XLVII. THE COMBINATIONS OF HAEMOGLOBIN WITH OXYGEN AND WITH CARBON MON-OXIDE. I.

By ARCHIBALD VIVIAN HILL (*Fellow of Trinity College, Cambridge*).

From the Physiological Laboratory, Cambridge.

(Received August 16th, 1913.)

In a previous paper Barcroft and Hill [1910] gave evidence to show that in dialysed solution haemoglobin has its smallest possible molecular weight, contains, namely, one atom of iron. In a later communication Hill [1910] suggested that the dissociation curves of oxyhaemoglobin in the presence of salts and carbon dioxide can be calculated from the hypothesis that these bodies tend to aggregate the large haemoglobin molecules into larger molecules, which then combine with oxygen according to the equation

If y be the percentage saturation of the haemoglobin with O_2 , and x the tension of the latter in the solution, it can be shown that this hypothesis leads to an equation for the dissociation curve of the type

$$y = 100 \frac{Kx^n}{1 + Kx^n} \dots\dots\dots (1),$$

where K is the equilibrium constant and n is a whole number > 1 .

This equation seems to suit all known dissociation curves of oxyhaemoglobin with a very high degree of accuracy, as numerous published and unpublished experiments of Barcroft [1913] and others (see e.g. Douglas, Haldane, J. S. and Haldane, J. B. S. [1912]) will show. In point of fact n does not turn out to be a whole number, but this is due simply to the fact that aggregation is not into one particular type of molecule, but rather into a whole series of different molecules: so that equation (1) is a rough mathematical expression for the sum of several similar quantities with n equal to 1, 2, 3, 4 and possibly higher integers.

The basis of the suggestion lies fundamentally in the idea that the

oxygenated haemoglobin molecules consist almost entirely of the fully saturated types

rather than of the partially saturated types Hb_2O_2 , Hb_3O_2 , Hb_4O_2 , etc. For the observed dissociation curves in the presence of salts exhibit a double curvature (i.e. are of an S-shape), differing therein from the dissociation curve of dialysed haemoglobin, which latter is a rectangular hyperbola always concave to the horizontal axis. The former dissociation curves start from the origin almost—if not quite—horizontally, bend upward at first with concavity facing the y -axis, pass through a point of inflexion, and then bend inward again with concavity towards the x -axis. This main property of the curves is simply represented physically by the idea that the aggregated haemoglobin molecule, taken e.g. to be Hb_3 , is oxidised into the form $\text{Hb}_3(\text{O}_2)_2$ by combination with two oxygen molecules simultaneously, and that the unsaturated molecule Hb_2O_2 either does not exist at all or exists in negligibly small quantities. For in the former case the dissociation curve is of the form $y = 100 \frac{Kx^2}{1 + Kx^2}$, the tangent to which at the origin is $y = 0$, i.e. the horizontal axis: and this approximates to the fact observed experimentally. While if the unsaturated Hb_2O_2 exists in appreciable quantities the dissociation curve involves an equation of the type $y = B \frac{Kx}{1 + Kx} + \text{etc.}$, where B is some constant less than 100, the tangent to which curve at the origin is $y = BKx$, a line going out at a slope as in the dissociation curve of dialysed haemoglobin, and not as observed when salts are present.

Now it seems at first sight almost unreasonable to suppose that Hb_3 would combine only with 2O_2 , i.e. with two molecules at once, and would not oxidise partially at first by combination with the single molecule O_2 : that in fact the oxidation goes on according to the equation

and not according to the equations

One could see no analogy for an equation of the former type when the latter scheme was so obviously possible. That the latter scheme does not immediately satisfy the facts is obvious from a short calculation, which however I believe will lead us to the clue to the whole matter. In equations (A) above, if K and K' be the equilibrium constants of the two reactions respectively, and

if the concentration of Hb_2O_2 be u , of $\text{Hb}_2(\text{O}_2)_2$ be v , of Hb_2 be w , and if x be the tension of oxygen and y the % saturation we find from the laws of mass-action

$$\frac{u}{wx} = K \text{ and } \frac{v}{ux} = K'.$$

Also since $(u + 2v)$ molecules of Hb are saturated with O_2 , and there are $(2u + 2v + 2w)$ molecules of Hb altogether,

$$y = 100 \frac{u + 2v}{2u + 2v + 2w},$$

which becomes

$$y = 100 \frac{Kx + 2K'Kx^2}{2Kx + 2K'Kx^2 + 2}.$$

when we put

$$u = Kwx \text{ and } v = K'xu = KK'x^2w.$$

The tangent at the origin to this curve is

$$2y = 100Kx,$$

a straight line which is not the horizontal axis; thus the scheme represented by equations (A) does not give a dissociation curve of the observed kind—coming in horizontally to the origin—*unless K is exceedingly small*.

If however K is exceedingly small the difficulty entirely disappears, and the equation becomes

$$2y = 100 \frac{KK'x^2}{KK'x^2 + 1},$$

which is of the type required to fit the dissociation curve, with tangent at the origin $y = 0$. (It should be noted that if K is small K' must be large, so that KK' is finite.)

Can we therefore assume that K is small, that in fact oxygen combines far more readily with Hb_2O_2 than it does with Hb_2 ? In other words, is it a justifiable assumption that the partially saturated body Hb_2O_2 is very unstable, and exists in almost inappreciable quantities: that in fact Hb_2O_2 once formed almost immediately combines with another O_2 to form the saturated compound $\text{Hb}_2(\text{O}_2)_2$? In support of this contention¹ one may argue that there is no transitional spectrum corresponding to unsaturated oxides, and therefore that, if there is any truth in the aggregation theory, the unsaturated oxides are present only in negligible quantities. The theory will obviously lead directly to a dissociation curve of the observed type, to which by adjusting the value of KK' it can be made to conform with some degree of accuracy. So far then as we are concerned with dissociation curves of oxyhaemoglobin, the equations to these can be deduced at once from the

¹ First made to Mr Barcroft by Mr W. H. Mills of the Chemical Laboratory, Cambridge.

aggregation hypothesis with the conception of gradual combination with oxygen according to the scheme

provided always that we may assume that O_2 combines much more readily with Hb_nO_2 to form $\text{Hb}_n(\text{O}_2)_2$, than with Hb_n to form Hb_nO_2 : *in other words that the partially saturated molecule is very unstable, is difficult to form and easy to combine further with oxygen.* For all such combinations lead to a dissociation curve with double curvature, which reaches the origin horizontally.

A plausible suggestion as to the reason why O_2 should combine much more readily with Hb_2O_2 than with Hb_2 may be put forward here. The haemoglobin molecules are probably aggregated by surface effects, by the neutralisation e.g. of the electric charges on their surfaces. It may be however that the free chemical "bonds" of Hb, with which reduced haemoglobin attaches itself to oxygen, are in addition combined with one another in the aggregated molecule, so that if reduced haemoglobin is $\text{Hb}=\text{}$ then the aggregated molecule is not merely Hb , Hb (two molecules just sticking together by surface forces) but rather $\text{Hb}=\text{Hb}$. In that case Hb_2O_2 , the half saturated molecule, has to be formed by the breaking loose of these bonds, and will be $=\text{Hb}$, $\text{Hb}=\text{O}_2$. This body one would naturally expect to be very unstable, as is $\text{Hb}=\text{}$ in the presence of oxygen. It would be very difficult for O_2 to form this compound, because it has to break down the $\text{Hb}=\text{Hb}$ molecule into $=\text{Hb}$, $\text{Hb}=\text{}$ in order to do it: and the compound $=\text{Hb}$, $\text{Hb}=\text{O}_2$ once formed, presenting as it does two unsaturated bonds, would immediately seize on another O_2 molecule becoming $\text{O}_2=\text{Hb}$, $\text{Hb}=\text{O}_2$. Whether this turns out to be the case or not, it presents a very clear physical conception as to why Hb_2O_2 should exist in only very minute concentrations.

So far then the aggregation hypothesis can be made to fit the observed facts very closely, and to have a reasonable physical explanation. Some doubt has been thrown on the particular form of it advanced in my earlier paper, by Douglas, Haldane and Haldane [1912]. These authors determined experimentally the CO-haemoglobin curve in the presence of a partial pressure of oxygen, and found it to be a rectangular hyperbola. According to the rough scheme, dealing with single molecules,

(and supposing the haemoglobin to be all combined with either CO or O_2) the equation to the CO-dissociation curve in the presence of a constant partial pressure of O_2 should be that of a rectangular hyperbola. But dealing with aggregated molecules, say Hb_2 , the formula

leads, by the laws of mass action, to an equation $\frac{y'x^2}{yx'^2} = K$, if x and x' be the

partial pressures of O_2 and CO , and y and y' be the % saturations with O_2 and CO . If all the haemoglobin is combined $y + y' = 100$, so that

$$\frac{y'}{100 - y'} = K \frac{x'^2}{x^2},$$

which represents a CO -dissociation curve of the type described above, containing a point of inflexion, and coming in horizontally to the origin. Thus this, and as a matter of fact all similar conceptions, of the balanced action between CO , O_2 and haemoglobin, led to the same type of equation, which gives a curve in no way like a rectangular hyperbola: so that the experiment of Douglas and Haldane led one to doubt the validity of the whole hypothesis of aggregation, in spite of the accuracy with which it fitted some of the facts. *Why is it that the CO - O_2 -haemoglobin dissociation curves, in the presence of a constant tension of O_2 , are unaffected by salts and CO_2 , when the O_2 -dissociation curves and the CO -dissociation curves are so very largely affected?* This very puzzling-question can, as a matter of fact, be quite simply answered by the aid of the conception sketched above.

Let us assume for simplicity that the haemoglobin exists as double Hb_2 molecules, and that according to the following scheme it combines gradually and by steps with CO and O_2 . The Hb_2 combines first with one O_2 or with one CO , the new molecule Hb_2O_2 or Hb_2CO combines further with O_2 , or CO , to form one of the three compounds

We then have the equations (in which the small letters represent concentrations and the k 's represent equilibrium constants)

From these, by the laws of mass action, we find

$$\begin{array}{ll} \text{I.} & \frac{y}{xy} = k_1, \quad \frac{u'}{x'y} = k_1', \\ \text{II.} & \frac{v}{ux} = k_2, \quad \frac{v'}{u'x'} = k_2', \\ \text{III.} & \frac{w}{ux'} = k_3, \quad \frac{w}{u'x} = k_3'. \end{array}$$

From equations I and II, multiplying up, we find

$$\text{IV.} \quad \frac{v}{x^2y} = k_1k_2, \quad \frac{v'}{x'^2y} = k_1'k_2',$$

and from equations I and III

$$\text{V.} \quad \frac{w}{xx'y} = k_1k_3, \quad \frac{w}{xx'y} = k_1'k_3',$$

so that $k_1k_3 = k_1'k_3' = k$ say.

Let ξ be the amount of haemoglobin per c.c. saturated with O_2 ; ξ' the amount saturated with CO. Then

$$\begin{aligned} \xi &= \frac{1}{2}(u + w) + v, \\ \xi' &= \frac{1}{2}(u' + w) + v'. \end{aligned}$$

Using the values of u , v and w given in equations I, IV and V we find

$$\begin{aligned} \xi &= \frac{1}{2}(k_1xy + kxx'y) + k_1k_2x^2y \\ \xi' &= \frac{1}{2}(k_1'x'y + kxx'y) + k_1'k_2'x'^2y \end{aligned} \quad \dots\dots\dots(\text{B}).$$

$$\therefore \frac{\xi'}{\xi} = \frac{x'}{x} \times \frac{\frac{1}{2}(k_1' + kx) + k_1'k_2'x'}{\frac{1}{2}(k_1 + kx') + k_1k_2x}.$$

Thus apparently the % saturations of the Hb with O_2 and CO are *not* in the direct ratio of the latter's tensions, as is necessary if the dissociation curve is to be a rectangular hyperbola. We come however to the hypothesis outlined above, viz. that the unsaturated Hb_2O_2 is very unstable and tends to break down either into Hb_2 or Hb_2O_4 . In this case we may assume that k_1 is very small, and k_2 very large, their product remaining finite. Similarly we have assumed that k_1' is very small and k_2' is very large. We therefore find, neglecting k_1' and k_1 ,

$$\frac{\xi'}{\xi} = \frac{x'}{x} \frac{\frac{1}{2}kx + k_1'k_2'x'}{\frac{1}{2}kx' + k_1k_2x} \quad \dots\dots\dots(\text{C}).$$

This, again, is not a direct proportion between ξ'/ξ and x'/x , as we should expect from Douglas and Haldane's observations [1912, see especially pp. 278, 290]. Looking however at the chemical equations (A) above, if we were to assume that CO has a very much higher affinity for the partially saturated $=Hb, Hb=O_2$ (i.e. Hb_2O_2) than has O_2 , i.e. that nearly all the Hb_2O_2 combines with CO rather than with O_2 ; and if we further assume that the partially saturated $=Hb, Hb=CO$ (i.e. Hb_2CO) has a much higher affinity for the CO than it has for the O_2 , i.e. that nearly all the Hb_2CO combines with CO rather than with O_2 , then the whole difficulty is immediately solved. For¹ in the first case we find that k_3x' is very much greater than k_2x , and

¹ It should be noted that the assumption is not simply that Hb_2O_2 has a greater affinity for a given pressure of CO than it has for the same pressure of O_2 : this, of course, we should expect. It is that the affinity of Hb_2O_2 for CO exceeds the affinity of Hb_2O_2 for O_2 more than the affinity of Hb for CO exceeds the affinity of Hb for O_2 : that the CO-affinity is even greater, relatively to the O_2 -affinity, in the aggregated molecule than it is in the simple.

in the second that $k_2'x'$ is much greater than $k_3'x$. These assumptions are in the highest degree reasonable, and making them, equation (C) above immediately gives the required result. Putting $k = k_1'k_3'$ in the numerator and $k = k_1k_3$ in the denominator (as shown above) we find

$$\frac{\xi'}{\xi} = \frac{x' k_1' \frac{1}{2} k_3'x + k_3'x'}{x k_1 \frac{1}{2} k_3x' + k_3x}$$

and assuming, as above, that $k_3'x$ can be neglected in comparison with $k_3'x'$, and that k_3x can be neglected in comparison with k_3x' , we find

$$\begin{aligned} \frac{\xi'}{\xi} &= \frac{x' k_1' k_3'x'}{x k_1 \frac{1}{2} k_3x'} \\ &= \frac{x' 2k_1'k_3'}{x k_1k_3}. \end{aligned}$$

That is to say, the CO- and O₂-saturation of the haemoglobin are proportional to their partial pressures, and if the haemoglobin be fully saturated with CO and O₂, i.e. putting $\xi + \xi'$ equal to a constant (say 100) we find, if $2k_1'k_3'/(k_1k_3) = K$,

$$\frac{\xi'}{100 - \xi} = \frac{x'}{x} \times K.$$

This is a rectangular hyperbola between ξ' , the % saturation with CO, and x' , the partial pressure of the CO. Thus on quite simple assumptions the theory is found to agree perfectly with the experimental observations of Douglas and Haldane.

Now Douglas and Haldane have brought to light some other experimental facts which seem at first sight to be difficult of explanation. They found [1912, p. 279] that the CO-O₂-haemoglobin dissociation curve, in the presence of a fixed tension of CO and a variable tension of O₂, was a rectangular hyperbola for large tensions of oxygen but bent back and fell again as the O₂-tension became very small and gradually vanished. Now, on our theory as advanced above, a very pretty explanation of this fact can be found, which may elicit some facts of further physiological interest. The assumption made above is practically that CO combines much more readily with Hb₂O₂ in the presence of O₂, than it does with Hb₂ in the presence of the same O₂. When the O₂-pressure is reduced beyond a certain limit very little Hb₂O₂ is formed, and therefore the CO is forced to combine with Hb₂ direct—which it does much less readily, and therefore the curve falls again. An analytical expression of this fact is given below. In order to find the dissociation curve in question we have to find the relation between ξ' , the CO %-saturation, and x , the O₂-tension. We know that the total amount of haemoglobin present can be represented by 100, so that, neglecting the small amounts of the unsaturated compounds present,

$$2\xi + 2\xi' + 2y = 100.$$

Putting into this equation the values of ξ and ξ' in terms of y given in equations (B) above, we find

$$2y = \frac{100}{1 + \frac{1}{2}(k_1x + k_1'x') + kxx' + k_1k_2x^2 + k_1'k_2'x'^2},$$

so that, from equation (B)

$$\xi' = \frac{100(k_1'k_2'x'^2 + \frac{1}{2}k_1'x' + \frac{1}{2}kxx')}{1 + \frac{1}{2}(k_1x + k_1'x') + kxx' + k_1k_2x^2 + k_1'k_2'x'^2}.$$

This relation between ξ' , the % saturation with CO, and x , the O_2 -tension, is too complicated for simple calculation and comparison with experimental results. It can however be shown that it may possess the general qualities of the experimental curve found by Douglas and Haldane [1912, p. 279]. It rises from the horizontal axis in much the same way as their curve, with diminishing O_2 -tension, and then reaches a maximum and falls again slightly as the O_2 -tension becomes very small and finally vanishes. This can be simply tested by finding the value of $\partial\xi'/\partial x$ at $x=0$, i.e. the slope of the curve at the point of zero O_2 -pressure. If $\partial\xi'/\partial x$ is positive, i.e. if the curve slopes upwards at first, ξ' first increases and then of course later decreases as x increases. If $\partial\xi'/\partial x$ is negative at $x=0$, then the curve slopes down from the beginning.

We should expect therefore, from Douglas and Haldane's results, that the value of $\partial\xi'/\partial x$ at $x=0$ should be >0 .

A simple calculation will show that

$$\frac{1}{100} \left(\frac{\partial\xi'}{\partial x} \right)_{x=0} = \frac{\frac{1}{2}kx' + \frac{1}{2}kk_1'x'^2 + kk_1'k_2'x'^3 - k_1k_1'k_2'x'^2 - \frac{1}{2}k_1k_1'x' - 2kk_1'k_2'x'^3 - kk_1'x'^2}{(1 + k_1'x' + 2k_1'k_2'x'^2)^2}.$$

Let us assume that x' is very large, i.e. that there is a large tension of CO. Then $(\partial\xi'/\partial x)_{x=0}$ is obviously negative, because of the preponderating negative value of the term $-2kk_1'k_2'x'^3$. Hence Douglas and Haldane's curve will *not* exhibit a fall of CO % saturation as the O_2 -tension is diminished to zero, if the CO-tension is high. The assistance provided to the combination of CO by the presence of a little unsaturated Hb_2O_2 is of no importance if the CO-tension is very high. Let us assume that x' is very small. Then the value of $\frac{1}{100} (\partial\xi'/\partial x)_{x=0}$ is

$$\frac{\frac{1}{2}kx' - \frac{1}{2}k_1k_1'x'}{(1 + k_1'x')^2},$$

neglecting squares or higher powers of x' . This quantity is therefore positive if $k > k_1k_1'$. Now $k = k_1k_3$. Hence we must have $k_3 > k_1'$. This we should expect to be the case. From the assumption already made, k_2' is very much greater than k_1' , CO combines much more readily with the partially saturated Hb_2CO than with the completely unsaturated Hb_2 : one

would expect it therefore to combine much more readily with the partially saturated Hb_2O_2 than with the completely unsaturated Hb_2 . In that case k_3 would be as much greater than k'_1 , as k'_2 is: and on the latter inequality the whole of the argument of this paper is based. Without therefore introducing any new hypothesis we have shown that, for a small constant CO -tension, the CO-O_2 -dissociation curve in the presence of an increasing tension of O_2 rises at first, reaches a maximum and falls slowly to zero, as in Douglas and Haldane's recorded observations.

In all the above considerations Hb_2 has been taken as the general type of the molecular cluster. Any other type of molecule, Hb_3 , Hb_4 , ..., can be considered of course in exactly the same way: the mathematical treatment would be more complicated, although exactly the same in type.

Douglas, Haldane and Haldane in their paper [1912, p. 296] put forward an ingenious theory to account for the several dissociation curves observed, both with O_2 and CO : and they gave an equation for a dissociation curve which fits the experimental points fairly closely, but not so closely as does the equation deduced from the hypothesis of this paper. In order to arrive at this equation they made various assumptions as to the exact values of the dissociation constants of the various reactions, for which, as they admit [1912, p. 301], there are no real *a priori* reasons: and there are other improbabilities in their assumptions, at least as great as in those made above in this paper. For example the idea that "the aggregated molecules do not give up or take up O_2 or CO without first splitting up into simple molecules" seems difficult to admit: and it is not easy to understand, if oxyhaemoglobin molecules are continually aggregating and reduced molecules are continually aggregating, how the oxyhaemoglobin and the reduced haemoglobin molecules do not aggregate together to form unsaturated clumps. It is true however that their theory does agree to a considerable extent with observed facts, even though they have made many rather difficult assumptions in the development of it: so that, even though the theory sketched in this paper seems to me to be more probable, I do not feel that one can as yet decide definitely between the two. It is noticeable however (as Barcroft shows) that the equation to the dissociation curve given here, and in the previous paper, does definitely suit the experimental facts better than the equation deduced from the hypothesis of Douglas, Haldane and Haldane. That a modification however of their assumptions would make their equation fit the experimental facts better than at present, seems to me very probable: so that a judgment between their theory and mine can scarcely be given on the facts hitherto discussed. At present however the theory and assumptions

given here seem to me to be simpler and easier, as well as to fit the experimental data better, than the hypothesis advanced by Haldane, Douglas and Haldane.

SUMMARY AND CONCLUSIONS.

The O_2 - and CO-dissociation curves of haemoglobin are known to differ, according as salts and CO_2 are present or are not. A theory to explain this has already been advanced, viz. that the simple molecules of haemoglobin are aggregated into molecular clusters. This theory explained many of the facts very exactly, but certain objections had been raised to it in particular by Douglas, Haldane and Haldane, especially in relation to the CO- O_2 -dissociation curves.

It is shown here that the theory is capable of including all the known facts in relation to CO- and O_2 -dissociation curves, if one is allowed to make certain simple assumptions as to the order of magnitude of the equilibrium constants in the several reactions involved. These are:

(i) that the half saturated molecules Hb_2O_2 and Hb_2CO are very unstable, and change at once into either Hb_2 or $Hb_2(O_2)_2$, $Hb_2(CO)_2$ or $Hb_2(CO)(O_2)$,

(ii) that the half saturated molecules Hb_2O_2 and Hb_2CO combine much more readily with CO than with O_2 .

The first of these assumptions can be explained as due to the fact that Hb_2 is $Hb = Hb$, while Hb_2O_2 is $= Hb, Hb = O_2$, with two unsaturated bonds, which tends to combine at once with more O_2 to form $O_2 = Hb, Hb = O_2$.

If these assumptions are justified, we may deduce that since CO combines much more readily with Hb_2O_2 than with Hb_2 , haemoglobin will take up more CO at a given tension, if a little O_2 is present than if O_2 is completely absent.

REFERENCES.

- Barcroft (1913), *Biochem. J.* **7**, 481.
 Barcroft and Hill (1910), *J. Physiol.* **39**, 411.
 Douglas, Haldane, J. S. and Haldane, J. B. S. (1912), *J. Physiol.* **42**, 296.
 Hill, A. V. (1910), *J. Physiol.* **40**, *Proceedings*, iv.