2 Mechanisms of Drug Action

William W. Fleming

RECEPTORS

A fundamental concept of pharmacology is that to initiate an effect in a cell, most drugs combine with some molecular structure on the surface of or within the cell. This molecular structure is called a *receptor.* The combination of the drug and the receptor results in a molecular change in the receptor, such as an altered configuration or charge distribution, and thereby triggers a chain of events leading to a *response.* This concept applies not only to the action of drugs but also to the action of naturally occurring substances, such as hormones and neurotransmitters. Indeed, many drugs mimic the effects of hormones or transmitters because they combine with the same receptors as do these endogenous substances.

It is generally assumed that all receptors with which drugs combine are receptors for neurotransmitters, hormones, or other physiological substances. Thus, the discovery of a specific receptor for a group of drugs can lead to a search for previously unknown endogenous substances that combine with those same receptors. For example, evidence was found for the existence of endogenous peptides with morphinelike activity. A series of these peptides have since been identified and are collectively termed *endorphins* and *enkephalins* (see Chapter 26). It is now clear that drugs such as morphine merely mimic endorphins or enkephalins by combining with the same receptors.

DRUG RECEPTORS AND BIOLOGICAL RESPONSES

Although the term *receptor* is convenient, one should never lose sight of the fact that *receptors are in actuality molecular substances or macromolecules in tissues that combine chemically with the drug.* Since most drugs have a considerable degree of *selectivity* in their actions, it follows that the receptors with which they interact must be equally unique. Thus, *receptors will interact with only a limited number of structurally related or complementary compounds.*

The drug–receptor interaction can be better appreciated through a specific example. The end-plate region of a skeletal muscle fiber contains large numbers of receptors having a high affinity for the transmitter acetylcholine. Each of these receptors, known as nicotinic receptors, is an integral part of a channel in the postsynaptic membrane that controls the inward movement of sodium ions (see Chapter 28). At rest, the postsynaptic membrane is relatively impermeable to sodium. Stimulation of the nerve leading to the muscle results in the release of acetylcholine from the nerve fiber in the region of the end plate. The acetylcholine combines with the receptors and changes them so that channels are opened and sodium flows inward. The more acetylcholine the end-plate region contains, the more receptors are occupied and the more channels are open. When the number of open channels reaches a critical value, sodium enters rapidly enough to disturb the ionic balance of the membrane, resulting in local depolarization. The local depolarization (end-plate potential) triggers the activation of large numbers of voltage-dependent sodium channels, causing the conducted depolarization known as an action potential. The action potential leads to the release of calcium from intracellular binding sites. The calcium then interacts with the contractile proteins, resulting in shortening of the muscle cell. The sequence of events can be shown diagrammatically as follows:

$$\text{Ach} + \text{receptor} \rightarrow Na^{+}\ \text{influx} \rightarrow \text{action potential} \rightarrow \text{increased free } Ca^{++} \rightarrow \text{contraction}$$

where Ach = acetylcholine. The precise chain of events following drug–receptor interaction depends on the particular receptor and the particular type of cell. The important concept at this stage of the discussion is that *specific receptive substances serve as triggers of cellular reactions.*

If we consider the sequence of events by which acetylcholine brings about muscle contraction through receptors, we can easily appreciate that foreign chemicals (drugs) can be designed to interact with the same process. Thus, such a drug would *mimic* the actions of acetylcholine at the motor end plate; nicotine and carbamylcholine are two drugs that have such an effect. *Chemicals that interact with a receptor and thereby initiate a cellular reaction are termed* agonists. Thus, acetylcholine itself, as well as the drugs nicotine and carbamylcholine, are agonists for the receptors in the skeletal muscle end plate.

On the other hand, if a chemical is somewhat less similar to acetylcholine, it may interact with the receptor but be unable to induce the exact molecular change necessary to allow the inward movement of sodium. In this instance the chemical does not cause contraction, but because it occupies the receptor site, it prevents the interaction of acetylcholine with its receptor. Such a drug is termed an *antagonist.* An example of such a compound is *d*-tubocurarine, an antagonist of acetylcholine at the end-plate receptors. Since it competes with acetylcholine for its receptor and prevents acetylcholine from producing its characteristic effects, administration of *d*-tubocurarine results in muscle relaxation by interfering with acetylcholine's ability to induce and maintain the contractile state of the muscle cells.

Historically, receptors have been identified through recognition of the relative selectivity by which certain exogenously administered drugs, neurotransmitters, or hormones exert their pharmacological effects. By applying mathematical principles to *dose–response relationships,* it became possible to estimate dissociation constants for the interaction between specific receptors and individual agonists or antagonists. Subsequently, methods were developed to measure the specific binding of radioactively labeled drugs to receptor sites in tissues and thereby determine not only the *affinity* of a drug for its receptor, but also the *density of receptors* per cell.

In recent years much has been learned about the chemical structure of certain receptors. The nicotinic receptor on skeletal muscle, for example, is known to be composed of five subunits, each a glycoprotein weighing 40,000 to 65,000 daltons. These subunits are arranged as interacting helices that penetrate the cell membrane completely and surround a central pit that is a sodium ion channel. The binding sites for acetylcholine (see Chapter 12) and other agonists that mimic it are on one of the subunits that project extracellularly from the cell membrane. The binding of an agonist to these sites changes the conformation of the glycoprotein so that the side chains move away from the center of the channel, allowing sodium ions to enter the cell through the channel. The glycoproteins that make up the nicotinic receptor for acetylcholine serve as both the walls and the gate of the ion channel. This arrangement represents one of the simpler mechanisms by which a receptor may be coupled to a biological response.

SECOND-MESSENGER SYSTEMS

Many receptors are capable of initiating a chain of events involving second messengers. Key factors in many of these second-messenger systems are proteins termed G *proteins,* short for guanine nucleotide–binding proteins. G proteins have the capacity to bind guanosine triphosphate (GTP) and hydrolyze it to guanosine diphosphate (GDP).

G proteins couple the activation of several different receptors to the next step in a chain of events. In a number of instances, the next step involves the enzyme adenylyl cyclase. Many neurotransmitters, hormones, and drugs can either stimulate or inhibit adenylyl cyclase through their interaction with different receptors; these receptors are coupled to adenylate cyclase through either a stimulatory (G_s) or an inhibitory (G_i) G protein. During the coupling process, the binding and subsequent hydrolysis of GTP to GDP provides the energy needed to terminate the coupling process.

The activation of adenylyl cyclase enables it to catalyze the conversion of adenosine triphosphate (ATP) to 3′5′-cyclic adenosine monophosphate (cAMP), which in turn can activate a number of enzymes known as *kinases.* Each kinase phosphorylates a specific protein or proteins. Such phosphorylation reactions are known to be involved in the opening of some calcium channels as well as in the activation of other enzymes. In this system, the receptor is in the membrane with its binding site on the outer surface. The G protein is totally within the membrane while the adenylyl cyclase is within the membrane but projects into the interior of the cell. The cAMP is generated within the cell (see Figure 10.4).

Whether or not a particular agonist has any effect on a particular cell depends initially on the presence or absence of the appropriate receptor. However, the *nature* of the response depends on these factors:

- Which G protein couples with the receptor
- Which kinase is activated
- Which proteins are accessible for the kinase to phosphorylate

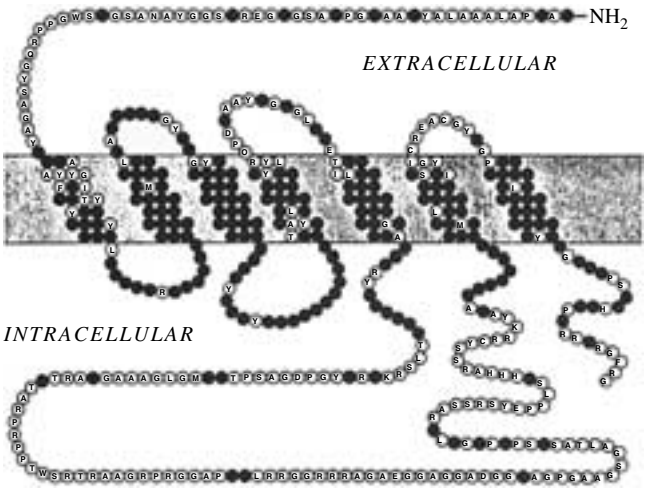

FIGURE 2.1
Primary structure of the human kidney α_2-adrenoceptor. The amino acid sequence is represented by the one-letter code. (Reprinted with permission from Regan JW et al. Cloning and expression of a human kidney cDNA. Proc Natl Acad Sci USA 85:6301, 1988.)

The variety of possible responses is further increased by the fact that receptor-coupled G proteins can either activate enzymes other than adenylate cyclase or can directly influence ion channel functions.

Many different receptor types are coupled to G proteins, including receptors for norepinephrine and epinephrine (α- and β-adrenoceptors), 5-hydroxytryptamine (serotonin or 5-HT receptors), and muscarinic acetylcholine receptors. Figure 2.1 presents the structure of one of these, the α_2-adrenoceptor from the human kidney. All members of this family of G protein–coupled receptors are characterized by having seven membrane-enclosed domains plus extracellular and intracellular loops. *The specific binding sites for agonists occur at the extracellular surface, while the interaction with G proteins occurs with the intracellular portions of the receptor.* The general term for any chain of events initiated by receptor activation is *signal transduction.*

THE CHEMISTRY OF DRUG–RECEPTOR BINDING

Biological receptors are capable of combining with drugs in a number of ways, and the forces that attract the drug to its receptor must be sufficiently strong and long-lasting to permit the initiation of the sequence of events that ends with the biological response. Those forces are *chemical bonds,* and a number of types of bonds participate in the formation of the initial drug–receptor complex.

The bond formed when two atoms share a pair of electrons is called a *covalent bond.* It possesses a bond energy of approximately 100 kcal/mole and therefore is strong and stable; that is, it is essentially irreversible at body temperature. Covalent bonds are responsible for the stability of most organic molecules and can be broken only if sufficient energy is added or if a catalytic agent that can facilitate bond disruption, such as an enzyme, is present. Since bonds of this type are so stable at physiological temperatures, the binding of a drug to a receptor through covalent bond formation would result in the formation of a long-lasting complex.

Although most drug–receptor interactions are readily reversible, some compounds, such as the anticancer nitrogen mustards (see Chapter 56) and other alkylating agents form relatively irreversible complexes. Covalent bond formation is a desirable feature of an antineoplastic or antibiotic drug, since long-lasting inhibition of cell replication is needed. However, covalent bond formation between environmental pollutants and cellular constituents may result in mutagenesis or carcinogenesis in normal, healthy cells.

The formation of an *ionic bond* results from the electrostatic attraction that occurs between oppositely charged ions. The strength of this bond is considerably less (5 kcal/mole) than that of the covalent bond and diminishes in proportion to the square of the distance between the ionic species. Most macromolecular receptors have a number of ionizable groups at physiological pH (e.g., carboxyl, hydroxyl, phosphoryl, amino) that are available for interaction with an ionizable drug.

The hydrogen atom, with its strongly electropositive nucleus and single electron, can be bound to one strongly electronegative atom and still accept an electron from another electronegative donor atom, such as nitrogen or oxygen, and thereby form a bridge (*hydrogen bond*) between these two donor atoms. The formation of several such bonds between two molecules (e.g., drug and receptor) can result in a relatively stable but reversible interaction. Such bonds serve to maintain the tertiary structure of proteins and nucleic acids and are thought to play a significant role in establishing the selectivity and specificity of drug–receptor interactions.

Van der Waals bonds are quite weak (0.5 kcal/mole) and become biologically important only when two atoms are brought into sufficiently close contact. Van der Waals forces play a significant part in determining drug–receptor specificity. Like the hydrogen bonds, several van der Waals bonds may be established between two molecules, especially if the drug molecule and a receptor have complementary three-dimensional conformations and thus fit closely together. The closer the drug comes to the receptor, the stronger the possible binding forces that can be established. Slight differences in three-dimensional shape among a group of agonists

and therefore slight differences in fit or strength of bonding forces that can be established between agonists and receptor form the basis for the *structure–activity relationships* among related agonists.

DYNAMICS OF DRUG–RECEPTOR BINDING

The drug molecule, following its administration and passage to the area immediately adjacent to the receptor surface (sometimes called the *biophase*), must bond with the receptor before it can initiate a response. Resisting this bond formation is a random thermal agitation that is inherent in every molecule and tends to keep the molecule in constant motion. Under normal circumstances, the electrostatic attraction of the ionic bond, which can be exerted over longer distances than can the attraction of either the hydrogen or van der Waals bond, is the first force that draws the ionized molecule toward the oppositely charged receptor surface. This is a reasonably strong bond and will lend some stability to the drug–receptor complex.

Generally, the ionic bond must be reinforced by a hydrogen or van der Waals bond or both before significant receptor activation can occur. This is true because unreinforced bonds are too easily and quickly broken by the energy of thermal agitation to permit sufficient time for adequate drug–receptor interaction to take place. The better the structural congruity (i.e., fit) between drug and its receptor, the more secondary (i.e., hydrogen and van der Waals) bonds can form.

Even if extensive binding has taken place, unless covalent bond formation has occurred, the drug–receptor complex can still dissociate. Once dissociation has occurred, drug action is terminated. For most drug–receptor interactions, there is a continual random association and dissociation. The frequencies of association and dissociation are a function of the affinity between the drug and the receptor, the density of receptors, and the concentration of drug in the biophase. *The magnitude of the response is generally considered to be a function of the concentration of the drug–receptor complexes formed at any moment in time.*

DOSE–RESPONSE RELATIONSHIP

To understand drug–receptor interactions, it is necessary to quantify the relationship between the drug and the biological effect it produces. Since the degree of effect produced by a drug is generally a function of the amount administered, we can express this relationship in terms of a *dose–response curve.* Because we cannot always quantify the concentration of drug in the biophase in the intact individual, it is customary to correlate effect with dose administered.

In general, biological responses to drugs are *graded;* that is, the response continuously increases (up to the maximal responding capacity of the given responding system) as the administered dose continuously increases. Expressed in receptor theory terminology, this means that *when a graded dose–response relationship exists, the response to the drug is directly related to the number of receptors with which the drug effectively interacts.* This is one of the tenets of pharmacology.

The principles derived from dose–response curves are the same in animals and humans. However, obtaining the data for complete dose–response curves in humans is generally difficult or dangerous. We shall therefore use animal data to illustrate these principles.

Quantal Relationships

In addition to the responsiveness of a given patient, one may be interested in the relationship between dose and some specified quantum of response among *all* individuals taking that drug. Such information is obtained by evaluating data obtained from a *quantal dose–response curve.*

Anticonvulsants can be suitably studied by use of quantal dose–response curves. For example, to assess the potential of new anticonvulsants to control epileptic seizures in humans, these drugs are initially tested for their ability to protect animals against experimentally induced seizures. In the presence of a given dose of the drug, the animal either has the seizure or does not; that is, it either is or is not protected. Thus, in the design of this experiment, the effect of the drug (protection) is *all or none.* This type of response, *in contrast to a graded response,* must be described in a noncontinuous manner.

The construction of a quantal dose–response curve requires that data be obtained from many individuals. Although any given patient (or animal) either will or will not respond to a given dose, a comparison of individuals within a population shows that members of that population are not identical in their ability to respond to a particular dose. This variability can be expressed as a type of dose–response curve, sometimes termed a *quantal dose–response curve,* in which the dose (plotted on the horizontal axis) is evaluated against the percentage of animals in the experimental population that is protected by each dose (vertical axis). Such a dose–response curve for the anticonvulsant phenobarbital is illustrated in Figure 2.2*A*. Five groups of 10 rats per group were used. The animals in any one group received a particular dose of phenobarbital of 2, 3, 5, 7, or 10 mg/kg body weight. The percentage of animals in each group protected against convulsions was plotted against the dose of phenobarbital. As Figure 2.2*A* shows, the lowest dose protected none of the 10 rats to which it was given, whereas 10mg/kg protected 10 of 10. With the intermediate doses, some rats were protected and some

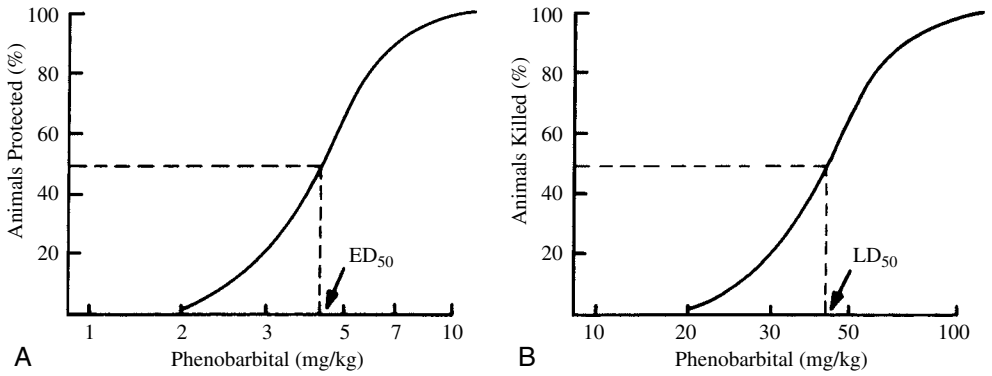

FIGURE 2.2

Quantal dose–response curves based on all-or-none responses. **A.** Relationship between the dose of phenobarbital and the protection of groups of rats against convulsions. **B.** Relationship between the dose of phenobarbital and the drug's lethal effects in groups of rats. ED_{50}, effective dose, 50%; LD_{50}, lethal dose, 50%.

were not; this indicates that the rats differ in their sensitivity to phenobarbital.

The quantal dose–response curve is actually a *cumulative plot* of the normal frequency distribution curve. The frequency distribution curve, in this case relating the minimum protective dose to the frequency with which it occurs in the population, generally is bell shaped. If one graphs the cumulative frequency versus dose, one obtains the sigmoid-shaped curve of Figure 2.2*A*. *The sigmoid shape is a characteristic of most dose–response curves when the dose is plotted on a geometric, or log, scale.*

Therapeutic Index

Effective Dose

The quantal dose–response curve represents estimates of the *frequency* with which each dose elicits the desired response in the population. In addition to this information, it also would be useful to have some way to express the average sensitivity of the entire population to phenobarbital. This is done through the calculation of an ED_{50} (effective dose, 50%; i.e., the dose that would protect 50% of the animals). This value can be obtained from the dose–response curve in Figure 2.2*A*, as shown by the broken lines. The ED_{50} for phenobarbital in this population is approximately 4mg/kg.

Lethal Dose

Another important characteristic of a drug's activity is its *toxic effect.* Obviously, the ultimate toxic effect is death. A curve similar to that already discussed can be constructed by plotting percent of animals killed by phenobarbital against dose (Fig. 2.2*B*). From this curve, one can calculate the LD_{50} (lethal dose, 50%). Since *the degree of safety associated with drug administration depends on an adequate separation between doses producing a therapeutic effect* (e.g., ED_{50}) *and doses producing toxic effects* (e.g., LD_{50}), one can use a comparison of these two doses to estimate drug safety. Thus, one estimate of a drug's margin of safety is the ratio LD_{50}/ED_{50}; this is the *therapeutic index.* The therapeutic index for phenobarbital used as an anticonvulsant is approximately 40/4, or 10.

As a general rule, a drug should have a high therapeutic index; however, some important therapeutic agents have low indices. For example, although the therapeutic index of the cardiac glycosides is only about 2 for the treatment and control of cardiac failure, these drugs are important for many cases of cardiac failure. Therefore, in spite of a low margin of safety, they are often used for this condition. The identification of a low margin of safety, however, dictates particular caution in its use; the appropriate dose for each individual must be determined separately.

It has been suggested that a more realistic estimate of drug safety would include a comparison of the lowest dose that produces toxicity (e.g., LD_1) and the highest dose that produces a maximal therapeutic response (e.g., ED_{99}). A ratio less than unity would indicate that a dose effective in 99% of the population will be lethal in more than 1% of the individuals taking that dose. Figure 2.2 indicates that Phenobarbital's ratio LD_1/ED_{99} is approximately 2.

Protective Index

The margin of safety is only *one* of several criteria to be used in determining a drug's clinical merit. Clearly, *the therapeutic index is a very rough measure of safety and generally represents only the starting point* in determining whether a drug is safe enough for human use. Usually, undesirable side effects occur in doses lower than the lethal doses. For example, phenobarbital induces drowsiness and an associated temporary neurological impairment. Since anticonvulsant drugs are intended to allow people with epilepsy to live normal

seizure-free lives, sedation is unacceptable. Thus, an important measure of safety for an anticonvulsant would be the ratio ED_{50} (neurological impairment)/ED_{50} (seizure protection). This ratio is called a *protective index*. The protective index for phenobarbital is approximately 3. It is easy to see that data derived from dose–response curves can be used in a variety of ways to compare the clinical usefulness of drugs. For instance, a drug with a protective index of 1 is useless as an anticonvulsant, since the dose that protects against convulsion causes an unacceptable degree of drowsiness. A drug with a protective index of 5 would be a more promising anticonvulsant than one with an index of 2.

Graded Responses

More common than the quantal dose–response relationship is the situation in which a single animal (or patient) gives graded responses to graded doses; that is, as the dose is increased, the response increases. *With graded responses, one can obtain a complete dose–response curve in a single animal.* A good example is the effect of the drug levarterenol (L-norepinephrine) on heart rate.

Results of experiments with levarterenol in guinea pigs are shown in Figure 2.3. The data are typical of what one might obtain from constructing complete dose–response curves in each of five different guinea pigs (*a–e*). In animal *a,* a small increase in heart rate occurs at a dose of 0.001 μg/ kg body weight. As the dose is increased, the response increases until at 1 μg/kg, the maximum increase of 80 beats per minute occurs. Further increases in dose do not produce greater responses. At the other extreme, in guinea pig *e,* doses below 0.3 μg/kg have no effect at all, and the maximum response occurs only at about 100 μg/kg.

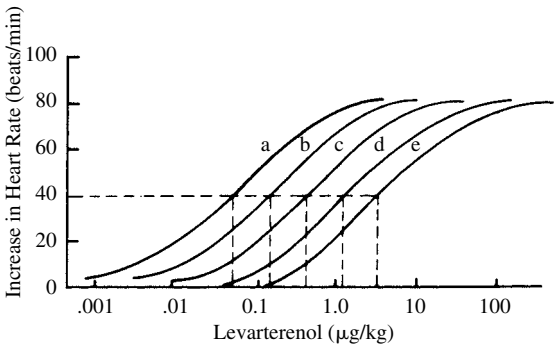

FIGURE 2.3
Dose-response curves illustrating the graded responses of five guinea pigs (a-e) to increasing doses of levarterenol. The responses are increases in heart rate above the rate measured before the administration of the drug. Broken lines indicate 50% of maximum response (horizontal) and individual ED_{50} values (vertical).

Since an entire dose–response relationship is determined from one animal, the curve cannot tell us about the degree of biological variation inherent in a population of such animals. Rather, variability is reflected by a *family* of dose–response curves, such as those given in Figure 2.3. The ED_{50} in this type of dose–response curve is the dose that produced 50% of the maximum response in one animal. In guinea pig *e,* the maximum response is an increase in heart rate of 80 beats per minute. Thus, 50% of the maximum is 40 beats per minute. From Figure 2.3, it can be seen that the dose causing this effect in guinea pig *e* is about 3 μg/kg. The average sensitivity of all of the animals to levarterenol can be estimated by combining the separate dose–response curves into a mean (average) dose–response curve and then calculating the mean ED_{50}. An estimate of the variation within the population can be indicated by calculating a statistical parameter, such as a confidence interval.

It is also possible to construct quantal dose–response curves for drugs that produce graded responses. To do so, one chooses a quantum of effect, for example, an increase in heart rate of 20 to 30 beats per minute above the control, or resting, rate. Doses of the drug are then plotted against the frequency with which each dose produces this amount of effect. The resulting graph has the same characteristics as the graph for the anticonvulsant activity of phenobarbital.

The doses in Figures 2.2 and 2.3 are on not an arithmetic but a logarithmic, or geometric, scale (i.e., the doses are displayed as multiples). This is more apparent in Figure 2.3 because of the greater range of doses. There are many reasons for the common practice of using geometric scales, some of which will become apparent later in this book. One important reason is that in most instances significant increases in response generally occur only when doses are increased in multiples. For example, in Figure 2.3, curve *e,* if one increased the dose from 10 to 11 or 12 μg/kg, the change in response would hardly be measurable. However, if one increased it 3 times or 10 times (i.e., to 30 or 100 μg/kg), one could easily discern increased responses.

The concept of the therapeutic index as a measure of the margin of safety has already been discussed. In the ratio LD_{50}/ED_{50}, the ED_{50} can be obtained from either quantal (Fig. 2.2*A*) or graded (Fig. 2.3) dose–response curves. In the latter case, it must be a *mean* ED_{50}, that is, the average ED_{50} obtained from several individuals.

Potency and Intrinsic Activity

Another drug characteristic that can be compared by use of ED_{50} values is *potency.* Figure 2.4 illustrates the mean dose–response curves of three hypothetical drugs that increase heart rate. Drugs *a* and *b* produce the same maximum response (an increase in heart rate of

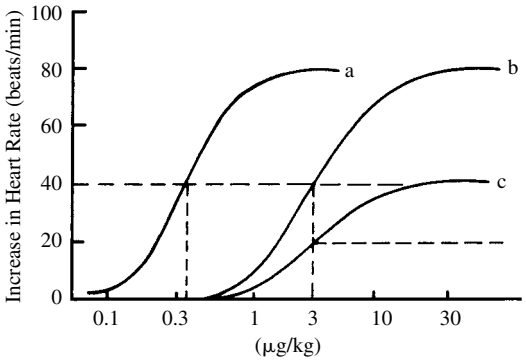

FIGURE 2.4
Idealized dose–response curves of three agonists (a, b, c) that increase heart rate but differ in potency, maximum effect, or both. Broken lines indicate 50% of maximum response (horizontal) and individual ED_{50} values (vertical).

80 beats per minute). However, the fact that the dose–response curve for drug *a* lies to the left of the curve for drug *b* indicates that drug *a* is more potent, that is, *less of drug* a *is needed to produce a given response.* The difference in potency is quantified by the ratio $ED_{50}b/ED_{50}a$: 3/0.3 = 10. Thus, drug *a* is 10 times as potent as drug *b.* In contrast, drug *c* has less maximum effect than either drug *a* or drug *b.* Drug *c* is said to have a lower *intrinsic activity* than the other two. Drugs *a* and *b* are full agonists with an intrinsic activity of 1; drug *c* is called a *partial agonist* and has an intrinsic activity of 0.5 because its maximum effect is half the maximum effect of *a* or *b.* The potency of drug *c,* however, is the same as that of drug *b,* because both drugs have the same ED_{50} (3 μg /kg). The ED_{50} is the dose producing a response that is one-half of the maximal response to that *same* drug.

It is important not to equate greater potency of a drug with therapeutic superiority, since one might simply increase the dose of a less potent drug and thereby obtain an identical therapeutic response. Such factors as the severity and frequency of undesirable effects associated with each drug and their cost to the patient are more relevant factors in the choice between two similar drugs.

EQUATIONS DERIVED FROM DRUG–RECEPTOR INTERACTIONS

It is important not to confuse the term *potency* with *affinity* or the term *intrinsic activity* with *efficacy.* The constants that relate an agonist A and its receptor R to the response may be represented as follows:

$$A + R \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AR \xrightarrow{k_3} \text{response}$$

Affinity is k_1/k_2, and efficacy is represented by k_3. Thus, affinity and efficacy represent kinetic constants that relate the drug, the receptor, and the response at the molecular level. Affinity is the measure of the net molecular attraction between a drug (or neurotransmitter or hormone) and its receptor. Efficacy is a measure of the efficiency of the drug–receptor complex in initiating the signal transduction process. In contrast, potency and intrinsic activity are simple measurements, respectively, of the relative positions of dose–response curves on their horizontal axes and of their relative maxima. Affinity is *one* of the determinants of potency; efficacy contributes *both* to potency and to the maximum effect of the agonist. Figure 2.4 shows that drug *c* has less efficacy (and less intrinsic activity) than either drug *a* or drug *b.* However, in contrast to intrinsic activity, no numerical value of efficacy can be calculated from the data presented. Unfortunately, the terms *potency* and *efficacy* are frequently used in a loose and misleading manner.

The mathematical relationship of response to efficacy and affinity is the following:

$$\frac{E_A}{E_m} = f\left\{\frac{e[A]}{K_A + [A]}\right\}$$

This equation states that the ratio of the response (E_A) to a given concentration of an agonist to the maximum response (E_m) of the test system, such as an isolated strip of muscle, is a function (f) of efficacy (e) times the concentration of the agonist ([A]) divided by the dissociation constant (K_A) plus the concentration of the agonist. K_A is the reciprocal of the affinity constant and, under equilibrium conditions,

$$K_A = \frac{[R][A]}{[RA]}$$

[R] is the concentration of free receptors and [RA] is the concentration of receptors bound to agonist. Although the details are beyond the scope of this textbook, it should be noted that by the use of combinations of agonists and antagonists, dose–response curves, and mathematical relationships, it is possible to estimate the dissociation constants of agonists and antagonists for a given receptor and to estimate the relative efficacy of two agonists acting on the same receptor.

DRUG ANTAGONISM

The terms *agonist* and *antagonist* have already been introduced. The several types of antagonism can be classified as follows:

1. Chemical antagonism
2. Functional antagonism

3. Competitive antagonism
   a. Equilibrium competitive
   b. Nonequilibrium competitive
4. Noncompetitive antagonism

Chemical Antagonism

Chemical antagonism involves a *direct chemical interaction between the agonist and antagonist* in such a way as to render the agonist pharmacologically inactive. A good example is the use of chelating agents to assist in the biological inactivation and removal from the body of toxic metals. *Chelation* involves a particular type of two-pronged attachment of the antagonist to a metal (the agonist). One chemical chelator, dimercaprol, is used in the treatment of toxicity from mercury, arsenic, and gold. After complexing with the dimercaprol, mercury is biologically inactive and the complex is excreted in the urine.

Functional Antagonism

Functional antagonism is a term used to represent the interaction of two agonists that act independently of each other but happen to cause opposite effects. Thus, indirectly, each tends to cancel out or reduce the effect of the other. A classic example is acetylcholine and epinephrine. These agonists have opposite effects on several body functions. Acetylcholine slows the heart, and epinephrine accelerates it. Acetylcholine stimulates intestinal movement, and epinephrine inhibits it. Acetylcholine constricts the pupil, and epinephrine dilates it; and so on.

Competitive Antagonism

Competitive antagonism is the most frequently encountered type of drug antagonism in clinical practice. *The antagonist combines with the same site on the receptor as does the agonist, but unlike the agonist, does not induce a response;* that is, the antagonist has little or no *efficacy.* The antagonist competes with the agonist for its binding site on the receptor. Competitive antagonists can fall into either of two subtypes, depending on the type of bond formed between the antagonist and the receptor. If the bond is a loose one, the antagonism is called *equilibrium competitive* or *reversibly competitive.* If the bond is covalent, however, the combination of the antagonist with the receptor is not readily reversible, and the antagonism is termed *nonequilibrium competitive or irreversibly competitive.*

If the antagonism is of the equilibrium type, the antagonism increases as the concentration of the antagonist increases. Conversely, the antagonism can be overcome (surmounted) if the concentration of the agonist in the *biophase* (the region of the receptors) is increased. This relationship can best be appreciated by examining dose–response curves, as in Figure 2.5. Curve *a* is obtained in the absence of the antagonist. Curve *b* is obtained in the presence of a modest amount of the antagonist. The curves are parallel, and the maximum effects are equal. The antagonist has shifted the dose–response curve of the agonist to the right. Any level of response is still possible, but greater amounts of the agonist are required. If the amount of the antagonist is increased, the dose–response curve is shifted farther to the right (curve *c),* still with no decrease in the maximum effect of the agonist. However, the amount of agonist required to achieve maximum response is greater with each increase in the amount of antagonist. Examples of equilibrium-competitive antagonists are atropine, *d*-tubocurarine phentolamine, and naloxone.

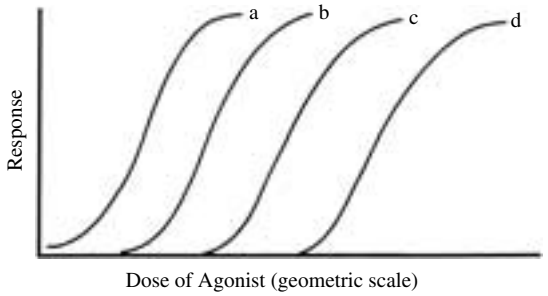

FIGURE 2.5
Idealized dose–response curves of an agonist in the absence (a) and the presence (b, c, d) of increasing doses of an equilibrium-competitive antagonist.

Of course, this continual shift of the curve to the right with no change in maximum as the dose of antagonist is increased assumes that very large amounts of the agonist can be achieved in the biophase. This is generally true when the agonist is a drug being added from outside the biological system. However, if the agonist is a naturally occurring substance released from within the biological system (e.g., a neurotransmitter), the supply of the agonist may be quite limited. In that case, increasing the amount of antagonist ultimately abolishes all response.

The effect of a nonequilibrium antagonist on the dose–response curve of an agonist is quite different from the effect of an equilibrium antagonist, as illustrated in Figure 2.6. As the dose of nonequilibrium antagonist is increased, the slope of the agonist curve and the maximum response achieved are progressively depressed. When the amount of antagonist is adequate (curve *d*), no amount of agonist can produce any response. The haloalkylamines, such as phenoxybenzamine, which form covalent bonds with receptors, are examples of nonequilibrium-competitive antagonists (see Chapter 11).

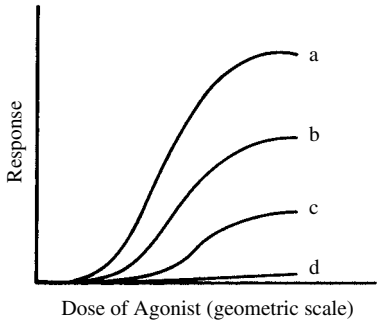

FIGURE 2.6
Idealized dose–response curves of an agonist in the absence (a) and the presence (b, c, d) of increasing doses of a non–equilibrium-competitive antagonist.

Noncompetitive Antagonism

In noncompetitive antagonism, the antagonist acts at a site beyond the receptor for the agonist. The difference between a competitive and a noncompetitive antagonist can be appreciated from the following scheme, in which two agonists, A and B, interact with totally different receptor systems, R_A and R_B, to initiate a chain of events leading to contraction of a vascular smooth muscle cell. X is a competitive antagonist, and Y is a noncompetitive antagonist.

$$\begin{array}{lcl} A + R_A & \searrow & \\ & & \text{Depolarization} \rightarrow \text{increased free calcium} \rightarrow \text{contraction} \\ B + R_B & \nearrow & \\ \uparrow & & \uparrow \\ X & & Y \end{array}$$

Antagonist X (competitive) has an affinity for R_B but not R_A. Thus, it specifically antagonizes agonist B. It does not antagonize agonist A. Antagonist Y acts on a receptor associated with the cellular translocation of calcium and inhibits the increase in intracellular free calcium. It will therefore antagonize the effects of both A and B, since they both ultimately depend on calcium movement to cause contraction.

The effect of a noncompetitive antagonist on the dose–response curve for an agonist would be the same as the effect of a non–equilibrium-competitive antagonist (Fig. 2.6). The practical difference between a noncompetitive antagonist and a nonequilibrium-competitive antagonist is *specificity.* The noncompetitive antagonist antagonizes agonists acting through more than one receptor system; the nonequilibrium-competitive antagonist antagonizes only agonists acting through one receptor system. The antihypertensive drug diazoxide is one of the few examples of therapeutically useful noncompetitive antagonists (see Chapter 20).

Study Questions

1. Receptors are macromolecules that
(A) Are designed to attract drugs
(B) Are resistant to antagonists
(C) Exist as targets for physiological neurotransmitters and hormones
(D) Are only on the outer surface of cells
(E) Are only inside of cells

2. All of the following are capable of initiating a signal transduction process EXCEPT
(A) Combination of an agonist with its receptor
(B) Combination of an antagonist with its receptor
(C) Combination of a neurotransmitter with its receptor
(D) Combination of a hormone with its receptor

3. Which of the following chemical bonds would create an irreversible combination of an antagonist with its receptor?
(A) Ionic bond
(B) Hydrogen bond
(C) Van der Waals bond
(D) Covalent bond

4. Potency is determined by
(A) Affinity alone
(B) Efficacy alone
(C) Affinity and efficacy
(D) Affinity and intrinsic activity
(E) Efficacy and intrinsic activity

ANSWERS

1. C. There are a large number of receptors in the body. Although many drugs are attracted to receptors, the receptors are not designed for that purpose. Antagonists also are attracted to receptors. Some receptors are on the cell surface, while others are found inside the cell.

2. B. An antagonist binds to a receptor and prevents the action of an agonist. Choice A is wrong because this combination does initiate a signal transduction process. C and D are incorrect because both neurotransmitters and hormones work through their appropriate receptor to initiate signal transduction.

3. **D.** A covalent bond is a strong and stable bond that is essentially irreversibly formed at normal body temperature. The other bonds are much weaker.
4. **C.** Potency is a useful measure of the comparison between two or more drugs. It does not equate to therapeutic superiority but rather is a measure of the size of the dose required to produce a particular level of response.

SUPPLEMENTAL READING

Brown BL and Dobson PRM (eds.). Cell Signaling: Biology and Medicine of Signal Transduction. New York: Raven, 1993.

Foreman JC and Johansen T. (eds). Textbook of Receptor Pharmacology. Boca Raton, FL: CRC, 1995.

Kenakin TP. Pharmacological Analysis of Drug–receptor Interaction. New York: Lippincott-Raven, 1993.

Kenakin TP, Bond RA, and Bonner TI. Definition of pharmacological receptors. Pharmacol Rev 1992; 44:351–362.

Ruffolo RR and Hollinger MA (eds.). G-Protein Coupled Transmembrane Signaling Mechanisms. Boca Raton, FL: CRC, 1995.

Ruffolo RR et al. Structure and function of α-adrenoceptors. Pharmacol Rev 1991;43:475–505.