

# Multi-Modal Visualization of Local Environment Data for Protein Structural Alignments

Marc Hansen, Erik Charp  
Computer Science Department, UCSC  
(mhansen,echarp)@cse.ucsc.edu

June 22, 1998

## Abstract

We present and evaluate an integrated visualization/sonification environment for analyzing protein structural alignments (superpositions of two or more protein structures in three-dimensional space). We explore how the use of sound can enhance the perception and recognition of specific aspects of the local environment at given positions in the molecular structure represented. Analysis of protein structures is an area in which it is often necessary to examine many variables simultaneously. This is one reason that we chose this field as our data domain.

Sonification presents the opportunity to improve understanding of the data both in terms of enhancing or disambiguating parameters which are already represented visually, and perhaps more importantly, in extending the number of variables that may be represented simultaneously, particularly by the inclusion of features that are difficult to represent visually. In order to maximize the ability of our sonifications to represent data, we used voices and melodic components that were very distinct. We also used several musical effects such as panning a voice to the left or right speaker, or changing a voice's volume to maximize the individuality of the sonification elements. In addition to enhancing the perception of data features presented visually, we also strove to make our sonifications pleasing and musically idiomatic.

**Key Words and Phrases:** sonification, audification, proteins, structural alignment, local environ-

ment.

## 1 INTRODUCTION

Our primary goals in this project were to sonify certain aspects of a protein's local environments, (i.e. the environment at one particular position in the protein), and to determine whether the sonifications succeeded in helping to make those features more evident to the subjects. We used musical patterns for our sonifications, and employed basic music theory as the basis for our sonification parameters.

There is an inherent trade-off in the application of music to data sonification: As with data visualization, when the number of parameters being sonified increases, so does the risk of creating a cluttered presentation in which it is difficult to detect individual features. One benefit of using sound for data presentation is that it is possible to sonify tremendous amounts of data while making the overall effect musically pleasing. For example, an orchestra may play a piece in which hundreds of instruments play different notes and melodies simultaneously with a beautiful effect. Unfortunately, there are few people who could listen to such a piece and clearly identify with a large degree of certainty what each instrument was playing at any given time. The benefits of using an orchestra as a sonification medium are diminished due to the subtlety necessary to produce a large layering of sound which remains coherent and pleasing. In order to maximize the ability of our sonifications to rep-

resent data, we used voices and melodic components that were very distinct. We also used several musical effects such as panning a voice to the left or right speaker, or changing a voices volume to maximize the individuality of the sonification elements.

A further goal of our project was to compare different sonification features in order to determine which are too subtle to be readily distinguished. In this way, future research will be able to focus more on those techniques which yield the most promise.

## 2 AUDIO-VISUAL DESIGN

### 2.1 Visual Mapping

Visually, the protein was displayed using the molecular graphics tool, Rasmol. [3]. A cartoon representation was chosen, in which the local secondary structure of the protein was explicitly represented (see Figure 1 ). In Rasmol's cartoon mode, the secondary structure is displayed as follows: Strands are depicted as flat ribbons with arrowheads on them. Helices are represented as ribbons wound into corkscrew shapes. Loops are shown as spaghetti-like thin tubes. The cartoon display mode is useful because it allows the overall features of the protein to be shown, while the details not being sonified are neglected. To indicate which section of the protein was being studied, we highlighted that portion in red. The remainder of the protein was colored light gray.

### 2.2 Auditory Mapping

The auditory mapping for this project was based on the idea of musical parts. We used synthesized sounds to emulate natural instruments such as the trumpet and drums.

In essence, our goal was to generate a musical composition out of the data associated with a given protein. Each residue within the protein structure was mapped to a musical interval of time in measures. The overall tempo of the music was constant. Further, we decided to exploit the musical qualities of melody, rhythm, timbre and dynamics, to create mappings of music to the values of local environment

variables. The music was composed of individual layers, or "parts": a melody, a drum part, a bass line, and a harmonic 'comping part (i.e. a rhythmic accompaniment consisting of the chords of the piece played on a keyboard instrument or guitar). Each of the four data parameters under investigation could take on three different values which were indicated by the use of distinct musical patterns. The mappings of the instrument layers to protein data is as follows:

bass line  $\Rightarrow$  secondary structure  
comp part  $\Rightarrow$  polarity  
drum part  $\Rightarrow$  exposure  
melody  $\Rightarrow$  goodness of fit

We added an aspect of intuition to the mappings from music within each layer to the appropriate environment variable value. For example, a polarity value of "more charged" produced a brighter, sharper and busier sounding drum part; an exposure value of "more buried" produced a duller, softer, and more sparse piano part. Thus, there was a set of musical patterns for each layer reflecting the range of values which the corresponding environment variable could take on. Again, there was an emphasis on making the musical nature of these patterns both strongly distinguishable, and suggestive of the values they represent.

### 2.3 Hardware and Software used

The graphical user interface for the project, (see Figure 1), was developed using the Forms Library for X [4]. The system ran on a Silicon Graphics *Octane* running the IRIX6.4 operating system. The workstation contained a 195 MHz MIPS R1000 processor and 128 MB of RAM. Attached to the workstation was a 19 inch color monitor also from SGI. The headphones we used were Sony Digital headphones.

We decided to use the Musical Instrument Digital Interface (MIDI) standard for the format of our musical files. The SGI machines support this specification and can read and play Standard MIDI Files through a MIDI file player.

For displaying the protein molecules, we used Rasmol version 2.6 running in 8 bit mode.

The software used to create midi files was Mark of the Unicorn Performer version 5.0, a multi-tracking

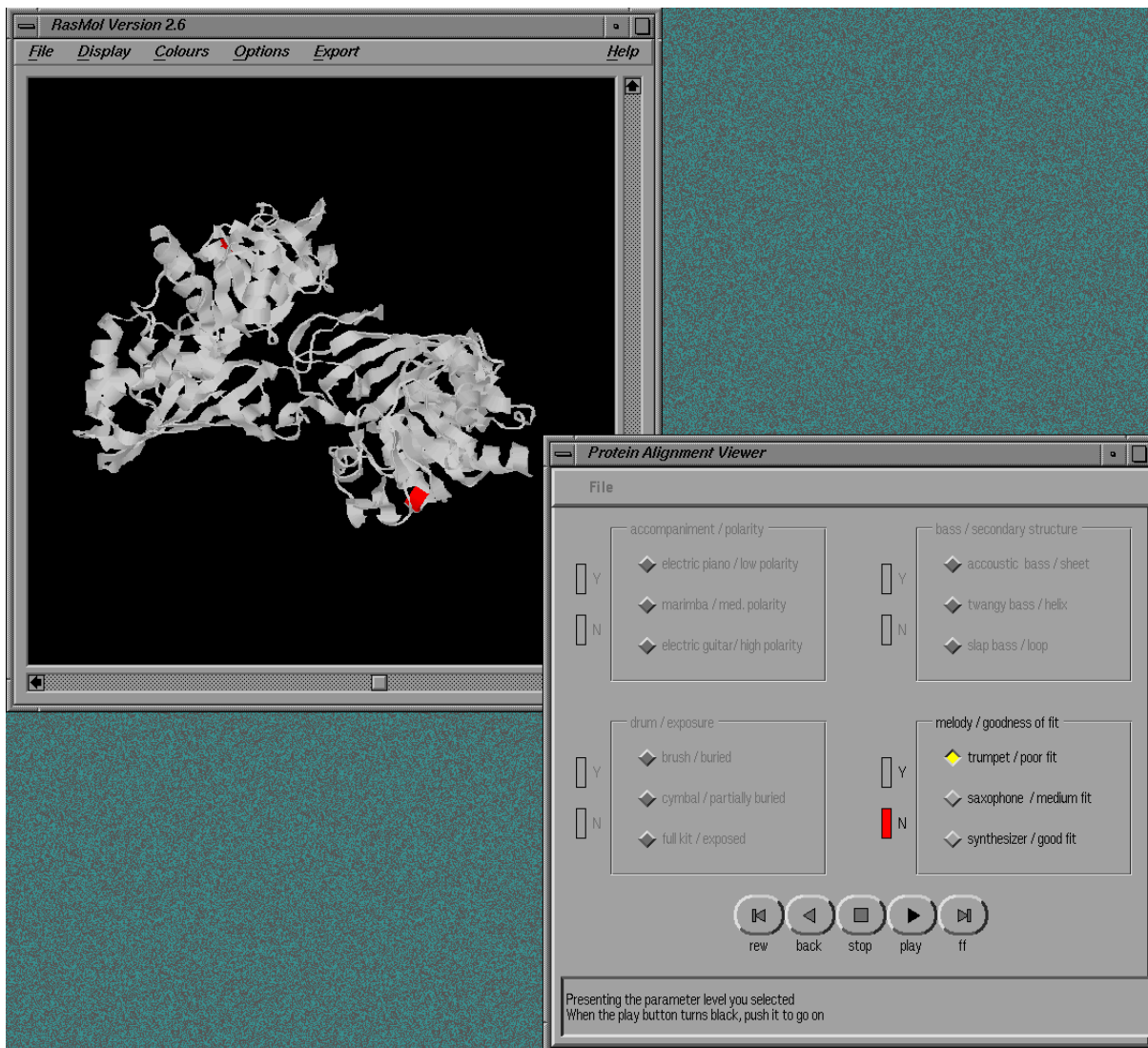


Figure 1: Graphical User Interface

sequencer for the Macintosh. The drum tracks were taken from the cd: XXX.

## 3 APPLICATIONS AND DATA

### 3.1 Description of application

The computational biology department of UC Santa Cruz is currently involved in a contest known as the Critical Assessment of Techniques for Protein Structure Prediction (or CASP for short). In this contest, researchers are given the sequences of proteins for whom no structure has yet been determined. These sequences, however, are released from laboratories where the structure is almost ready for release. In this way the researchers can test their structure prediction methods and compare against the true structures in a relatively short period of time.

In order to generate these predictions, our team will use the protein sequence of unknown structure (i.e. the target), to search a database of sequences with known structures. When a match is found, the sequences must be aligned relative to each other so that their sequence similarities are maximized. In addition, the target sequence must be “threaded” through the structure in order to determine whether the alignment makes sense physically, given the structure’s local 3d environment.

The protein structure features we chose to investigate are useful in assessing sequence-structure alignments as well as structure-structure alignments. Therefore, sonification techniques which yield additional information to those studying structure-structure alignments should also work well for the investigation of sequence-structure alignments.

### 3.2 Description and sources of data sets

The structural alignments were obtained from the FSSP (Families of Structurally Similar Proteins) database [2, <http://www2.ebi.ac.uk/dali/fssp/fssp.html>]. This database is generally accepted as containing excellent

structural alignments and is accessible via the world wide web. For each position in the protein serving as the base structure, the micro-environment can be determined using the suite of environment analysis tools developed by Bowie et al. [1] A comparison of the actual environment versus the environmental preference for the amino acid at each position can be made. If the amino acids are in environments they prefer, this will lend credence to the alignment.

## 4 EXPERIMENTAL DESIGN

### 4.1 Subjects, Collection Environment, and Tasks

We tested 18 subjects. Of these, 7 rated themselves as having better than adequate musical abilities. Two subjects rated themselves as having better than adequate experience with protein structures. Subjects were given headphones and allowed to work on a workstation reserved for the experiment. Subjects viewed or listened to protein data features, either one parameter at a time, or all four parameters simultaneously, and made selections on a graphical user interface, (see Figure 1), by clicking on radio buttons to indicate which value they detected for each parameter presented. The subjects’ responses were recorded, as was the length of time they took to answer.

### 4.2 Experimental Flow Detail

The experiment took a total time of about 45 minutes, and consisted of five phases:

1. Introduction: Subjects were given a two page overview of the experiment concept and purpose. A general explanation of the experiment followed, and the subjects were given a chance to ask questions. After starting the software, subjects were given brief instructions on the basic layout and functions of the relevant controls on the user interface. Subjects then put on a pair of high fidelity headphones.

2. Presentation of audio and visual mappings: Subjects pushed a “play” button on the graphical interface to cause it to simultaneously present a data

sonification and its equivalent visualization. Each of the four sonification parameters: bass, drums, accompaniment, and melody, were presented to each subject in random order using a latin square design. For each of the four sonification parameters the corresponding three levels were presented in order from first to last. The corresponding radio buttons were lit to show the subjects which level was currently being presented. The entire process was repeated twice for each of the four sonification parameters.

3. Training: As with the presentation phase, all training was conducted with both the visual and auditory stimuli presented together. A latin square was used to vary the order of the parameters used for training.

The following sets of values for each parameter were chosen to ensure each of the three possible values was presented at least once:

<i>part</i>	<i>level order</i>
Bass	2,3,1,3
Comp	1,2,3,3
Drums	3,2,1,1
Melody	1,3,2,2

The values in the set were presented to the subjects in random order.

After being presented with a sonification/visualization pair, the subjects were allowed to guess what information they had just received. Subjects indicated their guesses by clicking on a radio button. If the subject answered correctly, a green “Y” light turned on. An incorrect guess caused a red “N” light to turn on. In either case, the auditory and visual information corresponding to the subject’s guess were then presented, thereby allowing the subject to see and hear the difference (if any) between the data presented and the guess.

4. Testing: In the testing phase, a sonification and/or a visualization was presented and as in the training phase, the subject was expected to pick the pattern which was presented. Unlike the training phase, the presentation could consist of a single parameter or all four in combination. The testing section was divided into three modes: audio-only, visual-only, and audio + visual. A latin square was

used to vary the order of the modes between subjects. Each mode was tested through to completion, before proceeding on to the next mode. The following table was used to determine the order of presentation mode testing:

audio	visual	audio + visual
audio	audio + visual	visual
visual	audio + visual	audio
visual	audio	audio + visual
audio + visual	audio	visual
audio + visual	visual	audio

For each presentation mode, testing of individual parameters was followed by testing of all four parameters in combination.

4a. Testing of Individual Parameters:

The table of parameter levels discussed in the training section was also used here. A latin square was used to determine an ordering of the parameters such that 16 trials were created (four trials for each parameter) in random order. The values within each value set were also randomized.

4b. Testing of all Four Parameters Simultaneously:

The following table of eight parameter/value combinations was used to create the simultaneous parameter presentations:

bass 1	comp 2	drums 2	melody 3
bass 3	comp 1	drums 1	melody 2
bass 3	comp 1	drums 1	melody 1
bass 2	comp 1	drums 3	melody 3
bass 3	comp 2	drums 1	melody 1
bass 1	comp 3	drums 1	melody 3
bass 3	comp 2	drums 2	melody 1
bass 1	comp 1	drums 3	melody 2

The table of combinations was presented in random order. The subjects were therefore tested on eight trials of the four simultaneous parameters. The table was chosen from random samplings of actual protein data to represent a good parameter/value mixture of combinations.

5. Exit Questionnaire: Subjects were asked to complete a brief exit questionnaire for the purpose of obtaining feedback on qualitative aspects of the experiment. We were particularly interested in whether

the sonifications intuitively matched the data represented.

## 5 RESULTS

### 5.1 Overall Results

Overall, accuracy scores were much lower for the visual only mode than for either the audio only or audio + visual modes (see Figure 2). There was not much difference between the audio only and the audio + visual modes. In general, adding visual information to the audio did not increase accuracy. As expected, secondary structure had the strongest accuracy scores for the visual only mode, while goodness of fit had the lowest.

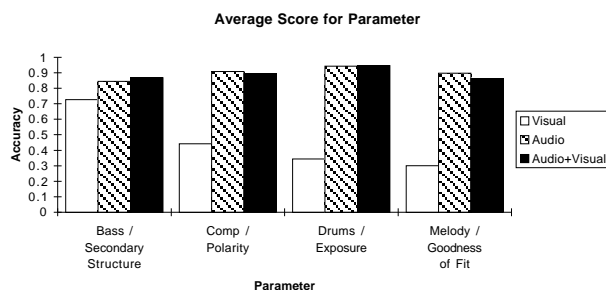


Figure 2: Overall Results

A protein alignment can be expected to show a better goodness of fit in the core regions of the proteins. For a correct alignment, regions with higher structure will tend to be more conserved. Therefore, goodness of fit should in general be higher in regions with secondary structure, and lower in the outer loop regions. This information should not be very difficult to detect visually, but one should not expect subjects without experience working with protein alignments to detect such features, and an explicit visual representation should probably have been used, or an explanation should have been made to the subjects to look for these types of cues. An unexpected finding was that the visual information for exposure of the local environment (how near a location is to the surface of the protein) was very difficult to detect visually.

### 5.2 Accuracy of Discrimination

One of our goals in this project was to see if we could develop a data to sound mapping in which the values of different variables could be distinguished while those individual sonifications were being played simultaneously. We expected to find a drop in accuracy when subjects had to pick out the data values from a simultaneous sonification versus a sonification of the individual voices. We found such a drop in all the sonification parameters except melody. Our interpretation of this finding is that melody is the easiest parameter to discriminate when multiple parameters are sonified simultaneously. The drops in accuracy that we did find for the other parameters were not nearly as steep as we expected, however (see Figure 3).

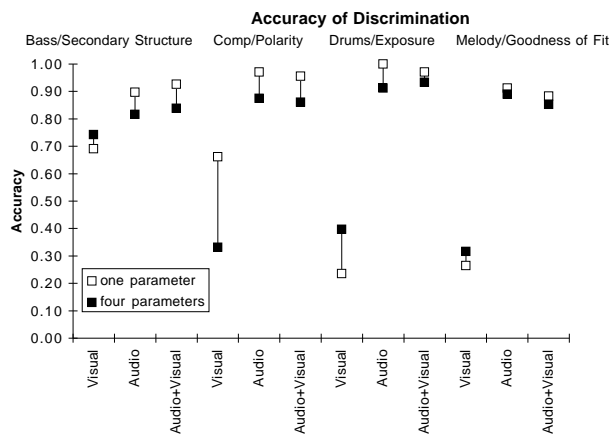


Figure 3: Accuracy of Discrimination

### 5.3 Effect of Experience with Protein Structures

Two of our subjects had prior experience in visualizing protein structures in the cartoon mode we used. As expected, these subjects had a very easy time determining the secondary structure of the highlighted location. Their accuracy levels for this task were near 100% in the visual and audio + visual modes. Surprisingly, their results showed no improvement over

the other subjects in any of the other trials (see Figure 4).

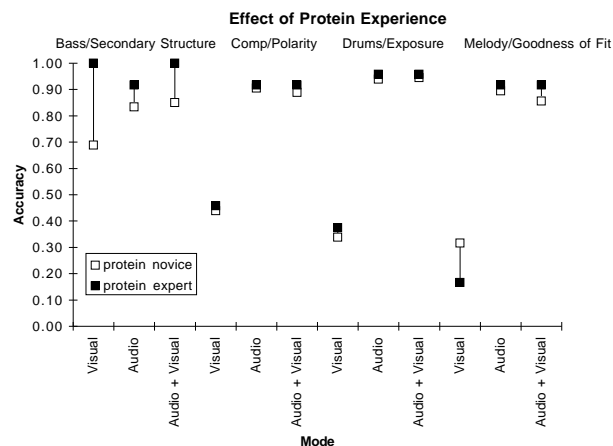


Figure 4: Effect of Experience with Protein Structures

## 5.4 Effect of Musical Ability

We expected that when it came to extracting information from the sonifications, subjects with a high self-rating in musical ability would do better than subjects who rated themselves lower. It was anticipated that these subjects would perform significantly better in the audio only mode. In fact, the only marked improvements we found were in the audio + visual mode (see Figure 5). The most striking improvement was found in the bass/secondary structure parameter. It would appear that this parameter was the most difficult to discriminate for non-musicians. This observation is supported by comments to that effect from several of our subjects. The finding that the musicians in the group showed the greatest improvement in the audio + visual mode is more difficult to explain. Perhaps the subjects with greater musical ability had an easier time discriminating the musical information and were therefore able to concentrate more on the visual information to extract additional information during the audio + visual phase. An analysis of response times in the audio only phase might help in determining the answer to this ques-

tion.

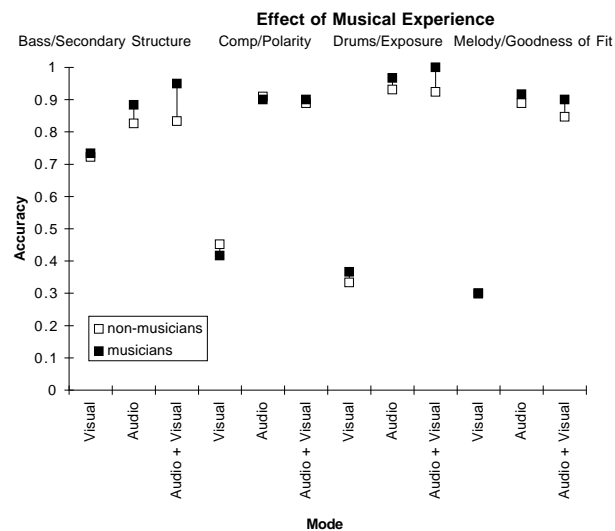


Figure 5: Effect of Musical Ability

## 5.5 Effectiveness of Individual Instruments

We were interested in determining whether certain instrument voices were easier to discriminate than others (see Figure 6). Some of our subjects commented that the easiest instrument to discriminate was the drum. Our results agree with this observation. The drum parameter had the overall highest score for accuracy, and the three drum voices, (brush, cymbal, and full kit), performed consistently well. The scores for the twangy bass performed the worst. The other two bass voices, (accoustic and slap), performed similarly well. In the accompaniment/polarity parameter, the electric piano voice scored better than the marimba and electric guitars which scored about equally well. In the melody/goodness of fit parameter, the synthesizer voice scored the highest, consistent with remarks from subjects that the trumpet and saxophone voices were difficult to distinguish from each other. The trumpet and saxophone voices performed about equally.

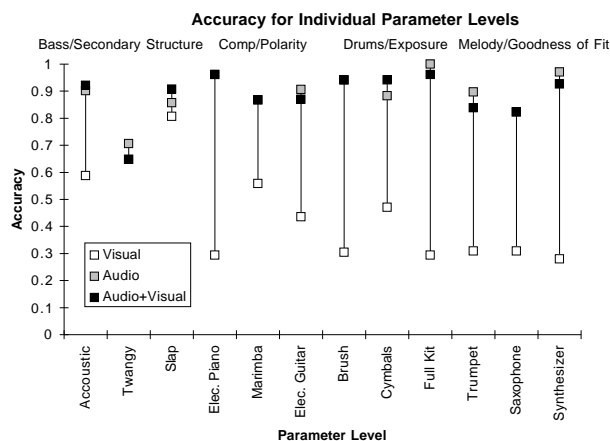


Figure 6: Effectiveness of Individual Instruments

## 6 CONCLUSIONS AND FUTURE WORK

Sonification appears to have a useful role in disambiguating data which may be unclear if presented only as visual information. Sonifications need not suffer serious degenerative effects from simultaneous presentation if care is taken to ensure that distinctive voices, rhythms, and melodic patterns are used for the different parameters and their levels. For sonifications based on musical patterns, melody has the benefit of standing out well when multiple sounds are presented.

### 6.1 Future Work

1. Currently we are sonifying four variables simultaneously. Through initial experimentation, we determined that panning the sonification parameters to the left or right speaker facilitated discrimination during simultaneous sonification. A natural extension of this finding would be to extend the sound output to quadrophonic instead of stereo. We believe that if each parameter is assigned to its own speaker, discrimination should be even further facilitated.

2. Most users would probably like to be able to customize the program: set their own instrument voices, volumes, left-right pan settings, and perhaps even musical patterns, etc. A library of musical patterns could also be provided.
3. It would be interesting to see if the 20 amino acids could be sonified uniquely. For example, almost everyone knows 20 different melodies. The melodies could be mapped to amino acids mnemonically. As a case in point, the amino acid Histidine's one-letter code is 'H'. Histidine could therefore be mapped to a melody beginning with the letter 'H', such as "Happy Birthday".
4. Sonifications could perhaps be extended by having two or more musical patterns per instrument voice.
5. The program could be extended for musicians by having additional levels mapped to similar instruments (e.g., brass instruments such as trumpet and trombone). Since the average listener would probably have difficulties with these mappings, the program could perhaps have an "expert" setting.
6. It would be useful to extend this project to sonify multiple alignments in addition to pairwise ones. Deriving the complex chord structures necessary to convey the quality of an alignment column would be a challenge.

### 6.2 Improvements in experimental design

1. Although in the simultaneous sonification tests we presented four times as much information as in the single sonifications, we only played the sounds for twice as long. It would be interesting to see if playing the sounds for four times as long as the single sonifications significantly increases accuracy.
2. A brief introduction to protein structure would probably have been helpful, along with information on the meanings of the data being mapped.



3. A more fair comparison between visualization and sonification would require additional visualization parameters, e.g., instead of just highlighting the location in red, color could be used to indicate goodness of fit.
4. It would probably be beneficial to train on the sonifications separately for the audio only, visual only, and audio + visual presentation modes. The same is true for training in simultaneous presentation mode.
5. It would be interesting to look at learning rates while using this program. Learning effects could be made a part of the experimental design by not turning off the red/green feedback lights during testing.

### 6.3 GUI improvements

1. Subjects were sometimes unable to tell the current mode of the program. The mode indication could be removed from the status line and made more prominent elsewhere on the interface, e.g., there could be check boxes to indicate whether visual and/or audio modes are on.
2. Subjects were also sometimes confused about which of the parameter groups required a response. Instead of simply enabling the radio button groups (i.e. changing their colors from gray to black), perhaps their borders could be colored bright yellow.

## ACKNOWLEDGEMENTS

We would like to thank Suresh Lodha for his helpful advice on experimental design, and for teaching the data sonification course which led to the work described here.

## DIVISION OF LABOR

Both authors contributed to the experimental design and to the mappings of data to sound. Marc Hansen was responsible for the visual data mappings, the

graphical user interface, and the data analyses. Eric Charp created the midi files used in the project, and also wrote the program for randomizing subjects using a latin square.

## References

- [1] James U. Bowie, Roland Luthy, and David Eisenberg. A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253:164–170, 1991.
- [2] Lisa Holm and Chris Sander. Fold classification based on structure-structure alignments of proteins (FSSP). *Nucleic Acids Research*, 26:316–319, 1998.
- [3] Roger Sayle and E.J. Milner-White. RasMol: Biomolecular graphics for all. *Trends in Biochemical Sciences*, 20:374–376, 1995.
- [4] T.C. Zhao. XFORMS home page. URL: <http://bragg.phys.uwm.edu/xforms>.