Vol 466|5 August 2010|doi:10.1038/nature09304

LETTERS

Predicting protein structures with a multiplayer online game

Seth Cooper[1], Firas Khatib[2], Adrien Treuille[1,3], Janos Barbero[1], Jeehyung Lee[3], Michael Beenen[1], Andrew Leaver-Fay[2]†, David Baker[2,4], Zoran Popović[1] & Foldit players

People exert large amounts of problem-solving effort playing computer games. Simple image- and text-recognition tasks have been successfully 'crowd-sourced' through games[1–3], but it is not clear if more complex scientific problems can be 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 non-scientists 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 methodology[4], while they compete and collaborate to optimize the computed energy. We show that top-ranked Foldit players excel at solving challenging structure refinement problems in which substantial backbone rearrangements are necessary to achieve the burial of hydrophobic residues. Players working collaboratively develop a rich assortment of new strategies and algorithms; unlike computational approaches, they explore not only the 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.

Although it has been known for over 40 years that the three-dimensional structures of proteins are determined by their amino acid sequences[5], 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—even small proteins have on the order of 1,000 degrees of freedom—and rugged due to unfavourable atom–atom repulsion that 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 side-chain conformations; gradient-based energy minimization; and energy-dependent acceptance or rejection of structure changes[6–8].

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 leader board of other players, and groups of players working together, competing in the same puzzle (Fig. 1, arrows 8 and 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 probably improve the structure (Fig. 1, arrows 1–5). Side chains are coloured by hydrophobicity and the backbone is coloured 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 β-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 and 7). They are also able to change the strength of the repulsion term to allow more freedom of movement. Available automatic moves—combinatorial side-chain 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 (Supplementary Fig. 1 and Supplementary Table 1): 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 (Supplementary Fig. 2)—in fact, few top-ranked players are professionally involved in biochemistry (Supplementary Fig. 3).

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 ten blind structures, including comparisons between the best-scoring Foldit predictions and the best-scoring Rosetta predictions using the rebuild and refine protocol[7], is given in Table 1. We found that Foldit

[1]Department of Computer Science and Engineering, University of Washington, Box 352350, Seattle, Washington 98195, USA. [2]Department of Biochemistry, University of Washington, Box 357350, Seattle, Washington 98195, USA. [3]School of Computer Science, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA. [4]Howard Hughes Medical Institute, University of Washington, Box 357370, Seattle, Washington 98195, USA. †Present address: Department of Biochemistry, University of North Carolina, CB 7260, Chapel Hill, North Carolina 27599, USA.

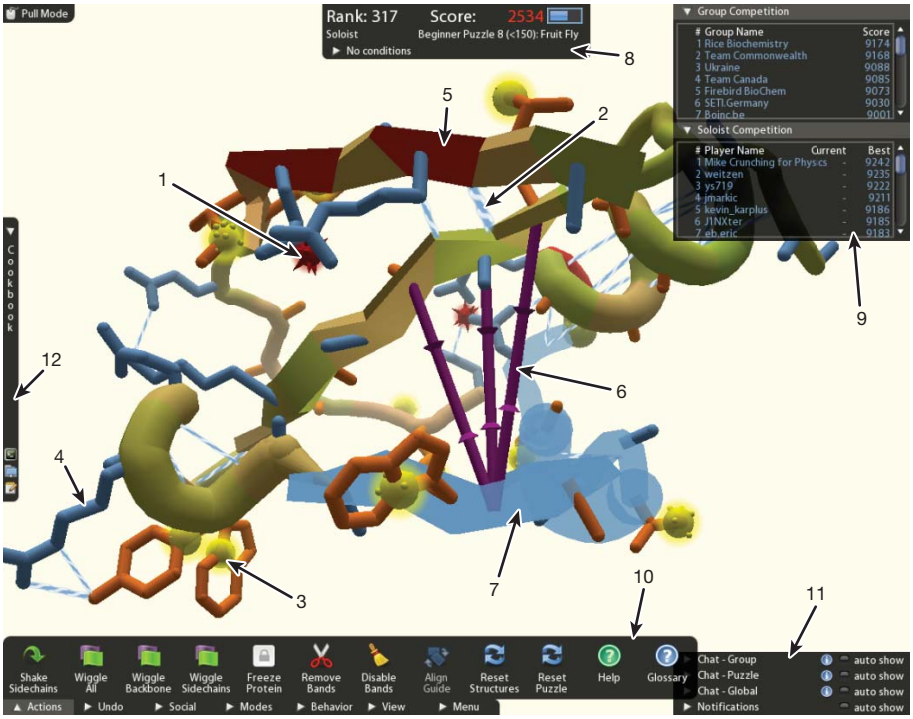

Figure 1 | Foldit screenshot illustrating tools and visualizations. The visualizations include a clash representing atoms that are too close (arrow 1); a hydrogen bond (arrow 2); a hydrophobic side chain with a yellow blob because it is exposed (arrow 3); a hydrophilic side chain (arrow 4); and a segment of the backbone that is red due to high residue energy (arrow 5). The players can make modifications including 'rubber bands' (arrow 6), which add constraints to guide automated tools, and freezing (arrow 7), which prevents degrees of freedom from changing. The user interface includes information about the player's current status, including score (arrow 8); a leader board (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).

players were particularly adept at solving puzzles requiring substantial backbone remodelling to bury exposed hydrophobic residues into the protein core (Fig. 2). When a hydrophobic residue points outwards into solvent, and no corresponding hole within the core is evident, stochastic Monte Carlo trajectories are unlikely to sample the coordinated backbone and side-chain shifts needed to bury the residue properly 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 remodelled loop (Fig. 2a, b), a rotated helix (Fig. 2c), two remodelled loops (Fig. 2d), and a helix rotation and remodelled loop (Fig. 2e).

Players were also able to restructure β-sheets to improve hydrophobic burial and hydrogen bond quality. Automated methods have difficulty performing major protein restructuring operations to change β-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 1.1 Å of the native structure, with the top-scoring Foldit prediction being 1.4 Å away. A superposition between the starting Foldit puzzle, the top-scoring Foldit solution, and model 1 of the native NMR structure 2kpo (Protein Data Bank) 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 unravelling the structure (Fig. 3c, bottom), with a corresponding unfavourable increase in energy (Fig. 3c, top). Players persisted with this reconfiguration despite the energy increase because they correctly recognized that the swap could ultimately lead to lower energies. In contrast, although the Rosetta

Table 1 | Blind data set

Puzzle ID	Foldit Cα r.m.s.d.	Rebuild and refine Cα r.m.s.d.	Native	Method	Number of residues	Figure(s)
986875	**1.4**	4.5	2kpo	NMR	99	3a–c, Supplementary 4
986698	**1.8**	3.7	2kky	NMR	102	3d, e
986836	**5.7**	6.6	3epu	X-ray	136	2c, Supplementary 6d
987088	**3.5**	4.3	2kpt	NMR	116	2a, b, Supplementary 6a, b
987162	**4.5**	5.2	3lur	X-ray	158	Supplementary 6c
987076	3.3	3.5	2kpm	NMR	81	2e, Supplementary 5c
986629	3.5	3.3	2kk1	NMR	135	Supplementary 5b
987145	2.6	2.3	3nuf	X-ray	105	2d, Supplementary 5a
986844	6.9	**5.8**	2ki0	NMR	36	Supplementary 10a
986961	10.6	**5.7**	2knr	NMR	118	Supplementary 10b

A listing of all the Foldit puzzles run in the blind data set. A Cα r.m.s.d. comparison to the native structure is given between the best-scoring model produced by Foldit players and the best-scoring model produced 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 after each puzzle ended) are represented by their Protein Data Bank (PDB) codes. Results from these Foldit puzzles can be accessed on the Foldit website by replacing ID with the corresponding Foldit puzzle ID in http://fold.it/portal/node/ID. 2kky, 2kpt, 2kpm, 2kk1 and 2knr were taken from the CASD-NMR experiment[10]. 2kpo was provided by N. Koga and R. Koga. 2ki0 and 3epu were found by searching for unreleased structures on the PDB website (http://www.rcsb.org/pdb/search/searchStatus.do). 3lur and 3nuf were provided by the Joint Center for Structural Genomics (JCSG). The location of figures containing results for each puzzle are provided in the last column.

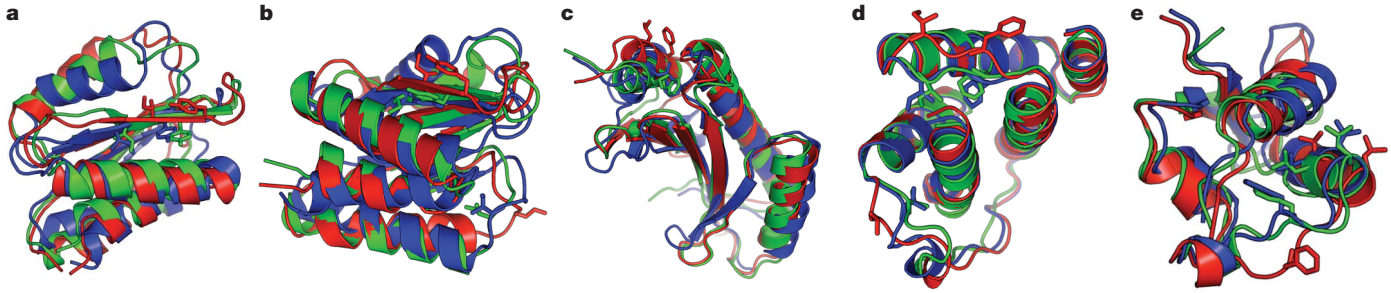

Figure 2 | Structure prediction problems solved by Foldit players. Examples of blind structure prediction problems in which players were successfully able to improve structures. Native structures are shown in blue, starting puzzles in red, and top-scoring Foldit predictions in green. **a**, The red starting puzzle had a register shift and the top-scoring green Foldit prediction correctly flips and slides the β-strand. **b**, On the same structure as above, Foldit players correctly buried an exposed isoleucine residue in the loop on the bottom right by remodelling 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 along with a remodelled loop that correctly buries an exposed phenylalanine. Images were produced using PyMOL software[11].

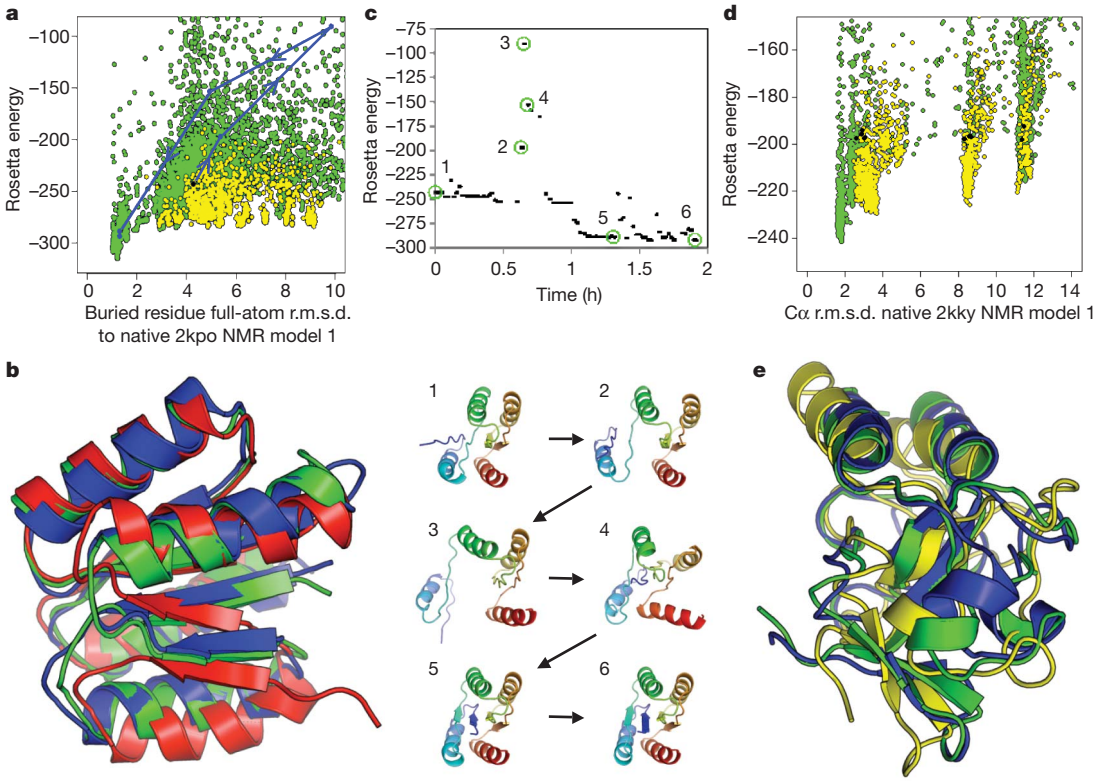

Figure 3 | Puzzles in which human predictors significantly outperformed the Rosetta rebuild and refine protocol. **a**–**c**, Puzzle 986875. **d**, **e**, Puzzle 986698. **a**, Comparison of Foldit player solutions (green) to the low-energy structures sampled in Rosetta rebuild and refine trajectories (yellow) for blind Foldit puzzle 986875 based on the recently determined structure of 2kpo. The *x* axis is the all-atom r.m.s.d. to 2kpo, and the *y* axis is the Rosetta energy. The starting Foldit puzzle was 4.3 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 structure, whereas the automated Rosetta protocol did not sample below 2 Å. 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 neighbour; 8% of all Foldit players were able to swap these strands correctly (Supplementary Table 2). **c**, A score trajectory with selected structures for the top-scoring player in puzzle 986875 over a 2-h window, showing how the player explores through high-energy conformations to reach the native state. The *y* axis shows the Rosetta energy and the *x* axis 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 (1) is near the starting puzzle structure, shown as the black dot in panel **a**. The following structures (2–6) are shown as blue dots in panel **a**. In structures 2–4, the player must explore higher energies to move the strand into place, shown by the blue lines. In structures 5 and 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 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 native structure. **e**, The native structure is shown in blue with the top-scoring Foldit prediction shown in green. The top-scoring Rosetta rebuild and refine prediction given the same ten starting models (shown in yellow) was unable to sample as close to the native structure as the Foldit players.

rebuild and refine protocol did sample some partially swapped conformations (Fig. 3a, leftmost yellow point), these were not retained in subsequent generations owing to their relatively high energies, resulting in the top-scoring Rosetta prediction being further from the native than the starting structure (Supplementary Fig. 5).

Human players are also able to distinguish which starting point will be most useful to them. Figure 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 ten starting models, the Rosetta rebuild and refine protocol was unable to get as close to the native structure as the top-scoring Foldit predictions.

Foldit players performed similarly to the Rosetta rebuild and refine protocol for three of the ten blind puzzles (Supplementary Fig. 6). They outperformed Rosetta on five of the puzzles (Fig. 3 and Supplementary Figs 5 and 7), including the two above cases where players performed significantly better. A larger set of successful solutions for similar, although non-blind, puzzles are described in Supplementary Figs 8–10. For two of the ten blind puzzles, the top-scoring Rosetta rebuild and refine prediction was numerically better than the Foldit solution (Table 1) but still basically incorrect (root mean squared deviation (r.m.s.d.) to native structure >5.7 Å) (Supplementary Fig. 11).

Despite the promising results described above, there 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 (Supplementary Fig. 12), players can have trouble reaching it from so far away (Supplementary Fig. 11a). 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 (Supplementary Fig. 14a), and within groups the exchange of solutions can help other members catch up to the leaders (Supplementary Fig. 14b).

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 side chains; 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 labour: 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 Supplementary Text, 'Player Testimonials' section and Supplementary Table 3 for more information). More in-depth analysis of player strategies should provide further insight into the basis for human achievement with Foldit and could lead to improved automated algorithms for protein structure prediction.

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

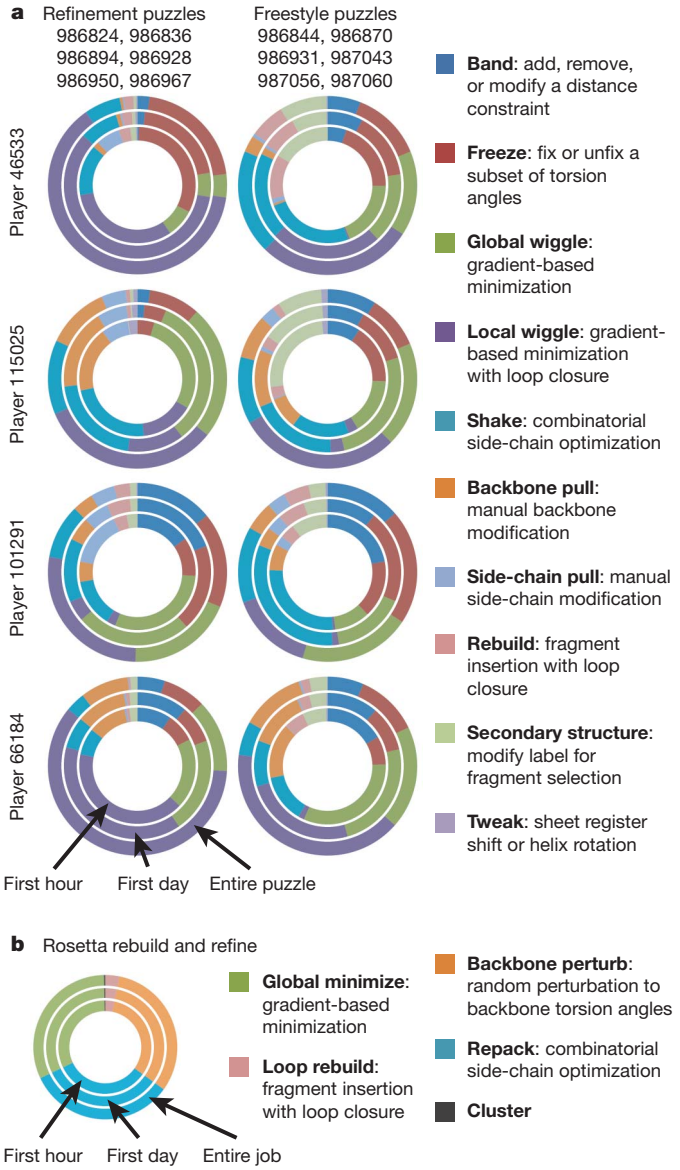

Figure 4 | Player move preferences. **a**, Different Foldit players take different approaches to solving the same problem. Each circle represents the move type frequencies used in the top-scoring solution produced by each player in different time frames: the inner circle denotes the first hour; the middle circle denotes the first day; and the outer circle denotes the puzzle's entire duration. Each colour 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 wiggle is heavily preferred by the end of puzzles but not by all players at the beginning. The move type preferences are very different from Rosetta's current best automated protocol, rebuild and refine, shown in **b**.

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 (Supplementary Fig. 4) revealed that although 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 shapes[9]. We expect generally that future scientific discovery games will also benefit from varied motivation sets.

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 three-dimensional structural problem solving can be used. Our results indicate that scientific advancement is possible if even a small fraction of the energy that goes into playing computer games can be channelled into scientific discovery.

Received 22 January; accepted 30 June 2010.

1. von Ahn, L. & Dabbish, L. Labeling images with a computer game. in *CHI '04: Proc. 2004 Conf. Human Factors Comput. Syst.*, 319–326 (ACM, 2004).
2. von Ahn, L., Liu, R. & Blum, M. Peekaboom: a game for locating objects in images. in *CHI '06: Proc. SIGCHI Conf. Human Factors Comp. Syst.*, 55–64 (ACM, 2006).
3. Westphal, A. J. *et al.* Non-destructive search for interstellar dust using synchrotron microprobes. In *X-ray Optics Microanalysis: Proc. 20th Int. Congr.* Vol. 1221, 131–138 (2010).
4. Rohl, C., Strauss, C., Misura, K. & Baker, D. Protein structure prediction using Rosetta. *Methods Enzymol.* **383,** 66–93 (2004).
5. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181,** 223–230 (1973).
6. Das, R. & Baker, D. Macromolecular modeling with Rosetta. *Annu. Rev. Biochem.* **77,** 363–382 (2008).
7. Qian, B. *et al.* High-resolution structure prediction and the crystallographic phase problem. *Nature* **450,** 259–264 (2007).
8. Bradley, P., Misura, K. M. S. & Baker, D. Toward high-resolution *de novo* structure prediction for small proteins. *Science* **309,** 1868–1871 (2005).
9. Yee, N. Motivations of play in online games. *J. CyberPsychol. Behav.* **9,** 772–775 (2007).
10. Rosato, A. *et al.* CASD-NMR: critical assessment of automated structure determination by NMR. *Nature Methods* **6,** 625–626 (2009).
11. DeLano, W. L. *The PyMOL Molecular Graphics System* (DeLano Scientific, 2002).

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

Acknowledgements We thank D. Salesin, K. Tuite, J. Snyder, D. Suskin, P. Krähenbühl, A. C. Snyder, H. Lü, L. S. Tan, A. Chia, M. Yao, E. Butler, C. Carrico, P. Bradley, I. Davis, D. Kim, R. Das, W. Sheffler, J. Thompson, O. Lange, R. Vernon, B. Correia, D. Anderson, Y. Zhao, S. Herin and B. Bethurum for their help. We would like to thank N. Koga, R. Koga and A. Deacon and the JCSG for providing us with protein structures before their public release. We would also like to acknowledge all of the Foldit players who have made this work possible. Usernames of players whose solutions were used in figures can be found in Supplementary Table 4. 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.

Author Contributions All named authors contributed extensively to development and analysis for the work presented in this paper. Foldit players (more than 57,000) contributed extensively through their feedback and gameplay, which generated the data for this paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Z.P. (zoran@cs.washington.edu) or D.B. (dabaker@u.washington.edu).