

What is the elemental composition of a cell?

One of the most interesting chemical asymmetries associated with life on Earth is the mismatch between the composition of cells on the one hand and of inanimate matter on the other. As a result of the rich and diverse metabolic processes that make cells work, living chemistry is largely built around carbon, oxygen, nitrogen and hydrogen, with these elemental components serving as the key building blocks making up the cell's dry weight.

The empirical composition on a per carbon basis for the dry weight of *E. coli* is $C:H_{1.77}:O_{0.49}:N_{0.24}$ (BNID 101800) which means that for every nitrogen atom in the cell there are about 2 oxygen atoms, 7 hydrogen atoms and 4 carbon atoms. Hence, the empirical composition can be approximated as $C_4:H_7:O_2:N_1$. For budding yeast the composition is similar, namely, $C:H_{1.61}:O_{0.56}:N_{0.16}$ (BNID 101801).

As noted above, it is interesting to compare the composition of cells to that at the Earth's crust. Strikingly, carbon and hydrogen, majority players in living matter are relatively rare in the Earth's crust. Carbon comes in as only the 17th most abundant element and hydrogen a slightly higher constituent coming in 10th place, way behind the major constituents, oxygen (60.5%), silicon (20.5%) and aluminum (6.2%). Similarly, in the atmosphere, the main carbon containing compound, CO_2 , makes up merely ~380 parts per million and extracting this dilute resource is the main reason for the need to water plants. Plants lose water when opening their stomata, small pores on leaves that are the channels for importing carbon dioxide molecules. This mundane process accounts for a staggering two thirds of humanity's water consumption. Hydrogen, which was prevalent in the early Earth's atmosphere before life's appearance was progressively lost to space as a result of its low mass meaning that thermal velocities alone provide enough kinetic energy to overcome the Earth's gravitational pull.

Can we say something more about the elemental compositions of living matter by thinking about the makeup of the key macromolecular constituents of cells? In particular, how might we infer the elemental ratios from our acquaintance with the cell components without consulting the empirical measurement? A bacterial cell has about 55% protein, 20% nucleic acids, 10% lipids and another 15% of various other components (by weight, BNID 101436). Exploiting the fact that the mass ratio of proteins to nucleic acids is about 3:1, let's see how far a few simple facts about these two dominant components can take us in estimating the elemental composition of a cell.

A nucleotide is composed of a phosphate (PO_4) and ribose ($C_5H_8O_2$) backbone and a base ($\sim N_5C_5O_1H_6$ – using guanine as our representative example). Thus the total chemical composition is $P_1N_5O_7C_{10}H_{14}$ with a total mass of about 350 Daltons (BNID 104886). An amino acid consists of a backbone with a peptide bond $-RC(O)NH-$ where the first group (R) is a carbon harboring a residue that on the average is crudely about 3 carbons, 1 oxygen and 6 hydrogens leading to a total elemental composition of $N_1C_5O_2H_8$ and a mass of about 110 Dalton (BNID 104877). If we focus our attention only on the protein and nucleic acid content of cells, we are now prepared to estimate

the overall composition of a cell. To reason this out, we recall that the mass of protein in a typical bacterium is roughly 3-fold larger than the mass of nucleic acids. Further, since nucleic acids have roughly three times the mass of amino acids, this implies that for every nucleotide there are roughly 9 amino acids. We need to evaluate the chemical composition of a mix of 9 amino acids and 1 nucleic acid resulting in the stoichiometric relation $9 \times (\text{NC}_5\text{O}_2\text{H}_8) + 1 \times (\text{PN}_5\text{O}_7\text{C}_{10}\text{H}_{14}) = (\text{C}_{55}:\text{O}_{25}:\text{H}_{86}:\text{N}_{14}:\text{P}_1)$. Normalizing by the number of nitrogen atoms this is $\text{C}_{3.1}:\text{H}_{6.1}:\text{O}_{1.8}:\text{N}_1$ pretty close to the empirical value of $\text{C}_4:\text{H}_7:\text{O}_2:\text{N}_1$ (or $\text{C}:\text{H}_{1.77}:\text{O}_{0.49}:\text{N}_{0.24}$). This estimate can be refined further if we include the next largest contributor to the cell mass, namely, the lipids that account for ~10% of that mass. These molecules are characterized by the general formula $\text{C}_N:\text{H}_{N-2N}:\text{O}_1$ depending on the length of the tail and the degree of saturation. Including lipids in our estimate will increase the proportion of C, H and less so O – producing the correct trends in line with empirical formula. The point of going to the trouble of estimating something that is already known through an empirical formula is that it serves as a critical sanity check of our understanding of the main biological components that determine the cell's composition.