DNA makes protein — makes music?

By Clare Sansom

All life scientists must now be familiar with the single-letter codes for both DNA and protein sequences. Any gene can be expressed as a sequence of the four letters, A, C, G and T, while the protein code contains 20 characters that represent the 20 amino acids. However, a simple sequence of letters cannot tell the whole story. There is an analogy here with music. A simple tune can be described as a sequence of note names, but just as a protein sequence cannot indicate the structure of the protein, or even the positions of disulphide bonds, so a tune needs a rhythm and dynamics, as well as the note sequence.

Some scientists and musicians have become fascinated by this similarity, and found ways to represent actual DNA and protein sequences in music. There are many sites on the web devoted to music generated from biological molecules: one of these, Linda Long's Molecular Music (www.molecularmusic.com) is described in this month's Best of the Web section.

But what attracts so many people to this rather esoteric field? Parts of DNA and protein sequences are often repeated with only small changes, and this imperfect repetition has echoes in the themes and variations found in classical music. Professor Susumu Ohno of the Beckman Research Institute was the first to propose that, "the repetition process governs both musical composition and DNA sequence construction". Others have simply noted that "when DNA sequences are converted to music, it sounds musical!"

Biologist Mary Ann Clark, of Texas Wesleyan University, has collaborated with artist John Dunn of Algorithmic Arts to develop algorithms to convert protein structures into musical sequences. Each of the 20 amino acids is set to be equivalent to a different note, giving enough for about three octaves of the classical Western diatonic scale. Their compositions reflect the structure of the proteins, as well as their sequences, by representing changes from helix to turn to strand by changes in the music's pitch and instrumentation. Also, hydrophobic amino acids were assigned notes that were lower in pitch than hydrophilic ones. Interestingly, Dunn and Clark found that "…some α- and β-regions also were marked by motifs whose sequences might not obviously repeat, but the general shaping of whose phrases did." This is of more than academic interest; they found that it was easier to pick out the similarity between human haemoglobin and a globin from a primitive three-eyed lizard, the tuatara, from their music than from a conventional sequence alignment.

Ross King, of the University of Wales, Aberystwyth, and Scottish pop musician Colin Angus have developed a program that makes music from the DNA sequence itself. This program has recently been re-written in Java, and is confusingly entitled 'Protein Music'. The program, documentation and example sequences are freely available for download from the Web. The treble line is composed simply from the four notes C, A, G and T, representing the nucleotides cytosine, adenine, guanine and thymine (in that order). The bass line does come from the protein sequence: each codon contributes one bass note, pitched according to the physical characteristics of the amino acid that it encodes.

Brent Hugh, an American pianist, has composed his 'Music of the Human Genome' from parts of the sequences of human chromosomes 1, 2, 3, 4, 5, and 22 and released it on CD. Each chromosome has its own distinctive melody and is represented by a different instrument: for example, chromosome 1 is represented by a flute sound and chromosome 2 by tubular bells. It is fortunate that he has no plans to set the whole genome to music because, as he notes, "Chromosome 22 has around 37 megabytes of data. That's enough genetic information to make a melody that lasts about 32 days."

All the music described so far has one feature in common: it is derived, in one way or another, from the single-letter protein or DNA code. Susan Alexjander, a composer based in Sacramento, CA, has collaborated with biologist David Deamer, to use the light absorption spectra from the four bases to generate musical notes. The raw absorption data is converted into frequencies within the range of the human ear and sent to a synthesizer; the sounds are then arranged into 'scales' (using a different scale for each base) and the music built up around this. Most of the intervals in her music are 'micro-tonic' (that is, they do not fit into the pattern of tones and semitones used in almost all Western music). Rick Weiss of the *New York Times* has described it as "Sometimes mysterious, sometimes playful… an evolutionary leap forward."

Clare Sansom
c.sansom@mail.cryst.bbk.ac.uk

Web links

Algorithmic Arts: http://algoart.com

Life Music (Dunn and Clark): http://mitpress2.mit.edu/e-journals/Leonardo/isast/articles/lifemusic.html

The Music of DNA (Susan Alexjander): http://www.healingmusic.org/SusanA/

Music of the Human Genome (Brent Hugh): http://artists.mp3s.com/artist_song/1301/1301401.html

Protein Music (King and Angus): http://www.aber.ac.uk/~phiwww/pm/

Best of the Web

http://www.molecularmusic.com

I'm no musician (I dabbled with the cello once, but had to give up when aged 7, as it was realized that I couldn't spread my fingers wide enough), but I do know what I like…

Molecular Music™ provides the discerning, ambient music-seeking artistic-type with a selection of tunes 'derived from the molecules of life'. They have been created by Dr Linda Long, a Research Fellow in Complementary Medicine at Exeter University, and are generated from the three-dimensional structures of proteins. "X-ray crystallography data (describing the three-dimensional positions of the amino acids in a protein molecule) are filtered and then mapped onto musical parameters such as pitch and amplitude," says Linda. "Data may be filtered to emphasize either small or large scale changes, thereby generating note sequences that describe protein structure on many different levels. In this way, characteristic patterns in protein structure such as helixes (heard as arpeggios) and β-sheets (heard as a succession of similar notes) emerge as recognisable musical note patterns from the three-dimensional structural data." Linda uses the music as a tool for teaching molecular modelling of complex protein structures, and also to generate music from herbs, medicinal plants and the human body for relaxation and therapeutic purposes.

There are two categories of MP3 files to try out on the website: *Music of the Plants* and *Music of the Body*; the CDs are price £9.99 and £13.95 (plus postage and packing) respectively.

The *Music of the Plants* sample files include pokeweed antiviral protein, myrosinase (from mustard) and plastocyanin (from parsley). They all have a sort of dreamy underbelly, with a piano plinking over the top, and are certainly of a relaxing and soothing nature. I find that quite odd — wouldn't you expect the pokeweed antiviral protein to be a bit more aggressive, in a sort of 'kill!' and 'get him!' kind of way? I think the samples could be longer, to give you more a feel for the music, rather than the snapshot that they are. But we know they exist to get you to buy the CD, and I'm afraid they didn't quite cut it for me. Better samples please!

The site has many examples and references that describe how music has assisted in the treatment of a number of conditions, including Alzheimer's disease, pain, depression and hormonal problems. Who am I to say that these musical pieces haven't got this kind of power? Tune in…

Gary Burd
gary.burd@biochemistry.org

