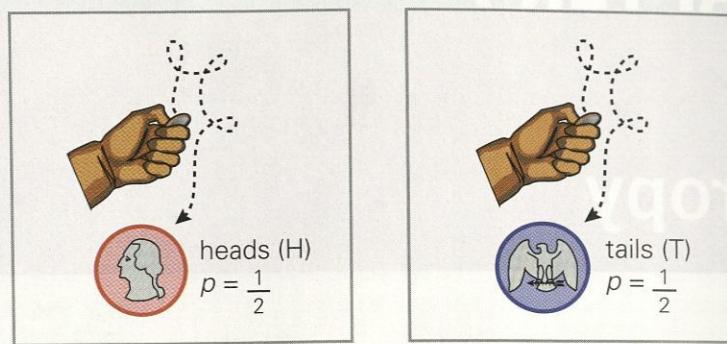


**Figure 7.2 The probability of the outcomes of coin tosses.** In an unbiased coin toss, the probability,  $p$ , of the outcome being heads (H) or tails (T) is equal to 1/2 in both cases.



Because calculations involving probabilities can be used to analyze many situations of interest in biology, including molecular diffusion.

Although we cannot predict the outcome of any particular sequence of coin tosses, the statistical outcome of the aggregate number of heads and tails in a large number of coin tosses can be predicted with high precision. The common-sense concepts introduced by studying coin tosses allow us to introduce the statistical definition of entropy, which is related in a simple way to probability. A guiding principle that emerges is that the direction of spontaneous change in a molecular system that is not at equilibrium is always in the direction of increasing entropy. This principle, also known as the **second law of thermodynamics**, means that entropy is maximal at equilibrium. As a result, we evaluate whether or not a system and its surroundings are at equilibrium by computing the entropy and determining whether its value is maximal.

In this chapter, we consider changes in entropy that arise solely from changes in molecular positions, without changes in energy. In Chapter 8, we study the interplay between entropy and energy, which leads to the crucial concept of free energy, which is discussed in Chapter 9.

## A. COUNTING STATISTICS AND MULTIPLICITY

### 7.1 Different sequences of outcomes in a series of coin tosses have equal probabilities

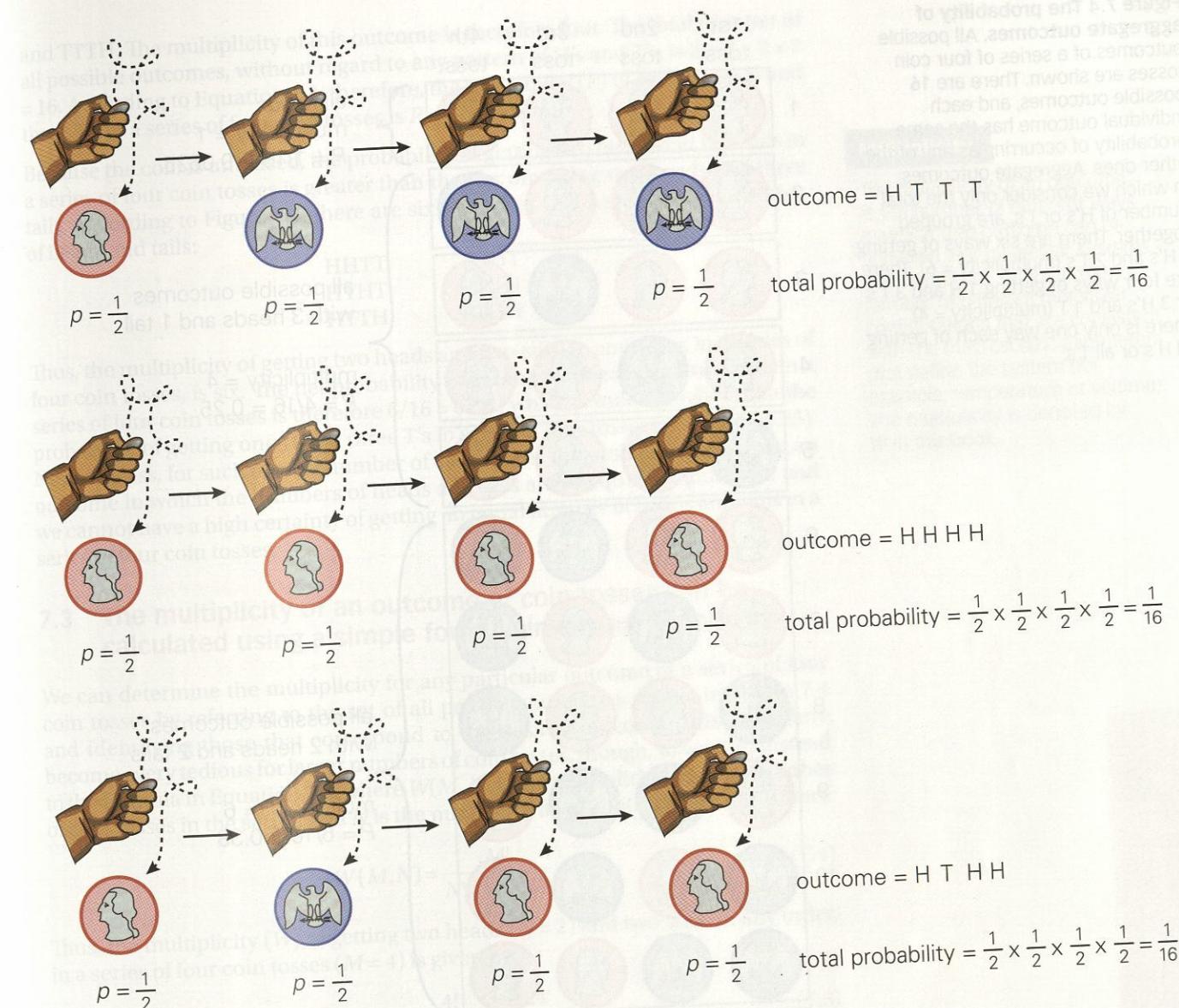
Consider a series of coin tosses in which the coin is unbiased—that is, the probability of the toss yielding a head (H) is equal to the probability of the same toss yielding a tail (T) (see Figure 7.2). What is the probability ( $p$ ) of obtaining first a head and then a tail in a series of two coin tosses? The probability of obtaining H in the first toss is 1/2 and, because the coin tosses are unbiased, the probability of obtaining T in the second toss is also 1/2. The combined probability of obtaining first a head and then a tail is the product of the individual probabilities:

$$p = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \quad (7.1)$$

Likewise, in a series of four coin tosses, the probability ( $p$ ) of obtaining the sequence HTHT is given by:

$$p = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16} \quad (7.2)$$

In fact, the probability of obtaining *any* particular sequence of outcomes in a series of four coin tosses, such as HHHH or HTHH, is the same (1/16, as shown



in Figure 7.3). As the number of coin tosses in the series increases, the probability of obtaining any particular sequence of heads and tails becomes smaller and smaller. At the same time, the probability of obtaining any particular sequence is always the same as that for any other sequence that contains the same number of coin tosses.

### 7.2 When considering aggregate outcomes, the most likely result is the one that has maximum multiplicity

Now, instead of being concerned with the result of any individual coin toss, let us consider the *aggregate* or net result of the series of coin tosses (Figure 7.4). What, for example, is the probability of getting one H and three T's in a series of four coin tosses? This probability ( $P$ ) is given by:

$$P = \frac{\text{number of desired outcomes}}{\text{total number of possible outcomes}} \quad (7.3)$$

where the number of desired outcomes (the number of different ways of achieving a particular outcome) is referred to as the **multiplicity** of the outcome. According

**Figure 7.3 The probability of an outcome in a series of coin tosses.** If a coin is tossed four times, the probability of any particular sequence of outcomes, such as the sequence HTTT, is

$$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16}$$

**Figure 7.4 The probability of aggregate outcomes.** All possible outcomes of a series of four coin tosses are shown. There are 16 possible outcomes, and each individual outcome has the same probability of occurring as any of the other ones. Aggregate outcomes, in which we consider only the total number of H's or T's, are grouped together. There are six ways of getting 2 H's and 2 T's (multiplicity = 6). There are four ways of getting 1 H and 3 T's or 3 H's and 1 T (multiplicity = 4). There is only one way each of getting all H's or all T's.

	1st toss	2nd toss	3rd toss	4th toss	
1	HHHH	HHHT	HHTH	HTHH	all heads multiplicity = 1 $P = 1/16 = 0.0625$
2	HHHT	HTHH	HTTH	THTH	
3	HHTH	HTTH	TTHT	TTTH	all possible outcomes with 3 heads and 1 tail
4	HTHH	TTHT	TTTH	HTHT	multiplicity = 4 $P = 4/16 = 0.25$
5	TTTH	TTHT	HTHT	HTHH	
6	TTHT	HTHT	HTTH	HHHT	all possible outcomes with 2 heads and 2 tails
7	HTHT	HTTH	HHHH	HHHT	multiplicity = 6 $P = 6/16 = 0.38$
8	HTTH	HHHH	HHHT	HTHH	
9	HHHH	HHHT	HTHH	HTTH	
10	HHHT	HTHH	HTTH	TTTH	all possible outcomes with 1 head and 3 tails
11	HTHH	HTTH	TTTH	TTHT	multiplicity = 4 $P = 4/16 = 0.25$
12	HTTH	TTTH	TTHT	HTHT	
13	TTTH	TTHT	HTHT	HTHH	all tails multiplicity = 1 $P = 1/16 = 0.0625$
14	TTHT	HTHT	HTHH	HHHH	
15	HTHT	HTHH	HHHH	HHHT	
16	HTHH	HHHH	HHHT	HTHH	all possible outcomes of a series of 4 coin tosses

to Figure 7.4, there are four possible outcomes of a series of four coin tosses that correspond to the desired outcome of one H and three T's: HTTT, THTT, TTHT,

and TTTH. The multiplicity of this outcome is therefore four. The total number of all possible outcomes, without regard to any pattern of H's and T's, is  $2 \times 2 \times 2 \times 2 = 16$ . According to Equation 7.3, therefore, the probability ( $P$ ) of seeing one H and three T's in a series of four coin tosses is  $P = 4/16 = 0.25$ .

Because the coin is unbiased, the probability of getting two heads and two tails in a series of four coin tosses is greater than that for obtaining one heads and three tails. According to Figure 7.4, there are six outcomes that have an equal number of heads and tails:

$$\begin{array}{ll} \text{HHTT} & \text{THHT} \\ \text{HTHT} & \text{THTH} \\ \text{HTHH} & \text{TTHH} \end{array}$$

Thus, the multiplicity of getting two heads and two tails, in any order, in a series of four coin tosses, is six. The net probability of getting two heads and two tails in a series of four coin tosses is therefore  $6/16 = 0.38$ , which is indeed higher than the probability of getting one H and three T's in series of four tosses (where  $P = 0.25$ ). Nevertheless, for such a small number of coin tosses, the probability of getting an outcome in which the numbers of heads and tails are unequal is significant, and we cannot have a high certainty of getting an equal number of heads and tails in a series of four coin tosses.

### 7.3 The multiplicity of an outcome of coin tosses can be calculated using a simple formula involving factorials

We can determine the multiplicity for any particular outcome of a series of four coin tosses by referring to the set of all possible outcomes shown in Figure 7.4 and identifying those that correspond to the desired outcome. This procedure becomes very tedious for larger numbers of coin tosses, though, so we turn instead to the formula in Equation 7.4, where  $W(M, N)$  is the multiplicity,  $M$  is the number of coin tosses in the series, and  $N$  is the number of heads obtained (in any order):

$$W(M, N) = \frac{M!}{N!(M-N)!} \quad (7.4)$$

Thus, the multiplicity ( $W$ ) of getting two heads ( $N = 2$ ) and two tails, in any order, in a series of four coin tosses ( $M = 4$ ) is given by:

$$W(4, 2) = \frac{4!}{2!(4-2)!} = \frac{24}{2 \times 2} = 6 \quad (7.5)$$

The answer given by Equation 7.5 exactly matches the situation that is illustrated in Figure 7.4.

How do we arrive at the formula in Equation 7.4? Imagine that we have  $M$  coins in a bag (Figure 7.5). Imagine, as well, that each coin in the bag is labeled distinctly (for example 1, 2, 3, ...,  $M$ ). We reach into the bag and randomly pull out one coin, then randomly pull out another coin and place it next to the first one, and so on until we have removed all  $M$  coins from the bag. In how many different ways can these  $M$  coins be pulled from the bag?

There are  $M$  possibilities for choosing the first coin, but once the first coin is selected and removed, there are only  $(M-1)$  choices for the second coin. Because each independent choice of the first coin can be combined with each independent choice of the second coin to get a different possible outcome, there are  $M \times (M-1)$  different ways in which the first two coins can be chosen. Likewise, there are  $M \times (M-1) \times (M-2)$  different ways in which the first three coins can be chosen and, by extension, the number of different ways in which all the coins can be chosen is given by the **factorial** of  $M$ :

$$M! = M \times (M-1) \times (M-2) \times \dots \times 3 \times 2 \times 1 \quad (7.6)$$

#### Multiplicity

The multiplicity of an outcome is the number of different ways in which that outcome can be achieved. For molecular systems, the multiplicity refers to the number of different molecular configurations that are consistent with the macroscopic parameters that define the system (for example, temperature or volume). The multiplicity is denoted by  $W$  in this book.

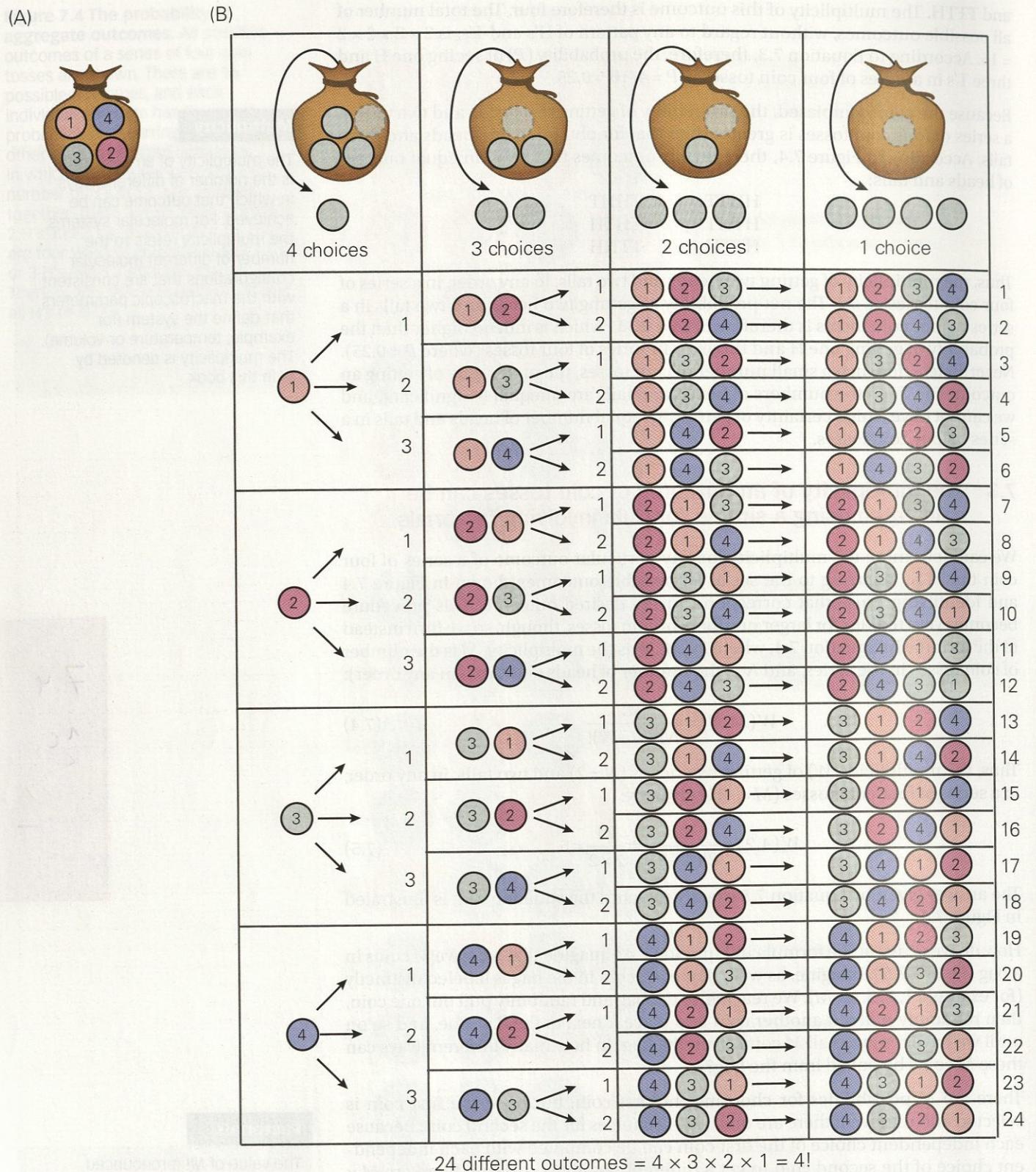
#### Factorials

The value of  $M!$  (pronounced "M factorial") is given by:

$$M! = M \times (M-1) \times (M-2) \times \dots \times 2 \times 1$$

For example, the value of  $5!$  is given by:

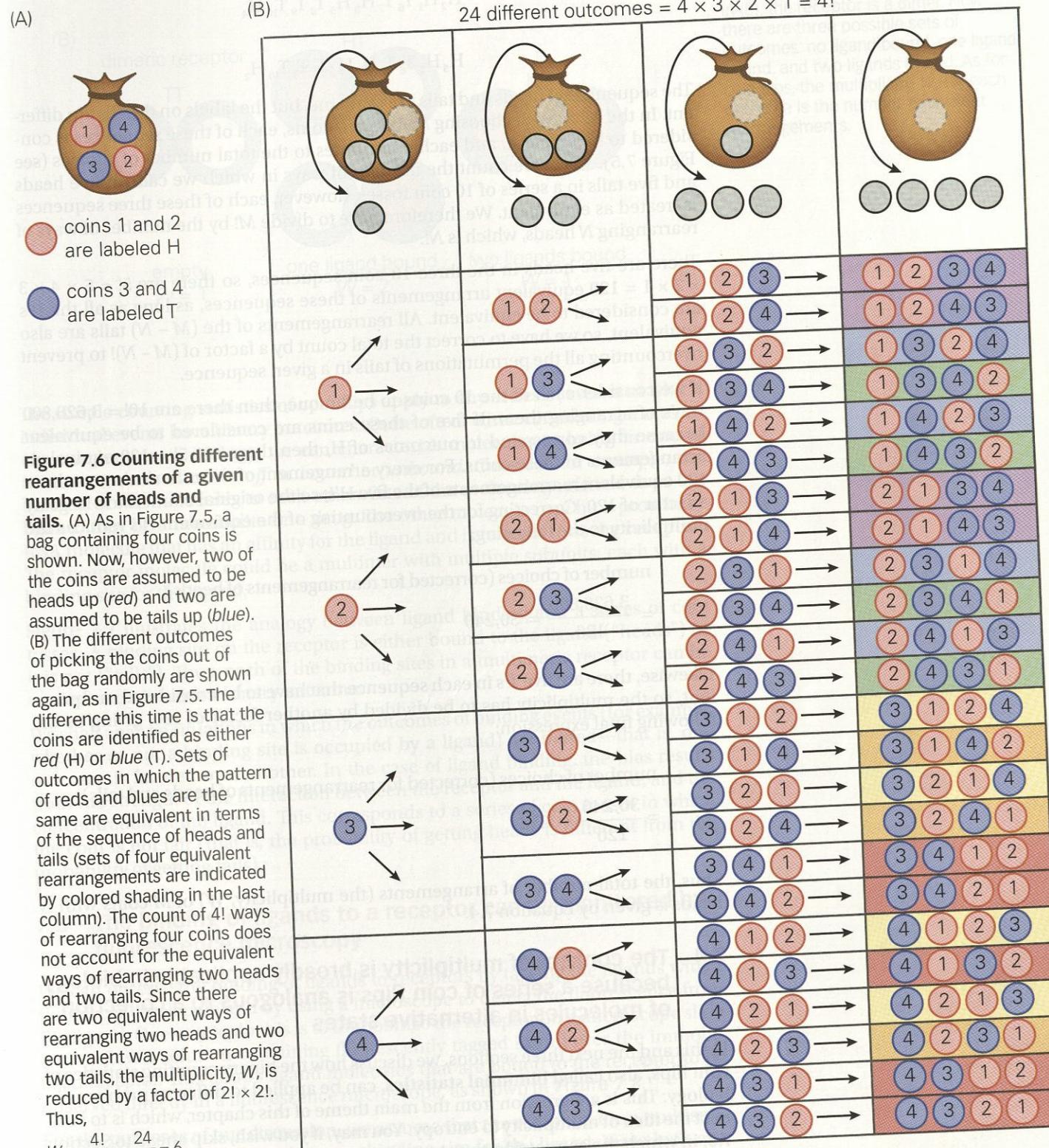
$$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$



**Figure 7.5** The number of ways of picking four distinct coins in random order. (A) Imagine a bag containing four different coins, labeled 1, 2, 3, and 4. A coin is selected by reaching into the bag and randomly picking a coin. (B) All possible outcomes of randomly picking four coins from the bag in sequence are shown. There are four ways of picking the first coin, three ways of picking the second coin, two ways of picking the third coin, and one way of picking the last coin in each sequence. There are a total of  $4 \times 3 \times 2 \times 1 = 4!$  different possible arrangements of the coins.

Up until now, we have ignored whether the coins have landed heads up or tails up when we pick them. Let us assume, however, that  $N$  of the coins in the bag are labeled "H" (for "heads"), and  $M - N$  are labeled "T" for "tails." That is, the coins are labeled  $H_1, H_2, \dots, H_N, T_{N+1}, T_{N+2}, \dots, T_M$  (Figure 7.6).

Figure 7.7 An analogy between receptor-ligand interactions and coin tosses. In these diagrams, the receptor is shown as a blue circle and the ligand as a smaller red circle. Two cases are shown here. In (A), the receptor is a monomer and there are two possible outcomes: either the ligand is not bound ("tails") or the ligand is bound ("heads" or "heads up").



**Figure 7.6** Counting different rearrangements of a given number of heads and tails. (A) As in Figure 7.5, a bag containing four coins is shown. Now, however, two of the coins are assumed to be heads up (red) and two are assumed to be tails up (blue). (B) The different outcomes of picking the coins out of the bag randomly are shown again, as in Figure 7.5. The difference this time is that the coins are identified as either red (H) or blue (T). Sets of outcomes in which the pattern of reds and blues are the same are equivalent in terms of the sequence of heads and tails (sets of four equivalent rearrangements are indicated by colored shading in the last column). The count of  $4!$  ways of rearranging four coins does not account for the equivalent ways of rearranging two heads and two tails. Since there are two equivalent ways of rearranging two heads and two tails, the multiplicity,  $W$ , is reduced by a factor of  $2! \times 2!$ . Thus,

$$W = \frac{4!}{2! 2!} = \frac{24}{4} = 6.$$

In the  $M!$  different arrangements of the coins, there are many sequences of heads and tails in which the labels (numbers) on the coins are different, but the overall pattern of heads and tails is the same. For example, the following outcomes from a series of 10 coin tosses have identical patterns of heads and tails:

$$H_1 H_2 T_6 T_7 H_3 H_4 T_8 T_9 T_{10} H_5$$

$$H_2 H_1 T_6 T_7 H_3 H_5 T_8 T_9 T_{10} H_4$$

$$H_5 H_3 T_6 T_7 H_1 H_4 T_8 T_9 T_{10} H_2$$

The sequences of heads and tails are the same, but the labels on the H's are different. In the  $M!$  ways of choosing  $M$  different coins, each of these sequences is considered to be different, and each contributes to the total number of choices (see Figure 7.5). When we count the number of ways in which we can get five heads and five tails in a series of 10 coin tosses, however, each of these three sequences is treated as equivalent. We therefore have to divide  $M!$  by the number of ways of rearranging  $N$  heads, which is  $N!$ .

There are five heads in the three 10-coin sequences, so there are  $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$  equivalent arrangements of these sequences, as long as all the H's are considered to be equivalent. All rearrangements of the  $(M - N)$  tails are also equivalent, so we have to correct the total count by a factor of  $(M - N)!$  to prevent overcounting all the permutations of tails in a given sequence.

If we consider each of the 10 coins to be unique, then there are  $10! = 3,628,800$  ways of arranging them. If five of these coins are considered to be equivalent because they correspond to outcomes of H, then there are  $5! = 120$  equivalent arrangements of these coins. For every arrangement of the other coins, we have 120 equivalent rearrangements of the five H's, so the original count is too large by a factor of 120. Correcting for the overcounting of the equivalent H's reduces the multiplicity to the following:

number of choices (corrected for rearrangements of heads)

$$= \frac{3,628,800}{120} = 30,240$$

Likewise, there are five T's in each sequence that have to be considered as equivalent, so the multiplicity has to be divided by another factor of 120, giving us the following final expression:

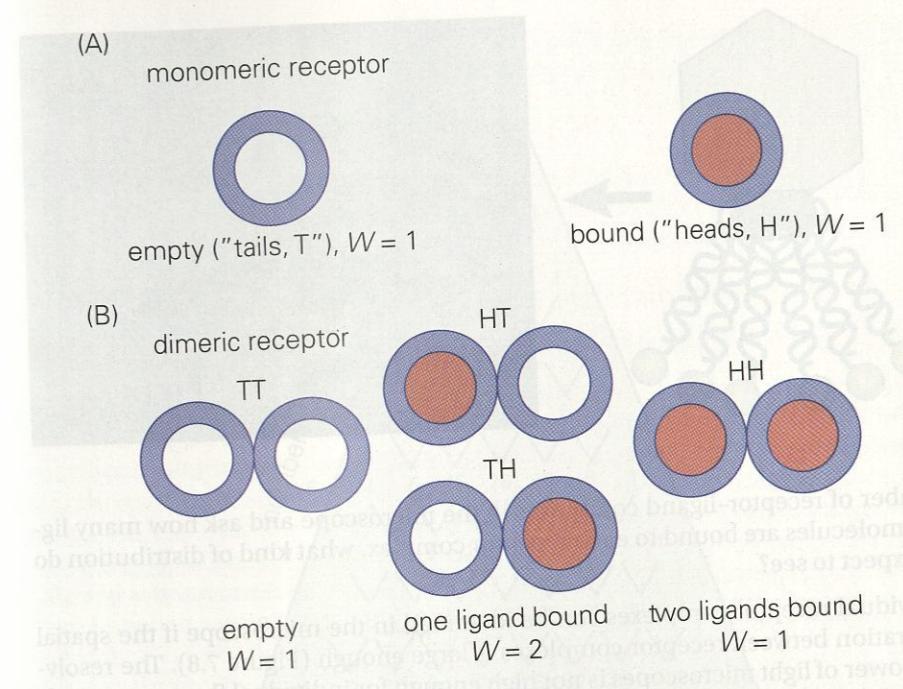
number of choices (corrected for rearrangements of heads and tails)

$$= \frac{30,240}{120} = 252$$

Thus, the total number of arrangements (the multiplicity,  $W$ ) of  $M$  coins with  $N$  heads is given by Equation 7.4.

#### 7.4 The concept of multiplicity is broadly applicable in biology because a series of coin flips is analogous to a collection of molecules in alternative states

In this and the next three sections, we discuss how the binary counting statistics of coin flips, also called **binomial statistics**, can be applied to different problems in biology. This is a digression from the main theme of this chapter, which is to connect the idea of multiplicity to entropy. You may, if you wish, skip ahead to Section 7.8, in which the connection to entropy begins to be made.



The simple counting statistics of coin flips apply to any series of events in which individual events have binary outcomes. Quite often, when considering a molecular mechanism in biology, we can characterize molecules as being in one of two states, such as *on* or *off* or *bound* or *unbound*. Consider, for example, the binding of two molecules to each other. One of the molecules, referred to as the **ligand**, may be a metabolite or a drug. The other molecule may be a protein or RNA molecule that has an affinity for the ligand and is referred to as the **receptor**. The receptor molecule could be a multimer with multiple subunits, each with a binding site for the ligand.

**Figure 7.7** illustrates the analogy between ligand binding and a series of coin tosses. A binding site on the receptor is either bound to the ligand ("heads") or it is not ("tails"). Thus, each of the binding sites in a multimeric receptor can be treated as a trial with a binary outcome.

We shall analyze situations in which the outcomes of binding events (for example, whether or not a binding site is occupied by a ligand) are biased—that is, one outcome is favored over another. In the case of ligand binding, the bias results from the strength of the interaction between the receptor and the ligand, and the concentration of the ligand. This corresponds to a series of coin tosses in which the coin is not fair (that is, the probability of getting heads is different from the probability of getting tails).

#### 7.5 The binding of ligands to a receptor can be monitored by fluorescence microscopy

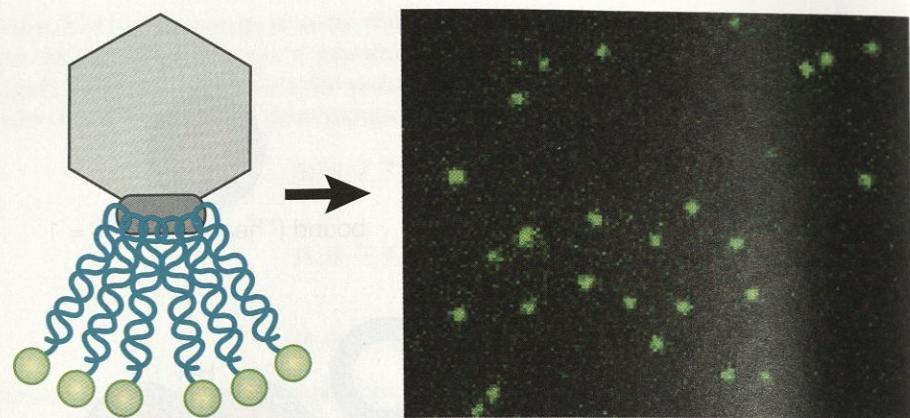
We can observe the binding of ligands to receptors by tagging the ligands with a fluorescent molecule and by using a microscope to detect the fluorescence from the ligand. One way to do this is to immobilize the receptor on a microscope slide and to flow a solution containing fluorescently tagged ligand over the immobilized receptor. This allows ligand molecules that are bound to the receptor to be observed directly in a fluorescence microscope, as shown in **Figure 7.8**. For the purposes of our discussion, we assume that the receptors form multisubunit complexes, with more than one binding site for the ligand. If we look at a

**Figure 7.7** An analogy between receptor-ligand interactions and coin tosses. In these diagrams, the receptor is shown as a blue circle and the ligand as a smaller red circle. Two cases are shown here. In (A), the receptor is a monomer and there are two possible outcomes: either the ligand is not bound ("tails" or T) or the ligand is bound ("heads" or H). In (B), the receptor is a dimer. Now there are three possible sets of outcomes: no ligand bound, one ligand bound, and two ligands bound. As for coin flips, the multiplicity,  $W$ , for each outcome is the number equivalent rearrangements.

#### Binomial coefficients

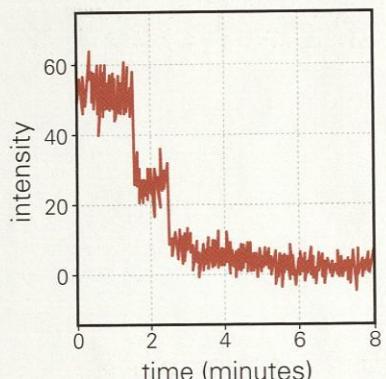
The values of the multiplicity of coin tosses are known as binomial coefficients. A simple diagrammatic representation of binomial coefficients is given in **Figure 7.8**.

**Figure 7.8 Imaging fluorescently tagged molecules.** In this experiment, a virus particle is the receptor. The virus contains six RNA molecules (blue structures) that are bound by dye molecules. Up to six dye molecules can bind to one virus headpiece. When viewed through a fluorescence microscope, each virus particle is observed as a spot of green fluorescence. (Adapted from H. Zhang et al., and P. Guo, *RNA* 13: 1793–1802, 2007. With permission from Cold Spring Harbor Press.)



number of receptor-ligand complexes in the microscope and ask how many ligand molecules are bound to each receptor complex, what kind of distribution do we expect to see?

Individual receptor complexes can be identified in the microscope if the spatial separation between receptor complexes is large enough (Figure 7.8). The resolving power of light microscopes is not high enough for individual fluorescent molecules within one complex to be detected individually. Nevertheless, the number of fluorescent molecules bound to each receptor can be counted by observing the stepwise reduction in fluorescence intensity as individual molecules become bleached by the light, as indicated in Figure 7.9. In the example shown in Figure 7.9, we infer that there are two fluorescent molecules within the field of view. This is because the fluorescence intensity is reduced in two steps to the background level.



**Figure 7.9 Counting the number of fluorescent molecules in a region of fluorescence.** One of the fluorescent spots in the sample shown in Figure 7.8 is observed in a fluorescence microscope, and the intensity of fluorescence is recorded as a function of time. As the molecules in the region absorb light, they eventually become bleached (that is, they no longer fluoresce), and the intensity is reduced. The observation of two stepwise reductions in intensity in this region shows that there are two fluorescent molecules in the region. (Adapted from H. Zhang et al., and P. Guo, *RNA* 13: 1793–1802, 2007. With permission from Cold Spring Harbor Press.)

## 7.6 Pascal's triangle describes the multiplicity of outcomes for a series of binary events

Consider a multimeric receptor complex with  $M$  binding sites for a ligand. The receptor complex can be bound to no ligand, or it can be bound to 1, 2, 3, ...,  $M$  ligand molecules. If we assume that each of the ligand binding sites in the receptor is independent of the others, then the probability of seeing a certain number of ligands bound to a receptor complex is proportional to the number of ways in which the bound ligands can be rearranged among the binding sites. This is given by the multiplicity  $W(M, N)$ , where  $N$  is the number of ligand molecules bound (number of positive outcomes) and  $M$  is the total number of binding sites in the receptor complex (total number of events).

Analogous to the number of ways of obtaining  $N$  heads in a series of  $M$  coin tosses, the multiplicity  $W(M, N)$  is given by:

$$W(M, N) = \frac{M!}{N!(M-N)!} \quad (7.7)$$

We can use Equation 7.7 to write down the multiplicities for various numbers of ligand molecules bound ( $N$ , the number of positive outcomes) to receptor complexes with different numbers of binding sites ( $M$ , the number of events). The multiplicities for receptors containing 1 to 10 binding sites (or events) are shown schematically in Figure 7.10. This diagram is known as **Pascal's triangle**. Each row in Pascal's triangle corresponds to a series of events with  $M$  trials, and the entries in the row are the multiplicities for  $N$  positive outcomes ( $N = 0, 1, 2, \dots, M$ ). The numbers in Pascal's triangle are known as **binomial coefficients**. The binomial coefficients exhibit many fascinating numerical relationships. For example, each binomial coefficient is the sum of the two numbers immediately above it, except for the numbers on the edges, which are all 1.

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ . If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is  $1-p$ .

If the probability of a binding site being occupied is  $p$ , then the probability of an empty site is

complex with seven binding sites or events, the multiplicity for having four ligands bound (or four positive outcomes) is given by:

$$W(7,4) = \frac{7!}{4!3!} = 35 \quad (7.8)$$

Note that the fifth entry in the seventh row of Pascal's triangle, corresponding to four ligands bound, is indeed 35 (see Figure 7.10).

### 7.7 The binomial distribution governs the probability of events with binary outcomes

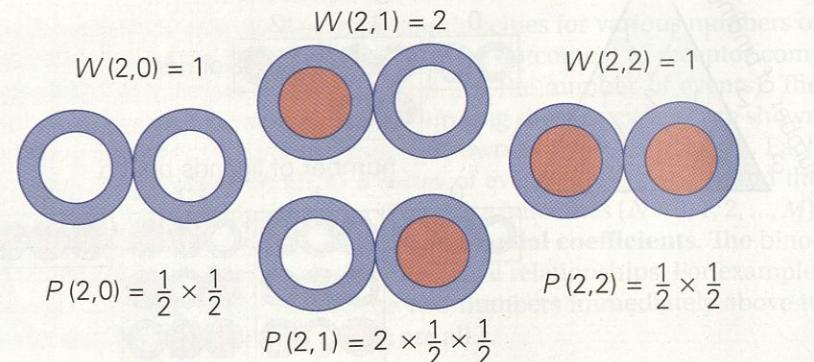
Recall from our discussion of coin flips that, if a coin is unbiased, then any particular series of outcomes has the same probability as any other series. If we ask, instead, about the probability of an *aggregate* outcome (for example, two heads in four coin tosses, without regard to order), then the probability is proportional to the multiplicity of the aggregate outcome. In the case of ligand binding, if we assume that the probability of a receptor molecule being bound to ligand is the same as it being empty, then the situation corresponds exactly to a series of unbiased coin tosses. What is the probability of finding a certain number of ligand molecules bound to a receptor complex? For example, if each complex is a tetramer with four binding sites, what is the probability of seeing receptor complexes with one, two, three, or four ligand molecules bound?

If we assume that the ligands bind independently of each other, then the probability  $P(M, N)$  of seeing  $N$  receptor molecules bound to ligand within a complex with  $M$  binding sites is given by:

$$P(M, N) = W(M, N) \times \left(\frac{1}{2}\right)^N \times \left(\frac{1}{2}\right)^{M-N} \quad (7.9)$$

Equation 7.9 is known as the **binomial distribution**. The first term on the right-hand side of Equation 7.9 is the multiplicity,  $W$ , which accounts for the number of equivalent rearrangements of the bound and empty receptors, as illustrated for a dimeric receptor in Figure 7.12. The second term accounts for the probability of seeing  $N$  ligands bound to the receptor complex, which is equivalent to seeing  $N$  heads in a series of coin tosses, which is  $(1/2)^N$ . If there are  $N$  ligands bound, then there are  $M - N$  empty receptors, and the probability for this occurring is  $(1/2)^{M-N}$ , which is given by the last term.

We do not expect, in general, that the probability of a binding site being empty will be the same as that of the site being occupied. The probabilities will depend on the interaction energy between the receptor and the ligand, with stronger interactions resulting in higher occupancy. The probability of binding will also depend on the concentration of the ligand—namely, the higher the concentration, the more likely that a receptor molecule will be bound rather than empty.



#### Binomial distribution

The probability distribution  $P(M, N)$  that determines the probability of obtaining  $N$  positive outcomes in a trial with  $M$  binary events is known as the binomial distribution. When the number of trials,  $M$ , is large, the binomial distribution is well approximated by a Gaussian distribution, as explained in Section 7.12.

**Figure 7.12** The probability of observing different numbers of ligands bound to a dimeric receptor. The diagram shows a dimeric receptor with no ligands bound (left) and with one or two ligands bound (middle and right). The probability of any receptor molecule being bound or empty is the same and is equal to 1/2. Receptor binding is assumed to be independent, and so the joint probability for a dimer is the product of the independent probabilities for each site. Each of the four configurations shown here has the same probability of occurring, but we are twice as likely to see receptor dimers with two ligands bound than with one or none because there are two ways in which one ligand can bind to a dimeric receptor.

A detailed analysis of binding equilibria is provided in Chapter 12. For now we shall assume that the probability of an individual receptor being bound is known, and simply focus on the distribution of bound and empty receptor molecules, given this probability.

If the probability of a binding site being occupied by ligand is  $p$ , then the probability that the site is unoccupied is  $(1 - p)$ . The probability  $P(M, N)$  of finding  $N$  binding sites occupied in a receptor complex with  $M$  binding sites is then given by:

$$P(M, N) = W(M, N) \times p^N \times (1-p)^{M-N} \quad (7.10)$$

Equation 7.10 is a modified form of the binomial distribution that applies to any situation where there is a bias in a binary outcome, as in coin flips involving a weighted coin.

To illustrate the application of Equation 7.10, we return to the experiment described in Section 7.5, in which virus particles are observed using a fluorescence microscope (see Figure 7.8). Each virus particle can bind up to six molecules of a green fluorescent dye. Individual virus particles are identified in the microscope, and the number of dye molecules bound to each one is counted.

The experimental conditions are such that the probability of a dye molecule being bound to any one binding site is known to be 0.7. We can calculate the expected distribution of dye molecules per virus particle by using the binomial distribution, as shown in Table 7.1. In this case, the probability,  $p$ , of a positive outcome (that is, binding site occupied) is 0.7 and that of a negative outcome,  $1 - p$ , is 0.3. The total number of binding sites,  $M$ , is 6, and the possible values of  $N$  (the number of ligands bound) range from 0 to 6.

The observed frequency distribution for the number of dye molecules per virus particle is shown in Figure 7.13, which compares these results with those calculated in Table 7.1. Note that the distribution peaks at  $N = 4$ —that is, the most probable outcome is that each virus particle is bound to four dye molecules. The agreement between the experimental data and the predictions from the binomial distribution is quite good for three to six ligand molecules bound. For one or two ligand molecules bound, the experimental results are smaller than the predicted values. This is because of the difficulty in identifying receptor complexes with small numbers of ligands bound, because such complexes exhibit lower levels of fluorescence and are therefore more difficult to count accurately.

**Table 7.1** Probability of seeing different numbers of ligand molecules bound to a receptor with six binding sites.

Number bound ( $N$ )	Multiplicity ( $W$ )	$p^N (1-p)^{M-N}$	$P = W \times p^N (1-p)^{M-N}$
0	1	$(0.3)^6 = 0.00073$	0.00073
1	6	$(0.7) \times (0.3)^5 = 0.00170$	0.0102
2	15	$(0.7)^2 \times (0.3)^4 = 0.00397$	0.0596
3	20	$(0.7)^3 \times (0.3)^3 = 0.00926$	0.1852
4	15	$(0.7)^4 \times (0.3)^2 = 0.02161$	0.3241
5	6	$(0.7)^5 \times (0.3)^1 = 0.05042$	0.3025
6	1	$(0.7)^6 = 0.11765$	0.1176

In this example, the probability,  $p$ , of a binding site being occupied is 0.7. The entries in the table show how the probability  $P(M, N)$  for different numbers of ligand bound is calculated using Equation 7.10. Here  $M$ , the total number of binding sites, is 6.  $N$ , the number of ligands bound, varies from 0 to 6.