mixing of labeled and unlabeled material which is required in the dilution methodology *is not essential for radioimmunoassay*. Indeed radioimmunoassay has been performed frequently in a way which is specifically designed to avoid uniform mixing i.e. unlabeled substance is allowed to occupy combining-sites on the antibody molecules prior to the addition of labeled substance and the mixtures are analyzed before complete reequilibration has taken place. (3) *In radioimmunoassay, the quantity of tracer need not be small compared to unlabeled substance present* (although the use of a large tracer may incur some loss of sensitivity or precision) nor, if the tracer is relatively large, is it even relevant to know the chemical amount of tracer. In contrast, a significantly large unknown quantity of tracer would invalidate application of the dilution principle.

Thus, it is evident that, except that marked molecules are employed in the application both of the isotope dilution and of the radioimmunoassay procedures, the two principles are essentially different.

It may be noted that, in addition to its use in radioimmunoassay, labeled insulin has been used as a tracer to measure the concentration of insulin antibodies and the kinetics of the insulin-antibody reaction by application of the dilution principle but that was done only after it was established that labeled and unlabeled insulin reacted identically with the antibodies that were employed (Berson and Yalow 1959).

A combination of the isotope dilution and radioimmunoassay procedures has been used to measure plasma insulin in the presence of endogenous insulin antibody (Berson and Yalow 1962). Labeled insulin was allowed to equilibrate between antibody-bound and free endogenous plasma insulin, and the equilibrium mixture was analyzed to determine the ratio between antibody-bound and free labeled hormone by the isotope dilution principle. The mixture was then subjected to ultracentrifugation and the supernatant fraction containing only free insulin (both endogenous and labeled forms) was then removed. The concentration of free unlabeled endogenous insulin was then independently measured by radioimmunoassay.

Appendix B. Mathematical considerations: Sensitivity and precision of radioimmunoassay

In the mathematical analysis which follows it is assumed that labeled and unlabeled hormone behave identically in the immune system and that the reaction follows the law of mass action (Berson and Yalow 1959; Yalow and Berson 1968, 1970a). However, it must be reemphasized that this condition is not necessary for the validity of the radioimmunoassay method, which requires only that the hormone to be measured and standard hormone behave identically

in the immune system. The analysis is carried out for only the simplest system involving the bimolecular reaction between a homogeneous univalent antigen and a simple homogeneous order of antibody combining-sites assuming that there are neither cooperative nor hindering effects between the two combining-sites of a single antibody molecule. These conditions do not necessarily obtain in any system, particularly since every antiserum contains a heterogeneous mixture of antibodies. Nevertheless, at high dilution of antiserum, a single order of combining-sites may be so predominantly responsible for the reaction observed that for practical purposes there is a close resemblance to a single order of sites.

The starting point is the primary reaction

$$\underset{F}{[Ag]} + [Ab] \underset{k'}{\overset{k}{\rightleftarrows}} \underset{B}{\overline{[AgAb]}}. \tag{1}$$

Here, $F \equiv [Ag]$, the molar concentration of uncomplexed antigen, $B \equiv \overline{[AgAb]}$, the molar concentration of complexed antigen or of antibody combining sites, and [Ab] is the molar concentration of uncomplexed antibody combining-sites. Then if K is the equilibrium constant and $[Ab^0]$,[1] is the molar concentration of total antibody combining-sites so that $[Ab^0] = [Ab] + \overline{[AgAb]}$, we have, from eq. (1),

$$K = \frac{k}{k'} = \frac{\overline{[AgAb]}}{[Ag]\,[Ab]} = \frac{B}{F([Ab^0] - [AgAb])} = \frac{B}{F([Ab^0] - B)}.$$

Thus,

$$B/F = K([Ab^0] - B). \tag{2}$$

In eq. (2), B/F is a linear function of B. A number of other linear relationships are also readily derived for reaction (1). In reality, however, all antisera are heterogeneous containing many different orders of antibody combining-sites. No simple linear relationship can be theoretically derived for such heterogeneous systems, although experimental data may, over certain ranges, closely approximate straight lines.

It is quite apparent from eq. (2) that the ability to detect a hormone concentration, [H], will depend on the relationship of K to [H]. If we denote the fraction of bound hormone by b so that $B = b[H]$, eq. (2) becomes

$$B/F = b/(1-b) = K([Ab^0] - b[H]). \tag{2a}$$

Now it is evident that B/F decreases as the concentration of bound hormone,

[1] Since antibody is generally divalent, we can also regard $[Ab^0]$ as twice the molar concentration of antibody molecules.

b[H], increases but if b[H] remains much smaller than [Ab^0] the reduction in B/F may be hardly detectable. It is then required that [Ab^0] be reduced (by dilution of the antiserum) at least to a value not much greater than [H], if [H] is to be readily detected. In the performance of an assay it is convenient to use an antiserum dilution that yields a B/F ratio of about 1.0 when tracer alone is present without added unlabeled hormone. If the concentration of tracer is negligible, then [H] → 0 and B → 0, but in any case we must have from eq. (2)

$$1 \leqq K[Ab^0].$$

Now, if we are to be able to dilute the antiserum sufficiently to meet the conditions, [Ab^0] = [H] and B/F = 1.0, the following must be satisfied

$$1 \leqq K[H],$$

$$K \geqq \frac{1}{[H]}.$$

Under the conditions chosen, at a hormone concentration, [H], it is easily shown that the B/F ratio would fall about 37½% if K were equal to the reciprocal of [H] but only about 5% if K were 10-fold lower. Thus, there is an inherent limitation in the sensitivity that can be achieved with any given antiserum, which is dependent on the K value characterizing the reaction of the predominating antibodies.

Conditions for optimal sensitivity

We define sensitivity as the slope of the dose-response relationship, b vs [H]. Other conditions remaining equal, therefore, optimal sensitivity will be obtained when db/d[H] is a maximum.

It follows from eq. (2a) that

$$\mathrm{db/d[H]} = -\frac{Kb(1-b)^2}{1+K[H](1-b)^2}. \tag{3}$$

We see from eq. (3) that for *any* given value of b, db/d[H] approaches a maximum as [H] → 0.[1] Thus, to find the conditions for maximal sensitivity we let [H] vanish and set the derivative of the right hand side of eq. (3) to zero. Thus,

$$\frac{d}{db}[-Kb(1-b)^2] = 0,$$

[1] This is not the same as saying that db/d[H] always approaches a maximum as [H] → 0. Indeed, it will be found that, under certain restraints (e.g. at high values of K[Ab^0]), db/d[H] is a maximum at some value [H] ≠ 0. The derivation in the text is directed at finding maximal db/d[H].

$$-K(1-b)^2+2Kb(1-b) = 0,$$
$$b = \tfrac{1}{3}, \qquad b=1.$$

There is a maximum at b = ⅓; a minimum at b = 1. Thus sensitivity is maximal at any b when [H] → 0 and is greatest at b = ⅓, when

$$\frac{db}{d[H]} = \frac{-4K}{27}. \tag{3a}$$

There has been some controversy concerning the proper definition of sensitivity (Ekins and Newman 1970). Sensitivity can be defined either as the minimal detectable concentration or the slope of the dose-response standard curve. Let us suppose that the slope of the dose-response curve is a 10% change in response per picogram and that the statistical error in the determination is 10%. Then the minimal detectable quantity would be about 1 picogram. If the slope is 10-fold greater, i.e., 100% change in response per picogram, then with the same 10% error in measurement, the minimal detectable quantity would be 0.1 pg. Thus, assuming the experimental error is unchanged, increasing the sharpness of the dose-response curve results in a reduction in minimal detectable quantity. Accordingly, we have defined sensitivity in terms of the slope of the dose-response curve.

Conditions for optimal precision

There are circumstances where a high degree of sensitivity is not required, as for example in the assay of glandular extracts. Say we wish to determine the purity of a certain hormone preparation with reference to some standard preparation defined to contain 1000 U/mg. We are then concerned with optimal precision in a sense that can be clarified by the following example. Suppose we prepare a stock solution of a hormone preparation at 1 mg/ml and assay the stock solution at dilutions of 1:10, $1:10^3$ and $1:10^5$. We might obtain the following values with reference to the standard preparation at the same dilutions: 100 ± 0.5 U/ml, 1 ± 0.1 U/ml, and 0.01 ± 0.005 U/ml, respectively, where the numbers following the ± signs represent some measure of the dispersion, say the standard error of the mean (SEM), the 95% or 99% fiducial limits or any other suitable statistical measure of confidence. It is evident that the *relative* dispersion is least in the assay carried out at 1:10 dilution. We might speak of the "relative uncertainties" of these assays as being 0.5%, 10% and 50% respectively, defining "uncertainty" according to the measure of dispersion employed. Clearly, for our purposes, the precision is greatest when the relative uncertainty is least. If then we define the absolute uncertainty of a hormone concentration [H] by Δ[H], the relative uncertainty is Δ[H]/[H]. Regardless of