in plot **a**. In structures c2-4 the player must explore higher energies to move the strand into place, shown by the blue lines. In structures c5-6 the player refines the strand pairing.

(d) Comparison of Foldit player solutions (green) to the low energy structures sampled in Rosetta rebuild and refine trajectories (yellow) for blind Foldit puzzle 986698 based on the

recently determined structure and sequence of 2kky. Foldit players were able to get the best Foldit score by correctly picking from multiple alternative starting Rosetta models (black) the model that was closest to the the native structure.

(e) The native structure is shown in blue with the top scoring Foldit prediction shown in green. The top Rosetta rebuild and refine prediction given the same 10 starting models (shown in yellow) was unable to sample as close to the native as the Foldit players.





Figure 4. Player move preferences

(a) Different Foldit players take different approaches to solving the same problem. Each circle represents the move type frequencies used in the top solution produced by each player in different time frames: the inner denotes the first hour, the middle denotes the first day, and the outer denotes the puzzle's entire duration. Each color represents a different type of move that can be made in the game. The left column reflects player move types for puzzles that start relatively close to the native topology. The right column reflects player move types for puzzles that start from a fully extended conformation. Each row represents a different Foldit player. Each player's preferred move types across each puzzle class are distinct from one another, yet a player's preferences are similar for both classes of puzzles. Also note that the move preferences change over the lifetime of a puzzle; local minimize is heavily preferred by the end of puzzles but not by all players at the beginning. The move types preferences are very different from Rosetta's current best automated protocol, rebuild and refine, shown in **b**.

Table 1

Blind data set

| Puzzle ID | Foldit CA-RMSD | Rebuild and refine CA-RMSD | Native | Method | Length | Figure(s) |
|-----------|----------------|----------------------------|----------|--------|--------|------------------------------|
| 986875 | 1.4 | 4.5 | 2kpo | NMR | 66 | 3 a-c , S4 |
| 986698 | 1.8 | 3.7 | 2kky | NMR | 102 | 3 d-e |
| 986836 | 5.7 | 6.6 | 3epu | X-ray | 136 | 2 c , S6 d |
| 987088 | 3.5 | 4.3 | 2kpt | NMR | 116 | 2 a-b , S6 a-b |
| 987162 | 4.5 | 5.2 | 3lur | X-ray | 158 | S6c |
| 987076 | 3.3 | 3.5 | 2kpm | NMR | 81 | 2e, S5c |
| 986629 | 3.5 | 3.3 | 2kk1 | NMR | 135 | S5b |
| 987145 | 2.6 | 2.3 | none yet | X-ray | 105 | 2 d , S5 a |
| 986844 | 6.9 | 5.8 | 2ki0 | NMR | 36 | S10a |
| 986961 | 10.6 | 5.7 | 2knr | NMR | 118 | S10b |
| | | | | | | |

by the Rosetta rebuild and refine protocol, given the same starting model(s). Solutions considerably better with one method than the other are indicated in bold. The solved structures (which were released A listing of all the Foldit puzzles run in the blind data set. A CA-RMSD comparison to the native is given between the best scoring model produced by Foldit players and the best scoring model produced after each puzzle ended) are represented by their PDB codes. Results from these Foldit puzzles can be accessed on the Foldit website by using the corresponding Foldit puzzle ID at http://fold.it/portal/ structures on the PDB website (http://www.rcsb.org/pdb/search/search/search/stuts.do). 3lur was provided by the JCSG, as well as the remaining structure that has not yet been released to the PDB. Figures node/ID. 2kky, 2kpt, 2kpm, 2kk1 and 2knr were taken from the CASD-NMR experimentxi. 2kpo was provided by Nobuyashu and Rie Koga. 2ki0 and 3epu were found by searching for unreleased containing results for each puzzle are provided in the last column.