We have presented a microscopic view of the chemical reaction between oxygen and hydrogen. The equation

\[
\ce{O2 (g) + 2 H2 (g) -> 2H2O (l)}
\]

represents the same event in terms of chemical symbols and formulas.

We also saw that plants form glucose from CO₂ and H₂O:

\[
\ce{6CO2 (g) + 6 H2O (l) + Light Energy -> C6H12O6 (s) + 6O2(g)}
\]

But how does a practicing chemist find out what is occurring on the microscopic scale? When a reaction is observed for the first time, little is known about the microscopic nature of the products. It is therefore necessary to determine experimentally the composition and formula of a newly synthesized substance.

The first step in such a procedure is usually to separate and purify the products of a reaction. The products above are easy to separate, because they are liquids or solids, while the reactants are gaseous. But how could you determine that the formula should be H₂O, and not H₂O₂? And plants produce mixtures of carbohydrates during photosynthesis, which must be separated by chromatography or other techniques before they can be identified. Glucose is perhaps the most common simple sugar, and it exists in two molecular forms, with the chemical formula C₆H₁₂O₆:

![the open chain form of glucose](image1)

![the cyclic form of glucose](image2)

Plants may produce other simple sugars with similar structures, like D-ribulose (C₅H₁₀O₅), D-fructose (C₆H₁₂O₆) and di- or polysaccharides (with two or more simple sugars linked together) like sucrose (C₁₂H₂₂O₁₁) or Maltose (C₁₂H₂₂O₁₁).
One answer involves quantitative analysis—the determination of the percentage by mass of each element in the compound. Such data are usually reported as the percent composition.

Example 1: Percent Composition

When 10.0 g oxygen reacts with sufficient hydrogen, 11.26 g of a pure compound is formed. Calculate the percent composition from these experimental data.

Solution

The percentage of oxygen is the mass of oxygen divided by the total mass of compound times 100 percent:

\[
\% \text{O} = \frac{\text{m}_{\text{O}}}{\text{m}_{\text{compound}}} \cdot 100\% = \frac{10.0 \text{ g}}{11.26 \text{ g}} \cdot 100\% = 88.81\%
\]

The remainder of the compound (11.26 g – 10 g = 1.26 g) is hydrogen:

\[
\% \text{H} = \frac{\text{m}_{\text{H}}}{\text{m}_{\text{compound}}} \cdot 100\% = \frac{1.26 \text{ g}}{11.26 \text{ g}} \cdot 100\% = 11.19\%
\]

As a check, verify that the percentages add to 100:

\[88.81\% + 11.19\% = 100\%\]

To obtain the formula from percent-composition data, we must find how many hydrogen atoms there are per oxygen atom. On a macroscopic scale this corresponds to the ratio of the amount of hydrogen to the amount of oxygen. If the formula is H\(_2\)O, it not only indicates that there are two hydrogen atoms per oxygen atom, it also says that there are 2 mol of hydrogen atoms for each 1 mol of oxygen atoms. That is, the amount of hydrogen is twice the amount of oxygen. The numbers in the ratio of the amount of bromine to the amount of mercury (2:1) are the subscripts of hydrogen and oxygen in the formula.

Example 2: Formula Determination

Determine the formula for the compound whose percent composition was calculated in the previous example.

Solution

For convenience, assume that we have 100 g of the compound. Of this, 88.81 g (88.81\%) is oxygen and 11.19 g is hydrogen. Each mass can be converted to an amount of substance...
\begin{align} & n_{\text{O}} = \text{88.81 g} \times \frac{1 \text{ mol O}}{16 \text{.00 g}} = \text{5.550 mol O} \\
& n_{\text{H}} = \text{11.19 g} \times \frac{1 \text{ mol H}}{1.008 \text{ g}} = \text{11.101 mol H} \end{align}

So the formula is H_{11.101}O_{5.550}, but we know there are no fractions of atoms, so we have to put this in a standard form. To do this, we divide the larger amount by the smaller:

\[ \frac{n_{\text{H}}}{n_{\text{O}}} = \frac{11.19 \text{ mol H}}{5.550 \text{ mol O}} = \frac{1.998 \text{ mol H}}{1 \text{ mol O}} \]

To the formula is H_{1.998}O_{1}, but again, there seems to be a problem. The ratio 1.998 mol H to 1 mol O also implies that there are 1.998 H atoms per 1 O atom. If the atomic theory is correct, there is no such thing as 0.998 atom; but remembering that our measurements are only good to three significant figures, we round 1.998 to 2.00 and write the formula as H$_2$O.

Example (PageIndex{3}): Formula of Compound

An oxide of hydrogen has the composition 94.07% O and 5.93% H. Find its formula.

**Solution**

Again assume a 100-g sample and calculate the amount of each element:

\begin{align} & n_{\text{O}} = \text{94.07 g} \times \frac{1 \text{ mol O}}{16.00 \text{ g}} = \text{5.88 mol O} \\
& n_{\text{H}} = \text{5.93 g} \times \frac{1 \text{ mol H}}{1.008 \text{ g}} = \text{5.88 mol H} \end{align}

The ratio is

\[ \frac{n_{\text{H}}}{n_{\text{O}}} = \frac{5.88 \text{ mol H}}{5.88 \text{ mol O}} = \frac{1 \text{ mol H}}{1 \text{ mol O}} \]

So the formula is H$_{5.88}$O$_{5.88}$, which is a 1:1 ratio within experimental precision. We would therefore assign the formula HO.

The formula obtained in Example 3 does not correspond to either of the two hydrogen oxides we have already discussed. Is it a third one? The answer is no because our method can only determine the ratio of H to O. The ratio 1:1 is the same as 2:2, and so our method will give the same result for HO or H$_2$O$_2$ (or H$_7$O$_7$, for that matter, should it exist). The formula determined by this method is called the empirical formula or simplest formula. Occasionally, as in the case of hydrogen peroxide, the empirical formula differs from the actual molecular composition, or the molecular formula. Experimental determination of the molecular weight in addition to percent composition permits calculation of the molecular formula.

Example (PageIndex{4}): Empirical and Molecular Formulas

A compound whose molecular weight is 28 contains 85.6% C and 14.4% H. Determine its empirical and molecular formulas.

**Solution**

\begin{align} & n_{\text{C}} = \text{85.6 g} \times \frac{1 \text{ mol C}}{12.01 \text{ g}} = \text{7.13 mol C} \\
& n_{\text{H}} = \text{14.4 g} \times \frac{1 \text{ mol H}}{1.008 \text{ g}} = \text{14.3 mol H} \end{align}

So the formula is H$_{5.88}$O$_{5.88}$, which is a 1:1 ratio within experimental precision. We would therefore assign the formula HO.

The formula obtained in Example 3 does not correspond to either of the two hydrogen oxides we have already discussed. Is it a third one? The answer is no because our method can only determine the ratio of H to O. The ratio 1:1 is the same as 2:2, and so our method will give the same result for HO or H$_2$O$_2$ (or H$_7$O$_7$, for that matter, should it exist). The formula determined by this method is called the empirical formula or simplest formula. Occasionally, as in the case of hydrogen peroxide, the empirical formula differs from the actual molecular composition, or the molecular formula. Experimental determination of the molecular weight in addition to percent composition permits calculation of the molecular formula.

Example (PageIndex{4}): Empirical and Molecular Formulas

A compound whose molecular weight is 28 contains 85.6% C and 14.4% H. Determine its empirical and molecular formulas.

**Solution**

\begin{align} & n_{\text{C}} = \text{85.6 g} \times \frac{1 \text{ mol C}}{12.01 \text{ g}} = \text{7.13 mol C} \\
& n_{\text{H}} = \text{14.4 g} \times \frac{1 \text{ mol H}}{1.008 \text{ g}} = \text{14.3 mol H} \end{align}
So the formula is $C_{7.13}H_{14.3}$, but to get integral subscripts we divide each by the smaller:

$$\frac{n_{\text{H}}}{n_{\text{C}}} = \frac{14.3 \text{ mol H}}{7.13 \text{ mol C}} = \frac{2.01 \text{ mol H}}{1 \text{ mol C}}$$

The empirical formula is therefore $\text{CH}_2$. The molecular weight corresponding to the empirical formula is

$$12.01 + 2 \times 1.008 = 14.03$$

Since the experimental molecular weight is twice as great, all subscripts must be doubled and the molecular formula is $\text{C}_2\text{H}_4$.

Example \(\PageIndex{5}\): Empirical Formula

D-Xylose contains 40.0% C, 6.71% H, and 53.29% O. What is its empirical formula?

**Solution**

\[
\begin{align}
& n_{\text{H}} = 6.71 \text{ g} \times \frac{1 \text{ mol H}}{1.008 \text{ g}} = 6.66 \text{ mol H} \\
& n_{\text{C}} = 40.00 \text{ g} \times \frac{1 \text{ mol C}}{12.01 \text{ g}} = 3.33 \text{ mol C} \\
& n_{\text{O}} = 53.29 \text{ g} \times \frac{1 \text{ mol O}}{16.00 \text{ g}} = 3.33 \text{ mol O} \\
\end{align}
\]

So the formula is $C_{3.33}H_{6.66}O_{3.33}$, and dividing all three by the smallest amount of substance we obtain $\text{CH}_2\text{O}$.

\[
\begin{align}
& \frac{n_{\text{C}}}{n_{\text{O}}} = \frac{3.33 \text{ mol C}}{3.33 \text{ mol O}} = \frac{1 \text{ mol H}}{1 \text{ mol O}} \\
& \frac{n_{\text{H}}}{n_{\text{O}}} = \frac{4.44 \text{ mol H}}{2.22 \text{ mol O}} = \frac{2 \text{ mol H}}{1 \text{ mol O}} \\
\end{align}
\]

Now we see where carbohydrates get their name. They all have formulas like $C_n(\text{H}_2\text{O})_m$, so they all look like hydrates (water containing compounds) of carbon. D-Xylose is actually $C_5(\text{H}_2\text{O})_5$ or $\text{C}_5\text{H}_{10}\text{O}_5$, Glucose is $\text{C}_6(\text{H}_12\text{O})_6$ or $\text{C}_6\text{H}_{12}\text{O}_6$, and Sucrose is $\text{C}_12(\text{H}_2\text{O})_{11}$ or $\text{C}_{12}\text{H}_{22}\text{O}_{11}$.

We have seen that experimental error can give mole ratios that are not exact units. Occasionally the ratio of amounts is farther from a whole number than can be explained by experimental error, as in the following example.

Example \(\PageIndex{6}\): Empirical Formula of Aspirin

Aspirin contains 60.0% C, 4.48% H, and 35.5% O. What is its empirical formula?

**Solution**

\[
\begin{align}
& n_{\text{H}} = 14.4 \text{ g} \times \frac{1 \text{ mol H}}{1.008 \text{ g}} = 14.3 \text{ mol H} \\
& n_{\text{C}} = 85.6 \text{ g} \times \frac{1 \text{ mol C}}{12.01 \text{ g}} = 7.13 \text{ mol C} \\
& n_{\text{O}} = 35.5 \text{ g} \times \frac{1 \text{ mol O}}{16.00 \text{ g}} = 2.22 \text{ mol O} \\
\end{align}
\]

Divide all three by the smallest amount of substance

\[
\begin{align}
& \frac{n_{\text{C}}}{n_{\text{O}}} = \frac{5.00 \text{ mol C}}{2.22 \text{ mol O}} = \frac{2.25 \text{ mol H}}{1 \text{ mol O}} \\
& \frac{n_{\text{H}}}{n_{\text{O}}} = \frac{4.44 \text{ mol H}}{2.22 \text{ mol O}} = \frac{4.44 \text{ mol mol}}{2.22 \text{ mol O}} \\
\end{align}
\]

We have seen that experimental error can give mole ratios that are not exact units. Occasionally the ratio of amounts is farther from a whole number than can be explained by experimental error, as in the following example.
Clearly there are twice as many H atoms as O atoms, but the ratio of C to O is not so obvious. We must convert 2.25 to a ratio of small whole numbers. This can be done by changing the figures after the decimal point to a fraction. In this case, .25 becomes \( \frac{1}{4} \). Thus 2.25 becomes \( \frac{9}{4} \), and

\[
\frac{n_{\text{C}}}{n_{\text{O}}} = \frac{2.25 \text{ mol C}}{1 \text{ mol O}} = \frac{9 \text{ mol C}}{4 \text{ mol O}}
\]

We can also write

\[
\frac{n_{\text{H}}}{n_{\text{O}}} = \frac{2 \text{ mol H}}{1 \text{ mol O}} = \frac{8 \text{ mol H}}{4 \text{ mol O}}
\]

Thus the empirical formula is \( \text{C}_9\text{H}_8\text{O}_4 \).

Once someone has determined a formula—empirical or molecular—it is possible for someone else to do the reverse calculation. Finding the weight-percent composition from the formula often proves quite informative, as the following example shows.

Example \( \PageIndex{7} \): Percent Composition of a Sugar

As we saw above, all carbohydrates have the general formula \( \text{C}_n(\text{H}_2\text{O})_m \). All simple sugars have the general formula \( \text{C}_n(\text{H}_2\text{O})_n \), so they all have the same percentage of C, H, and O.

a. Calculate the percent C in glucose. b. Show that the compound \( \text{C}_6(\text{H}_2\text{O})_4 \) cannot be a simple sugar by calculating its percent composition.

Solution

a. \( \text{C}_6(\text{H}_2\text{O})_6 \) contains 6 mol C, 12 mol H, and 6 mol O. The molar mass is thus

\[
M = (6 \times 12.011) + (12 \times 1.008) + (6 \times 16) = 180 \text{ g mol}^{-1}
\]

A 1-mol sample would weigh 180.56 g. The mass of 6 mol C it contains is

\[
\text{m}_C = 6 \text{ mol C} \cdot \frac{12.011 \text{ g}}{1 \text{ mol C}} = 72.066 \text{ g}
\]

Therefore the percentage of C is

\[
\text{%C} = \frac{72.066 \text{ g}}{180.56 \text{ g}} \cdot 100\% = 40.00\%
\]

The percentages of H and O are easily calculated as

\[
\text{m}_H = 12 \text{ mol H} \cdot \frac{1.008 \text{ g}}{1 \text{ mol H}} = 12.096 \text{ g}
\]

\[
\text{%H} = \frac{12.096 \text{ g}}{180.56 \text{ g}} \cdot 100\% = 6.70\%
\]

\[
\text{m}_O = 6 \text{ mol O} \cdot \frac{16.00 \text{ g}}{1 \text{ mol O}} = 96.00 \text{ g}
\]

\[
\text{%O} = \frac{96.00 \text{ g}}{180.56 \text{ g}} \cdot 100\% = 53.17\%
\]

Though not strictly needed to answer the problem, the latter two percentages provide a check of the results. The total 40.00\% + 6.70\% + 53.17\% = 100.00\% as it should.

Note that for any simple sugar,
\[
\%C = \frac{n(12.011)}{n(12.011) + n(18)} = 40\% 
\]

regardless of how many carbons it contains (simple sugars are trioses, \(C_3H_6O_3\), tetrose, \(C_4H_8O_4\), pentoses, and hexoses. The same goes for the percent H and O.

b. The molar mass of \(C_8(H_2O)_4\) is 168.15;

\[
\%C = \frac{(8 \times 12.011)/168.15}{100\%} = 57.14\%, \text{ not the 40\% characteristic of simple sugars.}
\]

Similarly,

\[
\%H = \frac{(8 \times 1.008) / 168.15}{100\%} = 4.80\% \text{ and the percent O is } 38.06\%.
\]

From ChemPRIME: 2.11: Formulas and Composition

---

**Contributors**

- Ed Vitz (Kutztown University), John W. Moore (UW-Madison), Justin Shorb (Hope College), Xavier Prat-Resina (University of Minnesota Rochester), Tim Wendorff, and Adam Hahn.