Special Issue on Immersive Analytics

# Auditory Analytics for pattern discovery in protein folding dynamics

Carla Scaletti[a], Kurt Hebel[a], Martin Gruebele[b,c,d,e]

[a]Symbolic Sound Corporation, Champaign, IL 61820

[b]Department of Chemistry, University of Illinois Urbana-Champaign, Urbana, IL 61801

[c]Center for Biophysics and Quantitative Biology, University of Illinois Urbana-Champaign, Urbana, IL 61801

[d]Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Urbana, IL 61801

[e]Department of Physics, University of Illinois Urbana-Champaign, Urbana, IL 61801

**Correspondence:** Carla Scaletti (carla@symbolicsound.com) and Martin Gruebele (mgruebel@illinois.edu)

**Keywords**: sonification, data sonification, multi-modal data analysis, protein folding reaction, hydrogen bonds, molecular dynamics simulation

## ABSTRACT

We introduce Auditory Analytics, a methodological framework that utilizes data sonification for scientific discovery. Auditory Analytics describes a cycle of collecting and deriving datasets, mapping data to audible signals (sonification), analytical listening, hypothesis formulation, and tool building, where human insights from any stage of the cycle can feed back into further iterations of the cycle in the form of new datasets, alternative mappings and new models of the original phenomenon. In Auditory Analytics, the remarkable capacities of the human auditory system to monitor complex soundscapes, track multiple sources, and extract meaningful information across multiple timescales are repurposed for exploring, interpreting, and analyzing data. To demonstrate its potential for uncovering relationships and dynamics in physical systems, we apply the Auditory Analytics methodology to the domain of protein-folding, investigating state transitions in a molecular dynamics simulation of the GTT

WW domain protein. Auditory Analytics led to the identification of distinct hydrogen bonding patterns that occur as the protein transits between folded and unfolded states and thus to a deeper understanding of the process of protein folding. A single, isolated data mapping — whether visual, auditory, haptic, mathematical, or verbal — provides an incomplete picture of reality; by adding the Auditory Analytics cycle to our portfolio of data interpretation tools, we can build a more complete picture of physical phenomena.

# 1 Introduction

Microscopes, radio telescopes, Geiger counters, and other scientific tools — like data sonification, visualization, and other data mappings — transpose measurements and observations of the natural world into ranges and timescales that lie within the dynamic range and bandwidth of the human somatosensory system.

However, a single laboratory instrument or isolated data mapping — whether visual, auditory, haptic, mathematical, or verbal — provides only an incomplete picture of reality. We propose the Auditory Analytics Cycle, a key component of Immersive Analytics, as one approach toward overcoming this limitation.

Auditory Analytics is an interactive cycle of collecting and deriving datasets, mapping data to audible signals, analytical listening, hypothesis formulation, and tool building; at all stages of the cycle, human insight feeds back into further iterations of the cycle as new datasets, alternative mappings and new models of the original phenomenon. We applied the Auditory Analytics methodology to investigate state transitions in a molecular dynamics simulation of the GTT WW domain protein.[1]

## 1.1 Visual Analytics

Visual Analytics, defined by Thomas and Cook as "the science of analytical reasoning facilitated by interactive visual interfaces"[2] is "an iterative process that involves information gathering, data preprocessing, knowledge representation, interaction and decision making" whose ultimate goal is "to gain insight into the problem at hand".[3] Whereas 'data visualization' is defined simply as representing data graphically, Visual Analytics is described as "an interface between humans and computers, providing (interactive) data representations to the user on demand. Such approaches can be used to understand data, to confirm or reject hypotheses, and to discover new knowledge."[4]

## 1.2 Immersive Analytics

By extension, Immersive Analytics[5] can be defined as analytical reasoning facilitated by interactive, multimodal interfaces that employ auditory, visual, haptic, and other sensory modalities to leverage human perceptual and pattern-matching abilities — capacities that have evolved in response to challenges posed by the natural world.

### 1.2.1 Multimodal all the way down

To accurately and reliably perceive their surroundings, terrestrial organisms fuse information from various sensor types; this multimodal integration, which extends along afferent pathways deep into the central nervous system, is crucial for navigating a dynamic 3D environment immersed in electromagnetic and pressure waves, chemical gradients, and the pervasive forces of gravity, contact, friction, and tension. As Asif Ghazanfar and Charles Schroeder write, "…the multisensory nature of most, possibly all, of the neocortex forces us to abandon the notion that the senses ever operate independently during real-word cognition."[6]

### 1.2.2 A role for the auditory system

While interactive visual interfaces have been central to Immersive Analytics since its inception, auditory, haptic, and olfactory interfaces have been less extensively explored. This, despite the evidence that all human perception is multisensory.

The remarkable capacities of the human auditory system for identifying and tracking multiple moving sound sources, detecting and localizing transient events in a 360° field surrounding the listener, and decoding complex modulated signals at multiple timescales make it a strong candidate channel for data transmission and interpretation — a compelling argument for the inclusion of interactive data sonification as a component of the Immersive Analytics tool kit.

## 1.3 The Human auditory system

Nearly every animal species can both detect and actively generate acoustic signals. The accurate detection, categorization, tracking, and interpretation of both self-produced and externally-generated acoustic signals confers a significant selective advantage.

### 1.3.1 Surviving in the dark

Compared to that of other species, the mammalian auditory system exhibits superior sensitivity, source localization, and discrimination over a wider range of frequencies, perhaps because the earliest mammals, being endothermic, were able to occupy a nocturnal niche to evade competition and predation from larger, diurnal animals.[7]

**Early warning system.** Functioning as an always-on, 360° early warning system, the mammalian auditory system uses sound pressure waves to sense the invisible, and, unhindered by ear-lids, can even rouse an animal from sleep. By way of contrast, the human field of view spans approximately 114° for binocular vision with about 40° of peripheral vision on each side; turning your head or closing your eyes interrupts your perception of a visually tracked object.

**Sound localization mechanisms.** The sound localization and tracking capabilities of mammals, including their ability to follow moving sources obscured from view, relies on inter-aural time and intensity differences (for azimuth cues), and the spectral filtering resulting from reflections off the shoulders and fine structure of the outer ear (for elevation and behind-the-head cues).

**Change detector.** As expected of an early warning system, the human auditory system, in common with other senses, has evolved to detect changes relative to a reference background: a ratio of two values, rather than a single, unchanging value. For example, while a static barometric pressure is not perceived as sound, a rapid pressure fluctuation of sufficient magnitude is perceived as a transient, broadband sonic event, commonly known as a 'click.'

**The ears tell the eyes where to look.** Vision and audition function cooperatively: often, for example, when an unexpected sound is detected (*e.g.* a deep growl), vision is used to further assess the situation, which then leads to higher level cognitive processes to plan subsequent actions.

**Auditory scene analysis.** The need to continuously monitor the environment, a selective pressure that persists to the present day**,** is reflected in the human capacity for Auditory Scene Analysis: the ability to extract multiple concurrent moving sound sources from a pair of one-dimensional audio signals (one from each ear).[8]

**Temporal and spectral integration.** Extracting auditory patterns from natural acoustic signals (which are distributed in both time and frequency and contaminated by extraneous ambient noise) requires integrating information in both the time and frequency domains.

Temporal and spectral integration in the auditory system enhances the signal-to-noise ratio by allowing neural activity incited by an input to accumulate over time (for long duration signals) or across frequencies (for broadband signals).

Integration can also serve to smooth over noisy rapid changes in audio signals, so the listener can form a unified percept of a sound source, even when it is interrupted by gaps or modulations on the order of tens or hundreds of milliseconds.

**Spectral shape (timbre) for identification.** 'Spectral shape' or the distribution of energy across all frequencies plays an important role in identifying a sound source and inferring some of its physical characteristics in terms of its resonant frequencies (*e.g.*, the high-pass hiss of a snake or the low-pass growl of a predator). The spectral shape or timbre of a sound source is often described as brighter or high-pass when its energy distribution is skewed toward higher frequencies and darker or low-pass when the energy is skewed toward the lower end of the spectrum.

### 1.3.2 Surviving in a social group

Arguably the strongest selective pressure on the human auditory system is the challenge of accurately perceiving and interpreting the sounds generated by other humans. For humans, the ability to accurately distinguish among auditory cues signaling danger (such as pain, fear, anger, screams) and those indicating safety or affiliation (such as laughter, infant-directed-prosody, or a sigh of relief) is crucial to survival and reproductive success.

Since speech is an audio signal whose frequency, amplitude, and spectrum are modulated to encode meaning as prosody, phonemes, syllables, and words, it is not unreasonable to expect that humans can also extract meaning from audio signals whose parameters are modulated by data streams: data sonification.

Most humans are adept at extracting meaning from audio signals — it's a skill that we've been practicing daily since infancy. By age 3-4 years, human children have already been analyzing the statistical distributions of the sounds they hear around them, categorizing sounds into phonemic categories, and using transitional probabilities and prosodic cues (pitch, timing and amplitude variation) to segment audio streams into meaningful units like syllables and words.[9] This skill of extracting meaning through analytical listening to sounds that evolve at multiple time scales can and does improve with practice, even with short-term training.[10]

**Prosody and indexical meaning.** Prosody — the continuous modulations in pitch, timing, and intensity that constitute the non-verbal aspects of human speech — conveys indexical meaning irrespective of which, if any, language is used. A familiar example is the exaggerated prosody one unconsciously adopts when addressing infants who are not yet expected to understand spoken language.[11] Babies and co-evolved domestic animals like dogs seem to understand the meaning of prosodic variations, even in cases where the symbolic words contradict that meaning.

Human prosody (in particular, modulations to the fundamental frequency: *F0*) can communicate indexical cues as to the speaker's identity, age, sex, social status and even predicted mating success,[12] and there is evidence that prosody may have its own vocabulary, semantics, and simple syntax.[13]

In addition to prosody, humans generate meaningful, non-linguistic audio signals — laughter, screams, cries of fear, moans of pain — that can be several orders of magnitude louder and have different pitch ranges than normal speech[14] and convey vital information to the listener by way of nonlinear spectra, frequency modulation and amplitude modulation.

**Sensitivity to continuous pitch variations.** In typical human speech signals, the fundamental frequency *F0* ranges from ~148–260 Hz for women and 81–153 Hz for men,[14] and the harmonics, vowel formants, noisy elements (like consonants) and extra-linguistic vocalizations extend to the full bandwidth of human hearing (a spectral analysis of a baby cry shows an *F0* ~500 Hz with strong harmonics up to 20 kHz).

**Spectral content and symbolic meaning**. While frequency, timing, and amplitude modulations to *F0* define the prosody of speech and convey indexical meaning, the symbolic meaning of speech is encoded by the vocal tract's time-varying filtering of the upper harmonics of *F0*, requiring high resolution in both timing and frequency and the capacity to detect parallel and divergent changes among groups of harmonics.[8,15]

**Multiple time scales.** The interpretation of meaningful audio signals (like speech, music and data sonification) requires integrating information across multiple timescales, ranging from the fine-structure of the waveform (with cycles ranging from 50 µs to 50 ms), to the median duration of individual phonemes (~60 ms), syllables (~200 ms) and words (~500 ms), and extending to even longer duration structures like phrases, sentences, and longer.[16]

# 2 Auditory Analytics Cycle

Auditory Analytics repurposes the remarkable capacities of the human auditory system — monitoring complex soundscapes, tracking multiple sources, and extracting meaningful information across multiple timescales — for exploring, interpreting, and analyzing data.

The Auditory Analytics Cycle shown in Figure 1 is a methodological framework — a process of iteratively traversing a cycle that includes collecting raw data, deriving subsidiary datasets to answer specific questions, mapping the derivative data to audible signals, interactive analytical listening leading to hypothesis formulation, and rapid prototyping of software/hardware tools for feeding insights back into the cycle in the form of new datasets, alternative mappings and models of the hypotheses.

## 2.1 Transforming and computing derivative data

Often, the first step is to extract or compute a derivative dataset from the raw observations. For example, an X-ray or electron diffraction pattern dataset (intensities as a function of angle and position) might first be recomputed as a dataset giving the ($x$, $y$, $z$) position of each atom, and this dataset may in turn be transformed into a data set including the van der Waals radii for each element type in order to visualize a protein molecule as a van der Waals surface or as a ribbon structure for interpretation by human analysts.

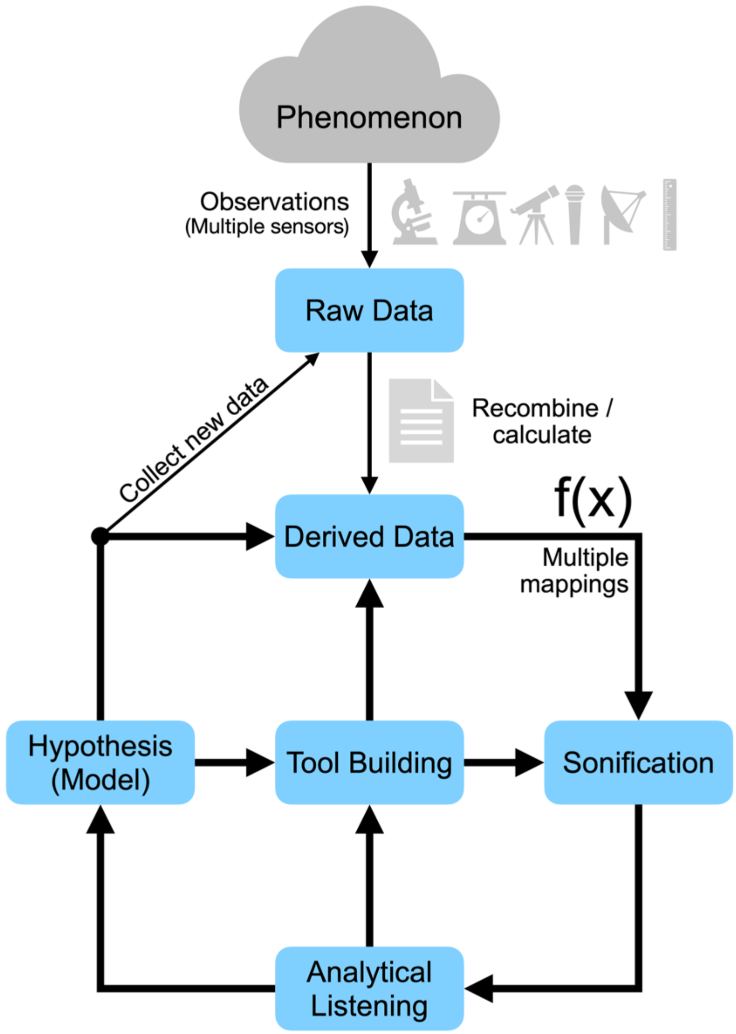



*Figure 1. The Auditory Analytics Cycle: A cycle of investigating a real-world phenomenon which includes data collection, deriving/reshaping the data, mapping data to sound signals, analytical listening, hypothesis formation, mediated by software/hardware tool building where human perception and cognition feed insights back into the loop.*

## 2.2 Data sonification

There are two components to data sonification: a technique (mapping data to sound), and an intent (to interpret, understand, or communicate aspects of the underlying phenomenon from which the data were sampled). Data sonification is "a mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purposes of interpreting, understanding, or communicating relations in the domain under study."[17]

In the context of Auditory Analytics, data sonification is embedded in a cycle of observation, analysis, tool-building, re-mapping, and model-building. Sonification — the mapping of data to audible signals — is a necessary step in the Auditory Analytics cycle, but sonification, in isolation, is not sufficient to define the Auditory Analytics process. The role of data sonification in the Auditory Analytics Cycle is to provide insight; sound is not the end goal but one stage in a cycle whose goal is to explore and to understand some aspect of the natural world.

In the words of science philosopher Alfred Korzybski, "A map is not the territory it represents, but, if correct, it has a similar structure to the territory, which accounts for its usefulness."[18] Since you cannot map everything, the art of data sonification lies in designing a mapping that preserves key relationships among data points and maps significant aspects of the data to perceptually salient parameters of sound.

A data sonification or visualization is useful only to the extent that it is inference-preserving. Observations that you make by listening to the data-driven sound should also hold true in the original phenomenon from which those data were sampled. An effective mapping captures the important relationships among data points, not just individual points in isolation.

A single data set, visualization, or sonification provides only an incomplete map of a phenomenon. Bruno Latour argues that scientific facts are progressively established through a cascade of evidence — such as texts, equations, and visualizations — which iteratively transform raw data and observations into a cohesive, cumulative argument.[19]

## 2.3 Analytical Listening

Analytical Listening is an active mode of listening characterized by feature extraction, pattern detection, and speculation as to the causes of the patterns (venturing into the Hypothesis node of Figure 1).

## 2.4 Centrality of Tool Building

Tool Building sits at the center of the cycle to enable the rapid, interactive exploration of an unknown territory. The answer to one question inevitably leads to another, and each new question may require a new tool devised specifically to answer that question.

Although standard mappings exist and can be used to address routine inquiries, exclusive adherence to a rigid set of standardized mapping techniques runs the risk of stifling innovation and missing out on the discovery of novel and potentially more effective mapping strategies.[20] Even in the long-established field of visualization, new mappings are regularly introduced to deal with specific or new problems in understanding data.

### 2.5 Exploratory, iterative, and interactive

In the words of biophysicist Martin Gruebele, "There is no research without interaction."

At its core, learning is a shift from passive consumption to active exploration and the ability to manipulate and question data.

Elements that support the interactive exploration of the unknown include:

1. Interactive visual/auditory displays that give a researcher agency to explore the space
2. Dynamic programming languages that support the rapid prototyping of new tools (*e.g.*, Smalltalk[21,22] and Python[23])
3. Algorithms that facilitate the discovery of unanticipated patterns in unlabeled data (Dimensionality-reduction, PCA, Clustering, Auto-encoders, GANs, and other unsupervised learning algorithms)
4. Explainable models for validating hypotheses

## 3 Case study

Proteins are essential to life, and an understanding of how and why a linear chain of amino acids reliably folds itself into a characteristic three-dimensional structure can provide insights into protein function, malfunction, and interaction with other molecules in living systems.

When proteins fold, specific secondary structures such as helices or beta sheets form, interlaced by protein-protein hydrogen bonds that organize shape. Hydrogen bonds form when a hydrogen atom is attracted by an atom such as oxygen or nitrogen in the protein. These secondary structures fold upon one another to produce tertiary structure or the 3D shape of a protein.[24] The solvent (usually an aqueous medium) forms additional protein-water hydrogen bonds with the surface of the protein, and water restructures itself near a protein surface to accommodate and assist folding.

When analyzing state transition passages in a molecular dynamics simulation of a small protein,[1] we traversed the Auditory Analytics cycle of Figure 1 multiple times, finding that:

1. **Analytical Listening drives Hypothesis and Tool Building:** Listening to the initial sonification of protein dynamics led to hypotheses about distinct timescales and the nature of transitions. This, in turn, drove the development of new tools like the $\mu \pm \sigma$ plot and the rarity function to better identify and analyze transition state passages.

2. **Sonification revealed patterns missed by visualization:** Our initial visualization of hydrogen bond dynamics was ineffective due to visual clutter and occlusion. Our solution — a 'Polyphonic Geiger Counter' sonification, mapping different bond types to distinct pitches and spatial positions — allowed us to hear a correlation between the speed of a transition and characteristic, distinctive hydrogen bonding patterns.

3. **The ears can tell the eyes where to look:** Listening to the sonification of hydrogen bond data revealed patterns that inspired the development of algorithms to automate the analysis, demonstrating the ways in which Auditory Analytics can guide further visual and computational exploration and tool building.

4. **Auditory Analytics can lead to discovery:** Mapping data to sound harnesses the highly evolved capacity of the human auditory system to track multiple auditory streams at multiple timescales.

## 3.1 Molecular dynamics simulation

Molecular dynamics (MD) simulation, a computational technique used in protein research since the 1970s,[25,26] models protein dynamics and structure by applying Newton's laws to update the position and velocity of each atom of the protein based on interatomic forces at femtosecond time intervals.

### 3.1.1 Raw Data

The focus of this study was not to predict the folded structure of the protein but to examine the transitions between the metastable folded and unfolded states and, in particular, the interaction of hydrogen-bond dynamics in connection with the rates of these 'transition state passages.'[1] To this end, we analyzed multiple folding and unfolding events in a molecular dynamics (MD) simulation of a small protein known as the GTT WW domain.[27] In the context of the Auditory Analytics Cycle in Figure 1, this MD simulation is a Model and the source of Raw Data from which we computed the multiple Derived Datasets discussed in the upcoming sections.

### 3.1.2 Computed Datasets (order parameters)

For each time step of an MD simulation, a protein's configuration space has dimension *3N* — three spatial coordinates for each of the *N* atoms in the protein — and the phase space, because it includes velocity, doubles this dimension to *6N*.

Given that a single WW domain protein comprises around 400 atoms (resulting in a *6N* of 2400), and that MD simulations also include tens of thousands of surrounding water molecules, the dimensionality of the phase space becomes exceedingly large — even more so, once you add the dimension of time.

To monitor the complex dynamics of an MD simulation, high-dimensional phase space trajectories are often projected onto one-dimensional order parameters. The reaction coordinate (the minimum free energy path) would be the ideal choice, but it is not directly observable, so geometrically computable proxies are used instead.

We selected three, geometric order parameters that require no *a priori* knowledge of the folded protein structure: SASA (solvent-accessible surface area), EHA (exposed hydrophobic area), and $R_g$ (radius of gyration).[1] The first two measure how much and what type of a protein's surface is exposed to the surrounding water molecules, and the third measures the size of the protein. Each can be calculated from the atomic coordinates using well-validated formulae available in standard molecular dynamics software.[28–30] In the Auditory Analytics Cycle of Figure 1, computing these order parameter trajectories corresponds to deriving new datasets from the Raw Data.

## 3.2 Sonification ($0^{th}$ order)

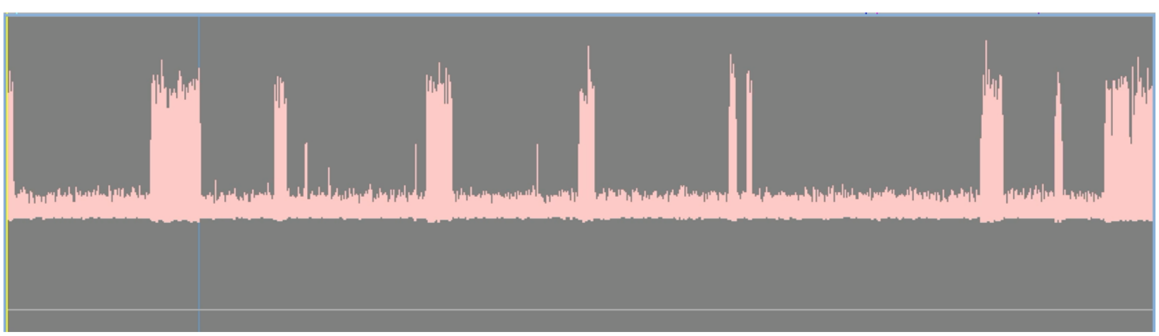
Sonification — designing and applying a mapping from data to some parameter of an audio signal — corresponds to the link labeled f(x) in Figure 1. In a $0^{th}$ order sonification mapping,[17] each row of a single column of a data file is mapped to the instantaneous amplitude of an audio signal.

Each time series dataset — $R_g$, SASA, EHA — contains approximately 3 million measurements at a sampling rate of 5 GHz, representing a total duration in simulation time of 653 µs. These time series were converted to audio signals by rescaling the numbers to 24-bit integers and saving them in a standard audio file format,[31] setting the sample rate to 48 kHz. During rescaling, we chose to maintain the sign and the offset of the data, so that visualizations of the audio waveform would be consistent with visualizations of the order parameter time series.

This mapping reduces the original 5 GHz sampling rate of the simulation by a factor of $\sim 10^5$ to 48 kHz (from the original 200 ps period to a period of ~21 µs), to obtain an audio signal whose duration is ~68 s and whose maximum bandwidth is 24 kHz.[32]

**Analytical Listening.** Analytical Listening is an active mode of listening characterized by feature extraction, observation, and speculation.

Listen, for example, to Video 1 — a mapping of $R_g$ to instantaneous amplitude, presented at a 48 kHz sample rate.



https://youtu.be/sFEFYRL4sP4

*Video 1. $R_g$ mapped to the instantaneous amplitudes of an audio signal at a sample rate of 48 kHz. Louder bursts correspond to unfolded protein states when atoms move freely in the solvent, while softer, quieter periods correspond to the folded state when the atom positions are more constrained.*

On first listening, this signal might be described as noise; however, with each subsequent listening, an analytical listener can extract progressively more information, make observations, and speculate as to the underlying causes in the original phenomenon.

Observations might include:

1. The audio fluctuations seem to contain loud amplitude bursts between softer time intervals (~3 s between bursts).
2. A larger percentage of the total time is spent in the low amplitude versus the high amplitude 'state'.
3. The transitions between small and large amplitude states happen so quickly that they sound instantaneous.
4. There is ~21 μs fine structure that sounds like a combination of very low amplitude broadband (white) noise mixed with low-pass filtered noise.

Possible explanations for the observed features include:

1. The low amplitude intervals (smaller $R_g$) correspond to folded states and the high amplitude intervals (larger $R_g$) correspond to unfolded states.
2. At the temperature chosen for the simulation, the folded state is more likely to be populated.
3. Transition state passages between the folded and unfolded states occur very quickly, typically < 20 ms in the audio file (corresponding to < 200 ns in the MD simulation).
4. The low-pass filtered noise could arise from the Brownian motion of the atoms in the simulation and the much softer broadband noise could be due to simulated thermal energy.

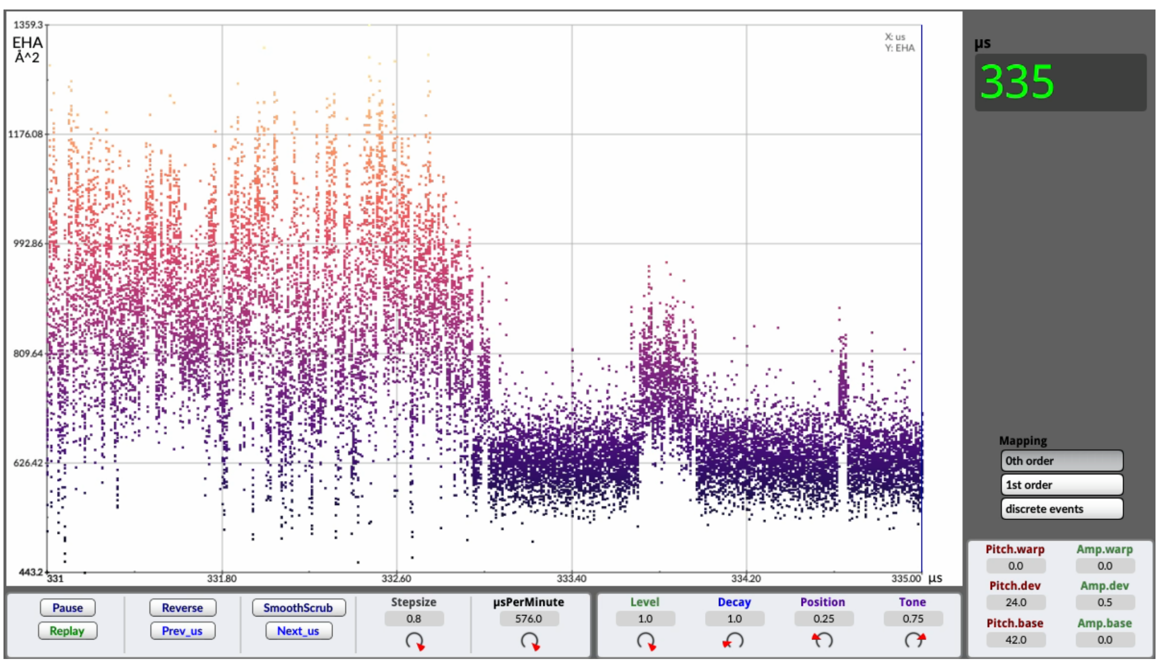
A listener to Video 1 can detect when the transitions occur and for how long the protein remains in its metastable unfolded state following each transition, thereby inferring something about the kinetics and equilibrium of folding. However, the actual transition state passages between folded and unfolded states happen so quickly (on the order of tens of milliseconds in the audio file) that it's difficult to hear fine-grained details of how the actual transition occurs.

As is typical in the Auditory Analytics cycle, Analytical Listening leads to Hypotheses, Models, and to further questions (“What’s going on in those transitions that are too brief to hear?”) which leads to Tool Building: devising an alternative mapping from data to sound to provide the fine-grain resolution needed to answer that question.

## 3.2 Zooming in on transitions

In a visual representation, zooming-in is typically achieved through axis scaling (magnification): smaller data increments are mapped to larger pixel ranges to increase their visibility.

The auditory equivalent of zooming-in would be to expand the timescale (i.e., to present fewer samples-per-second); this causes changes in the signal to occur more slowly, allowing the auditory system additional time to parse and interpret those changes (Video 2).




https://youtu.be/oX_xFkZHAi4

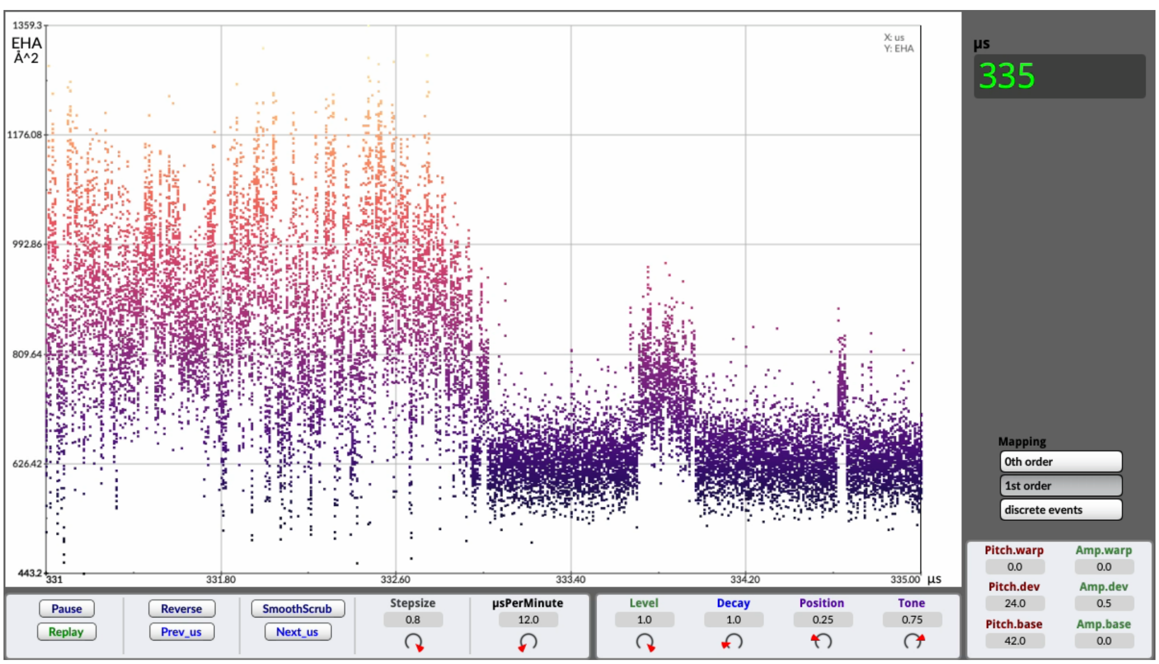
*Video 2. Zooming-in on a $0^{th}$ order mapping of EHA by reducing the sample rate (and its corresponding Nyquist low-pass filter) from 48 kHz to 1 kHz (a factor of 48). This is $5{\times}10^6$ times slower than the original 5 GHz sample rate of the MD simulation. Each row of data is mapped to an instantaneous amplitude.*

### 3.2.1 Continuous modulation

By leveraging the human ability to parse speech signals at multiple time scales, in data sonification, we can map data, not just to an audio waveform, but also to lower-rate (< 15 changes per second) modulations applied to the parameters of generative audio models.

This approach is related to the principle of amplitude modulation (AM) radio transmission, in which a high frequency electromagnetic carrier wave (100s of kHz) is amplitude modulated by another, lower rate, signal whose bandwidth lies with the range of human hearing (tens of kHz), and 'the modulator is the message'.

In parameter sonification (1st order mapping[17]), a data stream (or streams) modulates one (or more) parameters of a generative sound synthesis or processing model. The frequency of the audio signal generated by the model lies in the audible range, but its parameters are modulated at sub-audio rates (< 20 Hz), taking advantage of the auditory system's innate and trained capacities for speech recognition at those transmission rates (Video 3).




https://youtu.be/YxadQqZtfHg

*Video 3. Zooming-in on EHA by reducing the rate and mapping each row of data to continuous modulations of the pitch and amplitude parameters of a sustained oscillator sound synthesis model.*

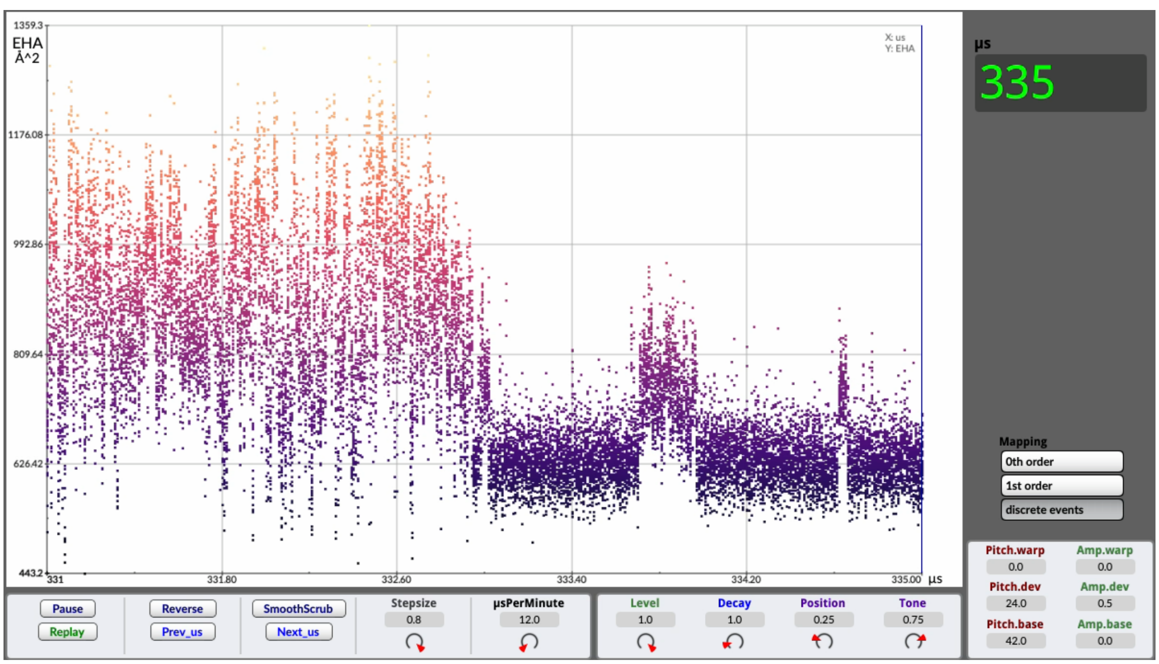
### 3.2.2 Discrete events

Alternatively, as heard in Video 4, each row of data (each time step) can initiate a discrete auditory event whose parameters are set by the value(s) in that row, relying on the human capacity for temporal integration to fuse the sequence of events into an auditory stream: a perceptual grouping of a sequence of sounds into a coherent whole.[8]

### 3.2.3 Analytic Listening to the transition sonifications

In Videos 3-4, unfolded states tend to be higher in pitch and have larger pitch variations, whereas the folded states tend to be lower in pitch with comparatively little variation. For each of the order parameters, we could hear that both mean value and the standard deviation were larger for unfolded states and smaller for folded states.

This is consistent with the expectation that an unfolded protein has greater conformational freedom as it moves through the solvent. Whereas, in the folded state, the protein is more compact, and the movements of its atoms are more constrained.




https://youtu.be/toYLlc5TsQc

*Video 4. Zooming-in on EHA by reducing the rate and mapping each row of data to the pitch and amplitude parameters of discrete but overlapping sonic events.*

## 3.3 Analyzing state transition passages

Now that we had a way to zoom-in on a single transition state passage, we returned to the Tool Building phase to study the transitions in more detail.

### 3.3.1 Trajectories through μ±σ space

Since changes in state appear to be accompanied by changes in the magnitude of both the mean and the standard deviation, we tried looking at the mean value of each order parameter with respect to its standard deviation. We computed the moving mean and standard deviation of each order parameter, using an adjustable-length window (500-125 ns) and selectable window shape (a Gaussian with a wider or narrower standard deviation).

When an order parameter trajectory is plotted in this way (Figure 2a), the folded states tend to cluster toward the lower left, the unfolded states toward the upper right, and the state transition passages span the middle area between folded/unfolded states. The positive slope from left to right suggests that there is a linear dependence of standard deviation on mean value.

To remove the linear dependence of standard deviation on mean and reduce the six dimensions (the mean and standard deviation of each of the three order parameters SASA, EHA, and $R_g$) to a 2D space, we used singular value decomposition (SVD) to perform a principal components analysis (PCA). The first two principal components (PC1 and PC2) account for 98% of the total variance and serve as the basis vectors for a new 2D space we call μ±σ space (see Supporting Information of Ref. 1).

In μ±σ space, folded states are concentrated on the left, unfolded states are toward the right, and the transitions are the trajectories in-between (Figure 2b).

As a plot of value versus how much that value fluctuates at that time, the μ±σ plot is related to the $f(t)$ vs $f'(t)$ phase portrait often used for describing the trajectories of a dynamical system through state space. Unlike the derivative, however, the standard deviation is strictly positive — appropriate for systems where the magnitude of fluctuations, rather than their sign, contains most of the physical meaning.

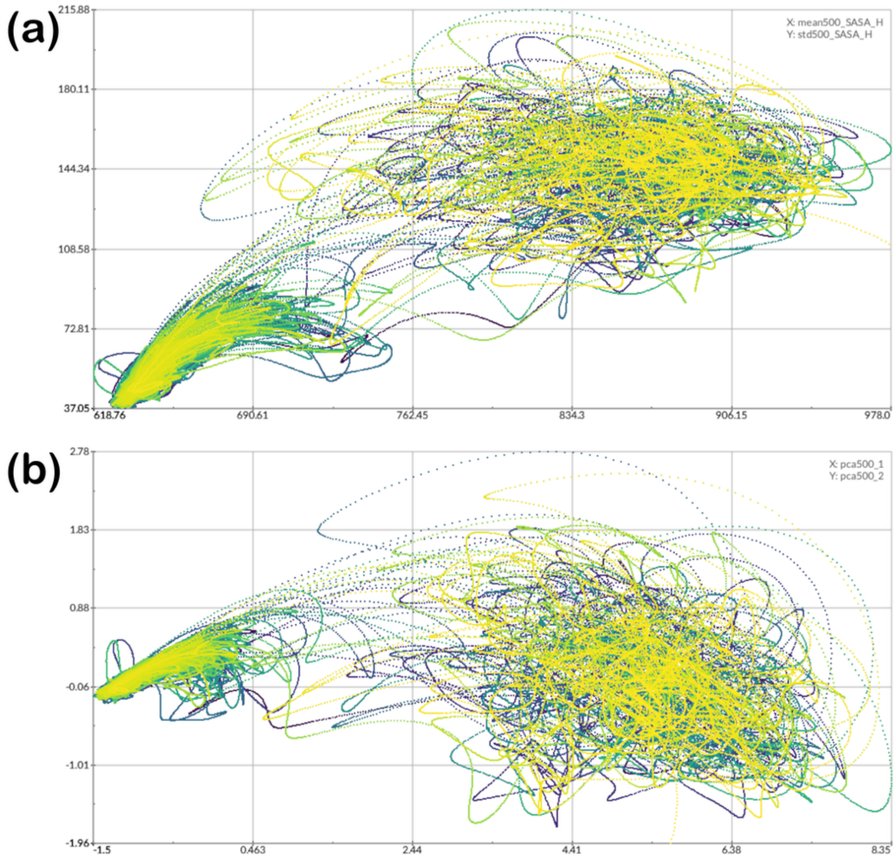



*Figure 2. (a) Mean versus standard deviation portrait of a single order parameter, EHA, using a 500 ns window, with time mapped to color. (b) A μ±σ portrait defined by the first two principal components of the mean and standard deviations of three order parameters. For this figure, the window length is 500 ns, but for the analysis we used a shorter window length (125 ns) to obtain a more detailed portrait of the state transition passages. (From Supporting Information of Ref. 1.)*

### 3.3.2 The Rarity function

To automatically detect the start and end times of transitions, we devised a rarity function based on the observation that the system spends most of its time in one of its metastable states; thus, it is rare for a trajectory to visit a transition state.

In a full trajectory through μ±σ space (Figure 2b), there is a concentrated cluster of small PC1 values (corresponding to the folded states on the left), a somewhat less concentrated cluster of large values (corresponding to the unfolded states on the right), and a sparsely populated region in the middle range of values (corresponding to the state transition passages between folded and unfolded states).

Sorting the PC1 values in ascending order results in a curve that can serve as a quantile density function (an inverse cumulative distribution function), where the steepest slope corresponds to the rarest PC1 values. We define the rarity function as the slope of the quantile density function at each value of PC1.

By applying a threshold to the Rarity function, we highlight the values of PC1 that are part of the state transition passage between folded and unfolded states. From this, we label the state transition passage time intervals in the trajectory time series (see Supporting Information Fig. S4 and SI Code and Data Archive of Ref. 1).

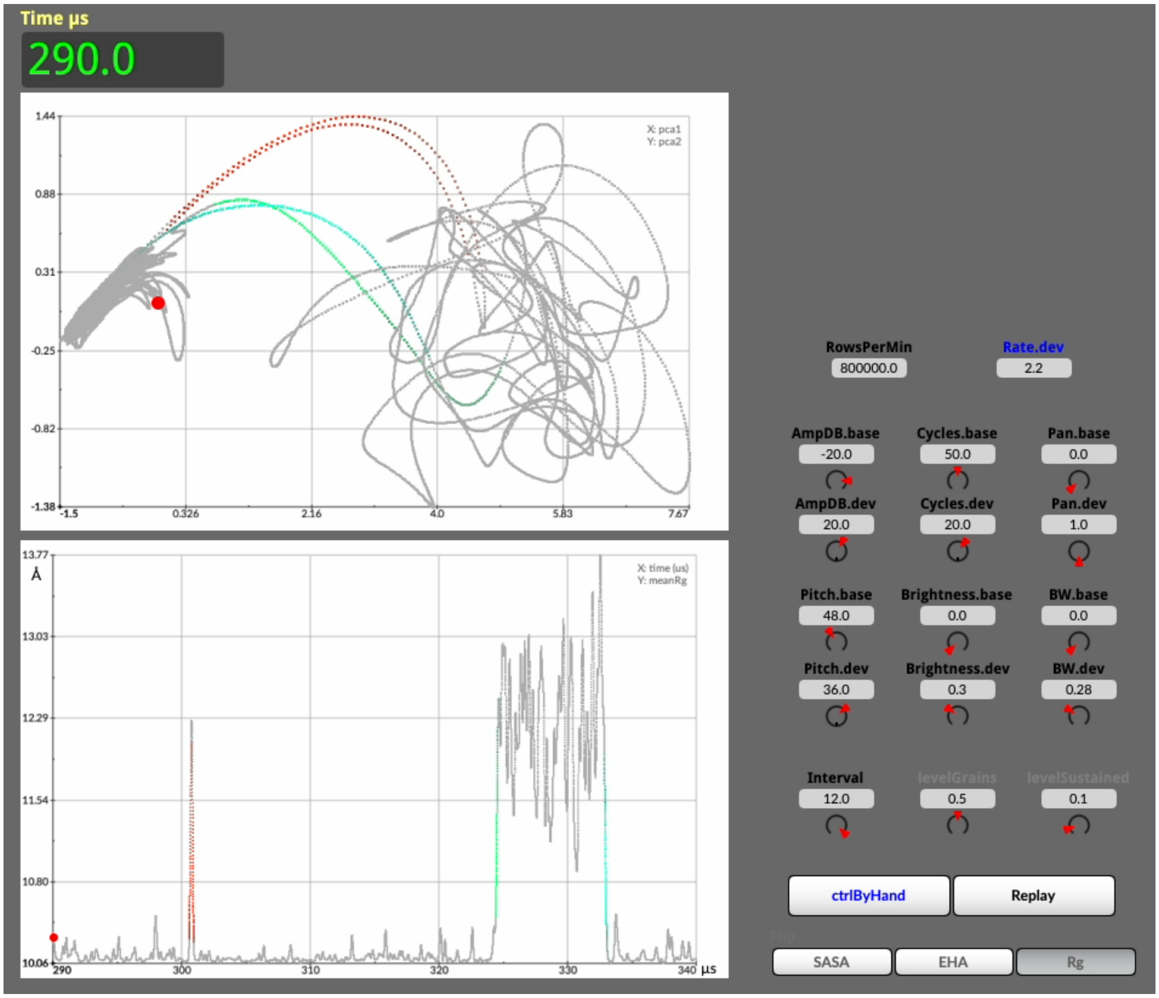
The full PC1 trajectory spends roughly 70% of the time in the unambiguously folded ensemble of states and about 10% of the time in the unambiguously unfolded states (in agreement with previously published results).[27] By setting different thresholds on the Rarity function, we can choose whether to categorize the remaining 20% of the states as Folded, Transition, or Unfolded time intervals.

### 3.3.3 Sonification of state transition passages in μ±σ space

After devising the μ±σ and Rarity function in the Tool Building step of the Auditory Analytics cycle (Figure 1), it was time to create new sonifications based on the newly derived variables.

Since the focus of this study was dynamics, we re-introduced the element of time by generating interactive μ±σ animations where time is shown by the position of a moving red dot on the μ±σ curve.

We examined multiple transition passages using several parameter-mapping sonifications and animations. For example, in Video 5, PC1 and PC2 are mapped to multiple sound parameters including pitch, amplitude, bandwidth and stereo position. Sonic events are triggered whenever the trajectory crosses a grid in PC1 or PC2. Rarity is used to stretch time during the transitions. In the visualization, Rarity is mapped to color saturation and the transition duration is mapped to the hue of the color.




https://youtu.be/ZnFvSn-WemI

*Video 5. A sonification and visualization of state transition passages through μ±σ space (upper) and $R_g$ time series (lower). PC1 and PC2 are mapped to multiple sound parameters including pitch, amplitude, bandwidth and stereo position. Sonic events are triggered whenever the trajectory crosses a grid in PC1 or PC2. Rarity is used to stretch time during the transitions. In the visualization, Rarity is mapped to color saturation and the duration of each transition is mapped to the hue of the color.*

During analytical listening to sonifications and animations of the time evolution of the protein folding reaction through μ±σ space, we observed (at least) two different types of transitions: 'Highway' — faster more direct pathways that cross between folded and unfolded states at higher PC2 values (e.g., orange transitions near time ~7 s in Video 5); and 'Meander' — slower, more meandering pathways that cross between folded and unfolded states at lower PC2 values (e.g., green transitions near times ~28 s and ~1 m 12 s in Video 5). Thus, while the μ coordinate reflects the distance between the folded and unfolded ensembles, the σ coordinate reflects heterogeneity of the transition state ensemble.

This led to speculation as to the cause of the heterogeneity and to a new hypothesis: that the different transition state passage times may be associated with different hydrogen bonding patterns.

## 3.4 H-bond dynamics

To test this hypothesis, we returned to the Tool Building phase of the Auditory Analytics Cycle to generate a new data set from the MD simulation — this one tracking H-bond formation between amino acid side chains, backbone and water molecules located in or near a small section of the WW loops.

### 3.4.1 H-bond datasets

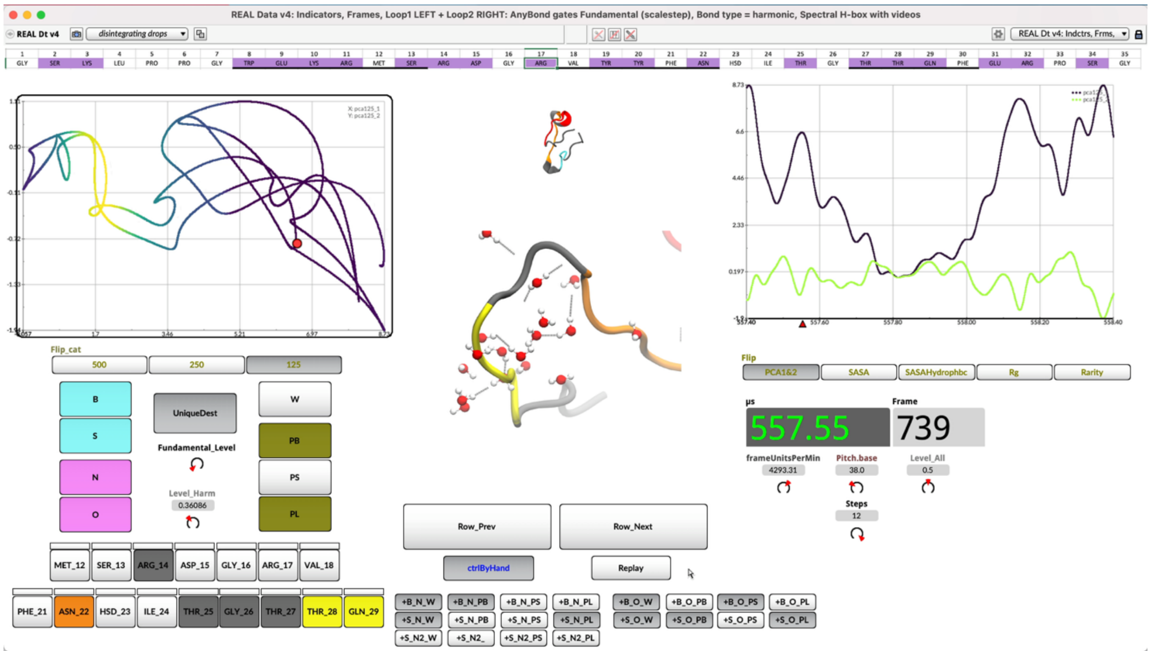
The hydrogen bonds selected for consideration were drawn from the regions surrounding loop 1 and loop 2 (residues M12 to V18 and F21 to Q29) based on prior findings that these regions are critical to the folding process, acting as early nucleation sites and orienting the two strands to form the beta-sheet.[24]

In the H-bond dataset, hydrogen bonds were represented discretely as either 0 (no bond) or 1 where a hydrogen bond was said to exist between three atoms (donor-hydrogen-acceptor) when the distance between the donor and acceptor in the MD simulation was less than 3 Å and the angle between the three atoms was greater than 140 degrees.

### 3.4.2 Visualization of H-bond data

We first mapped the H-bond data to an animated 3D visualization of the protein molecule, displaying the location of each hydrogen bond as a dashed line between the protein and a water molecule (depicted as red and white spheres).

In the visualization alone (Video 6, center), we were unable to perceive a connection between the H-bond dynamics, changes in $\mu \pm \sigma$, and the speed of the transition. It was difficult to visually track the appearance/disappearance of multiple hydrogen bonds, especially when the rotation of the molecule obscures sites of interest.




https://youtu.be/9D2ZPVB984s

*Video 6. Polyphonic Geiger Counter sonification and visualization of hydrogen bonds during state transition passages. Changes in the protein conformation (the large center animation) and the $\mu{\pm}\sigma$ trajectory (upper left) correspond to changes in the soundscape (indicating which hydrogen bonds are formed at each time step).*

To overcome some of the difficulties inherent in tracking multiple H-bonds by eye, we decided to take advantage of the auditory system's capacity for tracking multiple concurrent auditory streams, especially when each stream is situated in a unique frequency band and stereo source location.[8]

## 3.5 Polyphonic Geiger Counter sonification

To monitor the forming and breaking of various types of hydrogen bonds, we developed the Polyphonic Geiger Counter sonification in which each residue is mapped to a stereo location, each bond type is mapped to a distinct pitch, and each time a bond is formed in the data, a sound event is triggered: an amplitude-enveloped sine wave oscillator at the assigned pitch and stereo location. Similar to a classic Geiger counter, whose density of sound events corresponds to the relative probability density of beta particle decay in a given time interval, in the Polyphonic Geiger Counter, the density of sound events within a given time interval corresponds to the relative probability density of hydrogen bond formation in the corresponding MD simulation time interval.

Each residue number is assigned a pitch in half steps according to its distance from residue 12, starting from a base pitch at M12. The base pitch is adjustable during listening (in Video 6, it is 24 nn, three octaves below middle C or ~32.7 Hz).

A listener can adaptively control the rate and can shift the base pitch up or down to bring the bonds of interest into the range of highest acuity for that listener. This is analogous to interactively adjusting the color map of a visualization to bring out features by mapping them to colors to which the eye is most sensitive.

To further enhance the auditory event source separation, we used inter-aural intensity differences to position events associated with the seven residues identified as being in or near loop 1 in the left half of the stereo field from left (for M12) to center (V18), and we positioned the nine residues in or near loop 2 in the right half of the field from the center (F21) to right (Q29).

Using an interactive panel, a listener can select specific time intervals and combinations of bonds to listen for distinct changes in the soundscape as hydrogen bonds form or break during the folding or unfolding transition.

In Video 6, there is a clear temporal correspondence between changes in the 3D molecule conformation, the current position in $\mu\pm\sigma$ visual state space, and changes in the soundscape (which is based solely on the combination of hydrogen bonds formed at each time step).

### 3.5.1 Analytical listening

In Analytical Listening sessions, we could hear the soundscape changing as the protein traverses the $\mu\pm\sigma$ space; we could hear changes in the mix of hydrogen bonds when the hairpin forms or unravels; and we could hear a correlation between the speed of a transition and its characteristic hydrogen bonding patterns.

We split the 28 transitions among 8 members of our group (Gruebele, Scaletti, Hebel, Rickard, Danksagmüller, Taylor, Pogorelov, Russell) and asked each listener to label the transitions as Highway, Meander, or Ambiguous, and to take notes on what they heard before, during and after the transition.

When the responses were summarized, a correlation emerged between the total transition duration and the states of the H-bonds during Folded/Transition/Unfolded segments, confirming what we had heard during the group listening sessions (Supporting Information Table S3 of Ref. 1).

Motivated by these initial indications, we decided to introspect on how the listeners were summarizing the sonification so we could build a model to confirm the results and automate the process.

### 3.5.2 Modeling the human summarization and classification

In terms of the Auditory Analytics Cycle of Figure 1, we returned to the hypothesis-forming stage; however, this time the hypothesis was not about the underlying physical phenomenon, but about the process the human listeners used for assigning labels and summarizing the H-bond dynamics during the transitions.

In the H-bond data (derived from the MD simulation), hydrogen bonds can form, break, and re-form multiple times within each time segment. We noticed that the human listeners had unconsciously set a threshold for the proportion of time a bond was formed before they would judge it to be 'on' during that segment. This would be consistent with a model of continuous probability that an H-bond is formed during a time segment (rather than the discrete on/off representation in the data set).

When performing the transition labeling task, we noticed that the human listeners seemed to interpret smooth paths with large PC2 values through $\mu \pm \sigma$ space as 'Highway,' convoluted paths and low PC2 values as 'Meander,' and the remaining paths as 'Ambiguous.'

To automate this classification, we first applied k-means clustering to those same transition features in $\mu \pm \sigma$ space, revealing three distinct groups which, with one exception, aligned with clusters based on simple duration thresholds (where Highway transitions are the fastest, Meanders are slowest, and Ambiguous have intermediate durations). Furthermore, the clusters were highly consistent with subjective human labels (25/28 agreement); the exceptions were limited to the boundary between Highway and Ambiguous: three slow transitions that the algorithm grouped with the Highway cluster but that the humans labeled Ambiguous (see Fig. 2B of Ref. 1).

We then built a model using temporal coarse-graining (segmenting each transition into Folded, Transition, Unfolded time intervals) and thresholding to extract summaries directly from the H-bond data set, without mapping to sound first.

By comparing model-generated against human-generated summaries and varying the threshold parameter, we found the best match when the threshold was set to 0.13 (suggesting that the human listeners judged a bond to be formed if its value was 1 for > 13% of the time segment). In so doing, we may have stumbled upon the edge of a phenomenon often described in the literature: temporal integration and the statistical nature of auditory streaming.[33]

In the model-generated tables, the association between patterns of hydrogen bond formation and the duration of a transition state passage emerged even more strongly, with distinctive patterns associated with the four slowest Meander transitions as well as in several of the Ambiguous transitions (see SI Table S4 of Ref. 1).

### 3.5.3 The ears tell the eyes where to look

This is an example of data sonification leading from initial pattern discovery to the development of new analytical tools. In our Auditory Analytics framework, sonification is not an end-product but a catalyst for deeper inquiry.

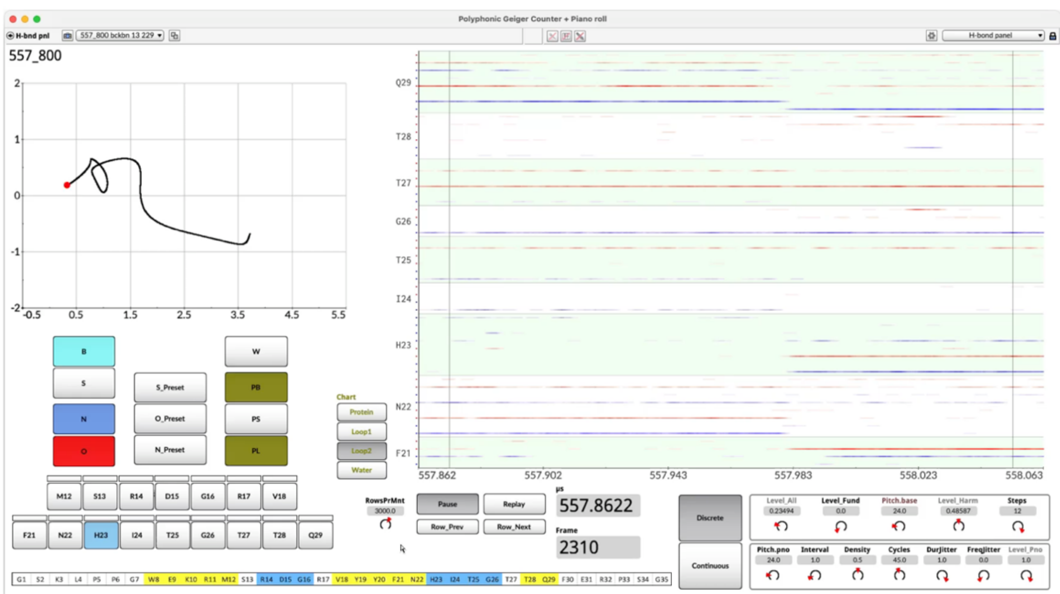
Significantly, it was only after hearing the patterns in the Polyphonic Geiger Counter sonification that we understood how to build a tool to extract those patterns directly from the data. This tool, in turn, generated new datasets amenable to traditional analysis, exemplifying how audition can help frame subsequent scientific questions. The ears tell the eyes where to look (and the mind decides where to go next).

## 3.6 Full circle

As a consequence of building the listening-summary model, we developed another visualization/sonification tool — the Piano Roll — which highlights an important aspect of hydrogen bond dynamics: co-operativity (i.e., groups of bonds that form or break simultaneously as the protein (un)folds).

In the Piano Roll sonification (Video 7), each of the tracked bond types and residues is associated with a sound generator with a unique pitch and, in the visualization, with the vertical position on the Y axis of a horizontal band of color (red for O, blue for N). The continuous probability that a hydrogen bond is formed at each time point is mapped to the amplitude of its associated sound generator and to the saturation value of its associated band of color. The result is a linear combination of sound generators with weights controlled by the time series data, and it functions as a time-varying sonic histogram.[34]

In the Piano Roll, the sound mapping leverages the auditory system's capacity for tracking multiple concurrent audio streams that have distinct frequencies or frequency ranges,[8] and the visual displays (the μ±σ portrait and the residue 'staff') provide an overview of the entire time evolution of the transition so the eye can scan ahead for upcoming events or look back at recently completed events.




https://youtu.be/XOFBGwwB6oU

*Video 7. An interactive control panel showing a single transition in μ±σ space (upper left) and on the H-bonds Piano Roll (right). A listener can switch between discrete (Polyphonic Geiger Counter) and continuous (Piano Roll) sound mappings and select different combinations of residues or bond types while monitoring the position of the protein in μ±σ space.*

With the Piano Roll, we had once again come full circle in the Auditory Analytics cycle: Sonification and Analytic Listening led to Hypothesis (how the listeners summarized what they heard) which drove Tool Building and led to an alternative Sonification and Visualization.

Using the Piano Roll we identified and analyzed differences in H-bond co-operativity patterns in Highways and Meanders and identified a candidate for a potential third class of transitions within the Ambiguous class. The Piano Roll made it easier to look and listen for correspondences between features in the $\mu \pm \sigma$ trajectory and the groups of H-bonds forming or breaking at specific time points in the transition state passage.

# 4 Discussion

Using Auditory Analytics, members of a multidisciplinary group comprising biophysicists, computational biologists, musicians, composers, sound designers, electrical engineers, and computer scientists were able to hear correlations between protein folding transition state passage duration and hydrogen-bonding patterns and to subsequently develop a quantitative analysis to confirm the correlation using more traditional analytical tools. We found that the transition state ensemble of WW domain is heterogeneous, with at least three categories of passages with distinct hydrogen bonding patterns – some that very efficiently form the folded state (or vice-versa), and some that require escape from incorrectly formed hydrogen bond patterns during that passage.

With respect to the role of hydrogen bond dynamics in the protein-folding reaction, the Auditory Analytics cycle suggests potential future directions including analyzing experimental and MD simulation data for additional proteins and polymers and creating a data-driven, explainable model linking hydrogen bonding heterogeneity with transition state passage durations.

As a tool for probing uncharted territories, Auditory Analytics has potential applications across a wide range of data-intensive fields, including computational physics and chemistry (materials science, quantum physics), life sciences[35] (synthetic biology, personalized medicine, biomolecular design, neural circuits, developmental dynamics), astrophysics, and signal processing (computational vision and audition, gravitational waves, image and signal processing).

Broader adoption of Auditory Analytics in the future could be bolstered by education — at the elementary level as a component of basic data literacy and at the post-secondary level in the form of training workshops for researchers — and infrastructure support — publishing sound and video as first-class figures in articles written for scientific journals and standardizing the availability of high-quality audio playback systems at conferences.[20]

# 5 Conclusion

Mathematician and computer scientist Richard W. Hamming wrote, "The purpose of computing is insight, not numbers."[36] Similarly, the purpose of Audio Analytics and, more generally, of Immersive Analytics, is not to generate animations, soundtracks, or virtual environments as ends in themselves, but rather to provide insight into some aspect of the system under study.

Recognizing that no single mapping can fully capture reality, Immersive Analytics offers a solution: an interactive cycle of multiple, multimodal mappings, analytical listening and interpretation, hypothesis formation, and validation through modeling.

Auditory Analytics, as a crucial component of Immersive Analytics, offers a powerful enhancement to scientific data analysis. By strategically mapping data to sound and engaging in iterative analytical listening and tool building, researchers can leverage the highly evolved capacities of the human auditory system to discover patterns and gain insights that may be difficult or impossible to discern through traditional methods alone. The protein folding case study is one example of how a multimodal approach can lead to novel findings and a deeper understanding of complex physical systems.

## Acknowledgements

The authors gratefully acknowledge the time granted to them to pursue this work by Symbolic Sound (for CS and KJH) and the University of Illinois (MG).

Thanks to Taras Pogorelov, Meredith Rickard, Premilla P. Samuel Russell, and Mayank Boob for generating trajectories, H-bond time series, and other data sets derived from the original MD simulation.[27] Thanks to composers Franz Danksamüller and Stephen A. Taylor for lending their ears during several analytical listening sessions.

## References


1. Scaletti C, Russell PPS, Hebel KJ, et al. Hydrogen bonding heterogeneity correlates with protein folding transition state passage time as revealed by data sonification. *Proc Natl Acad Sci*. 2024; **121**(22):e2319094121. doi:10.1073/pnas.2319094121.

2. Thomas JJ, Cook KA, eds. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. Los Alamitos, Calif: IEEE Computer Soc; 2005.

3. Keim DA, Mansmann F, Schneidewind J, Thomas J, Ziegler H. Visual Analytics: Scope and Challenges. In: Simoff SJ, Böhlen MH, Mazeika A, eds. *Visual Data Mining*. Vol 4404. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008:76-90. doi:10.1007/978-3-540-71080-6_6.

4. Preim B, Raidou R, Smit N, Lawonn K. An introduction to visual analytics. In: *Visualization, Visual Analytics and Virtual Reality in Medicine*. Elsevier; 2023:225-272. doi:10.1016/B978-0-12-822962-0.00018-3.

5. Klein K, Sedlmair M, Schreiber F. Immersive analytics: An overview. *It - Inf Technol*. 2022; **64**(4-5):155-168. doi:10.1515/itit-2022-0037.

6. Ghazanfar A, Schroeder C. Is neocortex essentially multisensory? *Trends Cogn Sci*. 2006; **10**(6):278-285. doi:10.1016/j.tics.2006.04.008.

7. Gerkema MP, Davies WIL, Foster RG, Menaker M, Hut RA. The nocturnal bottleneck and the evolution of activity patterns in mammals. *Proc R Soc B Biol Sci*. 2013; **280**(1765):20130508. doi:10.1098/rspb.2013.0508.

8. Bregman AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press; 1990. doi:10.7551/mitpress/1486.001.0001.

9. Kuhl PK. Early language acquisition: cracking the speech code. *Nat Rev Neurosci*. 2004; **5**(11):831-843. doi:10.1038/nrn1533.

10. Skoe E, Kraus N. Auditory Brain Stem Response to Complex Sounds: A Tutorial. *Ear Hear*. 2010; **31**(3):302. doi:10.1097/AUD.0b013e3181cdb272.

11. Grieser DL, Kuhl PK. Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Dev Psychol*. 1988; **24**(1):14-20. doi:10.1037/0012-1649.24.1.14.

12. Sell A, Bryant GA, Cosmides L, et al. Adaptations in humans for assessing physical strength from the voice. *Proc R Soc B Biol Sci*. 2010; **277**(1699):3509-3518. doi:10.1098/rspb.2010.0769.

13. Matalon N, Weinreb E, Freche D, et al. Structure in conversation: Evidence for the vocabulary, semantics, and syntax of prosody. *Proc Natl Acad Sci*. 2025; **122**(17):e2403262122. doi:10.1073/pnas.2403262122.

14. Pisanski K, Raine J, Reby D. Individual differences in human voice pitch are preserved from speech to screams, roars and pain cries. *R Soc Open Sci*. 2020; **7**(2):191642. doi:10.1098/rsos.191642.

15. Oganian Y, Bhaya-Grossman I, Johnson K, Chang EF. Vowel and formant representation in the human auditory speech cortex. *Neuron*. 2023; **111**(13):2105-2118.e4. doi:10.1016/j.neuron.2023.04.004.

16. Norman-Haignere SV, Long LK, Devinsky O, et al. Multiscale temporal integration organizes hierarchical computation in human auditory cortex. *Nat Hum Behav*. 2022; **6**(3):455-469. doi:10.1038/s41562-021-01261-y.

17. Scaletti C. Sound Synthesis Algorithms for Auditory Data Representations. In: Kramer G, ed. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Proceedings volume 18, Santa Fe Institute Studies in the Sciences of Complexity. Reading, Mass: Addison-Wesley; 1994:223-251.

18. Korzybski A, Strate L. *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*. Sixth edition. New York, New York, USA: Institute of General Semantics; 2023.

19. Latour B. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge (Mass.): Harvard University press; 1987.

20. *Sonification: A Tool for Research, Outreach and Inclusion in Space Sciences*. United Nations Office for Outer Space Affairs; 2023. https://www.unoosa.org/documents/pdf/Space4PersonswithDisabilites/UNOOSA_Special_Report_on_Sonification_2023.pdf. Accessed May 27, 2025.

21. Goldberg A. *Smalltalk-80: The Interactive Programming Environment*. Reading, Mass.: Addison-Wesley; 1984.

22. OpenSmalltalk. OpenSmalltalk. https://opensmalltalk.org. Accessed June 3, 2025.

23. Van Rossum G, Drake FL. *The Python Language Reference Manual: For Python Version 3.2*. Rev. and updated [ed.]. Bristol: Network Theory Ltd.; 2011.

24. Jäger M, Nguyen H, Crane JC, Kelly JW, Gruebele M. The folding mechanism of a β-sheet: the WW domain. *J Mol Biol*. 2001; **311**(2):373-393. doi:10.1006/jmbi.2001.4873.

25. McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. *Nature*. 1977; **267**(5612):585-590. doi:10.1038/267585a0.

26. Hollingsworth SA, Dror RO. Molecular Dynamics Simulation for All. *Neuron*. 2018; **99**(6):1129-1143. doi:10.1016/j.neuron.2018.08.011.

27. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How Fast-Folding Proteins Fold. *Science*. 2011; **334**(6055):517-520. doi:10.1126/science.1208351.

28. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015; **1-2**:19-25. doi:10.1016/j.softx.2015.06.001.

29. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph*. 1996; **14**(1):33-38. doi:10.1016/0263-7855(96)00018-5.

30. Phillips JC, Hardy DJ, Maia JDC, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys*. 2020; **153**(4):044130. doi:10.1063/5.0014475.

31. Audio file format. In: *Wikipedia*. ; 2025. https://en.wikipedia.org/w/index.php?title=Audio_file_format. Accessed May 26, 2025.

32. Nyquist–Shannon sampling theorem. In: *Wikipedia*. ; 2025. https://en.wikipedia.org/w/index.php?title=Nyquist%E2%80%93Shannon_sampling_theorem. Accessed June 4, 2025.

33. Barniv D, Nelken I. Auditory Streaming as an Online Classification Process with Evidence Accumulation. Snyder J, ed. *PLOS ONE*. 2015; **10**(12):e0144788. doi:10.1371/journal.pone.0144788.

34. Scaletti C, Craig AB. Using Sound to Extract Meaning from Complex Data. In: *Extracting Meaning from Complex Data: Processing, Display, Interaction II*. Vol 1459. SPIE; 1991:207-219. doi:10.1117/12.44397.

35. Braun R, Tfirn M, Ford RM. Listening to life: Sonification for enhancing discovery in biological research. *Biotechnol Bioeng*. 2024; **121**(10):3009-3019. doi:10.1002/bit.28729.

36. Hamming RW. *Numerical Methods for Scientists and Engineers*. 2nd ed. New York: Dover; 1986.