Heliyon 7 (2021) e07933

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Protein music of enhanced musicality by music style guided exploration of diverse amino acid properties

Nicole WanNi Tay [a,1], Fanxi Liu [a,1], Chaoxin Wang [b], Hui Zhang [c,***], Peng Zhang [d,**], Yu Zong Chen [d,e,*]

[a] Raffles Institution, 1 Raffles Institution Ln, 575954, Singapore
[b] Department of Computer Science, Kansas State University, Manhattan, KS, 66506, USA
[c] School of Arts, Minnan Normal University, Zhengzhou, 363000, China
[d] Bioinformatics and Drug Design Group, Department of Pharmacy, and Center for Computational Science and Engineering, National University of Singapore, 117543, Singapore
[e] Qian Xuesen Collaborative Research Center of Astrochemistry and Space Life Sciences, Institute of Drug Discovery Technology, Ningbo University, Ningbo, 315211, China

ARTICLE INFO

Keywords:
Protein music
Sequence-to-music mapping
Amino acid properties
Musicality

ABSTRACT

Inspired by the traceable analogies between protein sequences and music notes, protein music has been composed from amino acid sequences for popularizing science and sourcing melodies. Despite the continuous development of protein-to-music algorithms, the musicality of protein music lags far behind human music. Musicality may be enhanced by fine-tuned protein-to-music mapping to the features of a specific music style. We analyzed the features of a music style (Fantasy-Impromptu style), and used the quantized musical features to guide broad exploration of diverse amino acid properties (104 properties, sequence patterns and variations) for developing a novel protein-to-music algorithm of enhanced musicality. This algorithm was applied to 18 proteins of various biological functions. The derived music pieces consistently exhibited enhanced musicality with respect to existing protein music. Music style guided exploration of diverse amino acid properties enable protein music composition of enhanced musicality, which may be further developed and applied to a wider variety of music styles.

Key points

- Protein music has been composed from amino acid sequences for popularizing science, sourcing aesthetic melodies, recognizing and designing proteins, but its musicality has yet to rival human music
- Broader exploration of diverse amino acid properties can lead to a novel protein-to-music mapping algorithm of enhanced musicality
- For the first time, we used a specific music style to guide the exploration of diverse amino acid properties for protein music composition of enhanced musicality.

1. Introduction

Science and music are closely related. There have been quests for scientific explanations of music [1, 2]. Music typically contains identifiable acoustic features (accent, tempo, pitch range, etc.) in the context of the primary human behavior (love, healing, etc.), with the melodic and rhythmic bigrams largely falling into the power-law distributions [2]. The spectral and temporal organization of music is believed to arise from human biology [3] and to influence human physiology and psychology in complex ways [4]. DNA and protein sequences resemble the ordered sequences of music notes, which have inspired the development of DNA music [5] and protein music [6, 7, 8, 9, 10]. In particular, protein music has been composed for popularizing science [9], as sources of aesthetic melodies [9, 11], and for protein recognition [10] and design [12].

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: huizhangking@163.com (H. Zhang), pzhang@mail.rockefeller.edu (P. Zhang), chenyuzong@nbu.edu.cn (Y.Z. Chen).
[1] These authors contributed equally.

https://doi.org/10.1016/j.heliyon.2021.e07933
Received 22 March 2021; Received in revised form 19 June 2021; Accepted 2 September 2021

Protein sequences and music notes exhibit traceable analogies [6]. Music notes are cohesively assembled with characteristic recurring patterns [6], while protein sequences contain recurring secondary structures and tertiary folds guided by the principles of protein folding [13]. Driven by these analogies, various protein-to-music mapping algorithms have been constructed. Using amino acid frequency histograms and solubility, protein folding patterns have been translated into musical themes of short note ranges and fixed note lengths [7]. Multiple physicochemical properties of the amino acids have been synergistically mapped to the musical and instrument features for creating multi-instrument protein music [8]. The musicality of protein music has been enhanced by exploiting different note values (e.g. pitch, note length) and introducing chords [9]. Protein music of distinguished melodies has been derived by matching the contours of multiple amino acid properties to the musical motive of similar contours [10].

Despite these progresses, the musicality of protein music has yet to reach the level of human music. Music is of various styles with characteristic musical features [14]. Musicality may be enhanced by fine-tuned protein-to-music mapping guided by the features of specific music style. This mapping may be based on biologically-relevant amino acid properties, sequence patterns and variations, so as to robustly match but not overfit the musical features while maintaining biological implications. For instance, protein fold is recognizable by appropriate consideration of the amino acid physicochemical properties [15]. Protein remote homology is detectable by similarity analysis of the amino acid properties [16] and the recurring local sequence motifs (i.e. local amino acid patterns) [17]. Therefore, the rich amino acid properties and local sequence patterns provide diverse biologically-meaningful protein-to-music mappings, some of which likely lead to enhanced musicality. Moreover, single- and multi-amino acid insertions may lead to rewired regulations [18] and protein bindings [19]. Multi-location insertions occur in protein variable regions for novel structures [20] and loops for multi-domain protein evolutions [21]. Hence, introduction of amino acid insertions may facilitate the mapping to the difficult musical features. Some amino acids are of distinguished properties and function. For example, proline is an anomalous amino acid with its nitrogen atom covalently locked within a ring, and the proline-rich motifs are important for signaling and cognition [22]. The nitrogen atoms of arginine are modifiable for methylation, and arginine methylation epigenetically regulates gene transcription [23]. These distinguished amino acids may be employed for triggering certain sporadic music note variations (e.g. accidentals).

In this work, we analyzed the musical features of a specific music style, and used the quantized features of the music style to guide a broad exploration of diverse amino acid properties, sequence patterns and variations of biological relevances for a novel protein-to-music mapping algorithm. This algorithm was intended for producing protein music of enhanced musicality while maintaining biological implications. In principle, our method can be developed with respect to any music style. Here we focused on the *Fantasy-Impromptu* style, which is a style of solo piano compositions from the Romantic period [24]. Music of this style is generally characterized by chromaticism, chords, a wide keyboard range and elicits strong emotions [24]. A notable composition of this style is Chopin's *Fantaisie-Impromptu* in C-sharp minor, which is one of Chopin's most popular and most performed works, with multiple recordings online [25].

An analysis of *Fantaisie-Impromptu* [26] has revealed that it has a balanced structure of three sections. The first section presents a very strong characterized personality and forms a very important motif of the work on the left hand, where the whole left-hand part is constructed mainly on this motif structure. The second section displays a freer structure in D-flat major tonality, and the melody on the right hand has a slow and lyrical character. The third section begins exactly the same as the first section up to the hundred and eighteenth bar, followed by a C-sharp pedal ongoing for 9 bars. Such a balanced music structure makes it possible for mapping the structured amino acid properties along a protein sequence to the music of the *Fantasy-Impromptu* style. Nonetheless, generation of the pleasant-sounding protein music with respect to the *Fantasy-Impromptu* style is a challenging task used to test our new algorithm development method.

Our novel protein-to-music algorithm was based on the comprehensive analysis and quantization of the musical features of a specific style (e.g. the *Fantasy-Impromptu* style), and the quantitative profiling of 104 structural, physicochemical and binding properties of the amino acids of the proteins [27, 28, 29]. The quantitative profiles of the amino acid properties were screened against the quantized musical features of the music style for optimal protein-to-music matching. The exploration of 104 amino acid properties is far more than those explored by the existing algorithms, and the amino acid local patterns and insertions have not been explored for protein music before. These explorations significantly expand the protein-to-music matching capacity. Moreover, the clustering of the 104 amino acid properties provides the clustered alternative properties for fine-tuned protein-to-music mapping. The introduction of amino acid insertions enables the mapping to certain difficult musical features. These together facilitate more complex protein-to-music mapping for encompassing multiple musical features, music styles and synchronized play of complementary melody lines. The algorithm was applied on two groups of proteins for evaluating the musicality of the produced protein music. One group (11 proteins) is of biological functions conforming to the *Fantasy-Impromptu* themes, and another (7 proteins) is of vastly different functions (Table 1).

2. Methods

2.1. Music selection and quantization of musical features

Four music compositions (pieces) were selected to represent the *Fantasy-Impromptu* style. These are *Fantaisie-Impromptu in C-sharp minor Op* 66 by Frederic Chopin, *No. 2 from 7 Fantasia Op* 116 by Johannes Brahms, and *Wanderer Fantasy* and *Impromptu in C minor Op 90 no. 1 D8*99 by Franz Schubert [30]. These were quantized to extract the musical features of the *Fantasy-Impromptu* style by the following procedure: abc2midi program [31] was used for converting the MIDI files into the corresponding text files in a music notation known as abcnotation [32]. Six musical features, pitch, length, octaves, chords, dynamics and main theme, were quantitatively analyzed. In counting the percentage frequency of the 12 pitches in an octave, all pieces were transposed to the base key of *C major* and octaves were ignored in counting the percentage occurrence of each pitch. In counting the percentage occurrence of each note length, the tempo and number of beats of music per bar was disregarded and all note lengths were expressed as a multiple or fraction of the length of a crotchet. In counting the percentage of the notes that occur per octave, the middle C was taken as the fourth C from the left of the keyboard, the notes before the first C key are in the A0 octave, while the notes before the second C key are in the C1 octave, until the last note of the piano at C8. In counting the percentage occurrence of chords, the total number of chords in a piece was counted and the percentage frequency of occurrence of chords per piece was computed by dividing the number of chords against the total number of notes. In quantizing the dynamics, i.e. the loudness of notes, the frequency of changes in loudness was counted in two groups, a gradual change (crescendo or diminuendo) and a sudden change. Lastly, quantization was conducted on the first recurring melodic theme identified in each piece, to determine its characteristics for replicating such themes in the protein music.

2.2. Profiling of the amino acid properties

A total of 104 amino acid properties (Supplementary Table 1) were selected for quantitative and clustering analysis, which include 66 structural and physicochemical properties [27, 28] and 38 protein-binding

Table 1. Proteins selected for protein music generation.

Protein Name (Full Name, Protein Family)	Biological Function	Protein Uniprot ID	Reference
Protein with function that conforms to the Fantasy-Impromptu theme			
OXTR (Oxytocin receptor, G-protein coupled receptor 1 family)	A receptor of oxytocin, a hormone that improves social cognition and prosociality, which activates a G protein that regulate a phosphatidylinositol - calcium second messenger system.	P30559	[54]
CASA1 (Alpha-S1-casein, Alpha-casein family)	Important for calcium phosphate transport in the milk, reduces stress-related symptoms in females, sleep problems and general fatigue	P47710	[55]
NEGR1 (Neuronal growth regulator 1, Immunoglobulin superfamily)	Possibly acts as a trans-neural growth-promoting factor in regenerative axon sprouting in the mammalian brain, regulates mood	Q7Z3B1	[56]
TRPA1 (Transient receptor potential cation channel subfamily A member 1)	Involved in pain detection, possibly in cold and oxygen concentration perception, promotes anxiety and depression	O75762	[57]
botA (Botulinum neurotoxin type A, Peptidase M27 family)	Causes flaccid paralysis by inhibiting neurotransmitter acetylcholine release, slowed the processing of emotional language by inhibiting involuntary facial expression, cosmetic use affects processing of emotional language	P0DPI1	[58]
GH1 (Somatotropin, Somatotropin/prolactin family)	Stimulates IGF-1 secretion in the liver and other tissues, and amino acid uptake and protein synthesis in the muscle and other tissues, possibly involved in learning and the response to stressful experience.	P01241	[59]
Miraculin (Protease inhibitor I3 family)	Taste-modifying activity to convert sour stimuli to sweetness on the tongue	P13087	[60]
cocE (Cocaine esterase, CocE/NonD hydrolase family)	Hydrolyzes cocaine to benzoate and ecgonine methyl ester, as a protein-based therapy for cocaine overdose and addiction	Q9L9D7	[61]
GHR (Growth hormone receptor, Type I cytokine receptor family)	A receptor for pituitary gland growth hormone in regulating postnatal body growth, possibly involved in the memory and cognitive functions	P10912	[62]
FoxO6 (Forkhead box O6, Forkhead box family)	A transcriptional activator, possibly regulates memory consolidation and synaptic function	A8MYZ6	[63]
AQP1 (Aquaporin-1, MIP/aquaporin family)	Forms a water-specific channel that permits water to move in the direction of an osmotic gradient, provides an advantage during endurance exercise due to homeostatic mechanisms.	P29972	[64]
Protein with vastly different function			
SCMH1 (Polycomb protein SCMH1, SCM family)	Associates with polycomb group multiprotein complexes, the complex class is required to maintain the transcriptionally repressive state of some genes, regulates gene silencing for ordered development	Q96GD3	[65]
TP53 (Cellular tumor antigen p53, p53 family)	Involved in stress response and DNA protection, acts as tumor suppressor in many tumor types	P04637	[66]
Ricin (Ribosome-inactivating protein family)	One of the most poisonous naturally-occurring substances, acts as a glycosidase that removes a specific adenine residue from an exposed loop of the 28S rRNA leading to rRNA breakage.	P02879	[67]
rbcL (Ribulose bisphosphate carboxylase large chain)	Key enzyme in photosynthetic carbon fixation and photorespiration, catalyzes reactions of carboxylation of D-ribulose 1,5-bisphosphate, the primary event in carbon dioxide fixation, also catalyzes the oxidative fragmentation of the pentose substrate in the photorespiration process	P00875	[68]
GFP (Green fluorescent protein, GFP family)	Acts as an energy-transfer acceptor, involved in fluorescence emission	P42212	[69]
Ovalbumin (SERPINB14, Serpin family)	A non-inhibitory serpin, acts as a main protein in egg white	P01012	[70]
Concanavalin-A (Leguminous lectin family)	Plant lectin that specifically binds to certain structures in glycoproteins, also binds to mannose and metal ions.	P02866	[71]

properties [29, 33]. The selection of these 104 amino acid properties was based on the consideration that their unique property profiles may be mapped to certain musical features. For instance, the volume variations of amino acids of a protein sequence may be mapped to the variation of the pitches. The positions of the charged amino acids in a protein may trigger a sporadic change of musical feature. The clustered secondary structure patterns along a protein sequence may be linked to the octave patterns of a music piece. For comparative analysis of different properties, the values of the amino acid properties were normalized between -1 and 1 for those with both negative and positive values (e.g. hydrophilicity value), and between 0 and 1 for those with only positive or integer values (e.g. residue volume). Henceforth, all amino acid properties are in the normalized values. The quantitative profile of each amino acid property was given as the property values with respect to the 20 amino acids. The Spearman rank correlation [34] was used for measuring the similarity between the amino acid properties. An R package, heatmap (Stats, version 3.6.2) [35], was used for clustering the amino acid properties into groups of similar profiles, using the complete linkage.

2.3. The mapping of the amino acid properties to the quantized musical features

The general protein-to-music mapping strategy is as follows: First, for a specific musical feature (e.g. the relative frequency of occurrence of each pitch) of a music style, the variation of each property (e.g. composition) with respect to the 20 amino acids was compared with that of each musical feature with respect to the representative music pieces of the style. This was typically conducted by the comparison of the clustering patterns of the amino acid properties and those of the musical features. Secondly, the best matched amino acid property was selected for the subsequent generation of the protein-to-music mapping rule. Based on the matching between the best matched amino acid property and musical feature (e.g. amino acid composition and pitch), each of the 20 amino acids was assigned to a specific value of the musical feature (e.g. an amino acid alanine assigned to a pitch value of G).

Based on the clustering of the 104 amino acid properties, representative amino acid properties were selected from individual clusters. The quantitative profile of every representative amino acid property was screened against the quantized values of each of the six musical features (note pitches, lengths, octaves etc.) for finding the optimally mapped pairs of amino acid property and musical feature. The mapping was based on six considerations. First, the property values of individual amino acids or combined values of multiple amino acids were optimally matched with the quantized values of a musical feature (e.g. pitch). Second, the property values of the amino acids were mapped to the musical features based on the correlations with the physical characteristics (e.g. the larger sizes correspond to shorter note lengths, depicting the difficulty in moving larger objects). Third, individual amino acids of distinguished physicochemical characteristics and biological function were used for triggering sporadic note variations (e.g. accidentals). Fourth, neighboring amino acids were jointly considered for successively played musical features (e.g. successively played chords). Fifth, amino acid insertions were

introduced for the recurring music themes. Sixth, to avoid excessive simultaneous changes of several musical features at the same location, the mappings of the multiple musical features at an amino acid were conducted by joint consideration of the current, frontward, and next amino acid.

For each optimally mapped pair, the corresponding cluster of the representative amino acid property was screened for the property within the cluster that best matches the musical feature. Based on the mapping of the best matched amino acid property and quantized musical features, each of the 20 amino acids was assigned to a specific set of general musical features (i.e. a specific amino acid assigned to a particular pitch, note length etc.) and some selected amino acids were assigned to special musical features (e.g. chords). These were tabulated into musical feature mapping tables. For generating solo piano compositions played by both hands or multi-instrument music played by multiple instruments, the optimally-matched amino acid properties were substituted by the similar properties in the same cluster, leading to alternative musical feature mapping tables for different hands or instruments.

2.4. Protein selection and protein music generation

A total of 18 proteins were selected for music generation based on their biological functions (Table 1). One group of 11 proteins regulates human emotion, cognition, sensation or performance, which conforms to the usual themes of the *Fantasy-Impromptu* music. While they are involved in cognition-related physiological processes, these 11 proteins come from different protein families and are thus diverse in sequences and in sequence-derived series of amino acid properties. The second group of 7 proteins represents vastly different biological regulations (ordered development, disease, poison, photosynthesis, fluorescence, food protein, and glycoprotein binding) from different protein families. These 7 proteins are diverse both in functions and sequences. The sequences of the 18 proteins were obtained from the Uniprot database [36] (Supplementary Table 2). A Python code was used for converting each sequence into an abc file based on the musical feature matching tables. The abc file was then run through the abc2midi code [31] to generate the MIDI file. This file was then played in Musescore [37], which displays sheet music for manual analysis and troubleshooting along with audio analysis.

3. Results

3.1. The Fantasy-Impromptu musical features

The musical features of the four *Fantasy-Impromptu* pieces were analyzed based on the quantized feature values (Supplementary Table 3–7). Within one octave, there are 12 possible occurring notes. The most frequently occurring notes are C, G and E, at 16.1%, 15.7% and 10.0% frequency respectively (Figure 1). These three notes are part of the tonic chord, which is one of the most significant chords in most styles of tonal music. The note lengths were defined with respect to the crotchet note as one beat, which are either fractions or multiples of the crotchet note. The four pieces contain no note of length 1/16 or 1/8, the most frequently occurring note length is 1/4 at 51.3% frequency, and the longer note lengths showing clear trends of steeply decreasing frequency of occurrence (1/2, 1, 3/2, 2, 3, 4 at 24.2%, 10.4%, 9.0%, 3.9%, 0.8%, 0.3% respectively) (Supplementary Figure 1). The range of a piano spans from the lowest octave A0 to the highest octave C8. There is 0% note occurrence in octaves A0, C6, C7 and C8, and a clear trend of rapidly increasing frequency of note occurrences from C1 to C4 (2.2%–43.9%) followed by a steep drop to 5.5% frequency in C5 (Supplementary Figure 2).

A chord occurs when two, three or four notes are played simultaneously by the same hand. There is a large variation in the percentage occurrence of chords among the four music pieces (Supplementary Table 4), likely resulting from the varying preferences among the composers. The median percentage occurrence of the four music pieces, rounded off to 14%, was selected as the standard percentage occurrence for the protein music. The average percent distribution of the chords is 71.4%, 19.8%, and 8.8% for the 2-, 3-, and 4-note chords respectively (Supplementary Table 8–9). Dynamic markings at different points in the music piece define the volume of sound at the points. There are 8 dynamic markings, from ppp for the softest to fff for the loudest sound. The most commonly occurring dynamics are *pp, p, f,* and *ff* (Supplementary Figure 3). There is a wide variation in the percentage of gradual change in the loudness with respect to the total number of changes in the loudness. The median percentage of the change from soft to loud (50%) and from loud to soft (37.5%) was taken as the percentages of loudness changes. The four pieces have at least one melodic theme that recurs throughout each piece, either as an exact repetition of the theme, or as a rhythmic repetition where the sequence of the note lengths is preserved and the note pitches are varied.

3.2. The clustering of the amino acid properties

The 104 amino acid properties were clustered by using the heatmap package (Stats, version 3.6.2) [35], such that these properties were aligned into groups of similar distribution patterns with respect to the amino acids. Figure 2 shows part of the clusters. Noticeably, the amino acid properties of similar structural or physicochemical properties are clustered together. For instance, one cluster contains 8 properties, including propensity to be buried inside, mean fractional area loss, linker index, buriability, hydration free energy, unfolding Gibbs energy in water, normalized frequency of beta (i.e. β-Sheets), and average relative probability of beta. These properties are related to the protein burial phenomena. Specifically, β-sheets are parts of protein folding architecture that maximizes hydrogen bond formation, particularly within inner cores [38]. Linker amino acids are partially buried with an average normalized solvent accessibility of 26.7% [39]. Amino acid mean fractional area loss correlates with the degree of burial, the higher the burial the greater the area loss due to the partial shielding of the surrounding amino acids. Hydration is opposite to burial. Lower hydration corresponds to higher burial. The similar properties within individual clusters provide natural alternatives for optimal protein-to-music mapping and for music played by two-hands or multiple instruments.

3.3. The mapping of the amino acid properties to the Fantasy-Impromptu musical features

Based on the comparison of the quantitative profiles of the amino acid properties and the quantized *Fantasy-Impromptu* musical features, the following 11 amino acid properties were selected from the pool of 104 properties for generating a novel protein-to-music algorithm. These properties are the relative frequency of occurrence of the amino acid in proteins, amino acid composition, electron-ion interaction potential, net charge, bitterness, hydropathy, size, residue volume, rRNA binding propensity, mRNA binding propensity, and number of hydrogen bond donors. Based on these properties, a set of rules were devised for mapping the amino acids along a protein sequence to the musical features that lean towards the *Fantasy-Impromptu* style, which are described below.

3.4. Pitches and lengths

In order to conform to the solo piano playing style, both hands should not play in unison for the entire piece. Hence, two properties from the same cluster were used for synchronized mapping of the amino acids to the music pitches of the right- and the left-hand respectively. Specifically, the relative frequency of occurrence of amino acids and the amino acid composition was used for mapping the amino acids to the right- and left-hand pitches respectively (Table 2). Because these properties were derived from a large number of proteins, the mappings are generally applicable to diverse protein sequences. In mapping to the note length,

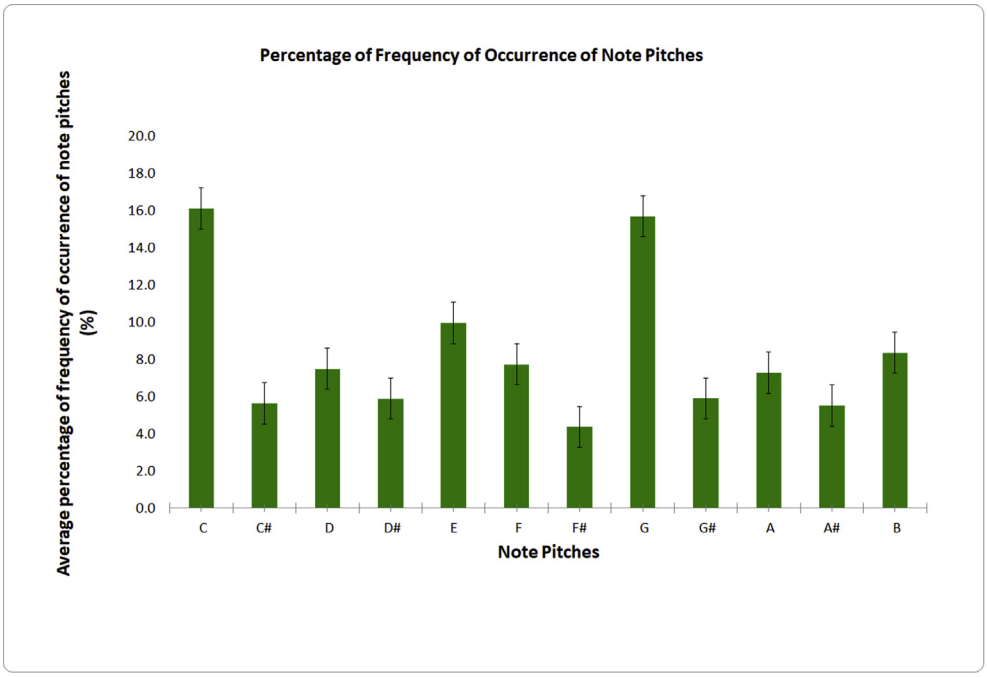

Figure 1. Percentage frequency of occurrence of note pitches in each octave of the four representative pieces (black keys denoted with "#").

there is a need to avoid a direct correlation with the changes of the pitches. Therefore, the mappings of the current amino acid to the note lengths were based on the corresponding properties of the frontward amino acid (i.e. the note length of the current amino acid was determined by the property of the frontward amino acid). The default length 1 was used for the first amino acid. Two properties from the same cluster were selected for the two hands. The amino acid size and residue volume were used for the right- and left-hand note lengths respectively (Table 2), while a larger size or volume corresponds to a shorter note length to depict the difficulty in moving larger objects.

3.5. Octaves

The octaves were commonly found to range from the first C note (C_1) to before the sixth C note (C_6). The right hand used mainly the C_3 to C_6 octaves while the left used C_1 to C_4. For mapping to the right-hand octaves, the 20 amino acids were first grouped based on their electron-ion interaction potential (Group 1: P, H, K, A, Y, W, Q, M, S, C, T, F, R, D; Group 2: L, I, N, G, V, E). Next, four trigger amino acids were selected based on the net charge property (arginine and lysine with +1 charge and aspartic acid and glutamic acid with -1 charge). To best fit the octave distribution of the *Fantasy-Impromptu* style, three rules were devised. Firstly, when encountering arginine or lysine, if the frontward amino acid is in Group 1, the octave of the notes of the subsequent amino acids is increased by one. Secondly, when encountering aspartic acid or glutamic acid, if the frontward amino acid is in Group 2, the octave of the notes of the subsequent amino acids is decreased by one. Thirdly, when encountering one of the four trigger amino acids while the notes of the subsequent amino acids are in an upper or lower capped octave, these notes are played one octave away from the previous one, simulating a bouncing off the wall effect.

For the left hand, the 20 amino acids were grouped based on their bitterness property (Group 1: A, R, N, D, C, Q, E, G, H, K, M, P, S, T; Group 2: I, L, F, W, Y, V). Four trigger amino acids were selected based on hydropathy (arginine and lysine with the most negative values, and isoleucine and valine with the most positive values). When encountering arginine or lysine, if the frontward amino acid is in Group 2, the octave of the notes of the subsequent amino acids is decreased by one. When encountering isoleucine or valine, if the frontward amino acid is in Group 1, the octave of the notes of the subsequent amino acids is increased by one. When encountering one of the four trigger amino acids while the notes of the subsequent amino acids are in an upper or lower capped octave, these notes are played one octave away from the cap.

3.6. Accidentals

Accidentals are an integral part of Romantic period music. To mimic this musical feature, two special amino acids were selected for triggering this musical feature in the two hands based on their distinguished property and the audio evaluation of the generated protein music. The distinguished amino acids for triggering the accidentals are proline and arginine for the right- and left-hand respectively. These triggering amino acids are of distinguished properties and functions [22, 23], and the triggering of the accidentals marks the recognition of these distinguished amino acids. When encountering a proline or arginine, the pitch of the right- or the left-hand note of the subsequent amino acid is raised by one semitone, which only lasts for that note and resets afterwards so that the subsequent amino acids are unaffected. This algorithm yields protein music that retains a degree of chromaticism while is pleasant to listen to.

3.7. Dynamics

To change the dynamics of protein music, the distinguished amino acid arginine was used as a trigger based on its number of hydrogen bond donors. When encountering an arginine, the dynamic of its frontward amino acid was changed as follows: First, the loudness was tallied with the number of hydrogen bond donors of the frontward amino acid (the loudness is unchanged if the number of hydrogen bond donors is 0). Then, the partition energy and partition coefficient of the frontward amino acid were tallied for a minor increase and decrease in the dynamic

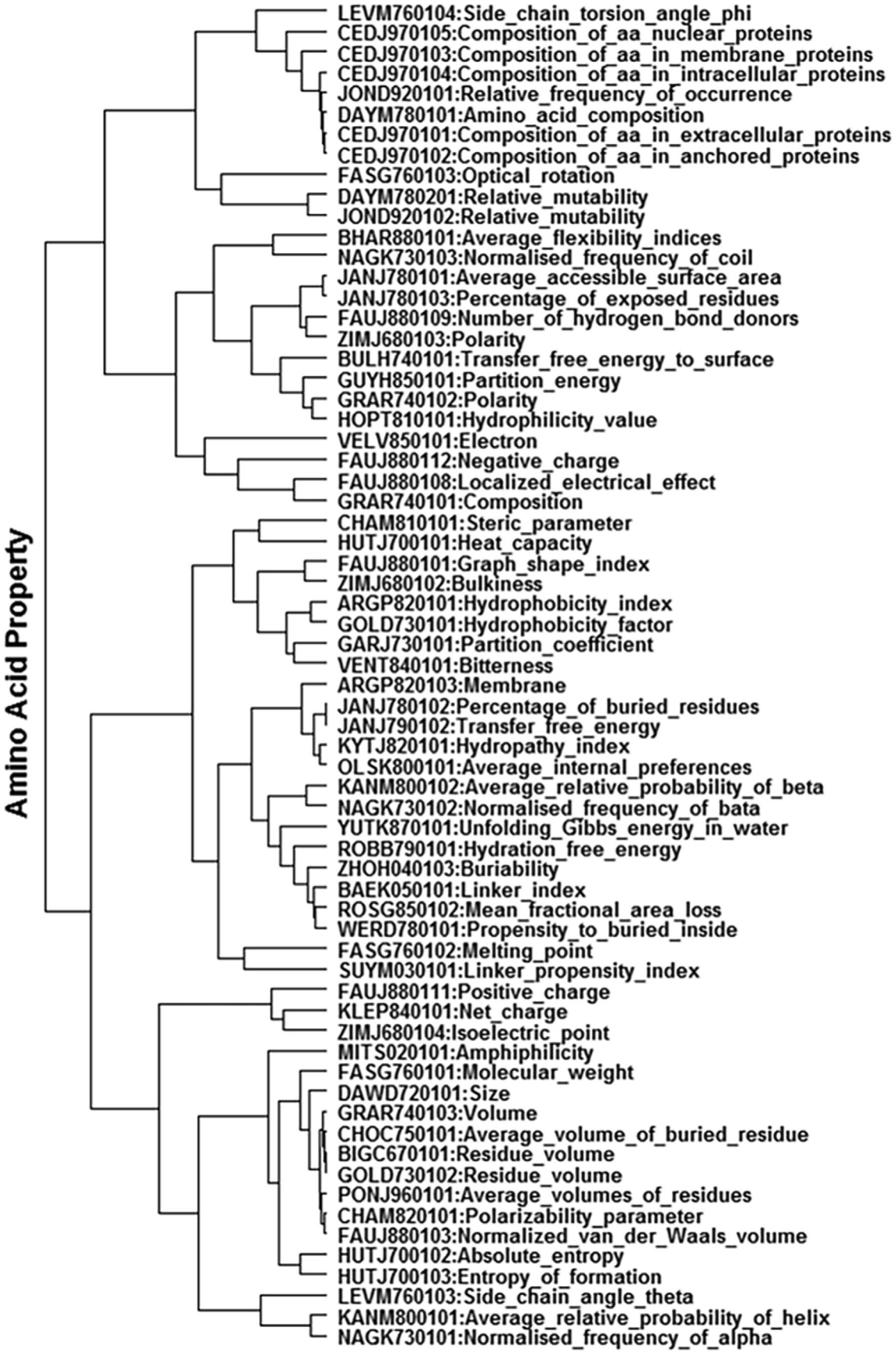

Figure 2. Hierarchical clustering of amino acid properties.

Table 2. The matching of the 20 amino acids to the pitch and note length in accordance with the *Fantasy-Impromptu* style.

Description	Matching relationship for the note pitch and length of the current amino acid						
Pitch of the right-hand	C	D	E	F	G	A	B
Matched amino acid based on relative frequency of occurrence of amino acid in proteins	T G D Y	L N	K Q	P R C	A I M E	V F H	S W
Pitch of the left-hand	C	D	E	F	G	A	B
Matched amino acid based on composition of amino acid in proteins	I F G P	A C H	Y K	V D	E Q N L	R M T	S W
Note length of the right-hand	2	3/2	1	1/2	1/4		
Matched amino acid before the current amino acid based on the size	G	D A	S C	N E T V	R K W Y F Q H M I L P		
Note length of the left-hand	2	3/2	1	1/2	1/4		
Matched amino acid before the current amino acid based on the residue volume	G	A S	C T	N D P V	Q E H I L K M F W Y		

level respectively. The combinations of these minor increases and decreases determine the net variations of the dynamic level.

3.8. Chords

Our analysis identified the rRNA binding propensity and the mRNA binding propensity as the best matched amino acid properties to the chords of the *Fantasy-Impromptu* style. Therefore, these two properties were used for mapping the amino acids to the chords of the right- and left-hand respectively (Table 3). Some chords are successively played as seen in harmonic chord progressions in existing music. To reproduce these chord progressions, special rules were devised for mapping the successive amino acids to these chord progressions, which are primarily based on the combinations of the amino acids mapped to the chords using their rRNA or mRNA binding propensity properties (Table 3). Of the multiple chords mapped to each amino acid, a random number generator was used to randomly select one of the chords. The probability of chord selection is as follows: If all possible chords are of equal number of notes, then each possible chord has an equal probability of being selected. If the chords are of 2- and 3-notes, the probability of selection is 0.7 and 0.3 for the 2- and 3-note chord respectively. If the chords are of 2-, 3-, and 4-notes, the probability is 0.7, 0.2 and 0.1 for the 2-, 3- and 4-note chord respectively. This is to account for the relative frequency of occurrence of different types of chords which was seen in the music pieces analyzed, where 2-note chords made up 70% of the total number of chords in the piece. With the use of a random number generator, this selection rule may generate protein music of varied chords.

3.9. Main theme

The main themes of the four pieces were found to exhibit the following characteristics: The themes typically start to appear early in the pieces. The total note lengths of these themes are 4, 8, or 16, with a longer length preferred, and the melodic interval between every two notes are primarily constrained to 0, 0.5, 1, 1.5, 2 or 2.5 tones. Based on these findings, a judge function was devised for screening the right-hand notes of the protein music to find the earliest appearing note pattern that optimally lean towards the above identified theme characteristics. After a theme was found, it was inserted into specific locations of the protein music triggered by certain amino acids at these locations. The insertion of the main theme simulates the multi-location insertion of multiple amino acids for novel structures [20] and evolution [21]. Four amino acids were selected as triggers based on the property of average volume of buried residue, which are tryptophan, tyrosine, phenylalanine, and arginine.

Table 3. The matching of the amino acids to the chord in accordance with the *Fantasy-Impromptu* style. The amino acids not included in the table are not mapped to any chords.

Chord for the right-hand	Main note	Matched amino acid based on rRNA binding propensity
C-E, C-E-G	C	D Y
F-A, F-A-C	F	P R C
G-C, G-C-E	G	M
A-C, A-C-F	A	F H
D-F, D-F-A	D	L
B-D, B-D-G, B-D-G-F	B	W
Successive chords for right-hand	Main note and main amino acid	Combination of the matched amino acids based on rRNA binding propensity
C-E-G-Bflat and F-A-C	G and M	If M is before P, R, or C, the chord for M is C-E-G-Bflat, the chord for P/R/C is F-A-C
A-C#-E-G and D-F-A	A and F H	If F/H is before L, the chord for F/H is A-C#-E-G, the chord for L is D-F-A If F/H then L then W, randomly choose either the successive F/H L rule or the individual L, W rule above
D-F#-A-C, D-F#-A and G-B-D	D and L	If L is before W, randomly select D-F#-A-C or D-F#-A as chord of L, the chord for W is G-B-D
Chord for the left-hand	**Main note**	**Matched amino acid based on mRNA binding propensity**
C-E, C-E-G	C	P F
D-F, D-F-A	D	C H
F-A, F-A-C	F	D
G-B, G-B-D, G-B-D-F	G	L
A-C, A-C-F	A	R M
B-D, B-D-G, B-D-F-G	B	W
Successive chords for left-hand	Main note and main amino acid	Combination of the matched amino acids based on mRNA binding propensity
D-F#-A-C, D-F#-A and G-B-D	D and C H	If C/H is before W, randomly select D-F#-A-C or D-F#-A as the chord of C/H, the chord for W is G-B-D
A-C#-E-G and D-F-A	A and R M	If R/M is before C/H, the chord for R/M is A-C#-E-G, the chord for C/H is D-F-A If R/M then C/H then W, randomly choose either the successive C/H W rule or the individual R/M C/H rule above

When encountering tryptophan or tyrosine, the theme was inserted between the current and the next amino acid in the original pitches of the theme, but using the octave of the current note. When encountering phenylalanine or arginine, the existing pitch remains unchanged but the note length is changed to that of the theme.

3.10. Composition and analysis of the protein music

The new protein-to-music mapping algorithm was used to compose the music for the 18 proteins (MIDI files for TP53, AQP1, botA in New Music, all MIDI files and codes can be found in http://www.bidd.group/protmusic/). Our work was a first attempt in protein-to-music mapping by music style guided exploration of diverse amino acid properties, patterns and variations for enhanced musicality. All composed protein music pieces are fairly pleasing to the ear, exhibiting fairly good musicality regardless of the specific sequence and function of the proteins. We further compared these music pieces with those of an existing algorithm by inputting the sequences of the 18 proteins (Supplementary Table 2) into the PROM web-server [8] (Old PROM Music TP53, AQP1, botA), which showed significantly enhanced musicality over those composed by the PROM web-server. While the new algorithm was guided by a specific music style (*Fantasy-Impromptu* style), its main objective was primarily for enhanced musicality, not necessarily for duplicating the music style per se. This main objective was preliminarily met.

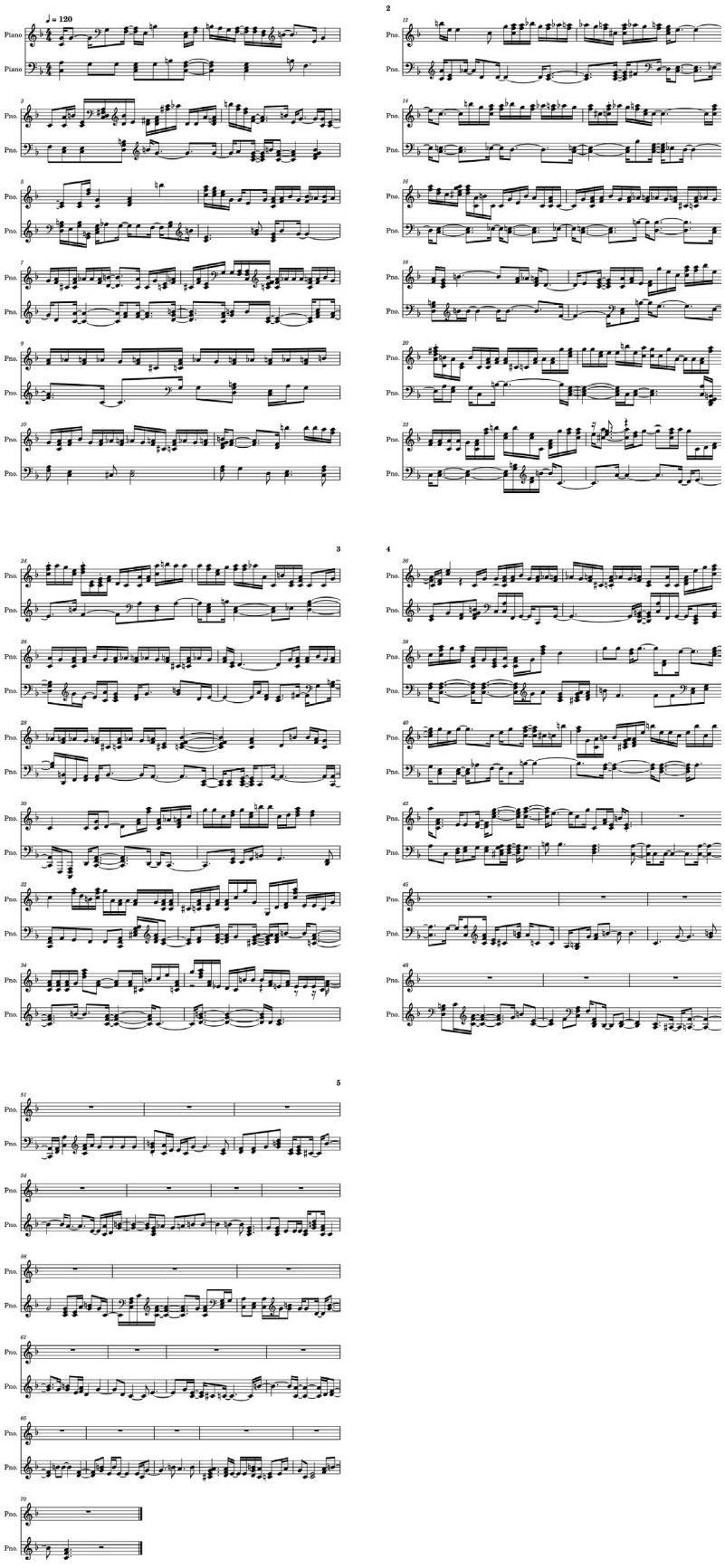

Figure 3. Sheet music for cellular tumor antigen TP53.

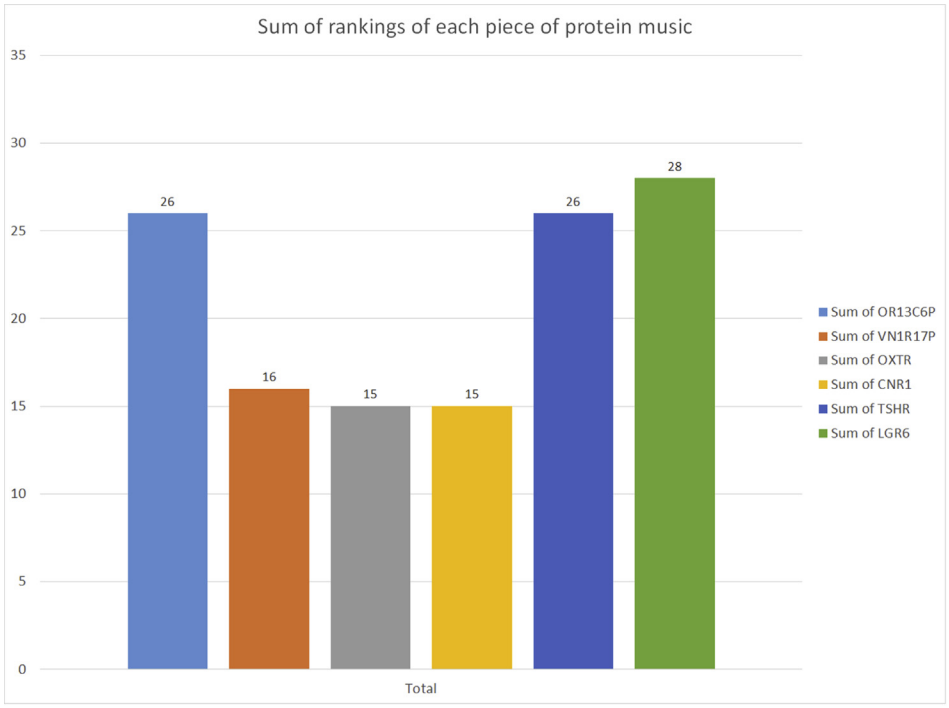

Figure 4. Sum of rankings of each piece of protein music.

Table 4. Rankings of each piece of protein music.

Respondent	OR13C6P (151)	VN1R17P (208)	OXTR (389)	CNR1 (472)	TSHR (764)	LGR6 (967)
Student 1	3	4	2	1	5	6
Student 2	4	6	2	1	3	5
Student 3	6	1	2	4	3	5
Student 4	6	3	1	2	5	4
Student 5	4	1	2	3	5	6
Teacher	3	1	6	4	5	2
Sum	26	16	15	15	26	28

Noticeably, the composed music pieces resemble some elements of the characteristics of the *Fantasy-Impromptu*, particularly with chromaticism, chords, and a wide keyboard range. Therefore, the new algorithm has certain capacity in capturing some essential features of the *Fantasy-Impromptu* style. However, these music pieces inadequately present the highly structured melody arrangements of the *Fantasy-Impromptu* style, particularly the movement and concentration of the energy and the smoothness and twists of the transitions (e.g. tension and release in the dynamics, orderliness in the chromatic scale, and harmonic progression of the chords). Further optimization of the mapping rules with respect to more appropriately selected pieces of a musical form may enable protein music composition that better resemble a music style. As expected, even under the same set of rules, the composed protein music pieces showed substantially varied melody lines and musicality, because of the differences in the protein sequences. By further adjustment of the protein-to-music mapping algorithms, small variations in the protein sequences may be amplified in the melody lines and become audibly distinguishable, leading to protein music unique to a specific genome type or person.

Although the musicality of all composed music pieces is significantly enhanced, there appears small variations of musicality among the composed music pieces. The music pieces of some proteins (e.g. Aquaporin-1 and Cellular tumor antigen p53) are slightly better sounding than others (e.g. oxytocin receptor and polycomb protein SCMH1). The reason for the small variations needs to be further investigated. Figures 3 and 4 presents the sheet music of TP53, which is a protein of important functions in DNA protection, tumor suppression and defensive responses to multiple stresses such as ionizing radiation, hypoxia, oxidative stress, and carcinogens [40]. To evaluate the effect of protein length on the musicality of protein music, we selected six proteins with varying sequence lengths from 151 to 967 amino acids. The protein music of these proteins were generated using our protein-to-music mapping algorithm. The musicality of these protein music pieces was putatively evaluated by human evaluation, specifically by a music teacher and five students from Raffles Institution Singapore. Each person ranked these music pieces from number 1 to number 6, with 1 being the piece with the greatest human-perceived musicality and 6 being the piece with the least. Table 4 shows the sequence length of the six proteins together with the rankings by the teacher and students, and the sum of the rankings by the teacher and five students, and Figure 4 shows the sum of the rankings for each piece. The ranking results suggest that there is no apparent length-dependence of the human-perceived musicality of the protein music.

4. Discussions

There may be several reasons for the deviation of the musicality of our generated protein music from that of the *Fantasy-Impromptu* style. One possible reason is that not all proteins or their properties are well represented by the datasets for deriving protein amino acid properties [27, 28] and protein-binding properties [29, 33]. This problem may be partially alleviated by expanded coverage of the training data or by proper definition of the proteins applicable by the protein-to-music mapping rules. In particular, the established protein property computing tools such as Profeat [41], Propy [42], and iFeature [43] may be expanded and combined for more extensive coverage of protein properties. Moreover, the deviation of the musicality of protein music from that of specific music style may be further reduced by the advanced artificial intelligence methods. Deep learning algorithms for recomposing the style of one audio data using the style of another have been developed [44] (arXiv:1711.11160v1). These algorithms may be potentially extended for the calibration of our derived protein music towards the *Fantasy-Impromptu* style or any other targeted style.

Protein music may have potential applications in music therapy. A recent investigation has shown that music therapy to some extent improves cognitive, psychological, and behavioral alterations in patients of Alzheimer disease [45]. Another recent study has demonstrated that music therapy increased the sleep quality of the nursing students of a University [46]. A third study has revealed that six weeks of music therapy reduces agitation disruptiveness and prevents medication increases in dementia patients [47]. Therefore, music therapy has observable effects on various nervous system disorders. Protein music in particular may play an important role in music therapy. First, by means of personal data privacy protection technology, the protein music may be produced based on the individual genotypes or genetic blueprints, making the protein music personally connected to the music therapy recipients. Secondly, the protein music may be produced in a music style preferred and selected by the therapy recipients.

Another potential application of protein music is the use of their features or melodies for the presentation and computational analysis of the profiles of amino acid properties of proteins, which are not only important for protein-to-music mapping, but also useful for assessing protein folds and functions. For instance, certain amino acid properties have been found to be conserved in evolution [48]. Another study has shown that selection of some amino acid properties improves the performance of protein fold classifications [49]. The protein music for the sequence sections dominated by these properties may thus indicate these evolutionary profiles. Moreover, protein-to-music mapping algorithm also includes features associated with amino acid insertions. It has been reported that amino acid insertion, small deletions, nonsense mutations and A to CC mutation of the REP-1 protein are responsible for the visual impairment of the Japanese choroideremia patients [50]. Therefore, the protein music features or melodies of amino acid insertions may be used for the presentation and analysis of the impairments induced by these insertions.

5. Conclusion

Composition of protein music of enhanced musicality is a challenging task and highly useful avenue for popularizing science [9] and sources of aesthetic melodies [9, 11]. On the basis of the analysis of the musical features of specific music style and diverse amino acid properties, novel algorithms can be developed for protein music composition of enhanced musicality, by means of the music style guided exploration of diverse amino acid properties, patterns and variations of biological relevance. While the preliminary objective of enhanced musicality can be met by the new algorithm, a substantial gap still exists between the musicality of protein music and human music. There is a need for further improvement of the protein-to-music mapping algorithms. In particular, algorithms may be further developed and optimized for music composition with more melodious and harmonious characteristics, which may be achieved by more extensive exploration of the musical features and the diverse amino acid properties, local and remote patterns, and variations. Secondly, considerations of different music styles and more variety of instruments may find more compatible mappings between the amino acid properties and the musical features of a specific style for protein music of high musicality. Thirdly, protein-to-music mapping and protein music composition can be facilitated by artificial intelligence (AI) [12]. The advanced AI technologies for algorithmic music composition [51], creative music synthesis [52], and protein folding [53] may be extended for the novel AI-based protein music composition.

Declarations

Author contribution statement

Nicole WanNi Tay, Fanxi Liu: Performed the experiments; Wrote the paper.

Hui Zhang: Analyzed and interpreted the data.

Chaoxin Wang, Peng Zhang: Contributed reagents, materials, analysis tools or data.

Yu Zong Chen: Conceived and designed the experiments.

Funding statement

This work was supported by the National Key R&D Program of China (2019YFA0905900), National Natural Science Foundation of China (91856126), and Singapore Academic Funds R-148-000-273-114.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2021.e07933.

References

[1] P. Ball, Science & music: facing the music, Nature 453 (7192) (2008) 160–162.
[2] S.A. Mehr, M. Singh, D. Knox, et al., Universality and diversity in human song, Science 366 (6468) (2019) pii: eaax0868.
[3] L. Trainor, Science & music: the neural roots of music, Nature 453 (7195) (2008) 598–599.
[4] J. Sloboda, Science and Music: the ear of the beholder, Nature 454 (7200) (2008) 32–33.
[5] A. Sanchez Sousa, F. Baquero, C. Nombela, The making of "the genoma music, Rev. Iberoam. De. Micol. 22 (4) (2005) 242–248.
[6] D.R. Hofstadter, Gödel, Escher, Bach, An Eternal Golden Braid, Basic Books, 1979, p. 520.
[7] J. Dunn, M.A. Clark, Life music: the sonification of proteins, Leonardo 32 (1) (1999) 25–32.
[8] L.Y. Han, Y.Z. Chen, PROM — Protein Music. http://bidd.group/html/music/index.html. (Accessed 16 May 2020).
[9] R. Takahashi, J.H. Miller, Conversion of amino-acid sequence in proteins to classical music: search for auditory patterns, Genome Biol. 8 (5) (2007) 405.
[10] R.P. Bywater, J.N. Middleton, Melody discrimination and protein fold classification, Heliyon 2 (10) (2016), e00175.
[11] X.J. Shi, Y.Y. Cai, C.W. Chan, Electronic music for bio-molecules using short music phrases, Leonardo 40 (2007) 137–141.
[12] C.H. Yu, Z. Qin, F.J. Martin-Martinez, M.J. Buehler, A self-consistent sonification method to translate amino acid sequences into musical compositions and application in protein design using artificial intelligence, ACS Nano 13 (7) (2019) 7471–7482.
[13] K.A. Dill, S. Bromberg, K. Yue, et al., Principles of protein folding-a perspective from simple exact models, Protein Sci. 4 (4) (1995) 561–602.
[14] R.B. Dannenberg, Style in music, in: S. Argamon, K. Burns, S. Dubnov (Eds.), The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning, Springer-Verlag Berlin Heidelberg, 2010, pp. 45–58.
[15] A. Sharma, K.K. Paliwal, A. Dehzangi, et al., A strategy to select suitable physicochemical attributes of amino acids for protein fold recognition, BMC Bioinf. 14 (2013) 233.
[16] B. Liu, X. Wang, Q. Chen, Q. Dong, X. Lan, Using amino acid physicochemical distance transformation for fast protein remote homology detection, PloS One 7 (9) (2012), e46633.
[17] K.F. Han, D. Baker, Recurring local sequence motifs in proteins, J. Mol. Biol. 251 (1) (1995) 176–187.
[18] M. Thomas-Chollier, L.C. Watson, S.B. Cooper, et al., A naturally occurring insertion of a single amino acid rewires transcriptional regulation by glucocorticoid receptor isoforms, Proc. Natl. Acad. Sci. U.S.A. 110 (44) (2013) 17826–17831.
[19] L. Orlichenko, R. Geyer, M. Yanagisawa, et al., The 19-amino acid insertion in the tumor-associated splice isoform Rac1b confers specific binding to p120 catenin, J. Biol. Chem. 285 (25) (2010) 19153–19161.
[20] H. Jiang, C. Blouin, Insertions and the emergence of novel protein structure: a structure-based phylogenetic study of insertions, BMC Bioinf. 8 (2007) 444.
[21] M. Scalley-Kim, P. Minard, D. Baker, Low free energy cost of very long loop insertions in proteins, Protein Sci. 12 (2) (2003) 197–206.

[22] B.K. Kay, M.P. Williamson, M. Sudol, The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains, FASEB J 14 (2) (2000) 231–241.
[23] R.S. Blanc, S. Richard, Arginine methylation: the coming of age, Mol. Cell 65 (1) (2017) 8–24.
[24] D.J. Grout, C.V. Palisca, A History of Western Music, sixth ed., W.W. Norton & Co., 2001, pp. 571–587.
[25] F. Chopin, M. Hinson, Fantaisie-Impromptu, Alfred Music, 2005, p. 3.
[26] M. Metin, The analysis of fantasie impromptu (Op. Posth. 66) in terms of structural and harmonical study on Chopin's fantasie impromptu, J. Educ. Fut. 3 (2013) (2013) 121–126.
[27] AAindex, Amino acid indices, substitution matrices and pair-wise contact potentials. https://www.genome.jp/aaindex/. (Accessed 16 May 2020).
[28] S. Kawashima, P. Pokarowski, M. Pokarowska, et al., AAindex: amino acid index database, progress report 2008, Nucleic Acids Res. 36 (Database issue) (2008) D202–D205.
[29] B. Ma, T. Elkayam, H. Wolfson, R. Nussinov, Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces, Proc. Natl. Acad. Sci. U.S.A. 100 (10) (2003) 5772–5777.
[30] IMSLP, International music score library project or petrucci music library. https://imslp.org/. (Accessed 16 May 2020).
[31] J. Allwright, S. Seymour, Guide to Abc2midi features. https://abcmidi.sourceforge.io/. (Accessed 15 June 2021).
[32] S. Seymour, Runabc. https://ifdo.ca/∼seymour/runabc/top.html. (Accessed 16 May 2020).
[33] J.J. Ellis, M. Broom, S. Jones, Protein-RNA interactions: structural analysis and functional classes, Proteins 66 (4) (2007) 903–911.
[34] J.H. McDonald, Handbook of Biological Statistics, third ed., Sparky House Publishing, Baltimore, Maryland, 2014.
[35] The R Core Team, RDocumentation, Stats, version 3.6.2. Maintained by R-core@R-project.org. https://www.rdocumentation.org/packages/stats/versions/3.6.2. Accessed 2018.
[36] The UniProt Consortium, UniProt: a worldwide hub of protein knowledge, Nucleic Acids Res. 47 (D1) (2019) D506–D515.
[37] Musescore, Musescore 2, The Musescore Developer Community, 2015.
[38] B. Wathen, Z. Jia, Protein beta-sheet nucleation is driven by local modular formation, J. Biol. Chem. 285 (24) (2010) 18376–18384.
[39] R.A. George, J. Heringa, An analysis of protein domain linkers: their classification and role in protein folding, Protein Eng. (Online) 15 (11) (2002) 871–879.
[40] E.R. Kastenhuber, S.W. Lowe, Putting p53 in context, Cell 170 (6) (2017) 1062–1078.
[41] H.B. Rao, F. Zhu, G.B. Yang, Z.R. Li, Y.Z. Chen, Update of PROFEAT: a web server for computing structural and physicochemical features of proteins and peptides from amino acid sequence, Nucleic Acids Res. 39 (Web Server issue) (2011) W385–390.
[42] D.S. Cao, Q.S. Xu, Y.Z. Liang, propy: a tool to generate various modes of Chou's PseAAC, Bioinformatics 29 (7) (2013) 960–962.
[43] Z. Chen, P. Zhao, F. Li, et al., iFeature: a Python package and web server for features extraction and selection from protein and peptide sequences, Bioinformatics 34 (14) (2018) 2499–2502.
[44] J. Chen, G.B. Yang, H. Zhao, M. Ramasamy, Audio style transfer using shallow convolutional networks and random filters, Multimed. Tool. Appl. 79 (21) (2020) 15043–15057.
[45] M. Gómez Gallego, J. Gómez García, Music therapy and Alzheimer's disease: cognitive, psychological, and behavioural effects, Neurologia 32 (5) (2017) 300–308.
[46] M. Kavurmaci, N. Dayapoğlu, M. Tan, Effect of music therapy on sleep quality, Alternative Ther. Health Med. 26 (4) (2020) 22–26.
[47] H.M. Ridder, B. Stige, L.G. Qvale, C. Gold, Individual music therapy for agitation in dementia: an exploratory randomized controlled trial, Aging Ment. Health 17 (6) (2013) 667–678.
[48] W.R. Rudnicki, T. Mroczek, P. Cudek, Amino acid properties conserved in molecular evolution, PloS One 9 (6) (2014), e98983.
[49] L. Zhang, L. Kong, A novel amino acid properties selection method for protein fold classification, Protein Pept. Lett. 27 (4) (2020) 287–294.
[50] M. Hayakawa, K. Fujiki, Y. Hotta, et al., Visual impairment and REP-1 gene mutations in Japanese choroideremia patients, Ophthalmic Genet. 20 (2) (1999) 107–115.
[51] O. Lopez-Rincon, O. Starostenko, G. Ayala-San Martín, Algoritmic Music Composition Based on Artificial Intelligence: A Survey, 2018 International Conference on Electronics, Communications and Computers (CONIELECOMP), IEEE, 2018, pp. 187–193.
[52] J.-P. Briot, G. Hadjeres, F.-D. Pachet, Deep Learning Techniques for Music Generation, Computational Synthesis and Creative Systems, Springer, 2019.
[53] J. Xu, Distance-based protein folding powered by deep learning, Proc. Natl. Acad. Sci. U.S.A. 116 (34) (2019) 16856–16865.
[54] J.A. Bartz, J. Zaki, N. Bolger, K.N. Ochsner, Social effects of oxytocin in humans: context and person matter, Trends Cognit. Sci. 15 (7) (2011) 301–309.
[55] J.H. Kim, D. Desor, Y.T. Kim, et al., Efficacy of alphas1-casein hydrolysate on stress-related symptoms in women, Eur. J. Clin. Nutr. 61 (4) (2007) 536–541.
[56] K. Noh, H. Lee, T.Y. Choi, et al., Negr1 controls adult hippocampal neurogenesis and affective behaviors, Mol. Psychiatr. 24 (8) (2019) 1189–1205.
[57] J.C. de Moura, M.M. Noroes, P. Rachetti Vde, et al., The blockade of transient receptor potential ankirin 1 (TRPA1) signalling mediates antidepressant- and anxiolytic-like actions in mice, Br. J. Pharmacol. 171 (18) (2014) 4289–4299.
[58] D.A. Havas, A.M. Glenberg, K.A. Gutowski, M.J. Lucarelli, R.J. Davidson, Cosmetic use of botulinum toxin-a affects processing of emotional language, Psychol. Sci. 21 (7) (2010) 895–900.
[59] C.P. Donahue, K.S. Kosik, T.J. Shors, Growth hormone is produced within the hippocampus where it responds to age, sex, and stress, Proc. Natl. Acad. Sci. U.S.A. 103 (15) (2006) 6031–6036.
[60] A. Koizumi, A. Tsuchiya, K. Nakajima, et al., Human sweet taste receptor mediates acid-induced sweetness of miraculin, Proc. Natl. Acad. Sci. U.S.A. 108 (40) (2011) 16819–16824.
[61] D. Narasimhan, J.H. Woods, R.K. Sunahara, Bacterial cocaine esterase: a protein-based therapy for cocaine overdose and addiction, Future Med. Chem. 4 (2) (2012) 137–150.
[62] F. Nyberg, Growth hormone in the brain: characteristics of specific brain targets for the hormone and their functional significance, Front. Neuroendocrinol. 21 (4) (2000) 330–348.
[63] D.A. Salih, A.J. Rashid, D. Colas, et al., FoxO6 regulates memory consolidation and synaptic function, Genes Dev. 26 (24) (2012) 2780–2801.
[64] M.A. Rivera, T.D. Fahey, Association between aquaporin-1 and endurance performance: a systematic review, Sport. Med. - Open 5 (1) (2019) 40.
[65] Y. Takada, K. Isono, J. Shinga, et al., Mammalian Polycomb Scmh1 mediates exclusion of Polycomb complexes from the XY body in the pachytene spermatocytes, Development 134 (3) (2007) 579–590.
[66] H.F. Horn, K.H. Vousden, Coping with stress: multiple ways to activate p53, Oncogene 26 (9) (2007) 1306–1316.
[67] M. Moshiri, F. Hamid, L. Etemad, Ricin toxicity: clinical and molecular aspects, Rep. Biochem. Mol. Biol. 4 (2) (2016) 60–65.
[68] H. Roy, S. Cannon, Ribulose bisphosphate carboxylase assembly: what is the role of the large subunit binding protein? Trends Biochem. Sci. 13 (5) (1988) 163–165.
[69] S.J. Remington, Green fluorescent protein: a perspective, Protein Sci. 20 (9) (2011) 1509–1519.
[70] Y. Sugimoto, S. Sanuki, S. Ohsako, et al., Ovalbumin in developing chicken eggs migrates from egg white to embryonic organs while changing its conformation and thermal stability, J. Biol. Chem. 274 (16) (1999) 11030–11037.
[71] D.K. Mandal, C.F. Brewer, Interactions of concanavalin A with glycoproteins: formation of homogeneous glycoprotein-lectin cross-linked complexes in mixed precipitation systems, Biochemistry 31 (50) (1992) 12602–12609.