HHS Public Access

Author manuscript

Biochem Mol Biol Educ. Author manuscript; available in PMC 2020 March 01.

Published in final edited form as:

Biochem Mol Biol Educ. 2019 March ; 47(2): 133–139. doi:10.1002/bmb.21208.

Creating Custom Foldit Puzzles for Teaching Biochemistry

Lorna Dsilva[#1], **Shubhi Mittal**[#1], **Brian Koepnick**[2], **Jeff Flatten**[3], **Seth Cooper**[*,1], and **Scott Horowitz**[*,4]

[1]College of Computer and Information Science, Northeastern University, Boston, MA 02115, USA

[2]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA

[3]Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, WA 98195, USA

[4]Department of Chemistry & Biochemistry and the Knoebel Institute for Healthy Aging, University of Denver, Denver, CO 80208, USA

[#] These authors contributed equally to this work.

Abstract

The computer game Foldit is currently widely used as a biology and biochemistry teaching aid. Herein, we introduce a new feature of Foldit called “custom contests” that allows educators to create puzzles that fit their curriculum. The effectiveness of the custom contests is demonstrated by the use of five distinct custom contests in an upper-level biochemistry class. The new custom contest feature can be implemented in classes ranging from middle school to graduate school to enable educators to best complement their current curriculum.

Keywords

Biochemistry; Puzzles; Web-Based Learning

INTRODUCTION

Foldit is a citizen science computer game [1] in which players tackle difficult biochemistry puzzles that are directly linked to laboratory advances. The developers of Foldit determine these puzzles, which are structured for maximal scientific impact. Gaming is now commonly used as a teaching tool across many educational fields, particularly for its ability to gain student interest in a subject [2]. Accordingly, educators worldwide have begun using Foldit in the classroom, as Foldit uniquely provides students with detailed hands-on experience in protein structure and function [3–5] in a game environment. This adoption by the education community comes despite Foldit not previously providing puzzles that integrate easily into general curricula, nor the ability for educators to create puzzles that fit their specific curricula. Here we introduce a new feature of Foldit called “custom contests”, in which individuals can create their own Foldit puzzle and provide it to a specific group of people,

[*]Corresponding Author scott.horowitz@du.edu, scooper@ccs.neu.edu.

such as a class. This development builds on previous features of Foldit including “template contests”, which allow users to administer their own puzzles, selecting from a small list of predefined simple examples, and Foldit Standalone [6], an application version of Foldit that allows users to load in their own structures but not run them as puzzles. Custom contests can be used to create folding or design puzzles containing a wide variety of molecules, including protein, RNA, DNA, and ligands including carbohydrates, enzyme substrates and cofactors. We have successfully used custom contests in an upper-level biochemistry class utilizing many of these different biological molecules.

CREATING AND IMPLEMENTING CUSTOM CONTESTS

For detailed step-by-step instructions on creating and implementing custom contests, please see the documentation on the Foldit website at https://fold.it/portal/customcontests. For detailed descriptions of the possible options for puzzle customization, see the supplementary information. In brief, custom contests are created through a web interface at the Foldit website (Figure 1). Puzzle files are uploaded either in .pdb format for simple puzzles containing only macromolecules, or .zip format for puzzles that require multiple files.

As an example, the chaperone puzzle (discussed below) has two possible starting structures, with the protein in either its folded or unfolded state. Each of these states is defined by a separate pair of setup and PDB files, coordinated through the numbering in the file name. Both of the puzzle setup files have the same information content, but a separate puzzle setup file is required for each starting structure. The “backbone_locked” term prevents the backbone angles of the indicated residues from changing. Locking the backbone of the chaperone increases the speed of the wiggle and shake functions within Foldit, making the puzzle more responsive to players and students. As a rule of thumb, it is best to keep the number of unlocked residues in Foldit puzzles to ~200 residues or less. Note that in the PDB files, the two proteins are specified to be separate chains, but their residue numbers are unique and continuous between the two proteins. The PDB files and setup files are compressed into a .zip archive and uploaded to the Foldit server to create the contest. Further examples of multiple-file puzzles are those that contain ligands or protein design choices.

Testing and creating puzzles for the classroom was facilitated by the Foldit Standalone program [6]. Foldit Standalone is version of the Foldit game that gives the puzzle designer control over setting up their puzzle, which aids in puzzle design and testing. Instructions for downloading and using Foldit Standalone can be found at http://fold.it/standalone. In brief, PDB files are loaded into Foldit Standalone along with puzzle setup files, Rosetta parameter files, and other needed files, producing the intended puzzle in a local environment. This step increases the speed of puzzle design, as it does not require the full upload and contest creation in order to test the puzzle. The puzzle designer can quickly see whether the puzzle causes errors in Foldit, whether the Foldit puzzle is responsive, the playability of the puzzle, and the effects of locking residues or adding protein design elements.

One common issue in Foldit puzzle creation is that the PDB file can only contain Rosetta-recognized residues. This requires editing PDB files downloaded from the PDB servers to remove excess ligands from crystallization. Necessary ligands can be accommodated using

Rosetta parameter files, which requires a degree of expertise in Rosetta to currently use (see https://www.rosettacommons.org/manuals/archive/rosetta3.5_user_guide/df/de9/preparing_ligands.html for details). Future developments of the Foldit Standalone program will make it possible to create Rosetta parameter files using a GUI tool. For educators new to Foldit, it is best to start with simpler puzzles that only contain protein or nucleic acid constituents.

After creating the custom contest, the contest administrator can share a link to the contest from the custom contest's page (Figure 2), which is then played in the Foldit client. During the contest, the administrator can see the progress of each student based on their Foldit score. The Foldit score is based on the Rosetta force field [7,8], and is therefore an accurate depiction of the feasibility of the student's biochemical choices. After the contest completes, the administrator can download the top structure from each student.

FOLDIT GAMEPLAY INTRODUCTION

Foldit is intended to be learnable and playable without significant external written instructions, as there are extensive tutorials that players work through to introduce the major moves and tools. Furthermore, a wiki created and maintained by Foldit players (http://foldit.wikia.com) describes many more advanced concepts. Briefly, to begin playing, the player first downloads and installs the free Foldit client program from https://fold.it, and makes an account. Upon logging into the client, the player is first prompted to complete the tutorial puzzles, which teach the player both basic move and tool types, as well as providing an introduction into the multiple varieties of puzzles possible within Foldit. The player has the option to take part in scientific puzzles posted by the Foldit developers, as well as to create and join contests. It is recommended that before designing puzzles, educators should play available scientific puzzles and investigate the Foldit wiki to become familiar with gameplay.

Foldit gameplay strategies vary from player to player, and can involve human interaction as well as scripting. In general, early gameplay strategies will utilize only interactive moves. However, recipes and scripting (http://foldit.wikia.com/wiki/101_-_Cookbook) can provide a more automated process that may or may not be in line with the educator's goals. Educators can disable scripts and recipes using the "allowed_macros" key, which forces students to solve the puzzle without the aid of scripts.

COURSE DESIGN AND IMPLEMENTATION WITH CUSTOM CONTESTS

An advantage of custom contests is the flexibility to create puzzles specifically designed to fit a curriculum. These puzzles could range considerably. For example:

1. For a non-expert course (at middle school, high school, or college level), educators could create simple puzzles focused on disease-causing proteins that will be easily relatable to the students. As an example for this type of puzzle, an instructor teaching the importance of protein folding using sickle cell anemia could create a puzzle with two copies of mutated hemoglobin, with the goal to assemble them into a dimer representing the polymerized form. In this example,

the instructor could provide the proteins in an orientation and conformation to facilitate quick dimerization, or could require the students to find the diseases-relevant conformation first.

2. For a biochemistry course, educators could create puzzles that fit standard biochemistry curricula, such as substrate binding puzzles of enzymes within glycolysis or the Krebs cycle. These puzzles could help students to visualize and understand the three dimensional qualities of the active site and how the reaction steps proceed.

3. For specialized topics courses, educators can design puzzles that specifically augment the subject of their course.

In winter 2018, we implemented an example of the third type of course at the University of Denver. Named “Topics: Debates in Biochemistry”, the course was a discussion-based seminar accompanied by Foldit custom contests. The course was open to fourth year undergraduate biochemistry majors and graduate students in the chemistry and biophysics programs. Foldit setup files and top solutions for each custom contest can be found at http://fold.it/dist/external/customcontests/Debates-in-Biochemistry-files.zip, and brief descriptions of the choices made in the setup of each puzzle, and a complete list of available puzzle options as well as the course syllabus scan be found in the Supporting Information.

Students were first required to complete the Foldit tutorials in the first two weeks of class, while the class discussed papers on the topics of the Rosetta force field and Foldit. Custom contests were then created and implemented to help teach the following topics:

- The role of phosphate in biology (phosphate).
- RNA structure and folding (RNA TAR).
- Protein folding and chaperone binding (chaperone).
- Enzyme catalysis (enzyme).
- Amyloid oligomerization (ABeta).

Topics being discussed in class were mirrored by concurrent Foldit puzzles on the same topic. For example, papers on the role of chaperones in protein folding that were discussed in class were accompanied by a Foldit puzzle featuring the same chaperone being discussed in class. The ordering of the topics was chosen to provide simpler contests earlier in the course, and more complicated contests later in course. Each contest was open for one week starting the Friday before the beginning of discussion, and ending at 5 PM Friday of the following week, the same day as the end of discussion.

In general, Foldit puzzles were not discussed during class. Puzzles were introduced to students via an announcement on the course webpage. An example announcement for the chaperone puzzle is shown below, and all announcements used to introduce puzzles can be found in the Supporting Information.

> Now for a more complicated Foldit puzzle. This contest goes with the two chaperone papers you’ll be reading this week that try to figure out how a chaperone

can help proteins fold. In the puzzle, you'll be getting the mini-GroEl chaperone discussed in the first paper, along with the protein that they studied folding (barnase). The chaperone will be mostly fixed in place, so you won't be able to move it much at all, but you'll have complete control over the folding protein. One note: the reset puzzle feature will be different in this puzzle, as when you use it, barnase will toggle between native and unfolded states. It will be up to you to determine what level of folding provides the optimal balance between folding and interaction with the chaperone.

The separate contests made use of specific customization options in Foldit. For example, in the enzyme catalysis puzzle, students were asked to mutate and optimize an enzyme active site to better bind a transition state analog. This puzzle differed considerably from the folding puzzles. As a counter-example, in the chaperone puzzle, the students were provided with the option of starting from a natively folded chaperone client or a fully unfolded client, and were required to optimize the level of protein folding and binding to the chaperone (Figure 3).

Three times over the course of the quarter, students were asked if they encountered any technical difficulties with Foldit. Students reported no technical difficulties with either the Foldit tutorial puzzles or the custom contests.

The educator can choose whether to grade based on the Rosetta energy function (Foldit score), or whether to ask students to try to achieve a specific structural or biochemical goal. In both cases, the educator is able to access the solutions using links on the puzzle homepage, which can be navigated to by clicking on the puzzle name on the contest homepage. The educator can then download the PDB files for these solutions and analyze them for the achievement of specific goals or for checking the top solutions. The top solutions are automatically linked in the puzzle homepage, but the shared solutions link will only appear once students have uploaded a solution to share with the puzzle administrator. The students save and share their solution from within the game by going to Menu>Open/Share Solutions>Save Current Solution, and then selecting their solution to upload and click Upload For Scientists. It can take up to a few hours after uploading for the solutions to appear in the puzzle homepage. In the class described here, top solutions were used for grading and not shared solutions.

Foldit grades were calculated as a combination of completion of the tutorials, participation in the Foldit puzzles, and Foldit score. For the contests, participation in the contest automatically netted the player 50% of the total possible grade for that contest. The remaining 50% was calculated using two different methods, and then combined to achieve a balanced grading profile. In the first method, the students' scores were normalized on a linear scale from 0 to 1 based on the student's Foldit score, with the lowest score earning a 0 and the highest earning a 1. In the second method, each student's score was taken as a percentage of the top student's score. These two methods were then combined together in a ratio of 1/3rd normalized score and 2/3rds percentage score to yield the competitive 50% of the Foldit grade. This competitive component was then added to the 50% participation component to determine the grade the puzzle.

RESULTS OF CLASS

In this first iteration, the class was taken by eight chemistry graduate students, five of whom had an interest in biochemistry, and one biophysics student. Eight of the nine students turned in anonymous course evaluations, which overall indicated that the students were of the opinion that they had learned a large amount of material and that the course was of excellent overall quality. Modifications suggested by students included integrating Foldit puzzles into class discussions. More detailed follow-up with students on how to better integrate Foldit suggested that a portion of class-time during each topic should be devoted to discussion of the Foldit puzzle, specifically covering the different students' solutions, and what constitutes a good puzzle solution for each specific case based on both energetics and the biology of the molecule. It was also suggested that students sometimes spent more time than felt necessary performing very small tweaks to their structures to try and pick up extra points, but that these small tweaks did not increase their learning.

The students' Foldit scores varied by puzzle (Table 1), both in the absolute scores and the range of scores over the class. The highest degree of variation in student was observed in the ABeta puzzle and the chaperone puzzle, whereas student scores were most tightly grouped in the two RNA puzzles (RNA TAR and phosphate). This tighter grouping likely reflects the more limited number of possible outcomes in the RNA puzzles.

Analysis of the student's top puzzle solutions reflected that the top scoring solutions often were biologically feasible structures. For example, the top scoring chaperone puzzle solution (Figure 4A) shows the chaperone and folding protein binding with the folding protein in a state that could mimic a late stage folding intermediate. This arrangement would yield the most stable structure of the complex by maximizing both intra- and inter-molecular interactions. As another example, the top scoring RNA TAR puzzle solution (Figure 4B) displays a feasible tertiary structure for an RNA molecule, made up of common RNA secondary structure elements. Lower student scores were usually associated with less-feasible structures. The lowest-scoring structures for the same puzzles, shown in Figures 4C and 4D, respectively, show less feasible structures. In the case of the chaperone puzzle (Figure 4C), the folding protein is less compact, displays less secondary structure, and has greater exposed hydrophobic residues. In the case of the RNA TAR puzzle (Figure 4D), the bottom-scoring solution does not display regular RNA secondary structure. These examples illustrate how Foldit's Rosetta-based score distinguishes between feasible and non-feasible solutions.

Although students were not specifically tested for whether the puzzles increased their learning, the instructor observed over the course of the quarter that the discussion complexity about biological macromolecules improved, with students asking more in-depth questions and providing more nuanced explanations for the physical basis of how each molecule operated. Due to the lack of class time devoted to the Foldit puzzles, students shared Foldit strategies with each other outside of class and helped each other to discover effective Foldit strategies. However, students expressed some frustration that they were not given more guidance in class.

DISCUSSION

The power of Foldit for education is now recognized by many in the biochemistry community. This article details how educators can now customize Foldit puzzles to fit their curricula. This new feature makes it possible for biology and biochemistry educators worldwide to innovate in their classrooms using Foldit as a tool. We anticipate that over time, many new and impactful course and activity designs will be made using Foldit custom contests. In its first iteration, the use of custom contests was implemented in an upper-level biochemistry class, “Topics in Biochemistry”.

Follow-up questions with students and the instructor’s observations will lead to several changes for the next iteration of the course. Foremost, class time each week will be devoted to the Foldit puzzle to discuss different students’ solutions and what constitutes a biologically meaningful solution. The first of these class time sessions will be devoted to providing the students with guidelines on effective Foldit strategies. Second, the grading will be altered to scale back the degree of competitiveness in the Foldit grading to reduce the incentive to spend large amounts of time performing small tweaks to gain minimal improvement. Third, the “allowed_macros” puzzle setup key will also be used to make the focus primarily on interacting with the biomolecules by hand. Fourth, the Foldit scores for each puzzle in this first iteration will be used to benchmark future grading. Finally, the syllabus will be updated to map learning activities to the course learning outcomes with the following addition:

COURSE LEARNING OUTCOMES

Present, discuss and analyze the science in each paper in the course.

Compare the points of view in the classroom debates.

Design and write a novel research grant independently.

The course reading, participation, presentations, and Foldit puzzles all contribute to course learning outcomes 1 and 2. The Foldit puzzles are intended to provide students an avenue for understanding the biochemistry that will contribute to their discussion.

These course learning outcomes are intended to fulfill the core goals of the Chemistry & Biochemistry graduate program to enable independent learning and advance fundamental understanding and communication in biochemistry.

In follow-up studies on how to improve the use of Foldit in class, the instructor will also add additional quantitative testing to evaluate the learning of the students due to Foldit.

In addition to its education potential, custom contests could be used by scientists to run Foldit puzzles on their own topics of interest independent of the Foldit developers, similar to the approach used by Open-Phylo [9].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors would like to thank the University of Denver students in Topics: Debates in Biochemistry in winter quarter of 2018 for their enthusiastic participation in the class. The authors would also like to thank the Foldit players who provided feedback on preliminary puzzles and Ronny George Mathew and Samicheen Khariwal for their work on the Foldit website. This work was supported by National Institutes of Health grants R00 GM120388 and 1UH2CA203780. This material is based upon work supported by the National Science Foundation under Grant No. 1629879.

References

[1]. Cooper S, Treuille A, Khatib F, Barbero J et al. (2010)Predicting protein structures with a multiplayer online game. Nature. 466, 756–760. [PubMed: 20686574]

[2]. McClarty, K. L. et al. in A Literature Review of Gaming in Education. Research Report (Pearson, 2012).

[3]. Stockman BJ, Asheld JS, Burburan PJ, Galesic A, Sikorski KF et al. (2014)Design and Characterization of a Zn2+-Binding Four-Helix Bundle Protein in the Biophysical Chemistry Laboratory. Journal of chemical education. 91, 451–454.

[4]. Farley PC. (2013)Using the Computer Game “FoldIt” to Entice Students to Explore External Representations of Protein Structure in a Biochemistry Course for Nonmajors. Biochemistry and Molecular Biology Education. 41, 56–57. [PubMed: 23382128]

[5]. Franco J. (2012)Online gaming for understanding folding, interactions, and structure. Journal of Chemical Education. 89, 1543.

[6]. Kleffner R, Flatten J, Leaver-Fay A, Baker D, Siegel JB et al. (2017)Foldit Standalone: a video game-derived protein structure manipulation interface using Rosetta. Bioinformatics. 33, 2765–2767. [PubMed: 28481970]

[7]. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J et al. (2011)ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. Meth. Enzymol. 487, 545–574. [PubMed: 21187238]

[8]. Alford RF, Leaver-Fay A, Jeliazkov JR, O’Meara MJ, DiMaio FP et al. (2017)The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. J. Chem. Theory Comput. 13, 3031–3048. [PubMed: 28430426]

[9]. Kwak D, Kam A, Becerra D, Zhou Q, Hops A et al. (2013)Open-Phylo: a customizable crowd-computing platform for multiple sequence alignment. Genome Biol. 14, r116. [PubMed: 24148814]

[10]. Zahn R, Buckle AM, Perrett S, Johnson CM, Corrales FJ et al. (1996)Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. Proc. Natl. Acad. Sci. U. S. A. 93, 15024–15029. [PubMed: 8986757]

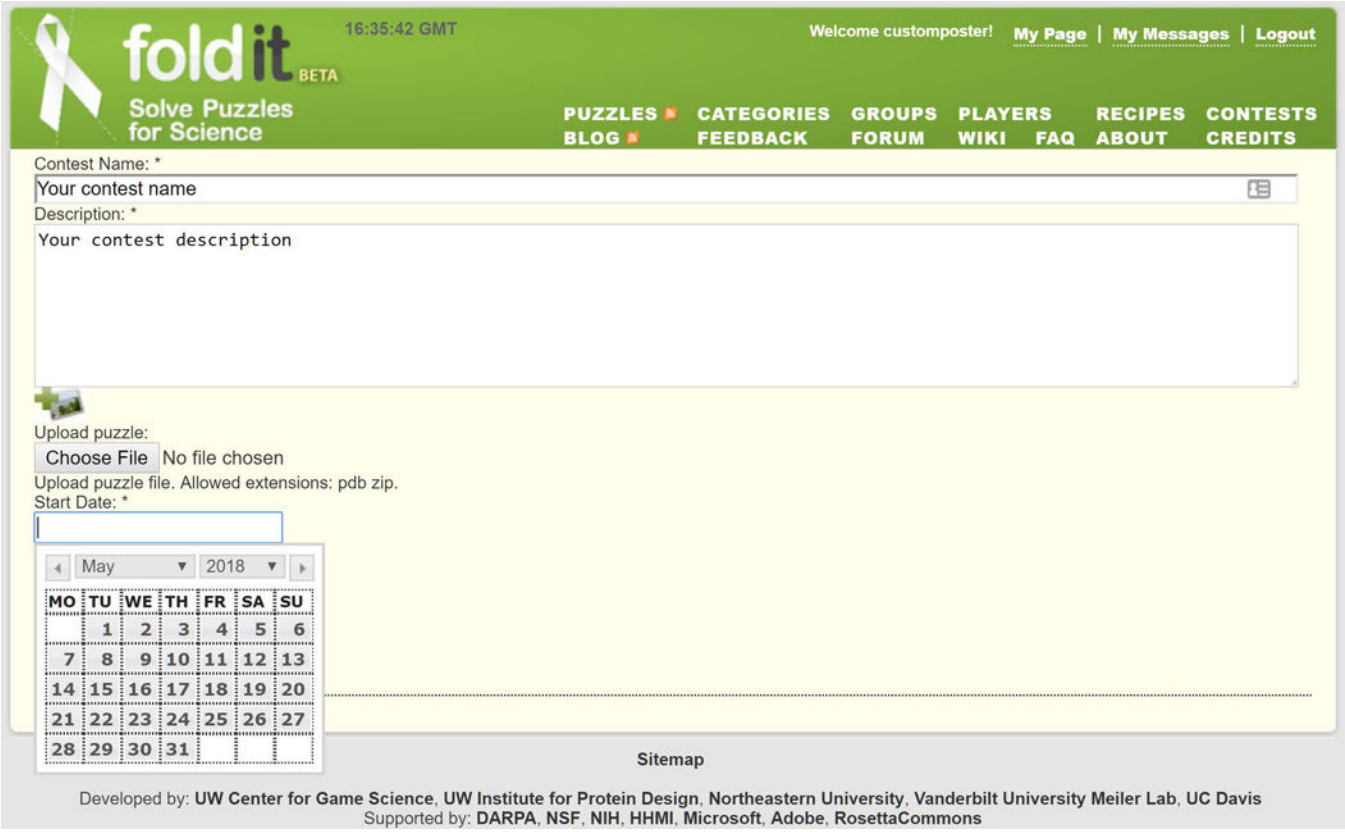

Figure 1.

Custom contest creation page. Contest name and description are required, as are source puzzle files (.pdb or .zip) and contest dates.

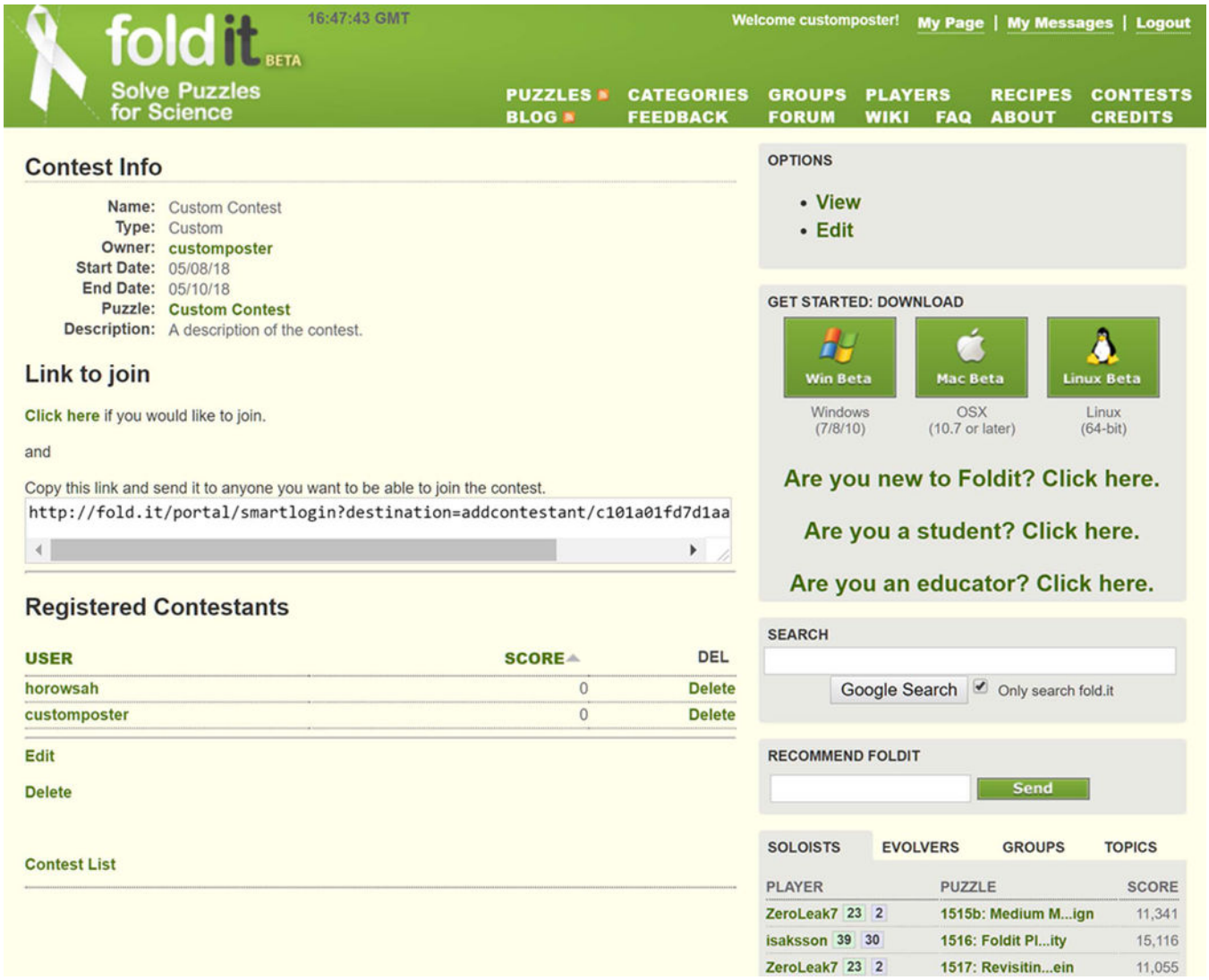

Figure 2.

Contest homepage, including link to join the contest. As contestants join, their user names and top Foldit scores will appear under the heading “Registered Contestants”. The edit link can be used to change contest dates, puzzle name, and puzzle description.

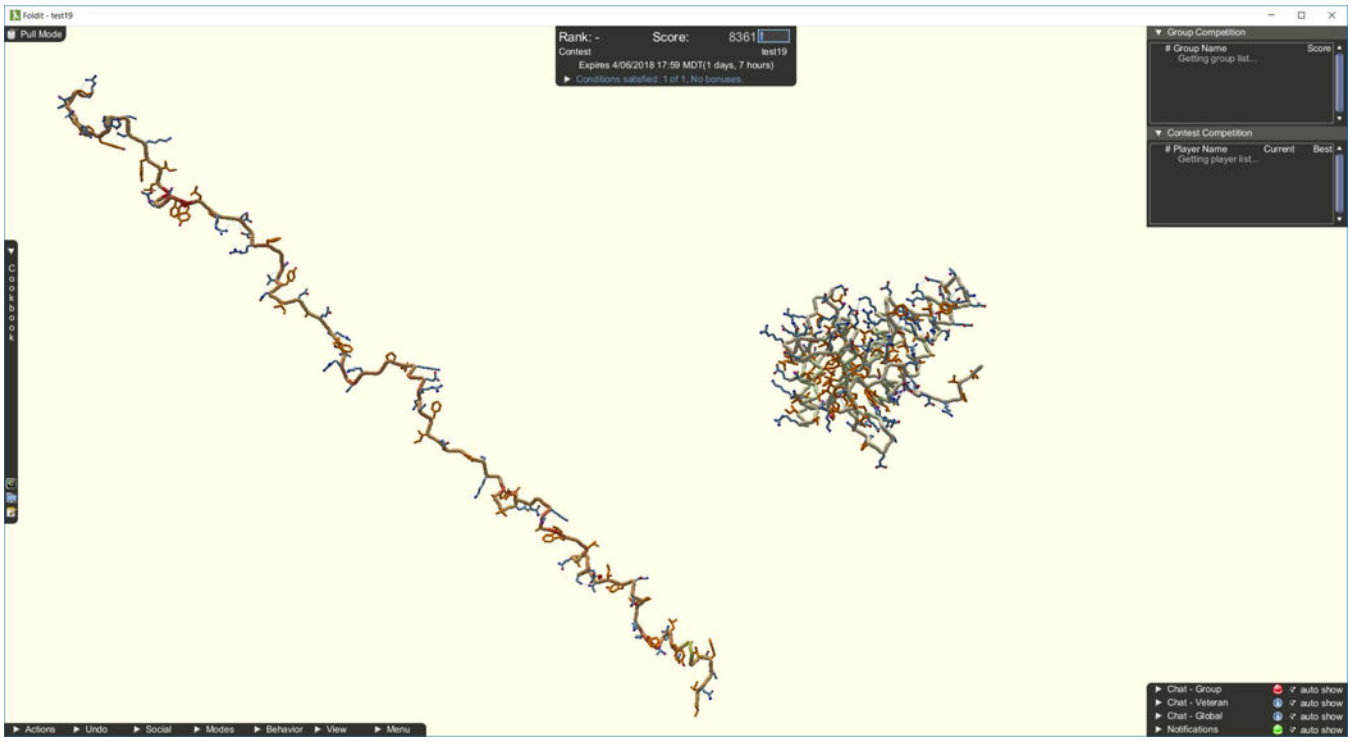

Figure 3.

Example custom contest puzzle used in class. In this puzzle, the mini-GroEL chaperone [10] (right) is placed in proximity to a protein client (left). By resetting the puzzle the students could toggle between the fully unfolded state and native state of the client as starting points for the puzzle.

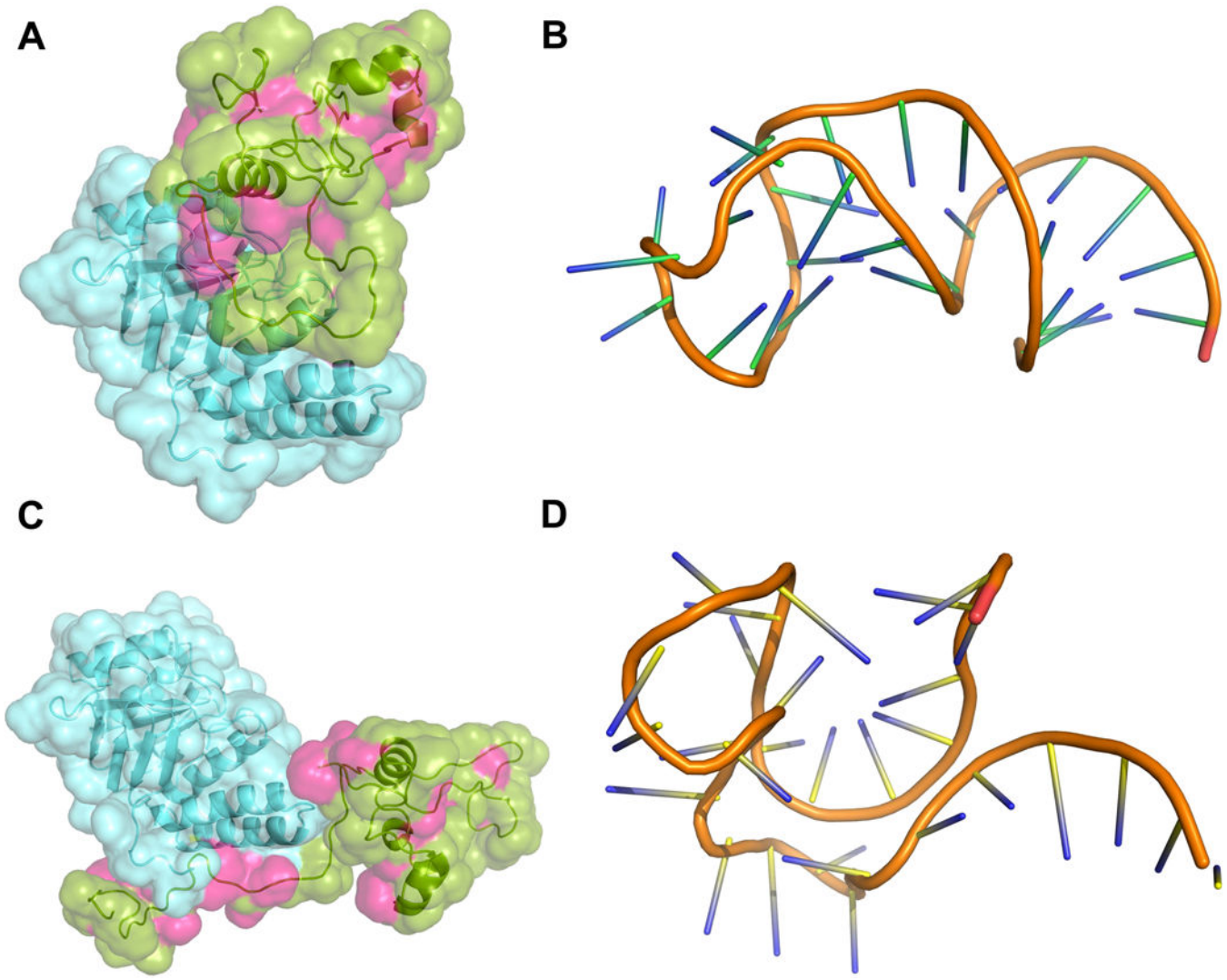

Figure 4.

Example top (A, B) and bottom (C, D) scoring solutions from class. (A & C) Chaperone puzzle top and bottom scoring solutions, respectively. The chaperone is rendered in blue, and the folding protein rendered in magenta for hydrophobic residues and green for non-hydrophobic residues. (B & D) RNA TAR puzzle top and bottom scoring solutions, respectively.

Table 1:

Foldit scores for puzzles used in class.

PUZZLE	HIGH SCORE	LOW SCORE	MEDIAN SCORE	MEAN SCORE ± SD
PHOSPHATE	8214	8031	8105	8104 ± 63
RNA TAR	8130	7500	8062	7985 ± 192
CHAPERONE	11004	9349	10690	10553 ± 482
ENZYME	9094	8484	8584	8668 ± 217
ABETA	8654	7582	8324	8298 ± 332