Anaesthesia, 2002, **57**, pages 348–356

REVIEW

Acid–base physiology: the 'traditional' and the 'modern' approaches

A. A. Sirker, A. Rhodes, R. M. Grounds and E. D. Bennett

Department of Intensive Care, St. Georges Hospital, London SW17 OQT, UK

Summary

The interpretation and understanding of acid–base dysfunction has recently been revisited. The 'traditional' approach developed from the pioneering work of Henderson and Hasselbalch and is still the most widely used in clinical practice. There are a number of problems identified with this approach, however. The 'modern' approach derives from Stewart's work in physical chemistry. In this review we describe the origins of the traditional approach and discusses related concepts. We then describe Stewart's approach, including how it is derived and how it may be used to classify acid–base derangements. The applications of Stewart's approach to clinical scenarios in intensive care is then discussed briefly before we examine some published clinical studies based on his work.

Keywords *Acid base balance*: Stewart; Henderson and Hasselbalch.

Correspondence to: Dr A. Rhodes
E-mail: arhodes@sghms.ac.uk
Accepted: 15 October 2001

The concentration of hydrogen ions within plasma and other aqueous solutions of the human body is maintained within very narrow limits. This control is needed because of the powerful effects these ions have on cellular function, particularly through alterations in hydrogen bonding and protein structure [1]. Enzymatically driven biological reactions and other intracellular processes are significantly affected by local hydrogen ion concentration $[H^+]$. Understanding how derangements of $[H^+]$ arise is therefore clinically important both for diagnosis and for guiding appropriate treatment.

The 'traditional' approach to interpreting acid–base disorders developed from the pioneering work of Henderson and Hasselbalch and is still the most widely used in clinical practice. One advantage of this approach is that it is relatively easy to understand and to apply in common clinical situations. However, problems have been identified with its interpretation. The use of HCO_3^- and $P_{a}CO_2$ to describe different types of acid–base disturbance has sometimes led to the supposition that these two variables are independently adjusted factors that ultimately determine pH. This implies that the dissociation equilibrium for carbonic acid is used as the control system for setting pH (thereby determining the position of the equilibrium for other buffer pairs in the body).

The 'modern' approach to acid–base disorders was initially proposed by Stewart in the early 1980s [2, 3]. Stewart used the fundamental principles of physical chemistry to elucidate factors that must determine $[H^+]$ in biological solutions. Using this mechanistic approach, he derived three independent variables that ought to be the sole final determinants of pH and explained how other factors, including the bicarbonate concentration $[HCO_3^-]$, would be dependent on these three independent variables.

This review describes the origins of the traditional approach and related concepts. We then describe Stewart's approach, including its deviation and use for classifying acid–base derangements. The application of Stewart's approach to clinical scenarios in intensive care is discussed briefly before we examine some published clinical studies based on his work.

The traditional approach

The Law of Mass Action states that the velocity of a chemical reaction is proportional to the active concentrations of the reactants. Many reactions in biological systems are reversible and so reach equilibrium. The equilibrium constant indicates to which side of the reaction the

equilibrium point lies. The value of the equilibrium constant depends on a number of factors including the temperature. In a reversible chemical reaction,

$$A + B \Leftrightarrow C + D$$

$$K = \frac{[C] \times [D]}{[A] \times [B]}$$

where K denotes the equilibrium constant. The '[ ]' terms here should strictly represent activities (i.e. effective concentrations), rather than concentrations. However, for ease of discussion, concentrations are used in this review.

Hence for an acid HE,

$$HE \Leftrightarrow H^+ + E^-$$

$$K_E = \frac{[H^+] \times [E^-]}{[HE]}$$

The K_E will be large for a strong acid (i.e. one which almost fully dissociates) and smaller for a weak acid (which dissociates to a lesser extent).

In 1909, Henderson applied the Law of Mass Action to the equilibrium reaction for carbonic acid. He then substituted a function of the derivable carbon dioxide concentration [CO_2] in place of the unmeasurable carbonic acid concentration [H_2CO_3] and then rearranged the equation to allow calculation of the pH [4]. His equation expressed in algebraic form is:

$$CO_2 + H_2O \Leftrightarrow H_2CO_3 \Leftrightarrow H^+ + HCO_3^-$$

$$K_1 = \frac{[H^+] \times [HCO_3^-]}{[H_2CO_3]}$$

therefore

$$[H^+] = \frac{K_1 \times [H_2CO_3]}{[HCO_3^-]}$$

$$K_2 = \frac{[CO_2] \times [H_2O]}{[H_2CO_3]}$$

therefore

$$[H_2CO_3] = \frac{[CO_2] \times [H_2O]}{K_2}$$

and if [H_2O] is large enough to be considered a constant

$$[H_2CO_3] = K_3 \times [CO_2]$$

Substituting for H_2CO_3

$$[H^+] = \frac{K_4 \times [CO_2]}{[HCO_3^-]}$$

where [CO_2] = dissolved molecular carbon dioxide which can be calculated from $P_{a}CO_2$ using Henry's law and K_1, K_2, K_3 and K_4 are numerically different constants.

In that same year, Sorensen introduced the pH scale [5]. Plasma pH measurement became available over the next few years [6]. It was already known that changes in plasma [HCO_3^-] reflected the accumulation of non-volatile acids such as lactic and keto-acids. With the availability of pH measurement (involving glass electrodes or indicator dyes), it became clear that changes in $P_{a}CO_2$ also directly affected pH. This led to the concepts of metabolic vs. respiratory causes of acid–base derangements. Hasselbalch re-arranged the Henderson's equation in logarithmic form in 1916 [7].

$$pH = pK + \log \frac{[HCO_3^-]}{[CO_2]}$$

and introduced pCO_2 into the equation in place of [CO_2].

$$pH = pK + \log \frac{[HCO_3^-]}{S_{CO_2} \times pCO_2}$$

where S_{CO2} is the solubility coefficient for carbon dioxide and pK is the negative logarithm of the equilibrium constant K_4 above.

From this followed the concept of carbon dioxide and HCO_3^- as variables that are adjusted as a control system to correct derangements of [H^+] and which, in doing so, set the position of other buffer-pair equilibria. The organs systems involved in this putative control system were as follows.

The respiratory system

The principal acidic product of cellular metabolism is carbon dioxide [8]. This travels down a concentration gradient from cells into the interstitial fluid and blood to be expired via the lungs. Chemoreceptors in the medulla of the brainstem and the carotid and aortic bodies respond to alteration in the [CO_2] of cerebrospinal fluid and the [H^+] and pCO_2 of arterial plasma, respectively [9]. Minute ventilation is thereby increased when plasma [CO_2] or [H^+] increase.

The kidneys

Non-volatile acids are produced to a far lesser extent than volatile (carbonic) acid. Sources of non-volatile acids include metabolism of methionine and cystine in dietary proteins and the incomplete metabolism of carbohydrates and fats. The free protons (H^+) are removed very rapidly from body fluid by reaction with buffers. The pK of the bicarbonate system is 6.1, whereas the normal pH of extracellular fluid is $\approx$7.4 and that of mean intracellular fluid $\approx$6.9 [8]. A buffer system functions most effectively

when its p*K* lies close to the pH of the compartment in which it is present. However, the bicarbonate system is viewed as the most important buffer system in the body because it is so plentiful [10].

The amount of available buffer would be depleted by the continual production of acid were it not for the ability of the kidney to reclaim and regenerate bicarbonate. Bicarbonate filtered by glomeruli, combines with a free proton secreted by the renal tubular cell to produce H_2CO_3. This is then converted by carbonic anhydrase to carbon dioxide which diffuses into the renal cell (along its concentration gradient because of increasing luminal pCO_2). Inside the cell the pathway is reversed, with HCO_3^- passing back into the blood and H^+ being secreted into the lumen to retrieve more bicarbonate. However, in addition, bicarbonate regeneration occurs because of carbon dioxide production within the tubular cell by cellular metabolism. The carbon dioxide is converted to H^+ and HCO_3^-. This HCO_3^- then diffuses into the blood, and the H^+ is passed into the tubular lumen where it combines with an anion B^- and is excreted in the urine. (B^- represents a nonbicarbonate anion; the two important examples in this context are phosphate and ammonia). This process produces new bicarbonate for buffering in the blood.

The role of ammonia in the kidney in the conventional approach is worth further comment [10]. Ammonia is produced by hepatic deamination of amino acids and then incorporated into urea and glutamine. Glutamine is taken up by renal tubular cells and hydrolysis releases NH_4^+, which is in equilibrium with NH_3. NH_3 diffuses into the tubular lumen where it combines with H^+ to produce NH_4^+ which then combines with Cl^- to be excreted in the urine. This allows excretion of H^+ in the urine but this process has required the production of another H^+ within the renal cell hence the overall benefit is unclear.

The gastrointestinal tract

The stomach's parietal cells secrete Cl^- and H^+ into the gastric lumen. This H^+ is derived, via carbonic anhydrase, from carbon dioxide produced within the cell. HCO_3^-, which is also produced in the process, passes into blood, supposedly the mechanism causing the recognised post-prandial 'alkaline tide' [11]. The pancreas' secretion of bicarbonate-rich fluid, with net movement of HCO_3^- from plasma into pancreatic cells and thereafter into the duodenal lumen, corrects this alkaline tide.

Red blood cells

Carbon dioxide in plasma diffuses into red blood cells along a concentration gradient and, in a reaction catalysed by carbonic anhydrase, is converted to bicarbonate which then diffuses back into plasma. The H^+ generated in the same process is buffered by combination with haemoglobin. In order to maintain electrical neutrality within all compartments, movement of a bicarbonate ion out of the cell is balanced by movement into the cell of a Cl^- ion from plasma (the chloride shift) [9].

Assessing the metabolic component

It is clear from the carbonic acid equilibrium reaction that a change in P_aCO_2 will cause a change in $[HCO_3^-]$. Thus value of $[HCO_3^-]$ *per se* cannot be used as an indicator of the metabolic contribution to any disorder unless this is taken into account. A number of methods of assessing the metabolic component have been devised. In 1948, Singer & Hastings [12] suggested the concept of 'buffer base' as the sum of all the plasma buffer anions, i.e. bicarbonate plus the nonvolatile weak acid buffers. In 1960, Astrup *et al.* [13] devised the 'standard bicarbonate', which was the value of $[HCO_3^-]$ once pCO_2 was standardised to 40 mmHg. Also in 1960, Siggaard-Andersen & Engel [14] proposed the term 'base excess'. This was defined as the concentration of titratable H^+ required to return the pH to 7.4, while the pCO_2 was maintained at 40 mmHg by equilibration. Data collection studies in Danish volunteers in the 1950s allowed the empirical development of nomograms for the value of *in vitro* base excess once the temperature, pH, pCO_2 and haemoglobin concentration (Hb) of a blood sample were known. These nomograms were used in computerised blood gas analysers in the 1960s. In 1977, Siggaard-Andersen published the 'Van Slyke' equation, which was derived from known physicochemical relationships and enabled calculation of the base excess from the variables pH, $[HCO_3^-]$ and the Hb [15]. This has subsequently been validated by others as showing good agreement with nomogram results when applied *in vitro* over a wide spectrum of pCO_2 values [16]. This equation is now widely used in blood gas analysers.

The anion gap

Another recognised limitation of the traditional approach is its inability to separate the various possible causes of metabolic acidoses. There is practical benefit in being able to divide these conditions into smaller groups in order to facilitate diagnosis and treatment. The 'anion gap' concept was developed for this purpose. It is derived from the principle of electroneutrality and is calculated by $([Na^+] + [K^+]) - ([Cl^-] + [HCO_3^-])$ where [ ] represents the millimolar concentration. Its value is usually positive and reflects anions which are not accounted for in the above expression, such as proteins (the most quantitatively important in healthy subjects), phosphate, sulphate [17], as well as other unidentified anions [18]. More

recent reviews of the 'anion gap' have cited many potential confounding factors that may limit its usefulness [19]. Particularly relevant to critically ill patients is the concentration of plasma protein, as reduced plasma albumin concentrations result in a reduction in the expected or 'normal' anion gap.

Stewart's physicochemical approach

Stewart's approach employs fundamental principles of physical chemistry on increasingly complex solutions in order to derive variables that actually determine $[H^+]$. The principles involved are:

1 Electroneutrality. In aqueous solutions in any compartment, the sum of all the positively charged ions must equal the sum of all the negatively charged ions.

2 The dissociation equilibria of all incompletely dissociated substances, as derived from the law of mass action, must be satisfied at all times.

3 Conservation of mass. The amount of a substance remains constant unless it is added, removed, generated or destroyed. The relevance is that the total concentration of an incompletely dissociated substance is the sum of concentrations of its dissociated and undissociated forms.

Stewart's analysis involves examining the various components which constitute human fluids and applying the above principles. The relevant components are:

1 water,
2 strong ion solutions in water,
3 weak acid (or 'buffer') solutions in water,
4 solutions containing carbon dioxide.

Stewart creates a model of human solutions by adding each of these constituents in turn.

Pure water

Water has a high dielectric constant, so causing substances which contain electrostatic bonds (including water itself) to dissociate into component ions. The degree of dissociation of water itself is slight but is of great importance in Stewart's analysis.

The equilibrium can be written as:

$$[H^+]\,[OH^-] = K_W\,[H_2O]$$

where K_W is the dissociation constant.

And because $[H_2O]$ is so much larger than $[H^+]$ and $[OH^-]$, it too can be considered constant. Hence

$$[H^+]\,[OH^-] = K'_W$$

where K'_W is a dissociation constant (different value to K_W).

The value of K'_W depends on temperature and, in solutions other than pure water, also on osmolarity and ionic strength of the solution. However, it is termed a constant as its value is not affected by any of the other terms in the equation.

Strong ion solutions in water

These are solutions that contain ions that are effectively fully dissociated in biological solutions, because their dissociation equilibria have pK values far removed from the local pH. They are always present at the concentration at which they were added and do not participate in any reactions within the solution. In mammalian extracellular solutions, the most abundant strong ions are Na^+ and Cl^-. Other strong ions include K^+, Mg^{2+}, Ca^{2+} and SO_4^{2-}.

In a solution which contains Na^+, Cl^- and K^+ ions in water, the principle of electroneutrality dictates that:

$$[Na^+] + [K^+] + [H^+] - [Cl^-] - [OH^-] = 0$$

Using the water dissociation equilibrium, which must still be satisfied,

$$[OH^-] = K'_W/[H^+]$$

which can then be substituted into the last equation. Hence

$$[Na^+] + [K^+] - [Cl^-] + [H^+] - (K'_W/[H^+]) = 0.$$

This equation can be solved for $[H^+]$ to give the quadratic equation:

$$[H^+] = \sqrt{K'_W + \frac{([Na^+]+[K^+]-[Cl^-])^2}{4}} - \frac{([Na^+]+[K^+]-[Cl^-])}{2}$$

The term strong ion difference (SID) can be used to conveniently denote $([Na^+] + [K^+] - [Cl^-])$. It can be seen that $[H^+]$ is a function of SID in the above expression. For a solution containing additional strong ions, the term SID could then be more generally defined as the sum of all the strong cation concentrations in the solution minus the sum of the strong anion concentrations in the solution. The ionic concentrations need to be expressed in milliequivalents per litre to account for the different ionic charges.

The important conceptual point is that $[H^+]$ and hence pH are dependent variables in this system and their values depend on the SID. The SID is an independent variable because it is imposed externally on the system and can be primarily altered by other factors.

Weak acid together with strong ions in solution with water

A weak acid HA will, by definition, dissociate only partially into its constituent ions, and will satisfy the equilibrium:

$$[H^+]\,[A^-] = K_A\,[HA]$$

where K_A is the dissociation constant for this acid.

Following the principle of conservation of mass, if HA and A^- take part in no other reactions in the solution, the sum of [HA] and [A^-] will remain constant:

$$[\mathrm{HA}] + [A^-] = [A_{\mathrm{TOT}}]$$

where A_{TOT} is the total weak acid.

The system must still satisfy the water dissociation equilibrium $[H^+][OH^-] = K'_W$ and the principle of electroneutrality $SID + [H^+] - [OH^-] - [A^-] = 0$.

The formula to calculate $[H^+]$ involves solving a third order (cubic) polynomial equation (shown below) that conforms to the simultaneous requirements to satisfy the four algebraic statements immediately above. It can be shown that the formula for $[H^+]$ is a function of SID, $[A_{TOT}]$, K_A and K'_W only [3].

$$[H^+]^3 + (K_A + \mathrm{SID}) \times [H^+]^2 + (K_A \times (\mathrm{SID} - A_{\mathrm{TOT}}) - K'_W) \times [H^+] - K_A \times K'_W = 0.$$

The important point is that in this solution SID and $[A_{TOT}]$ are independent variables whilst $[H^+]$, $[OH^-]$, $[A^-]$ and [HA] are dependent variables. The only way that $[H^+]$ or any of the other dependent variables can change is because of a change in either SID or $[A_{TOT}]$ or both.

Solutions containing carbon dioxide, in addition to weak acids and strong ions in water

In bodily solutions, carbon dioxide is present at an externally regulated partial pressure.

Four molecular species are added whenever carbon dioxide is added to an aqueous solution, namely dissolved molecular carbon dioxide [denoted $CO_2(d)$], carbonic acid (H_2CO_3), bicarbonate ions (HCO_3^-) and carbonate ions (CO_3^{2-}).

Henry's Law for gases allows dissolved molecular carbon dioxide to be related to pCO_2 via the formula:

$$[CO_2(d)] = S_{CO_2} \times pCO_2$$

where S_{CO_2} is the solubility coefficient for carbon dioxide.

The value of S_{CO_2} varies with temperature, other solute concentrations and ionic strength of the solution.

Because carbonic acid is in equilibrium with $CO_2(d)$ and H_2O, the dissociation equilibrium for carbonic acid can be written, with the $CO_2(d)$ term substituted, as:

$$[H_2CO_3] = L \times (S_{CO_2} \times pCO_2) \times [H_2O],$$

where L is another equilibrium constant.

As $[H_2O]$ is so large as to be virtually constant, this expression can be rewritten, with the constants combined into one:

$$[H_2CO_3] = L' \times pCO_2,$$

where L' is another constant (numerically different to L).

Because carbonic acid dissociates to H^+ and HCO_3^-, this dissociation can be combined with the previous equation to give:

$$[H^+]\,[HCO_3^-] = M \times pCO_2,$$

where M is a constant.

The final equilibrium to be considered is the relationship between $[HCO_3^-]$, $[H^+]$ and $[CO_3^{2-}]$.

This is defined by

$$[H^+]\,[CO_3^{2-}] = N \times [HCO_3^-],$$

where N is a constant.

In a solution containing carbon dioxide, weak acids and strong ions in water, the formula to calculate $[H^+]$ relies on solving the following simultaneous equations:

(1) Water dissociation equilibrium $[H^+]\,[OH^-] = K'_W$
(2) Weak acid dissociation equilibrium $[H^+]\,[A^-] = K_A\,[HA]$
(3) Conservation of mass for weak acid A $[HA] + [A^-] = [A_{TOT}]$
(4) Bicarbonate–carbon dioxide equilibrium $[H^+]\,[HCO_3^-] = M \times pCO_2$
(5) Bicarbonate-carbonate equilibrium $[H^+]\,[CO_3^{2-}] = N \times [HCO_3^-]$
(6) Electroneutrality $SID + [H^+] - [HCO_3^-] - [A^-] - [CO_3^{2-}] - [OH^-] = 0$

The expression describing these simultaneous requirements involves fourth-order polynomials and its solution reveals that $[H^+]$ is a function of SID, $[A_{TOT}]$, pCO_2 and the constants K'_W, K_A, M and N.

$$[H^+]^4 + (\mathrm{SID} + K_A) \times [H^+]^3 + (K_A \times (\mathrm{SID} - A_{\mathrm{TOT}}) - K'_W - M \times pCO_2) \times [H^+]^2 - (K_A \times (K'_W + M \times pCO_2) - N \times M \times pCO_2) \times [H^+] - K_A \times N \times M \times pCO_2 = 0.$$

In other words, in this solution which approaches human arterial plasma in its constitution, the *only* independent variables which determine pH are SID, $[A_{TOT}]$ and pCO_2. Any change in pH must be because of a change in one or more of these.

Consequently, [HA], $[A^-]$, $[HCO_3^-]$, $[CO_3^{2-}]$, $[OH^-]$ and $[H^+]$ are the dependent variables. If one of the independent variables changes in value, all of these dependent variables in that body compartment will consequently change. The fact that HCO_3^- and H^+ are both dependent variables emphasises the differences between the traditional and Stewart's approaches. If the Stewart approach is valid, it is misleading to think of bicarbonate as being specifically regulated to manipulate pH as bicarbonate cannot be individually or primarily altered.

Applying the model

There are certain points to note when applying this model to human plasma. First, P_aCO_2 is an independent variable in an 'open' system, i.e. where the total carbon dioxide is not fixed because it is in equilibrium with alveolar gas. However, this does not strictly apply to venous blood and fluid within the tissues, where the system is closed and the total carbon dioxide content rather than P_aCO_2 is the independent variable. Second, the independent variable $[A_{TOT}]$ is composed in plasma mainly of serum proteins and inorganic phosphate. Figge *et al.* [20, 21] have shown that albumin normally accounts for almost all the acid–base effects of plasma proteins, with globulins playing a negligible role. They calculate that in normal human serum at pH 7.4, with [total protein] = 70 $g.l^{-1}$ and [albumin] = 43 $g.l^{-1}$, the charge attributable to proteins is $\approx$ 12 $mEq.l^{-1}$ [20]. Figge's group developed a mathematical model of the contribution of albumin to $[A_{TOT}]$ by firstly assigning dissociation constants to all the ionisable groups on the albumin molecule (derived from the amino acid sequence of albumin). A computer program was then used to adjust the value of the dissociation constants so as to produce the best fit for the model to measured data from solutions containing albumin as the only protein. This optimised model was then compared with measured data using human serum and in fact matched well, suggesting that other serum proteins such as globulins usually make little contribution to the total anionic charge. However, the caveat remains that only serum with normal albumin/globulin ratios (1.3–2.0) was examined and hence the above conclusions may be incorrect in some situations, e.g. severe hypoalbuminaemia and hyperglobulinaemias.

Implications of Stewart's theory

One implication which is stressed in Stewart's original study is that the movement of hydrogen ions between solutions (by ion channels or pumps) will not actually affect local hydrogen ion concentration. If one considers a membrane separating two body fluid compartments, for each compartment the value of $[H^+]$ depends solely on the value of the independent variables in that compartment. Directly adding or removing H^+ to or from one of the compartments will not alter the value of any of the independent variables present and hence $[H^+]$ will be maintained at the same value as previously by a change in the dissociation of water to reverse any $[H^+]$ fluctuations. The water dissociation equilibrium is able to provide an essentially inexhaustible source or sink for H^+ ions.

A related question then is how the body normally maintains different pH values in adjacent fluid compartments separated by a membrane. Clearly, if Stewart's approach is correct, this must involve manipulating one of the three independent variables which determine $[H^+]$. Carbon dioxide diffuses freely across all membranes in the body, hence carbon dioxide cannot ordinarily be used to regulate pH. Proteins do not cross intact biological membranes. Phosphate is regulated by mechanisms in the gut and kidneys to maintain calcium homeostasis rather primarily than for acid–base regulation. Consequently, SID is considered the mechanism for generating pH differences in adjacent compartments.

The specific mechanisms that regulate the SID between compartments include the following.

The kidneys

The kidneys are the most important regulators of SID for acid–base purposes. The concentration of strong ions in plasma can be altered by adjusting absorption from glomerular filtrate or secretion into the tubular lumen from plasma. However, plasma $[Na^+]$ is used in the control of intravascular volume and plasma $[K^+]$ needs to be closely controlled to ensure appropriate cardiac and neuromuscular function. Hence, Cl^- appears to be the strong ion that the kidney uses to regulate acid–base status without interfering with other important homeostatic processes. For example, in the compensation for a respiratory acidosis, the excretion of H^+ in the urine is in itself not important. Instead, the removal of Cl^- in the urine (as opposed to its resorption back into plasma) will increase the value of the SID in plasma and thus help return plasma pH towards normal. The importance of ammonia when using Stewart's approach is that the weak ammonium cation allows the urinary excretion of the chloride anion without loss of any strong cations. In the correction of an alkalosis, resorption of additional Cl^- by renal tubular cells will reduce the plasma SID and therefore lower plasma pH.

The gastrointestinal tract

In the stomach, there is net movement of Cl^- ions from plasma into gastric parietal cells and then into the lumen of the stomach. This is increased after meals and the resulting reduction in the plasma SID (because no strong cations are involved in this process) can account for the postprandial 'alkaline tide'. Normally this is corrected by movement of chloride in the opposite direction in the duodenum. If there is loss of gastric fluid from the body, alkalosis develops because Cl^- is removed from plasma in far greater amounts than it is returned, thereby increasing plasma SID.

The pancreas secretes fluid which is low in Cl^- and has a high (very positive) SID. In order to generate this, the plasma returning from the pancreas has a lower SID than that arriving and this assists in the correction of the postprandial 'alkaline tide'.

Red blood cells

The previously described 'chloride shift' defends plasma pH against a decrease caused by increasing plasma carbon dioxide. This Cl^- movement between compartments raises plasma SID and hence will help return plasma pH towards normal.

Classification of acid–base disorders using Stewart theory

A natural division of acid–base disorders when using Stewart's approach is based on derangements of the independent variable(s). As with the conventional approach, respiratory acidoses or alkaloses are those in which the first independent variable affected is the pCO_2. A change in the plasma SID may then occur as a compensatory response.

'Metabolic' acidoses might be considered to arise from conditions that cause either a reduction in the plasma SID or increase in $[A_{TOT}]$. Conversely, 'metabolic' alkaloses could be defined as those which produce either a primary increase in the plasma SID or a decrease in $[A_{TOT}]$. However, whereas there appears to be complex regulation of SID for acid–base purposes, no such mechanisms are known to control $[A_{TOT}]$ for this purpose. For this reason, changes in $[A_{TOT}]$ are not classified as 'acid–base' disorders by some authors [22] despite their effect on pH.

Stewart's approach allows an understanding of the commonly seen metabolic acidosis following the administration of large volumes of normal saline. The concentration of chloride in plasma increases to a greater extent than that of sodium when normal saline is given because normal saline (unlike plasma) contains sodium and chloride in equal amounts. This leads to a reduction in the value of plasma SID and a consequent decrease in pH.

Two terms that have been used previously in the classification of acid–base derangements are 'contractional alkalosis' and 'dilutional acidosis'. These terms may lead to the erroneous idea that change in extracellular fluid volume alone might cause acid–base disturbances. Stewart's approach implies that changes in plasma or extracellular fluid volume alone will not change the value of any of the three independent variables and hence cannot affect pH. However, if the change in volume is accompanied by a change in the proportional water content of plasma, the SID will change. For example, if water is removed from plasma, the concentration of strong cations and strong anions is increased in equal proportion. This increases the SID by the same proportion and so causes an increase in pH. However, this effect may be complicated by compensatory mechanisms that regulate plasma volume.

The strong ion difference gap (SIG) as an indicator of unmeasured strong ions

The term SID has been used thus far to describe the difference between the concentrations of the strong cations and strong anions in a fluid compartment. In health, the strong ions present in plasma are Na^+, K^+, Mg^{2+}, Ca^{2+} (cations) and Cl^- (anion). Because these anions are all easily measured, the term 'apparent SID' (SID_{app}) is used:

$$SID_{app} = [Na^+] + [K^+] + [Mg^{2+}] + [Ca^{2+}] - [Cl^-]$$

Using the law of electroneutrality as applied to a solution whose constitution mimics that of plasma,

$$SID + [H^+] - [OH^-] - [HCO_3^-] - [CO_3^{2-}] - [A^-] = 0 \text{ (as stated earlier)}$$

The major weak anions (A^-) in plasma are provided by albumin and inorganic phosphate. The $[H^+]$, $[OH^-]$ and $[CO_3^{2-}]$ in the above expression are present in $nmol.l^{-1}$ and $\mu mol.l^{-1}$ concentrations rather than the $mmol.l^{-1}$ concentrations of SID and A^-. Hence an approximation for the electroneutrality equation can be produced by disregarding the charge contributions of the $[H^+]$, $[OH^-]$ and $[CO_3^{2-}]$ ions. Hence

$$SID - [HCO_3^-] - [Alb^{x-}] - [Pi^{y-}] \approx 0$$

where the last two terms refer to the milliequivalent concentration of the anionic contribution from serum albumin and inorganic phosphate, respectively. So, another expression for the SID in plasma, which does not make any assumption about which strong cations or anions constitute the SID, may be written

$$SID \approx [HCO_3^-] + [Alb^{x-}] + [Pi^{y-}].$$

This version of the SID is termed the 'effective SID' (SID_{eff}) and may encompass strong ions other than those covered in the term SID_{app}. The calculation of SID_{eff} involves

1 the measured concentration of HCO_3^-,

2 the calculated concentration of $[Pi^{y-}]$, obtained from an empirically derived formula involving the measurable independent variable $[Pi_{TOT}]$ and the known pH [20],

3 the calculated $[Alb^{x-}]$. This is obtained from Figge's (revised) model which indicates the charge contribution from albumin at a given pH. Their experimental data (to which they fit their model) demonstrate that there appears to be a linear relation between pH and the anionic charge contribution of protein in serum (see, Fig. 2, Ref. 21).

The quantitative difference between SID_{eff} and SID_{app} is termed the strong ion gap or SIG:

$$SIG = SID_{app} - SID_{eff}.$$

The SIG indicates the presence of unmeasured strong anions (if its value is positive) or strong cations (if its value is negative). The 'normal' value of the SIG is zero [23], implying that there are very few strong ions other than Na^+, K^+, Ca^{2+}, Mg^{2+} and Cl^- in the plasma of healthy subjects (though lactic acid levels are elevated in vigorously exercising subjects [24]). The anions which may result in an elevated SIG are those which will also cause an 'elevated anion gap' or 'normochloraemic' metabolic acidosis, e.g. lactic acid, keto-acids, sulfates, which accumulate in renal impairment and various exogenous acids, e.g. formic acid following methanol ingestion. The SIG is similar in concept to the anion gap, once the latter has been corrected for anionic contributions of albumin and phosphate. In critically ill patients, a number of other 'unidentified' strong anions can also often be detected by calculation of the SIG.

Clinical studies using the Stewart approach

McAuliffe and colleagues studied hypoproteinaemic patients on an American surgical intensive care unit [25]. They did not study patients with conditions associated with a high anion gap (e.g. renal failure, cardiovascular instability and hyperglycaemia). Nor did they study patients with deranged serum $[Na^+]$ in order to remove any possible effects of osmolarity on acid–base. All other measured strong cation and anion concentrations were within normal limits, i.e. these patients had a normal SID_{app}. Over the course of 6 months they identified eight patients who satisfied all their criteria. In each of these patients a significant 'metabolic' alkalosis was identified (by a positive base excess and an elevated standard bicarbonate). The calculated levels of 'anion gap' were normal after correction for hypoproteinaemia. The authors concluded that these patients had a primary hypoproteinaemic alkalosis, in keeping with the predicted results of Stewart's work. Kellum has pointed out that critically ill patients with hypoalbuminaemia often have a normal pH, reduced SID and a normal base excess [26]. In this situation, the reduced SID can be considered a physiological response to hypoalbuminaemia (in order to keep pH normal) rather than constituting part of a complex mixed metabolic disorder [27]. This point is also emphasised by proponents of the 'base excess' concept such as Siggard-Andersen & Fogh-Andersen [22].

Gilfix and colleagues examined 21 patients with a mixture of medical and surgical problems in an intensive care unit. They examined the use of three different measures in the assessment of unidentified ions in these patients [28]. They looked at the SIG, the anion gap and also what they define as a 'base excess gap'. The last of these takes into account the fact that any deviation in $[Na^+]$, $[Cl^-]$ or $[Alb^{x-}]$ from normal values will produce a base excess or deficit. Relatively simple formulae quantify the base excess/deficit produced by these factors and necessarily involve setting some arbitrary limits on reference values.

The difference between the base excess attributable to the above factors and the actual measured base excess is termed the 'base excess gap' and will reflect the contribution to base excess from 'other charged species' present. This study indicated that there is a tight correlation between the detection of unmeasured anions or cations using the SIG and the base excess gap approaches. The authors suggest that the base excess gap may therefore provide a good estimation of the SIG and that its calculation can be performed relatively easily at the bedside rather than requiring a computer.

The relationship between SIG and the traditional anion gap is less strong in the study by Gilfix *et al*. Kellum and colleagues examined data on ICU patients with sepsis and also those with severe liver disease [29]. They showed that the relationship between the SIG and the anion gap becomes much stronger if one corrects the anion gap for the charge contribution from protein and phosphate in the specfic patient. This is not altogether surprising, as the formulae for calculating SIG and corrected anion gap encompass the same molecular species although the numerical values are derived in different ways. Regardless of which method is used, it is clear that unknown and unidentified anions are frequently present in the plasma of critically ill patients and contribute significantly to the metabolic acidoses often seen in these patients.

Conclusion

This review describes the important features of both the conventional and more modern approaches to the analysis and interpretation of acid–base disturbances. The traditional approach evolved from the work of nineteenth and early twentieth century researchers and thus led to a pragmatic classification system based on those variables which were, at the time, easily measured. This approach remains by far the most widely used in everyday clinical practice and, as discussed above, with only minor adjustments can provide the same practical information as that obtained using Stewart's approach. The importance of Stewart's work is the insight it provides by identifying the factors that actually determine pH. It promotes review of the mechanisms by which organ systems such as the kidney might act to regulate the body's acid–base status and how, in illness, derangements of pH may arise. Stewart's work has its critics [22], who particularly dislike the way in which quantitative derangements in any strong ion or weak acid (including proteins) are potentially regarded as metabolic acidoses or

alkaloses. In addition, the cumbersome nature of some of the calculations needed in Stewart's work has limited its use in clinical settings. Nevertheless, Stewart's work has highlighted the importance of rigorously distinguishing between independent and dependent variables in pH setting in order to understand the primary derangement, its secondary consequences and thus allow appropriate direction of therapy.

References

1 Mathews CK, van Holde KE, Ahern KG. *Biochemistry*, 3rd edn. Oregon State University, Corvalis: Benjamin-Cummings, 1999.

2 Stewart PA. How to understand acid–base. *A Quantitative Primer for Biology and Medicine*. New York: Elsevier, 1981.

3 Stewart PA. Modern quantitative acid–base chemistry. *Canadian Journal of Physiology and Pharmacology* 1983; **61**: 1444–61.

4 Henderson LJ. Das Gleichgewicht zwischen Sauren und Basen im tierischen Organsimus. *Ergebnisse der Physioogie Biologischen Chemie und Experimentellen Pharmakologie* 1909; **8**: 254–325.

5 Sorensen SPL, Enzymstudien II. Mitteilung. Uber die Messung und die Bedeutung der Wasserstoffionenkonzentration bei enzymatischen Prozessen. *Biochemische Zeitschrift* 1909; **21**: 131–304.

6 Hasselbalch KA, Lundsgard C. Elektrometrische Reactionsbestimmung des Blutes bei Korpertemperatur. *Biochemische Zeitschrift* 1912; **38**: 77–91.

7 Hasselbalch KA. Die Berechnung der Wasserstoffzahl des Blutes aus der freien und gebundenen Kohlensaure desselben und die Sauerstoffbindung des Blutes als Funktion der Wasserstoffzahl. *Biochemische Zeitschrift* 1916; **78**: 112–44.

8 Levinsky NG. Acidosis and alkalosis. In: Isselbacher KJ, Braunwald E, Wilson JD, Martin JB, Fauci AS, Kasper DL, eds. *Harrison's Principles of Internal Medicine*, 13th edn. New York: McGraw-Hill, 1994; 253–62.

9 Widdicombe J, Davies A. *Respiratory Physiology*. London: Edward Arnold, 1991.

10 Mayne PD. *Clinical Chemistry in Diagnosis and Treatment*. London: Edward Arnold, 1994.

11 Moore EW. The alkaline tide. *Gastroenterology* 1967; **52**: 1052–4.

12 Singer RB, Hastings AB. An improved clinical method for the estimation of disturbances of the acid–base balance of human blood. *Medicine (Baltimore)* 1948; **27**: 223–42.

13 Astrup P, Jorgensen K, Siggard-Andersen O, *et al*. Acid–base metabolism: new approach. *Lancet* 1960; **i**: 1035–9.

14 Siggard-Andersen O, Engel K. A micro method for determination of pH, carbon dioxide tension, base excess and standard bicarbonate in capillary blood. *Scandinavian Journal of Clinical and Laboratory Investigation* 1960; **12**: 172–6.

15 Siggard-Andersen O. The Van Slyke equation. *Scandinavian Journal of Clinical and Laboratory Investigation* 1977; **37**: 15–20.

16 Morgan TJ, Clark C, Endre ZH. Accuracy of base excess – an *in vitro* evaluation of the Van Slyke equation. *Critical Care Medicine* 2000; **28**: 2932–6.

17 Emmet M, Narins RG. Clinical use of the anion gap. *Medicine* 1977; **56**: 38–54.

18 Gabow PA, Kaehny WD, Fennessey PV, *et al*. Diagnostic importance of the anion gap. *New England Journal of Medicine* 1980; **303**: 854–8.

19 Salem MM, Mujais SK. Gaps in the anion gap. *Archives of Internal Medicine* 1992; **152**: 1625–9.

20 Figge J, Rossing TH, Fencl V. The role of serum proteins in acid–base equilibria. *Journal of Laboratory and Clinical Medicine* 1991; **117**: 453–67.

21 Figge J, Mydosh T, Fencl V. Serum proteins and acid–base equilibria: a follow-up. *Journal of Laboratory and Clinical Medicine* 1992; **120**: 713–19.

22 Siggard-Andersen O, Fogh-Andersen N. Base excess or buffer base (strong ion difference) as a measure of a non-respiratory acid–base disturbance. *Acta Anaesthesiologica Scandinavica* 1995; **39**: 123–8.

23 Kellum JA. Determinants of blood pH in health and disease. *Critical Care* 2000; **4**: 6–14.

24 Lindinger MI, Heigenhauser GJF, McKelvie RS, *et al*. Blood ion regulation during repeated maximal exercise and recovery in humans. *American Journal of Physiology* 1992; **262**: R126–36.

25 McAuliffe JJ, Lind LJ, Leith DE, Fencl V. Hypoproteinaemic alkalosis. *American Journal of Medicine* 1986; **81**: 86–9.

26 Kellum JA. Recent advances in acid–base physiology applied to critical care. In: Vincent JL, ed. *Yearbook of Intensive Care and Emergency Medicine*. Heidelberg: Springer-Verlag, 1998; 579–87.

27 Wilkes P. Hypoproteinaemia, SID and acid–base status in critically ill patients. *Journal of Applied Physiology* 1998; **84**: 1740–8.

28 Gilfix BM, Bique M, Magder S. A physical chemical approach to the analysis of acid–base balance in the clinical setting. *Journal of Critical Care* 1993; **8**: 187–97.

29 Kellum JA, Kramer DJ, Pinsky MR. Strong ion gap: a methodology for exploring unexplained anions. *Journal of Critical Care* 1995; **10**: 51–5.