Spring 2021

# Protein Composition: Translating Amino Acid Sequences into Music

Melody Campbell
*Western Washington University*

# Protein Composition: Translating Amino Acid Sequences into Music

Melody G. Campbell*

***Department of Chemistry, Western Washington University***

**ABSTRACT:** The musical translation of proteins is an emerging field of research dedicated to increasing our current understanding of protein structure and disease states. Various methods have been developed to represent proteins in a way that is both musically pleasing and scientifically accurate. These endeavors have the incidental benefit of broadening the accessibility of science to a wider audience, including young children and those who are visually impaired. Multimodal learning has recently gained favor among academics and provides an immersive, interactive environment in which students can learn. Here, transmembrane channel-like protein 1 (TMC1) is translated into a musical composition utilizing a variation of a method previously published in 2007. TMC1 plays an integral role in the mammalian hearing process and is therefore a fitting choice for musical translation. Additional analysis of TMC1 structure using musical conversion may further advance our understanding of the role of TMC1 in the hearing process.

## INTRODUCTION

### Musical Translation of Proteins

Translating proteins into musical compositions is an exciting new area of research currently studied at the Massachusetts Institute of Technology[1,2]. This program has built a musical library of numerous different proteins for use in training an artificial intelligence (AI) program to recognize and understand protein structure[1]. Based on the observed parameters, the AI creates original musical arrangements that can be translated into novel, theoretically functional proteins. Such endeavors are aimed at better understanding protein folding and mutation for use in disease research. Not only is this approach a more intuitive way to think about protein structures, but it also has the added benefit of sharing science with a wider audience, specifically those without a scientific background and those who are visually impaired.

Several methods have emerged for the translation of proteins into music. In 2007, Takahashi and Miller assigned 7 distinct chords (combinations of three or more notes) to the 20 endogenous amino acids, pairing like amino acids by assigning different inversions of the same chord to the paired residues[2]. Yu *et al* utilized a different approach in 2019, using the vibrational frequencies of each amino acid as the basis for note assignment[1]. This method is currently employed in training the aforementioned AI software to recognize protein structural features. This study utilizes a modified approach based on the method pioneered by Takahashi and Miller to translate the human TMC1 protein into a musical composition.

### Auditory and Vestibular Sensation

Numerous genetic mutations are known to greatly affect auditory and vestibular senses, which in turn can drastically impact the quality of life of an affected individual. As a result, these senses are the subject of extensive research throughout the scientific community. The organs responsible for such sensations are housed in the mammalian inner ear, and their anatomy and mechanical function are well-understood. However, little is known about the exact biochemical mechanisms of these processes. Deciphering the precise pathways involved in the sensation of sound and bodily acceleration is instrumental in the treatment (and eventual eradication) of auditory and vestibular dysfunction.

The human ear is divided into three sections: the outer ear, the middle ear, and the inner ear. Sound is funneled through the external auditory canal of the outer ear and vibrates the tympanic membrane, which in turn transmits the vibrations through the three small bones of the middle ear to the cochlea[3]. The inner ear is comprised of three continuous organs, each responsible for sensing a specific phenomenon: the cochlea, which senses sound; the vestibule, which senses linear acceleration; and the semicircular canals, which sense angular acceleration[3]. Sound and acceleration sensation is reliant upon mechanotransduction: the translation of mechanical stimuli to electrical impulses. This is accomplished within specialized hair cells within the inner ear. Hair cells consist of bundles of stereocilia, which are arranged in three to six rows in a stepwise fashion and linked together at the tips[4]. The stereocilia bend in response to mechanical stimulation via pressure waves (sound) or the movement of surrounding fluid (acceleration)[4]. The bending of these cells causes ion channels within the stereocilia to open and depolarize the membrane, sending an action potential along the vestibulocochlear nerve to the brain, where perception occurs (**Figure 1**)[3]. Recent studies have identified the protein complex responsible for mechanotransduction, as well as a key player in that complex: transmembrane channel-like protein 1 (TMC1).

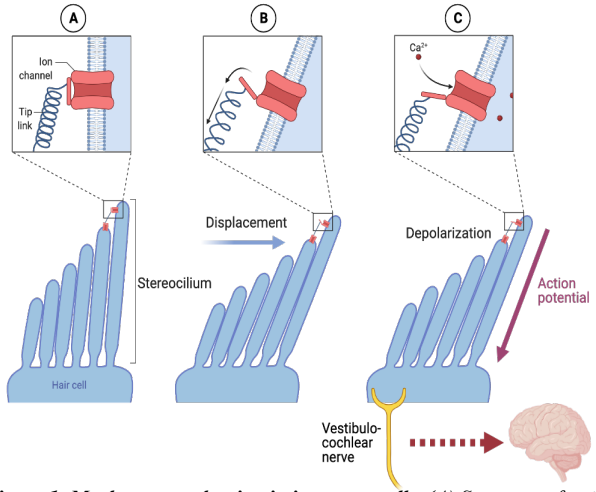



***Figure 1. Mechanotransduction in inner ear cells.*** ***(A)*** *Structure of a single hair cell.* ***(B)*** *Mechanical deflection of stereocilia opens a gated ion channel,* ***(C)*** *allowing for the generation and transduction of an electrical signal. Created using BioRender*.*

### Transmembrane Channel-like Protein 1

Transmembrane channel-like protein 1 (TMC1) is an essential protein for mammalian hearing, demonstrated by the complete hearing loss of *Tmc1* knockout mice[5]. TMC1 forms the pore within the mechanotransduction complex through which calcium ions travel[6]. It is estimated that TMC1 mutations are responsible for 4-8% of genetic hearing loss[7]. There are at least 40 mutations of TMC1 that cause non-syndromic deafness (i.e., hearing loss unaccompanied by other symptoms such as impaired balance)[7]. While some of these mutations are nonsense mutations resulting in a truncated protein, 75% of TMC1 mutations resulting in non-syndromic hearing loss are simple, nonconservative missense mutations[8]. As demonstrated by TMC1, even a miniscule alteration in protein structure can lead to drastic changes on the organismal level, similar to how one wrong note in a song can modify the character of an entire composition.

In this study, the structure of TMC1 will be portrayed as a musical composition. The resulting melody will represent the DNA sequence, amino acid sequence, and overall membrane topology of the protein. Two known mutations leading to non-syndromic deafness will be analyzed through this method.

## RESULTS

### Generation of Musical Translation Method

In order to translate a protein into music, there must be a clear method for assigning musical elements to proteinaceous components. The musical properties examined in this study include notes, rhythm, and dynamics. Each of these factors are integral to creating a composition resembling a plausible musical arrangement rather than a collection of cacophonous sounds. This endeavor is not wholly focused on the musical aspect, however. The protein itself is the model for the piece, and any translation method implemented must provide an accurate representation of the protein. Therefore, any musical elements assigned must have origins in the protein sequence.

The first feature to address is the musical notes that will be portrayed in the piece. The primary sequence of TMC1 serves as the basis for note assignments. Since there are 20 endogenous amino acids and only 13 unique notes in any given octave, combinations of notes (i.e., chords) had to be explored. Within a musical scale, there are seven scale degrees. These scale degrees serve as the root note for 7 basic triad chords. These triad chords can be expanded with the addition of different notes. For instance, a C major chord contains the notes C-E-G, and a C major 7th chord is comprised of the notes C-E-G-B. By including both triad and seventh chords, the number of possible chords that can be assigned to amino acids increases from 7 to 14. This still leaves an excess of 6 amino acids, however, requiring the assignment of suspended chords to fill the void. Since triad and seventh chords differ from one another by a single note within each scale degree, amino acids were paired such that those with similar structural properties were assigned to chords containing shared notes. For example, the constitutional isomers Leucine and Isoleucine were assigned the same base chord (C major), with one amino acid retaining the base assignment (Leu = C) and the other receiving the corresponding seventh chord (Ile = Cmaj7). The following amino acids were paired according to structural similarity: Leucine-Isoleucine, Alanine-Glycine, Serine-Threonine, Asparagine-Glutamine, Aspartate-Glutamate, Methionine-Cysteine, Arginine-Lysine-Histidine, and Phenylalanine-Tyrosine-Tryptophan. Both Valine and Proline did not receive structural partners in the above scheme. Chords were assigned to amino acids in order of amino acid frequency within the protein, such that amino acids that appear more frequently were assigned more common chords (**Figure 2**). See *Methods* section for further detail on chord assignment.

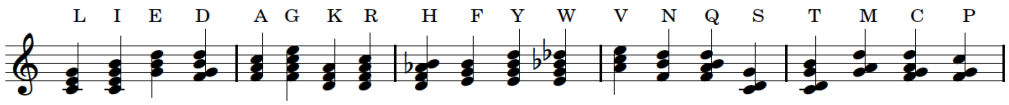


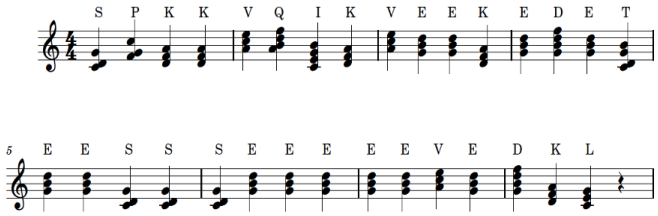

***Figure 2. Musical notation of the 20 different amino acids.*** *Leu=C, Ile=Cmaj7, Glu=G, Asp=G7, Ala=F, Gly=Fmaj7, Lys=Dm, Arg=Dm7, His=Ddim7, Phe=Em, Tyr=Em7, Trp=Edim7, Val=Am, Asn=Bdim, Gln=Bm7b5, Ser=Csus2, Thr=Csus2+7, Met=Gsus2, Cys=Gsus2+7, Pro=Fsus2. Residues labeled by single letter abbreviations.*

A collection of chords in a specific order does not constitute a musical composition; rhythm must be incorporated in order to impart greater musicality upon the arrangement. The DNA sequence that codes for TMC1 was used as the basis for note duration assignments. Codons that occurred most frequently were assigned the shortest note durations (eighth notes), whereas codons occurring least frequently were assigned the longest note durations (whole notes). As codon frequency decreased, note length increased from eighth notes to whole notes. Ranges of codon frequency were chosen such that the number of residues represented by each note assignment increases as frequency increases (**Table 1**). Utilization of the DNA sequence to assign rhythmic elements allows for greater variety within the musical piece (**Figure 3**), and also allows certain chords to be represented at different note lengths. For example, Cysteine is coded for by two different codons that occur at different frequencies. One codon occurs infrequently and therefore is represented by a whole note, whereas the second codon occurs more frequently and is represented by a half note (**Table S1**). See *Methods* section for further detail.

**Table 1.** Note duration based on codon frequency.

| **Frequency /1000** | **[0-9)** | **[9-15)** | **[15-25)** | **25+** |
|---|---|---|---|---|
| **Note Assignment** | Whole | Half | Quarter | Eighth |
| **Occurrence /760** | 38 | 125 | 256 | 340 |

**Arrangement 1**




**Arrangement 2**

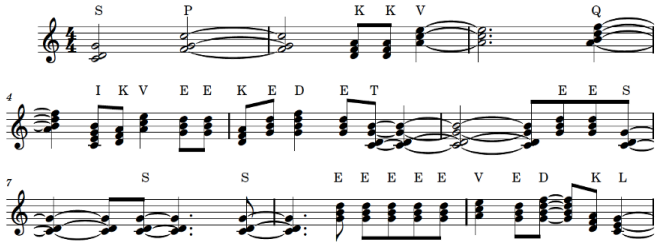



***Figure 3. Comparison between first and second arrangements.*** *Arrangement 1 assigns chords based on amino acid sequence but does not account for rhythm. Arrangement 2 utilizes the same note assignments, but also dictates rhythm based on codon frequencies.*

Finally, in order to produce a realistic musical composition, dynamics had to be introduced. Membrane topology served as the basis for dynamics assignment. As a transmembrane protein, TMC1 alternates between three domains: cytoplasmic, transmembrane, and extracellular. Cytoplasmic domains were assigned a *forte* dynamic (loud), transmembrane domains were assigned a *mezzo forte* dynamic (moderately loud), and extracellular domains were

assigned a *piano* dynamic (quiet, **Table 2**). See *Methods* section for further details.

**Table 2.** Membrane topology dynamics assignment.

| **Domain** | Cytoplasmic | Transmembrane | Extracellular |
|---|---|---|---|
| **Dynamics** | *Forte* | *Mezzo forte* | *Piano* |

The final composition pairs all three of these elements together to produce a musical arrangement both representative of TMC1 and musically pleasing to the ear (**Figure S1)**. The composition of the protein is identifiable throughout the musical sequence. For instance, transmembrane segments of the protein are represented in *mezzo forte* dynamics with a higher proportion of chords indicating hydrophobic amino acids; C-terminal polyadenylation is easily recognizable by the repetition of a single chord at the finale of the piece. Close analysis of the resulting musical composition provides unique insight into the structure of TMC1.

### Musical Representation of Protein Mutations

As previously mentioned, there are over 40 known mutations in TMC1 that cause genetic hearing loss in humans. Most of these mutations are missense mutations, though there are several known nonsense mutations as well as deletions. These mutations and their effects on TMC1 can be modeled with the same musical translation method utilized above. This study examined two known missense mutations: M654V and D572N.

M654V is an autosomal recessive mutation which causes congenital deafness (DFNB7). At the 654$^{th}$ position in the protein, a Valine residue is substituted for Methionine. This mutation is located at the DNA position 1960, in which an Adenine in the codon ATG is mutated to a Guanine, forming the codon GTG. According to the chord assignments detailed in the *Methods* section, methionine is represented by a G suspended 2$^{nd}$ (Gsus2) chord, whereas Valine is represented by A minor (Am). Not only does the amino acid change, but the frequency of the codon is also altered. ATG has a frequency of 43.36, whereas GTG has a frequency of 17.08 (**Table S1**). Therefore, the Methionine coded for by ATG was originally represented as a Gsus2 eighth note; however, the substituted Valine coded for by GTG is now represented by an Am quarter note (**Figure 4**). As a result, the cadence of this section of the protein is altered, and the rhythm of the remaining composition is permanently distorted. This mirrors the structural changes within TMC1 as a result of the M654V mutation, which renders the protein nonfunctional. Representing this mutation with music highlights the detrimental effects of even a simple missense mutation.

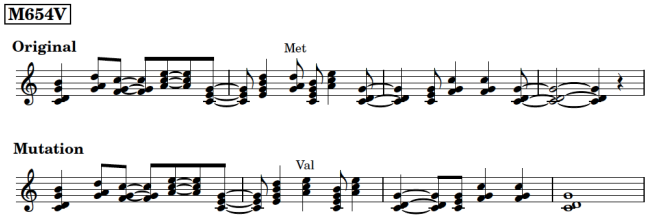



***Figure 4. Comparison between original musical sequence and M654V musical sequence****. Musical translation as in Figure 2. The chord representing the mutated residue is labeled with the corresponding three letter abbreviation.*

Even seemingly innocuous mutations such as an acidic residue to its related amide are injurious. D572N is an autosomal dominant mutation that causes progressive hearing loss (DFNA36). This mutation of Aspartic Acid to Asparagine results from a DNA mutation at the 1714$^{th}$ nucleotide, changing a GAC codon to an AAC codon. Because these amino acids are similar in structure, the differences in their chord structures are minimal, varying by the removal of a single note. Aspartic Acid is represented by a G dominant 7$^{th}$ (G7) chord (G-B-D-F), whereas Asparagine is represented by a B diminished (Bdim) chord (B-D-F). Additionally, the GAC and AAC codons occur with a similar frequency (17.08 and 22.34, respectively), thus the note duration coded for does not change. Therefore, the Aspartic Acid coded for by GAC was originally represented as a G7 quarter note; however, the mutated Asparagine coded for by AAC is now represented by a Bdim quarter note (**Figure 5**). The loss of a single note in the musical translation demonstrates how such minor changes in DNA and protein sequence can lead to drastic effects on the organismal level.

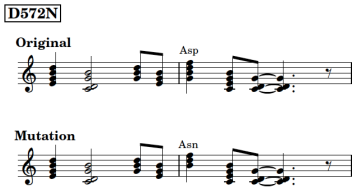



***Figure 5. Comparison between original musical sequence and D572N musical sequence****. The chord representing the mutated amino acid is labeled with the corresponding amino acid.*

### TMC1 Topology Resembles Classical *Rondo* Form

Further analysis of the membrane topology of TMC1 demonstrates that the pattern of domains follows a similar pattern to the classical music form, *rondo*. The *rondo* form follows a strict arrangement of motifs, often denoted as "ABACA." The "A" section, designated the *refrain*, is the principal theme of the song and is repeated between each alternating section, or *episodes*[9]. TMC1 follows an altered version of this pattern, with the intervening transmembrane domain resembling the musical refrain, and the extracellular and cytoplasmic domains representing the "B" and "C" episodes, respectively. The comparison diverges upon deeper analysis, as each episode within the *rondo* form is played in a different key depending on when in the sequence it is performed. While this key change was not investigated here, it would be an interesting endeavor to attempt to create a nature-based musical composition with a greater resemblance to classical music.

### Further Endeavors

Musical translation of proteins offers a unique learning experience, especially for young children, those who are visually impaired, and individuals previously unexposed to protein science. While efforts are currently concentrated on translations resembling classical music, translation methods resembling contemporary works could engage students lacking training in classical music.

In addition to its pedagogical value, musical translation of proteins can also further our understanding of biological processes. The exact role that TMC1 plays in the mechanotransduction complex and overall audition process is not yet fully defined. While numerous labs are currently investigating biochemical methods for deciphering the functions of TMC1, musical conversion may offer an alternative view. Close analysis of the resulting musical composition by both scientists and AI programs provides the potential for insight into the protein structure that may go unnoticed under traditional sequence analysis.

This simple translation of TMC1 is merely a first step, however. Additional analysis of the musical components of this composition, as well as application of varying translation techniques could further our understanding of TMC1 and its role in the hearing process. Furthermore, proteins do not act in isolation; rather, their function is context specific. Coding biological context into musical translations could provide insight into protein functions. Such analyses may aid in clarifying the characteristics of TMC1 important to the biochemical mechanisms underlying mechanotransduction, which could allow for improved therapeutics targeting hearing loss.

## METHODS

### Amino Acid Frequency Dictates Chord Assignments

Each amino acid was assigned a combination of 3-4 individual notes to form a chord within the musical key of C major (**Table 3**). Chord assignment was determined by amino acid frequency within TMC1. Residues with the highest frequencies were assigned the most "common" chords in the chosen key. The three chords used

most often in musical compositions are the I (*root* or *tonic*), V (*dominant*), and IV (*subdominant*) triads. In the key of C, these translate to C major, G major, and F major, respectively. These three triads were assigned to the three most abundant residues in TMC1: Leucine (L), Glutamic acid (E), and Alanine (A). For each of these residues, a structurally similar amino acid was assigned a variation of the given chord. Isoleucine, structurally similar to Leucine, was assigned a C major 7th (Cmaj7) chord, differing from the chord assigned to Leucine by the addition of a single note. A similar process was applied to Glutamic Acid (G major), paired with Aspartic Acid (G dominant 7th, G7), as well as Alanine (F major), paired with Glycine (F major 7th, Fmaj7).

Following assignment of the three most common chords, subsequent chord assignment occurred in ascending order: ii triad (*supertonic*, D minor, Dm), iii triad (*mediant*, E minor, Em), vi triad (*submediant*, A minor, Am), and vii° triad (*leading tone*, B diminished, Bdim). The amino acids assigned to these chords were determined in a descending order of frequency. Similar to above, residues with similar structures or properties were assigned chord variations differing by a single note.

The fourth most frequent amino acid, Lysine, was assigned the chord D minor (ii triad). Lysine shares similarities with both Arginine and Histidine, all of which are positively charged. Of these three residues, Arginine had the second highest frequency after Lysine, and was therefore assigned the chord D minor 7th (Dm7). Histidine required an additional ii chord assignment and was therefore assigned D diminished 7th (Ddim7), differing from Dm7 by a flat fifth note. Assignments for the next most frequent amino acid, Phenylalanine, and its related residues Tyrosine and Tryptophan were assigned to the iii chord in a similar fashion. The following is the result: Phenylalanine (E minor, Em), Tyrosine (E minor 7th, Em7), and Tryptophan (E diminished 7th, Edim7).

Of the remaining amino acids, the next highest frequency belonged to Valine, which was assigned the vi chord (Am). Valine was not paired with any other amino acid, as its somewhat similar residues, Leucine and Isoleucine, were already paired and shared a higher degree of similarity with each other than with Valine.

The next highest frequency belonged to Glutamine, which was assigned to the vii° chord (B diminished, Bdim). Its close relation, Asparagine, was assigned to a similar chord, B minor 7th flat 5 (Bm7b5). The B diminished 7th (Bdim7) chord was not utilized, as this is merely a different inversion of Ddim7, which had previously been assigned.

With five remaining residues to assign, chord choices had to shift from the exhausted major, minor, and diminished chords to the suspended second chords. The order of suspended chords assigned follows that of the major chords assigned: I, V, and IV. Therefore, the next most abundant amino acid, Serine, was assigned a C suspended 2nd (Csus2) chord. Its paired amino acid, Threonine, was assigned the chord C suspended 2nd add 7 (Csus2+7). Subsequently, Methionine was assigned G suspended 2nd (Gsus2); and its related residue, Cysteine, was assigned G suspended 2nd add 7 (Gsus2+7). Finally, Proline was assigned to F suspended 2nd (Fsus2).

**Table 3**. Chord assignments based on amino acid frequency.

| Amino acid | Frequency | Chord degree | In C major |
|---|---|---|---|
| Leucine | 10.4% | I | C |
| Isoleucine | 5.3% | I maj7 | Cmaj7 |
| Glutamic Acid | 8.9% | V | G |
| Aspartic Acid | 3.9% | V 7 | G7 |
| Alanine | 7.5% | IV | F |
| Glycine | 5.7% | IV maj7 | Fmaj7 |
| Lysine | 7.5% | ii | Dm |
| Arginine | 5.0% | ii m7 | Dm7 |
| Histidine | 0.7% | ii dim7 | Ddim7 |
| Phenylalanine | 7.0% | iii | Em |
| Tyrosine | 3.4% | iii m7 | Em7 |
| Tryptophan | 2.6% | iii dim7 | Edim7 |
| Valine | 5.8% | vi | Am |
| Asparagine | 5.3% | vii° | Bdim |
| Glutamine | 2.0% | vii° m7b5 | Bm7b5 |
| Serine | 5.1% | I sus2 | Csus2 |
| Threonine | 3.7% | I sus2+7 | Csus2+7 |
| Methionine | 4.3% | V sus2 | Gsus2 |
| Cysteine | 1.7% | V sus2+7 | Gsus2+7 |
| Proline | 4.2% | IV sus2 | Fsus2 |

This method of chord assignment ensures that the resulting musical composition exhibits common chords more frequently than less common chords. As a result, the composition resembles music more so than if uncommon chords were more prevalent. It also allows for amino acids of similar properties to sound similar and therefore be recognizable within the sequence.

### Codon Frequency Determines Note Duration

Codon frequency was used to assign note durations. The frequency of all codons within the *Tmc1* was calculated out of 1000, using the JavaScript program Sequence Manipulation Suite[10]. A frequency range of [0-9) was assigned to a whole note. Frequencies of [9-15) were assigned to half notes. Those with frequencies of [15-25) were assigned to quarter notes. Finally, frequencies of $\geq 25$ were assigned eighth notes. Assignments were generated such that a greater number of amino acids would be assigned to shorter note lengths, such that the resulting musical composition could be kept to a reasonable length.

### Membrane Topology Directs Musical Dynamics

Membrane topology was utilized to determine changes in dynamics throughout the piece. TMC1 adheres to the following pattern of domains: cytoplasmic, transmembrane, extracellular, transmembrane, cytoplasmic, et cetera. Given that the transmembrane domains intervene between each change in domain, this topology was chosen to be depicted as *mezzo forte*, an intermediate dynamic. Cytoplasmic domains were assigned *forte* (loud), while extracellular domains were assigned *piano* (quiet). The resulting composition therefore oscillates between *forte* and *piano* with intermediate volumes in between.

### Musical Composition and Notation

Audio tracks were generated using GarageBand software[11]. Individual notes were entered manually into a software instrument MIDI track, made publicly available through Apple's royalty free licensing. No prerecorded loops or tracks were utilized. Instrumentation included a pipe organ MIDI track, a grand piano MIDI track, and a basic drum track to act as a metronome to keep time. Audio mixing was performed in Logic Pro X[12]. Resulting .mp3 files are included in supplementary materials. Finally, representative sheet music of the composition was generated using MuseScore3 software[13]. MIDI files from GarageBand were imported and manually edited for enhanced readability. Each domain begins at the start of a measure; therefore, transitions between domains often include time signature changes to accommodate.

## SUPPLEMENTARY MATERIAL

**Table S1.** Summary of codon note duration assignments.

The following materials are available on Microsoft OneDrive:

**Capstone_Audio.mp3**
Audio recording of the musical translation of TMC1 (mp3)

**Capstone_Presentation.mp4**
Video of Senior Symposium Presentation (mp4)

**Capstone_Presentation.pdf**
Presentation slides used for Senior Symposium (PDF)

**Capstone_Score.pdf**
Sheet music representing the musical translation of TMC1 (PDF)


## ACKNOWLEDGMENTS

*Figure 1 adapted from "Mechanoelectrical Transduction in Hair Cells," by BioRender.com (2021).

I would like to thank my advisor, Dr. Nick Galati, Assistant Professor of Biology, for all his help throughout this process, as well as my father, Scott Campbell, for his assistance with audio mixing.

## REFERENCES


1. Yu, C.-H., Qin, Z., Martin-Martinez, F. J. & Buehler, M. J. A Self-Consistent Sonification Method to Translate Amino Acid Sequences into Musical Compositions and Application in Protein Design Using Artificial Intelligence. *ACS Nano* **13,** 7471–7482 (2019).
2. Takahashi, R. & Miller, J. H. Conversion of amino-acid sequence in proteins to classical music: search for auditory patterns. *Genome Biol.* **8,** 405 (2007).
3. Marieb, E. N., Wilhelm, P. B. & Mallatt, J. in *Human Anatomy* (ed. Beauparlant, S.) 542–551 (Pearson, 2017).
4. McPherson, D. R. Sensory hair cells: an introduction to structure and physiology. *Integr. Comp. Biol.* **58,** 282–300 (2018).
5. Nakanishi, H. *et al.* Tmc2 expression partially restores auditory function in a mouse model of DFNB7/B11 deafness caused by loss of Tmc1 function. *Sci. Rep.* **8,** 12125 (2018).
6. Pan, B. *et al.* TMC1 forms the pore of mechanosensory transduction channels in vertebrate inner ear hair cells. *Neuron* **99,** 736–753.e6 (2018).
7. Askew, C. *et al.* Tmc gene therapy restores auditory function in deaf mice. *Sci. Transl. Med.* **7,** 295ra108 (2015).
8. Ballesteros, A., Fenollar-Ferrer, C. & Swartz, K. J. Structural relationship between the putative hair cell mechanotransduction channel TMC1 and TMEM16 proteins. *Elife* **7,** (2018).
9. Ewen, D. in *The Home Book of Musical Knowledge* (Prentice-Hall Inc., 1954).
10. Stothard, P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* **28,** 1102, 1104 (2000).
11. *GarageBand*. (Apple, 2004).
12. *Logic Pro*. (Apple, 1993).
13. Schweer, W. *MuseScore3*. (MuseScore Developer community, 2011).


**Table S1.** Summary of codon note duration assignments arranged by amino acid.

| Amino Acid | Codon | Frequency /1000 | Note Duration |
|---|---|---|---|
| A | GCG | 3.94 | Whole |
| | GCA | 30.22 | Eighth |
| | GCT | 26.28 | Eighth |
| | GCC | 14.45 | Half |
| C | TGT | 5.26 | Whole |
| | TGC | 11.83 | Half |
| D | GAT | 22.34 | Quarter |
| | GAC | 17.08 | Quarter |
| E | GAG | 35.48 | Eighth |
| | GAA | 53.88 | Eighth |
| F | TTT | 40.74 | Eighth |
| | TTC | 28.91 | Eighth |
| G | GGG | 10.51 | Half |
| | GGA | 21.02 | Quarter |
| | GGT | 9.20 | Half |
| | GGC | 15.77 | Quarter |
| H | CAT | 2.63 | Whole |
| | CAC | 3.94 | Whole |
| I | ATA | 1.31 | Whole |
| | ATT | 22.34 | Quarter |
| | ATC | 28.91 | Eighth |
| K | AAG | 31.54 | Eighth |
| | AAA | 43.36 | Eighth |
| L | TTG | 23.65 | Quarter |
| | TTA | 9.20 | Half |
| | CTG | 19.71 | Quarter |
| | CTA | 11.83 | Half |
| | CTT | 11.83 | Half |
| | CTC | 27.60 | Eighth |
| M | ATG | 43.36 | Eighth |
| N | AAT | 30.22 | Eighth |
| | AAC | 22.34 | Quarter |
| P | CCG | 1.31 | Whole |
| | CCA | 18.40 | Quarter |
| | CCT | 17.08 | Quarter |
| | CCC | 5.26 | Whole |
| Q | CAG | 10.51 | Half |
| | CAA | 9.20 | Half |
| R | AGG | 17.08 | Quarter |
| | AGA | 17.08 | Quarter |
| | CGG | 2.63 | Whole |
| | CGA | 9.20 | Half |
| | CGT | 1.31 | Whole |
| | CGC | 2.63 | Whole |
| S | AGT | 9.20 | Half |
| | AGC | 11.83 | Half |
| | TCG | 1.31 | Whole |
| | TCA | 10.51 | Half |
| | TCT | 5.26 | Whole |
| | TCC | 13.14 | Half |
| T | ACA | 17.08 | Quarter |
| | ACT | 6.57 | Whole |
| | ACC | 13.14 | Half |
| V | GTG | 17.08 | Quarter |
| | GTA | 6.57 | Whole |
| | GTT | 15.77 | Quarter |
| | GTC | 18.40 | Quarter |
| W | TGG | 26.28 | Eighth |
| Y | TAT | 15.77 | Quarter |
| | TAC | 18.40 | Quarter |